



HAL
open science

De quoi la ‘justice prédictive’ est-elle le nom ? Algorithmes, décision et jugement

Laurence Dumoulin

► **To cite this version:**

Laurence Dumoulin. De quoi la ‘justice prédictive’ est-elle le nom ? Algorithmes, décision et jugement. Science politique. Université Paris-Saclay, 2022. tel-04227420

HAL Id: tel-04227420

<https://shs.hal.science/tel-04227420v1>

Submitted on 3 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

De quoi la 'justice prédictive'
est-elle le nom ?
Algorithmes, décision et jugement

Habilitation à diriger des recherches
de l'Université Paris-Saclay

présentée et soutenue à Guyancourt,
le 12 décembre 2022, par

Laurence DUMOULIN

Habilitation à diriger des recherches

Composition du jury

Laurie BOUSSAGUET

Professeure des universités en science
politique, Université de Rouen,
European University Institute, Florence

Rapportrice

Nicolas DODIER

Directeur de recherche INSERM en
sociologie

Directeur d'études EHESS

CEMS-EHESS Paris

Rapporteur

Jérôme PELISSE

Professeur des universités en
sociologie,

CSO-Sciences Po Paris

Rapporteur

Philippe BEZES

Directeur de recherche CNRS en science
politique

CEE- Sciences Po Paris

Examineur

Jacques de MAILLARD

Professeur des universités en science
politique

CESDIP-UVSQ-Univ. Paris Saclay

Garant

Isabelle SAYN

Directrice de recherche CNRS en droit
Centre Max Weber, Université de Lyon

Examinatrice

Remerciements

Une partie des recherches présentée ici a été soutenue par le CDP Data-Institute et la chaire « Sociétés algorithmiques » de l'Institut d'intelligence artificielle de Grenoble-UGA (MIAI), coordonnée par Gilles Bastin, que je tiens à remercier. C'est dans ce cadre que Florent Frisque, alors stagiaire de M1 puis de M2 à Pacte et maintenant doctorant sous ma direction, a réalisé l'important travail de constitution et de première analyse de la base de données de *tweets* utilisée dans ce travail. Sa contribution a été décisive et je ne saurais assez le remercier. Si ce manuscrit a été conçu comme le bilan d'étape d'un travail de recherche en cours, commencé il y a plusieurs mois, personnel, inédit et encore partiel, il dessine un programme à poursuivre dans les années qui viennent — notamment à travers le projet d'ANR Just-IA, qui a commencé en mars 2021, que je coréalise avec Christian Licoppe. Je me réjouis que cette collaboration, commencée il y a fort longtemps sur un autre projet, ait trouvé des prolongements inattendus et s'engage vers de nouveaux horizons.

Pour personnel qu'il soit, ce travail s'inscrit, à bien des égards, dans des compagnonnages, coopérations, collectifs à géométrie variable, dont il a bénéficié. Je reste bien évidemment seule responsable des erreurs et approximations qu'il pourrait comporter.

La maturation d'un manuscrit d'HDR est un processus lent et complexe, nourri d'échanges d'autant plus fructueux qu'ils se nouent dans un climat d'exigence intellectuelle et de bienveillance. Je remercie Paul S. et Laurence L. qui n'y sont pas pour rien. Je remercie aussi infiniment les relecteurs et relectrices du projet d'HDR, plusieurs fois remanié avant de prendre enfin vie, de chapitres ou de passages de ce manuscrit : Anthony Amicelle, Jacques Commaille, Martine Kaluszynski, Jean-Noël Retière, Vincent Rivollier, Cécile Vigour ; et en particulier Claire de Galembert, avec qui le compagnonnage, ces dernières années, a été si stimulant et joyeux, même à distance. Jacques de Maillard a accepté la tâche ingrate d'être garant de cette HDR, c'est-à-dire de contribuer à ce qu'elle parvienne à bon port... dans des délais raisonnables. La mission n'était pas impossible, mais il fallait bien toutes les qualités intellectuelles et humaines qui sont les siennes, l'humour et le tact dont il est coutumier, pour la mener à bien. Ma gratitude envers lui est aussi sincère qu'immense.

Les membres du jury de soutenance m'ont fait le plaisir d'accepter de discuter ces premiers résultats, à une période, la fin de l'année civile, où les soutenances ne manquent pas et où tout le monde court après le temps. Mes remerciements n'en sont que plus grands, à la hauteur aussi de ce que la discussion avec elles et eux m'apportera, j'en suis sûre.

Véronique Curély a assuré la relecture et la mise en forme du tapuscrit avec beaucoup de professionnalisme. Elle a, de plus, fait preuve d'une souplesse qui n'a d'égal que l'élasticité du temps et l'étirement des échéances qui ont marqué ma piètre gestion du calendrier initialement posé. Je lui en suis sincèrement reconnaissante.

J'ai eu la chance jusqu'ici de prendre part à des collectifs où il fait bon travailler, y compris dans le cadre de la direction de Pacte où j'ai trouvé beaucoup de solidarité et de compréhension à l'égard de mon agenda personnel d'HDR. Je ne peux qu'éprouver une bonne dose de gratitude envers les membres du Directoire. Mes remerciements très directs vont aussi à Céline Gaufreteau, documentaliste à Pacte, qui m'a aidée à (re)constituer le dossier de mes publications. Sait-on que les chercheurs et enseignants-chercheurs ne reçoivent pas toujours une copie de l'article ou du chapitre qu'ils ont écrit et qu'il leur faut parfois remuer ciel et terre pour l'obtenir ?

Enfin, parce que les chercheuses et chercheurs ne sont pas de purs esprits, mais développent leurs activités dans un quotidien pluridimensionnel, où la séparation entre le professionnel et le privé souffre souvent d'importants défauts d'étanchéité qu'aucun maçon, plombier ou autre corps de métier ne saurait résoudre — aussi talentueux soit-il —, cette HDR n'aurait pu aboutir sans un contexte familial et amical propice. Merci à mes proches pour leur patience et leur compréhension face à mes absences et défaillances... Je remercie en particulier Emmanuel, Gabriel et Adèle d'avoir accepté les contraintes et vicissitudes liées à ce projet d'HDR, parmi lesquelles soirées, week-ends, vacances, et un été entier, plus studieux que récréatifs...

Sommaire

Table des sigles.....	5
Introduction générale	7
Chapitre 1 : Les algorithmes dans la justice, de l'effervescence des expériences à l'analyse de leurs effets.....	37
Chapitre 2 : La 'justice prédictive', controverse autour d'une notion floue.....	83
Chapitre 3 : La traduction de la politique d' <i>open data</i> dans le secteur de la justice	143
Chapitre 4 : Un nouveau répertoire de réforme en action : la fabrique d'une expérimentation de 'justice prédictive' auprès de magistrats	193
Point(s) de suspension	233
Du jugement « produit fini » au jugement comme action équipée	233
Conclusion générale	269
Bibliographie	273
Table des graphiques, encadrés et tableaux.....	303
Table des matières	307

Table des sigles

AAI : Autorité administrative indépendante

ADIJ : Association pour le développement de l'informatique juridique

AFC : Analyse factorielle des correspondances

AFJE : Association française des juristes d'entreprise

AFSP : Association française de sciences politiques

AGD : Administrateur général des données

ANR : Agence nationale de la recherche

API : Agence pour le patrimoine immatériel de l'État

API : *Application Programming Interface*

APPI : application des peines, probation, insertion

CAQDAS : *Computer-assisted qualitative data analysis software*

CEDH : Cour européenne des droits de l'homme

CEPEJ : Commission pour l'efficacité de la justice

CNB : Conseil national des barreaux

CNIL : Commission nationale de l'informatique et des libertés

CNRS : Centre national de la recherche scientifique

COMPAS : *Correctional offender management profiling for alternative sanctions*

DASC : Direction des affaires civiles et du sceau

DAGE : Direction de l'administration générale de l'équipement

DAVC : Diagnostic à visée criminologique

DILA : Direction de l'information légale et administrative

DINSIC : Direction interministérielle du numérique et du système d'information et de communication

DINUM : Direction interministérielle du numérique

DISIC : Direction interministérielle du système d'information et de communication de l'État

DMI-TCAT : *Digital methods initiative Twitter capture and analysis toolset*

DSI : Direction des systèmes informatiques

DSJ : Direction des services judiciaires

ENM : École nationale de la magistrature

IA : Intelligence artificielle
ITT : Incapacité temporaire de travail
MAAD : Modes algorithmiques d'analyse des décisions
MARC : Modes alternatifs de règlement des conflits
MARL : Modes alternatifs de règlement des litiges
MARD : Modes alternatifs de règlement des différends
NOÉ : Nouvelles opportunités économiques
NTIC : Nouvelles technologies de l'information et de la communication
ODR : *Online Dispute Resolution*
OGP : *Open Government Partnership*
OQTF : Obligation de quitter le territoire français
PC : Prestation compensatoire
PGO : Partenariat pour un gouvernement ouvert
RGPD : Règlement général sur la protection des données
PQN : Presse quotidienne nationale
PQR : Presse quotidienne régionale
QPC : Question prioritaire de constitutionnalité
SGMAP : Secrétariat général pour la modernisation de l'action publique
SIAF : Service interministériel des archives de France
SPIP : Service pénitentiaire d'insertion et de probation
TAL : Traitement automatique du langage
TTR : Traitement en temps réel
UHNJ : Union nationale des huissiers de justice
UJA : Union des jeunes avocats

Introduction générale

De quoi la ‘justice prédictive’ est-elle le nom ? C’est de cette question qu’est né le présent mémoire original d’habilitation à diriger des recherches (HDR). L’expression ‘justice prédictive’ est en effet apparue à la fin des années 2000 et a semblé faire soudainement florès dans la littérature juridique et le débat public. Or, elle n’a rien d’évident ou de naturel. Bien au contraire, elle nous amène immédiatement à nous interroger. Que recouvre ce terme et que signifierait une justice devenue prédictive ? Peut-on même prédire la justice ?

Ce manuscrit apporte de premières réponses à ces questions. Il a été conçu comme le bilan d’étape d’un travail de recherche en cours¹, commencé il y a plusieurs mois. Il est original, au sens de personnel et inédit, mais encore partiel. Il dessine un programme à poursuivre dans les années qui viennent — notamment à travers le projet ANR Just-IA² que je pilote et dans lequel il s’inscrit³. Après avoir rappelé la mise en données de la société puis ses manifestations dans la justice, la problématique et le cadre d’analyse adopté sont présentés. L’approche retenue se situe au croisement entre sociologie de l’action publique et des administrations, sociologie de la justice et sociologie de l’innovation. La ‘justice prédictive’ apparaît comme une controverse qu’il convient d’interroger et qui en dit long sur la manière dont la justice et le jugement sont pensés.

I / La « société des algorithmes »

Les sociétés contemporaines sont marquées par un tournant algorithmique qui touche un grand nombre d’activités et de secteurs, marchands, mais aussi publics, au point que l’on parle désormais de « sociétés des algorithmes » (Burrell et Fourcade, 2021). Nationalement et internationalement, l’intelligence artificielle (IA) est positionnée comme un enjeu majeur sur les plans économique et géopolitique (Villani, 2019 ; Thibout, 2019), mais également politique, démocratique et éthique (Benbouzid et Cardon, 2022 ; O’Neil, 2016 ; Pasquale, 2015 ; Stern, 2018). Des critiques sur la mise en données de la société se multiplient (Sadin, 2018).

1.1. *Big data* et algorithmes

Le tournant algorithmique est généralement associé à la montée en puissance du *big data*, c’est-à-dire au développement exponentiel de la quantité, de la diversité, de la rapidité

¹ Une partie des recherches ici présentées a été soutenue par le Data-Institute et la Chaire « Société algorithmique » de l’Institut d’intelligence artificielle de l’Université de Grenoble (UGA).

² Projet Just-IA « Décisions équipées. Genèse, usages et effets des algorithmes de “justice prédictive” sur la chaîne de traitement des litiges », programme ANR 2020 (PRC). Commencé en mars 2021, ce projet est coréalisé avec Christian Licoppe. Certains *workpackages* sont réalisés séparément et d’autres, ensemble.

³ Les chapitres 1 à 4 font partie du *workpackage* 4 « Travail sur les discours autour de la justice prédictive : genèse, développement et controverses » dont j’ai la responsabilité.

des données et des traces digitales produites par les activités numériques (Bastin et Tubaro, 2018 : 376-379). Ces données massives requièrent des moyens de circulation, stockage et traitement de l'information inédits, réalisables seulement par des outils informatiques très sophistiqués qui démultiplient les capacités de calcul. Le *big data* est ainsi inséparable du développement des algorithmes qui permettent de traiter ces données (Cardon, 2015). Un algorithme est défini comme une série d'instructions formulées pour résoudre un problème (Abiteboul et Doweck, 2017), qu'il soit relatif au recueil des données, à leur traitement, ou bien à la restitution des résultats produits.

Les pratiques individuelles et collectives sont de plus en plus guidées par des algorithmes, mus par les multiples « métriques de l'Internet » (Cardon, 2013a), celles des réseaux sociaux⁴. Un « capitalisme de plateforme » (Abdelnour et Bernard, 2018), dit encore capitalisme numérique, se met en place ; il recompose les formes d'organisation des marchés, dont certains intermédiaires classiques sont exclus ou marginalisés, au profit de nouveaux, les acteurs non professionnels et surtout les plateformes elles-mêmes (Beuscart, Dagiral et Parasie, 2019 : chap. 7). Les marchés du tourisme, des mobilités, de la culture sont bousculés par l'économie collaborative, de même que le journalisme et l'information (Christin, 2020). Les algorithmes sont le bien le plus précieux des géants du Net qui en protègent jalousement les méthodes d'élaboration et les codes sources — en ouvrant toutefois inégalement l'accès à leurs données via les interfaces de programmation d'applications (API), afin d'étendre leur rayonnement (à propos de Twitter, voir Francony, 2018). À l'ombre des plateformes, se met en place un « *digital labor* » (Casili, 2019), plus ou moins visible, qui représente une transformation importante, mais non uniforme du travail et des conditions de son exercice.

1.2. *E-gouvernement, open data* et modèles prédictifs dans l'action publique

Les institutions et pouvoirs publics sont acteurs de ce tournant algorithmique. Depuis les années 1990, à travers des politiques nationales, européennes et internationales de développement de l'*e-gouvernement* et de l'*open data* (cf. *infra*), ils ont d'abord étendu l'infrastructure de réseaux et systèmes, puis fourni des données et services en ligne (pour l'OCDE, voir Ubaldi, 2013). Les technologies numériques ont été encouragées comme moyen d'améliorer l'efficacité des organisations bureaucratiques (Margetts 2003 ; Dunleavy *et al.*, 2005 ; Margetts et Dunleavy, 2013) et leurs relations avec le public (West, 2005). Dans différents secteurs d'action publique aujourd'hui, la santé, les politiques sociales, l'enseignement ou bien encore l'administration fiscale, des plateformes numériques de dématérialisation des procédures, services et prestations, sont en place. Associées à des dispositifs de visioconférence, elles permettent d'interagir à distance avec les destinataires de l'action publique, et de moduler les formes de l'échange avec patients, allocataires, étudiants ou contribuables. La sophistication des technologies et du matériel, la prolifération des pratiques numériques — globalement et dans certaines franges de la société en particulier — l'expérience d'une pandémie mondiale lors de laquelle le confinement a été massivement

⁴ À propos de Facebook, voir Bastard *et al.*, 2017 ; sur les moteurs de recherche concernant le *PageRanking* de Google, voir Cardon, 2013b

utilisé, ont considérablement accru la diffusion de ces manières de faire et d'un gouvernement à distance.

Le tournant algorithmique se manifeste de différentes façons dans le secteur public. Il est observable à travers le recours à des logiciels d'appariement entre une ressource et plusieurs individus (Simioni et Steiner, 2022), pour gérer le délicat problème de l'attribution de certains biens matériels ou immatériels, en tenant compte de structures de choix complexes et de critères multiples. En ce qui concerne l'affectation scolaire en France, l'appariement entre les choix des bacheliers et les places disponibles dans les nombreuses formations de l'enseignement supérieur français, est en grande partie déléguée au logiciel Parcoursup, successeur d'APB (admissions post-bac) (Frouillou, Pin et Van Zanten, 2022). De même, dans le secteur de la santé, les listes de priorité et de compatibilité entre donneurs et receveurs sont gérées via des logiciels d'attribution des organes à des personnes en attente de greffe de cœur (Hénin, 2021) ou de transplantation rénale (Steiner, 2022).

Les modèles prédictifs, associant théories, algorithmes et bases de données de grande ampleur, sont, eux aussi, utilisés dans l'action publique pour répondre « à des situations d'incertitude par l'évaluation d'un risque, en proposant aux décideur·euse·s des scénarii pour le futur entre lesquels il s'agit de trancher » (Boullier et *al.*, 2021 : 11). Leur rôle est maintenant bien documenté en matière de politique économique, d'environnement, d'énergie (Aykut et Nadaï, 2019), ou de régulation des produits chimiques (Boullier, Demortain et Zeeman, 2019). La crise sanitaire de la Covid-19 a mis en scène le rôle crucial de ces modélisations dans la gestion de la pandémie et l'anticipation de ses conséquences pour la santé publique. Dans différents pays, les décisions publiques de confiner ont notamment reposé sur l'examen de scénarii relatifs à la diffusion du virus et à ses effets en matière de mortalité (Boullier et *al.*, 2021 : 6 ; Pawson, 2021). Ces modèles prédictifs sont également utilisés pour répartir les forces de police dans l'espace urbain (Benbouzid, 2018 ; Brayne, 2020). Les algorithmes de détection des bagages abandonnés (Neyland, 2019), comme ceux qui surveillent les transactions financières pour repérer des activités suspectes liées à du blanchiment d'argent, reposent quant à eux sur des « ensemble[s] d'instructions ordonnées, étape par étape, et structurées par des conditions initiales (de type SI) devant mener à des conséquences (de type ALORS) » (Amicelle, 2021 : 38).

Les infrastructures informationnelles et de calcul (Star et Ruhleder, 1996) jouent ainsi un rôle important non seulement sur le plan des pratiques sociales, mais aussi sur celui de la production de l'action publique et des manières de gouverner (Arsène et Mabi, 2021). Dans une perspective foucauldienne, Antoinette Rouvroy et Thomas Berns (2013) ont avancé la notion de « gouvernementalité algorithmique » pour désigner l'émergence d'un nouveau mode de gouvernement des conduites qui repose sur la dissolution de l'individu, éclaté en une multitude de données éparpillées, mais également sur la réduction du monde à *ce qui est*, écartant les potentialités, les possibles, et donc les sources d'imprévisibilité et d'incertitude. C'est à la fois le sens du recours à des modèles prédictifs, leurs promesses, leur paradoxe et leur faiblesse. « Les modèles sont construits et utilisés pour déterminer des choix lorsque les données sont douteuses, coûteuses, ou non encore produites. Ce constat simple invite à s'intéresser à la fragilité inhérente aux modèles : même si leur déploiement semble sans

limite, ils restent des colosses aux pieds d'argiles, toujours susceptibles d'être fragilisés par de nouvelles données venant contredire leurs prédictions. » (Boullier et *al.*, 2021 : 12) En outre, ils sont traversés par une contradiction fondamentale : ils dessinent des scénarii qu'il s'agit précisément d'empêcher d'advenir.

II / La justice dans la société algorithmique

Comment ce tournant algorithmique s'opère-t-il dans la justice, domaine classiquement fortement régulé par des logiques organisationnelles, institutionnelles et professionnelles ? Quels types d'algorithmes y sont implantés ? Par qui, pour quoi faire, avec quels effets et quels cadrages ? Les institutions et professionnels de justice ne sont pas épargnés par cette évolution profonde (Aas, 2005 ; Hutton, 2013). Celle-ci n'est pas non plus radicalement nouvelle (Harcourt, 2007, 2011a, 2011b ; Hutton, 1995) même si son ampleur et ses effets restent à identifier précisément.

2.1. 'justice prédictive' et intelligence artificielle

La 'justice prédictive' est une manifestation du développement des algorithmes dans la justice. Ce terme fleurit au fil de nombreux ouvrages, colloques et journées d'étude, articles et numéros spéciaux de revues, en droit, théorie et philosophie du droit⁵. Il désigne des plateformes en ligne⁶ qui proposent de « prédire » les décisions de justice ou bien encore de « réduire l'aléa judiciaire ». Ces plateformes fonctionnent à partir de bases de données regroupant une grande quantité de décisions de justice, dont le traitement est assuré par différentes méthodes mathématiques, statistiques et/ou informatiques et dont les résultats sont accessibles à travers différents types de représentations visuelles réalisables par des outils dynamiques de visualisation des données. Elles permettent de simuler différentes décisions de justice possibles sur un cas donné ; d'évaluer leur probabilité respective, en référence à un nombre important de cas comparables déjà jugés. Elles s'adressent en priorité aux avocats, juristes d'entreprise et compagnies d'assurances, dans le cadre de l'élaboration de leurs stratégies juridiques, et plus marginalement aux magistrats, dans leurs activités de jugement. En France, une expérimentation a été tentée en matière civile auprès de magistrats de deux cours d'appel (Licoppe et Dumoulin, 2019) : il s'agit de l'expérimentation anonymisée dans ce manuscrit sous le nom de [PréviCompute]⁷.

2.2. La 'justice prédictive' à la croisée des chemins

La 'justice prédictive' se situe à la rencontre de trois principaux mouvements : l'ouverture des données publiques, qualifiée d'*open data* ; le développement de l'intelligence artificielle et l'ubérisation du droit, qui interagissent dans la « transformation numérique du droit » (Cassar, 2020).

⁵ Tellement nombreux que je ne les cite pas ici : cf. le premier chapitre.

⁶ Par exemple : [Anajur] : <http://Anajur.com>, consulté le 9 mars 2017 ; [PréviCompute] : <https://www.prévicompute.com/>, consulté le 12 mars 2017. Les crochets indiquent que les noms ont été pseudonymisés.

⁷ Il s'agit d'un alias. Pour une explication de l'anonymisation, cf. *infra*.

2.2.1. La transformation du régime de publicité des données de justice

Cette offre de 'justice prédictive' s'inscrit dans le mouvement de l'*open data* ou, plus exactement, elle est rendue possible par la transformation du régime de publicité des données publiques en général, et juridiques en particulier. En 2016, la loi dite Lemaire pour une République numérique⁸ proclame en effet le caractère public, gratuit et ouvert des décisions de justice⁹, ouvrant la voie à la diffusion d'un volume considérable d'entre elles¹⁰. En cela, elle représente à la fois un prolongement et une rupture ([Desjardins], [Bourg], 2017 : 1483 ; Deffains, 2019 : 52).

Elle marque indubitablement un pas de plus dans un processus d'ouverture et de diffusion des données juridiques, déjà engagé de longue date. La création du site Légifrance¹¹ remonte en effet à 1999. Il associe un moteur de recherche à une base de données qui rassemble les textes législatifs et réglementaires de droit positif, et une sélection de décisions de juridictions judiciaires et administratives (partiellement anonymisées). Il est accessible à tous, à la différence de JuriCa, qui est une base de données de la Cour de cassation, créée en 2005¹², et qui recense la totalité des arrêts des cours d'appel en version non anonymisée (Barthe, 2008). JuriCa est accessible aux magistrats via l'intranet du ministère de la Justice et un accès payant est possible pour les éditeurs juridiques, qui l'intègrent ensuite dans des bouquets d'information juridique, à destination des professionnels du droit. Ces derniers peuvent aussi avoir accès à certaines décisions, à travers le paiement d'une redevance directement auprès de la Cour de cassation. L'*open data* des décisions de justice proclamée par la loi Lemaire intervient donc dans un paysage déjà peuplé de différents dispositifs d'accès aux décisions de justice, inégalement ouverts¹³. Mais en instaurant le modèle de la gratuité, il bouscule l'arrangement marchand sur les données de justice, en place jusque-là.

De plus, l'opérationnalisation de l'*open data* ne se fait pas en un jour, car le principe de la publicité et du libre accès entre en conflit avec d'autres principes, au premier rang desquels celui de la protection des données personnelles. Les jugements et arrêts comprennent en effet de nombreuses informations identifiantes, *a fortiori* si elles sont croisées entre elles. C'est une des raisons qui expliquent que la mise à la disposition effective du public des décisions de justice tarde à se concrétiser¹⁴. Elle est échelonnée entre 2021 et

⁸ Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique.

⁹ Art. 20 et 21 de la loi n° 2016-1321 du 7 octobre 2016 pour une République numérique.

¹⁰ Estimées en 2016 à 3 800 000 décisions civiles et pénales ; 130 000 décisions des juridictions administratives (Cadiet, 2018 : 26).

¹¹ Arrêté du 6 juillet 1999 relatif à la création du site Internet Légifrance, modifié en octobre 2002, à la suite du décret n° 2002-1064 du 7 août 2002 relatif au service public de la diffusion du droit par l'Internet.

¹² Décret n° 2005-13 du 7 janvier 2005 modifiant le code de l'organisation judiciaire (partie réglementaire) et relatif au service de documentation et d'études de la Cour de cassation.

¹³ Pour une description des liens entre les différentes bases de données disponibles et le versement de certaines décisions vers Légifrance, voir (Barthe, 2008).

¹⁴ Avec la loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice (art. 33) et le décret n° 2020-797 du 29 juin 2020 relatif à la mise à disposition du public des décisions de juridictions judiciaires et administratives.

2025, suivant la matière et les juridictions, soit jusqu'à neuf ans après la promulgation de la loi Lemaire¹⁵, si le calendrier annoncé est respecté.

Pour les juridictions de l'ordre judiciaire, c'est la Cour de cassation qui est responsable de garantir et d'organiser l'accessibilité des décisions de justice en format numérique. Le moteur de recherche Judilibre de son site Internet donne accès à toutes les décisions qu'elle a rendues depuis le 1^{er} octobre 2021 et à environ 480 000 arrêts, rendus depuis 1947. Il n'y a pas de sélection, contrairement à ce qui se passe pour Légifrance. Judilibre comprend également toutes les décisions civiles, sociales et commerciales rendues publiquement par les cours d'appel depuis le 15 avril 2022, auquel s'ajoutent 82 000 décisions rendues antérieurement et déjà diffusées sur Légifrance¹⁶. Les décisions civiles et pénales de première instance devraient être versées en septembre 2025 et celles relatives aux procédures criminelles et à l'appel en matière pénale en décembre 2025¹⁷. À l'heure où nous écrivons ces lignes, l'*open data* des décisions de justice est donc en cours de réalisation.

2.2.2. Les promesses floues de l'intelligence artificielle

Les discours sur la 'justice prédictive' associent celle-ci à l'intelligence artificielle (IA) et en particulier à l'apprentissage profond, la méthode la plus radicalement nouvelle d'IA. Or le flou règne sur ce qu'est exactement l'IA. On considère généralement qu'elle naît dans les années 1960, portée par deux grandes théories informatiques en concurrence : l'approche symbolique ou cognitive et l'approche connexionniste (Cardon, 2019 : 385-398). Ces deux paradigmes sont « les héritiers d'un ensemble commun de concepts élaborés dans les années 40 par les fondateurs de la théorie du calcul, de l'informatique, et de la simulation cognitive » (Memmi, 1990 : 44) ; après quoi, ils se sont dissociés et sont entrés en concurrence. Depuis lors, ils ont connu des vagues de succès et de revers (Cardon, 2019 : 387).

L'intelligence artificielle symbolique repose sur un système de règles formelles construites par les humains, souvent par simulation du fonctionnement de la cognition humaine. Les systèmes experts exploités notamment en droit (Bourcier, 1995) relèvent de ce type d'IA : ils consistent à décomposer et formaliser les mécanismes décisionnels humains par des arbres de décision, qui sont ensuite traduits dans des langages mathématiques et informatiques. Les logiciels utilisent ces systèmes de règles et les appliquent pour résoudre des problèmes donnés. Ce type d'IA repose sur une conception formelle du langage et de la cognition, la manipulation et la reproduction théoriques de systèmes symboliques, au sens où la forme syntaxique et le contenu sont dissociés, les raisonnements sont épurés, etc. En retour, les résultats de la simulation réalisée par ordinateur sont parfaitement explicables puisqu'ils reposent sur des systèmes logiques explicites et utilisés en amont pour coder et alimenter les calculs informatiques.

¹⁵ Arrêté du 28 avril 2021 pris en application de l'article 9 du décret n° 2020-797 du 29 juin 2020 relatif à la mise à la disposition du public des décisions des juridictions judiciaires et administratives.

¹⁶ <https://www.courdecassation.fr/toutes-les-actualites/2022/04/29/open-data-les-decisions-civiles-sociales-et-commerciales-des-cours>, mis en ligne le 29 avril 2021, consulté le 16 août 2022.

¹⁷ <https://www.vie-publique.fr/en-bref/279750-open-data-le-calendrier-de-lacces-en-ligne-aux-decisions-de-justice>, mis en ligne le 6 mai 2021, consulté le 16 août 2022.

L'IA connexionniste s'appuie en revanche sur un tout autre paradigme ; elle met l'accent sur l'apprentissage par la machine elle-même d'un modèle qui est issu de manière inductive de l'analyse d'une quantité massive de données. Le *machine learning* (généralement traduit par apprentissage machine en français alors qu'il s'agirait plutôt d'apprentissage automatique ou automatisé) relève de ce modèle. Aucune règle formelle de raisonnement n'est introduite dans l'ordinateur. Cette technique consiste à laisser la machine apprendre par elle-même en la confrontant à un grand jeu de données, avec la possibilité d'introduire une part variable de supervision humaine, destinée à corriger les éventuelles erreurs et favoriser l'apprentissage le plus fin possible.

Cette méthode a connu un succès nouveau dans les années 2000-2010 avec l'augmentation des capacités de calculs des ordinateurs et l'accès à de gros volumes de données numériques (Cardon, Cointet et Mazières, 2018). L'apprentissage profond (*deep learning*) en est l'incarnation la plus emblématique. Il fonctionne à partir d'une infrastructure que l'on appelle un réseau de neurones, car « il s'agit de reproduire mathématiquement, de façon formelle, le fonctionnement d'un neurone qui s'excite lorsqu'un seuil électrique a été atteint par les différents flux qui lui viennent de ses synapses » (Cardon, 2019 : 5). Cette technique a été conçue dès les années 1960 (avec une machine fonctionnant à l'image d'un cerveau humain), mais jugée irréaliste et abandonnée ensuite au profit de l'IA symbolique. Aujourd'hui, ce type de machine apprenante refait surface, avec une plus grande sophistication, des puissances de calcul inédites d'ordinateurs dont les mémoires ont été étendues et qui ont à leur disposition des quantités massives de données numériques. La traduction automatique de textes (via des applications comme DeepL) et la reproduction du style littéraire ou musical d'un auteur ou compositeur célèbre sont des illustrations de ce que les algorithmes d'apprentissage machine peuvent réaliser.

L'IA symbolique et l'IA connexionniste ne reposent pas sur les mêmes présupposés, cadres théoriques et méthodes, et n'abordent donc pas la résolution des problèmes de la même manière. L'exemple classiquement donné est celui de la reconnaissance d'un chat dans une image : dans l'IA symbolique, des règles et régularités sont intégrées dans le logiciel (un chat est une forme qui a deux yeux, une queue, quatre pattes, etc.) alors que dans l'IA connexionniste, la machine est confrontée à des quantités massives d'images de chats ce qui lui permet de se constituer un modèle mathématique de ce qu'est un chat (y compris par essai-erreur dans le cas de la supervision). De ce fait, les enjeux de ces différents types d'AI sont très disparates, en particulier en ce qui concerne la transparence et l'explicabilité des résultats obtenus. Les résultats de l'IA connexionniste ne sont pas complètement explicables ni parfaitement reproductibles. La machine est alimentée en données d'entrées ; un objectif est fixé en sortie (est-ce un chat ou pas ?). Entre les deux, des milliers de couches de neurones formels s'activent ou pas en fonction du poids des synapses qui les relient, comme pour le fonctionnement d'un cerveau humain. Un algorithme règle les coefficients de synapses pour former une grande matrice de chiffres correspondants au modèle (du chat) et reconnaître s'il y en a un ou pas dans l'image (Cardon, 2019 : 393-394). Les calculs réalisés par la machine ne suivent pas toujours exactement le même chemin, ce qui confère au résultat une part d'aléatoire et à la machine un côté boîte noire. Si les réseaux de neurones sont efficaces, on

ne peut pas expliquer pourquoi, car c'est un fonctionnement inductif. Les informaticiens le savent et le reconnaissent. Si cela ne porte pas à conséquence pour reconnaître un chat, il en va différemment pour d'autres actions, par exemple produire des décisions publiques relatives à des individus. Les questions de transparence et d'explicabilité sont en effet cruciales. Rappelons qu'à ce jour, le Règlement général sur la protection des données (RGPD) de l'Union européenne prévoit d'importantes limitations à la prise de décisions publiques entièrement automatisées sur la base d'un profilage et de calculs algorithmiques¹⁸ — quand elles s'appliquent aux personnes.

Au sein de l'IA connexionniste et de l'apprentissage machine, les spécialistes distinguent de plus l'IA forte de l'IA faible selon le niveau d'autonomie et d'ampleur de l'intelligence de la machine. Une IA faible est capable de réaliser une tâche qui, pour complexe qu'elle soit, reste bien délimitée. Les assistants vocaux des smartphones peuvent reconnaître une voix parmi d'autres, comprendre et interpréter le langage pour répondre à des requêtes et le faire de mieux en mieux — au fil du temps et à mesure de la diversité des situations rencontrées. Il s'agit toutefois de tâches précises, exécutées à partir d'instructions, et effectuées sans émotion, conscience ou intentionnalité. En revanche, une IA forte serait une forme d'intelligence qui gagnerait en sensibilité, voire en conscience. Elle se rapprocherait ainsi de l'intelligence humaine, au sens où elle pourrait accomplir des tâches multiples et qui appellent toute la gamme des compétences cognitives de l'intelligence humaine. Non advenue à ce jour, elle est l'objet de recherches scientifiques et d'investissements considérables, par exemple avec les voitures autonomes. Sans assistance aucune et dans des environnements complexes et fluctuants, ces véhicules dotés d'IA devraient interpréter des situations et prendre des décisions en conséquence (freiner ou accélérer, etc.). L'IA forte est très discutée, car il est probable qu'elle gagnerait aussi en autonomie à l'égard des humains.

Les algorithmes englobés sous le terme de 'justice prédictive' entretiennent des liens hétérogènes avec ces différentes formes d'IA. Les concernant, l'offre reste assez opaque. Il est difficile de cerner sur quelles technologies ces algorithmes reposent exactement, d'autant que leurs créateurs ont intérêt à se positionner tels des pionniers et à surfer sur l'engouement pour la révolution technologique de l'IA. Cela est facilité par le fait que leurs clients potentiels n'ont pas forcément la culture numérique et informatique leur permettant de saisir en quoi ce qu'ils font est vraiment nouveau, plus performant et/ou véritablement de l'IA. Comme souvent, la dimension marketing est déterminante.

Dans ce contexte, en juillet 2019, le Conseil national des barreaux (CNB) a commandité auprès d'une « société de conseil en transformation digitale »¹⁹ une étude comparée des principaux systèmes disponibles de 'justice prédictive' sur le marché national²⁰. Cette cartographie confirme la diversité des dispositifs proposés par les différents opérateurs.

¹⁸ Article 22 du règlement (UE) 2016/679, adopté par le Parlement européen et le Conseil le 27 avril 2016. Il est entré en vigueur dans les vingt-sept pays membres à partir de mai 2018.

¹⁹ Sopra Steria Next, <https://www.soprasteria.fr/services/conseil>, consulté le 25 février 2021.

²⁰ <http://lemondedudroit.com/decryptages/72373-cnb-preconisations-actions-legaltech-jurimetrie.html>

Certes, dans l'ensemble, « à partir de [la] masse de données [disponible], ils génèrent des éléments statistiques mis en regard de la saisie des caractéristiques d'un cas renseigné » (Degos, 2020 : 2). Ils reposent sur des moteurs de recherche qui fouillent dans les bases de données de décisions de justice (de taille inégale suivant les systèmes, cf. *infra*). En revanche, certains moteurs sont textuels ([Jurissimo], [Prédidroit], [Legdata]) : ils travaillent à partir de recherches exprimées en langage naturel et s'appuient donc sur les avancées de l'IA connexionniste en matière d'analyse du langage. D'autres sont quantitatifs ([LegalDonnées+], [PréviCompute]), au sens où, après avoir caractérisé en quelques mots-clefs l'affaire, une série de filtres contextualisés est disponible pour affiner les recherches parmi les décisions de justice remontées. Dans le second cas, il s'agit d'outils basés sur des modèles statistiques et non sur de l'IA — ni symbolique ni connexionniste. Enfin, il existe une offre de modules de simulation ([Anajur]) qui mobilisent de l'apprentissage profond à une étape du processus ainsi que des méthodes statistiques à d'autres.

Ces distinctions, pour importantes qu'elles soient, ne disent toutefois rien de l'efficacité ou de la performance de ces différents outils. Elles pointent simplement la diversité des techniques de traitement et d'analyse utilisées et l'écart entre celles réellement employées et le discours consistant à présenter ces algorithmes comme l'avènement d'une justice robotisée ou relevant d'une IA forte. Le débat sur la 'justice prédictive' est ainsi amarré à celui de l'IA, mais parfois à l'excès, du point de vue de la réalité des techniques mises en œuvre, comme nous le verrons plus loin. Ces algorithmes sont certes les héritiers de courants de recherche qui s'intéressent à la modélisation des décisions. En ce sens, ils partagent le paradigme prédictif né du tournant probabiliste qu'a connu l'histoire des sciences (Hacking, 1990) et une « raison modélisatrice » (Boullier et al., 2021 : 11-12) centrée indirectement sur la quantification de scénarii pour le futur.

2.2.3. L'ubérisation du droit par des startup de la legaltech

D'un point de vue de sociologie économique, l'ubérisation correspond à la notion émergente de « capitalisme de plateforme » (Abdelnour et Bernard, 2018). Dans le cas du droit, l'ubérisation passe par la montée en puissance de *startup* qualifiées de *legaltech*, abréviation de l'anglais *legal technology*. La notion de *startup* est elle-même assez floue et fonctionne comme un mot d'ordre (Quijoux et Saint-Martin, 2020). « Parfois utilisé[e] pour désigner des entreprises innovantes, d'autres fois pour désigner un modèle d'entreprise risqué ou pour mettre l'accent sur des entrepreneur-ses jeunes et héroïques » (Flécher, 2019 : 41), elle ne correspond à aucune catégorie juridique et fait partie d'un discours enchanté sur l'innovation. A la suite de M. Flécher, le terme est entendu ici comme recouvrant des « entreprise[s] innovante[s], censée[s] croître rapidement et financée[s] par des levées de fonds" (*ibid.* 41). En l'occurrence, il s'agit d'entreprises du numérique qui innovent dans le secteur du droit. Le terme d'ubérisation, qui se réfère à la plateforme Uber, souligne surtout que ces nouveaux venus concurrencent les positions établies de certains professionnels installés — les avocats et les notaires à l'image des chauffeurs de taxi — en proposant aux clients des services plus simples et moins chers, permis par la dématérialisation, la standardisation de documents juridiques (par exemple les contrats) commercialisés à des

coûts très avantageux et les nouvelles technologies comme la *blockchain*²¹ (Rhattat, 2019). Le site Demanderjustice.com suggère d’ester en justice pour le compte de particuliers, de gérer leur procédure amiable puis contentieuse pour des litiges de faible montant (et qui ne requièrent pas procéduralement une représentation par avocat). Il est dans une relation principalement numérique avec ses clients via la plateforme et différentes technologies associées : formulaires en ligne, *chats*, échanges téléphoniques, etc.

Le développement d’une offre de ‘justice prédictive’ s’inscrit donc dans un paysage de bouleversement du droit et de remise en cause de la juridiction de professionnels (Abbott, 1988) comme les avocats (Milburn, 2003) ou les notaires (Delmas, 2019). Ils tentent de protéger celle-ci par des plaintes et actions en justice à l’encontre de plusieurs sociétés entrantes sur le marché du droit²².

2.3. Prédiction et justice actuarielle

L’émergence de la ‘justice prédictive’ intervient dans un contexte qui est préparé, travaillé par le développement de la justice actuarielle (Feeley et Simon, 1992 et 1994 ; Mary, 2001). Cette dernière est une notion proposée par Malcom Feeley et Jonathan Simon dans les années 1990, en lien avec celle de « nouvelle pénologie » et dans le cadre des théories de la société du risque (Beck, 2008 ; Castel, 1995). Elle décrit une philosophie pénale et des dispositifs utilisés principalement aux États-Unis, dès les années 1980.

Ce nouveau paradigme de la peine consiste en un « passage d’une pénologie axée sur l’individu, sa punition et/ou son traitement, à une pénologie axée sur la gestion de groupes à risques, leur surveillance et leur contrôle, afin de réguler les niveaux d’une délinquance considérée comme un risque normal dans la société. L’objectif ne serait dès lors plus d’éliminer ce risque, mais de le rendre tolérable, de le circonscrire dans des limites sécuritaires acceptables. Parallèlement à une attitude plus répressive, un nouveau discours émergerait, qui [...] adopte, dans une perspective managériale, un langage actuariel (la probabilité et le risque) appliqué à des populations (distributions statistiques) » (Mary, 2001 : 35).

Le crime et les déviances sont perçus comme n’importe quel risque qu’il convient de gérer sur un mode probabiliste. Le raisonnement est de type actuariel. Le Trésor de la langue française indique que le terme actuariat est employé avec ce sens à compter des années 1870²³. C’est une « technique appliquant les méthodes de la statistique et du calcul des probabilités aux problèmes d’assurance, de prévoyance, d’amortissement²⁴ ». Elle est accomplie par des actuaires, dont l’activité consiste « à faire des calculs de probabilités à partir de renseignements statistiques. Ces calculs sont le plus souvent destinés à établir des taux de primes d’assurance en tenant compte de la fréquence des risques courus : mortalité, maladie, accidents, incendies, naufrages, vols » (Blumberg, 2022). L’enjeu est celui de la prédictibilité

²¹ Ce service de cryptage et certification de processus sécurise les transactions.

²² C’est le cas de [Jurissimo] ou du site Demanderjustice.com, toutes les deux poursuivies par le Conseil national des barreaux (CNB) et l’Ordre des avocats de Paris, voir (Chassagnard-Pinet, 2017).

²³ <https://www.cnrtl.fr/definition/actuaire>, consulté le 22 août 2022.

²⁴ Selon Le Robert, dictionnaire en ligne, <https://dictionnaire.lerobert.com/definition/actuariat>, consulté le 22 août 2022.

d'un avenir à la fois inconnu et marqué par la survenance possible de risques. L'incertitude est rendue maîtrisable par le recours à des méthodes statistiques et à des calculs de probabilités²⁵.

Dans ses recherches sur la logique actuarielle en matière pénale, Bernard Harcourt (2007) a mis en évidence l'importance de travaux sociologiques comme ceux d'Ernest Burgess pour la formulation d'un paradigme prédictif relatif à la récidive. Sociologue de l'École de Chicago dans les années 1920, Burgess recense les facteurs sociaux et pénaux qui influencent la récidive des personnes condamnées puis les regroupe dans des tests statistiques de manière à faciliter et rationaliser la prise de décisions touchant la libération conditionnelle de détenus. La logique repose donc sur l'analyse de bases de données passées afin d'en tirer des régularités, utilisables pour orienter les futures décisions à prendre. L'individu n'est plus vu comme un cas unique, mais comme membre d'une catégorie statistique dont les comportements et les risques sont connus. L'analyse des données individuelles permet le profilage, c'est-à-dire le classement de ceux-ci en groupes et sous-groupes.

Évaluer une personne au stade pré ou post-sentenciel consiste, dans ce cadre, à traduire ses caractéristiques sociodémographiques, son rapport à la délinquance, etc. en score pour quantifier son risque de récidive. Plusieurs types de facteurs (relatifs au profil de la personne, à son passé judiciaire, à sa situation actuelle, etc.) sont pris en compte à travers différentes générations d'outils d'évaluation des risques (*risk assessment tool*) qui se sont succédé au plan international depuis les années 1990. Certains sont exclusivement actuariels, c'est-à-dire scorés, tandis que d'autres hybrident une logique quantifiée avec le jugement clinique sur la personne (Gautron et Dubourg, 2014 ; Gautron, à paraître). Cette évaluation des risques est un des supports de la prise de décision, par exemple concernant la condamnation ou les modalités d'application des peines (remise en liberté ou pas, bracelet électronique ou pas, etc.).

Largement utilisés dans les pays anglo-saxons, ces outils ont suscité d'importantes résistances en France (Gautron et Dubourg, 2014). En 2008, la tentative de l'administration pénitentiaire d'implanter un instrument formalisé d'évaluation des risques de récidive — le diagnostic à visée criminologique (DAVC) — s'est soldée par « une mise en œuvre partielle et disparate [...] au sein des services pénitentiaires d'insertion et de probation, puis [par] son abandon progressif » (*ibid.* : 13). S'ils restent peu pratiqués, ils tendent à se développer pour des populations pénales et des risques particuliers. L'évaluation du niveau de radicalisation djihadiste des détenus repose sur des grilles d'évaluation de ce type (de Galembert, 2016 : 62-66).

Ces grilles d'évaluation peuvent être incorporées dans des logiciels. Le cas de Compas, souvent cité par la littérature²⁶, est bien connu. L'acronyme signifie *Correctional offender*

²⁵ T. Porter (2020[1995] : chap.5) a étudié la manière dont la profession d'actuaire, d'abord appuyée sur des savoirs d'expérience et expertises au cas par cas, a d'abord résisté à la rationalisation des tables d'assurance pour préserver son pouvoir.

²⁶ (Angqin *et al.*, 2016) ; (Defferrard & Papineau 2017) ; (Desmoulin-Canselier & Le Metayer, 2020) ; (Garapon & Lassègue, 2018).

management profiling for alternative sanctions (outil de profilage pour la gestion des délinquants en matière correctionnelle en vue de sanctions alternatives). Cet outil algorithmique d'évaluation des risques calcule, à partir de caractéristiques sociodémographiques et pénales, statiques et dynamiques, les probabilités de récidive d'un individu. Il lui attribue un score, destiné à informer le processus de prise de décision pénale, au stade pré ou post-sentenciel (Harcourt, 2007 ; Christin et Brayne, 2020). Ce logiciel, même s'il est inégalement utilisé dans les juridictions américaines, fait partie de l'équipement cognitif dont disposent les acteurs de la chaîne pénale (Christin, 2018).

L'affaire *Loomis c/ Wisconsin* a contribué à la médiatisation de ce logiciel (pour un résumé, voir *inter al.* Desmoulin-Canselier et Le Metayer, 2020). En 2013, Erik Loomis a été condamné à six ans de prison ferme pour avoir tiré des coups de feu sans avoir fait de blessés. Lorsqu'il demande une libération conditionnelle auprès d'une juridiction correctionnelle du Wisconsin, celle-ci refuse en raison du risque de récidive qu'il présente, évalué à l'aide du logiciel *Compas*. L'avocat du condamné fait appel d'abord devant la Cour d'appel puis la Cour suprême de l'État du Wisconsin avec l'argument selon lequel il ne dispose d'aucun moyen de contester l'évaluation réalisée, faute d'information sur le fonctionnement de l'algorithme. La société privée éditrice de *Compas* argue alors de son droit de propriété sur le logiciel et de son intérêt à ne pas divulguer ses procédés (codes sources, etc.) au risque de perdre tout avantage compétitif sur ses concurrents. Les deux juridictions ont successivement validé la ligne de défense de la société (*Equivant, ex-Northpointe*). La Cour suprême des États-Unis, qui aurait pu se saisir de l'affaire, ne l'a pas fait. Cette controverse juridique s'est donc soldée par la reconnaissance de la suprématie des intérêts commerciaux de l'entreprise privée concernée sur les droits de la défense et les enjeux de transparence et d'explicabilité envers les juges et les justiciables. Ce cas a suscité une levée de boucliers du côté de nombreux commentateurs et juristes, aux États-Unis et au plan international. Il a été l'occasion de pointer les risques liés à l'utilisation d'algorithmes propriétaires pour fonder des arrêts et jugements aux effets directs sur les droits et libertés des personnes. L'importance de la régulation publique des algorithmes n'en a été que davantage soulignée (Lacour et Piana, 2019). Le droit européen, vu les exigences du RGPD, est apparu plus protecteur que le droit étatsunien (Malapert, 2020).

Une large controverse sur le rôle des algorithmes dans la cité se dessine alors. Quels en sont les risques et les pouvoirs réels ? De nombreux scientifiques, parfois eux-mêmes *data scientists* devenus lanceurs d'alerte (O'Neil, 2016) se sont mobilisés pour attirer l'attention des politiques et du public sur le fonctionnement en boîte noire des dispositifs algorithmiques et ses conséquences au plan individuel et collectif. Le caractère opaque des dispositifs calculatoires questionne sur les risques encourus : protection des données personnelles, contrôle de la qualité des données utilisées et des usages développés, effets en matière de biais et d'erreurs. Les enjeux de subordination du politique à la technique, de remise en cause des fondements de la démocratie sont également soulignés dès lors que les décisions publiques, et ici judiciaires, n'apparaissent plus maîtrisées ni explicables aux citoyens.

III / Ce que nous disent les controverses : question de recherche et problématique

La ‘justice prédictive’ est donc au carrefour d’une série d’évolutions considérables qui la dépassent largement. Dans ce contexte, les superlatifs ne manquent pas pour souligner son caractère profondément et durablement disruptif. Elle est présentée par certains comme symptôme d’une « rupture anthropologique » et d’une « révolution graphique » (Garapon et Lassègue, 2018). Qu’en est-il exactement ? En quoi ces affirmations sont-elles fondées ? Que nous dit l’émergence de la ‘justice prédictive’ en tant qu’expression et en tant qu’ensemble de pratiques sociales, sur la justice et sur la manière dont l’action publique y est élaborée ? Ces questions sont au cœur du présent manuscrit.

3.1. La ‘justice prédictive’, un objet carrefour

La démarche de recherche que je mène se situe au croisement de la sociologie politique de l’action publique et de la sociologie du droit et de la justice, en incorporant des acquis de la sociologie de l’innovation et de la sociologie pragmatique.

Il est aisé de comprendre l’intérêt que représente l’objet ‘justice prédictive’, compte tenu de l’actualité et l’acuité des discours sur les transformations que provoquerait le développement des algorithmes. Cet objet carrefour permet en outre une approche à l’interface entre droit et sciences sociales de même qu’à l’intérieur de la sociologie et de la science politique. C’est un objet qui nécessite de croiser différentes littératures : la sociologie des réformes administratives, la sociologie du droit et de la justice, la sociologie de l’innovation et les *Science and Technology Studies* (STS), et même par certains côtés, la sociologie économique. Il faudrait encore ajouter l’approche sociohistorique, sauf à considérer qu’elle est consubstantielle de la sociologie de l’action publique. L’attention à l’historicité des processus d’action publique (Laborier et Tröm, 2003) est déterminante dans l’analyse d’interactions qui se déploient dans le présent, mais à partir de condensés non uniformes du passé et d’anticipations variées d’un futur qui reste à dessiner.

Le croisement de ces approches vise trois objectifs. D’abord réaliser un état de l’art des travaux extrêmement nombreux, particulièrement en droit, et mettre en évidence les tensions normatives qu’ils expriment autour de la ‘justice prédictive’ et de ses effets possibles sur le droit et le système judiciaire. Ensuite, documenter le sujet avec des données empiriques originales, pour analyser et comprendre le développement de cette notion, sa genèse et sa trajectoire, et les abondants discours qu’elle suscite, dans un contexte où peu de travaux empiriques existent. Enfin opérer un déplacement analytique et proposer un cadre original de la rencontre entre les dispositifs algorithmiques étudiés et l’activité de juger.

Dans cette perspective, la ‘justice prédictive’ est une expression à examiner de plus près et à déconstruire. Quel cadrage du rapprochement entre justice et algorithmes opère-t-elle ? Quels sont les acteurs qui discutent cette notion ? Que dit son émergence de la manière dont les mondes du droit et de la justice se représentent l’activité de juger d’une part, les décisions de justice d’autre part, la jurisprudence enfin ? Comment cette question est-elle reliée à celle, en réalité connexe, de l’*open data* des décisions de justice ? Pour saisir tout ce qui est en jeu ici, les premiers gestes scientifiques à accomplir consistent à cartographier les acteurs et communautés, sans préjuger de leur force ou de leurs validités respectives ; à

repérer la circulation des acteurs — notamment entre public et privé et à dresser la chronologie de ce qui se déroule dans ces mondes. La sociologie des problèmes publics (Neveu, 2015, Henry, 2021) et la sociologie de l'innovation et des sciences et techniques (Callon, Latour et Akrich, 2006) fournissent les outils analytiques pour dénaturiser la 'justice prédictive' et déplier les différentes conceptions de l'*open data* des décisions de justice.

L'approche par les controverses porte sur les objets encore chauds ; donne accès aux points qui font débat, processus d'intéressement et de problématisation. Croiser la sociologie de l'innovation et la théorie de l'acteur réseau avec la sociologie plus classique de l'action publique et les approches de politiques publiques est heuristique pour saisir des boîtes noires et des possibles pas encore refermés. C'est dans cette direction que j'ai travaillé dans mes précédentes recherches sur la visioconférence (Dumoulin et Licoppe, 2017), mais aussi sur la vidéosurveillance, en insistant sur l'enjeu de traiter de manière symétrique les différents acteurs d'une histoire, « perdants » et « gagnants » ainsi qu'en étudiant des innovations qui ne sont pas encore consolidées et peuvent encore avorter (Douillet et Dumoulin, 2015). Je m'inscris dans un mouvement que d'autres collègues alimentent également (Cordella, 2007 ; Cordella et Bonina, 2012 ; Cordella et Hesse, 2015 ; Robert et Dufresne, 2015).

3.2. Controverses sur la 'justice prédictive'

La 'justice prédictive' ne fait pas consensus. Les débats autour d'elle peuvent être analysés comme l'émergence d'une controverse dont la genèse et la trajectoire méritent d'être éclaircies.

Les controverses sont parties intégrantes du processus de validation scientifique. C'est par l'échange d'arguments que les pairs évaluent le caractère exact ou erroné d'une affirmation théorique, d'une démonstration. Mais, de plus en plus, les controverses sortent des espaces et forums spécialisés pour déborder dans l'espace public. La notion de controverse est maintenant bien implantée en science politique, où elle est utilisée pour comprendre les mécanismes de mise en visibilité et de mise en politique des problèmes publics (Barthe, 2006 ; Borrás, 2008). Les controverses sont « des séquences de discussion et d'affrontement entre des points de vue divergents sur un sujet. Elles sont aujourd'hui considérées comme des temps d'exploration et de stabilisation des enjeux durant lesquels la diversité des dimensions, la pluralité des acteurs engagés et des voies d'action possibles sont envisagées avant la clôture politique » (Lascoumes, 2019 : 170). Dans la controverse, les acteurs définissent leur place et leur rôle, affinent leur positionnement, sélectionnent et affutent leurs arguments. Ils repèrent d'autres acteurs, créent des alliances, les recomposent. La controverse peut dans certains cas donner lieu à la constitution de « forums hybrides » (Barthe, Callon et Lascoumes, 2001) où différents publics et différents types d'expertises et de savoirs se côtoient, se confrontent, et parfois s'articulent à travers des procédures de concertation.

L'enjeu d'étudier ces moments d'expression des désaccords a d'abord été souligné par la sociologie des sciences et techniques – et parfois discutée aussi, voir (Lamy, 2017) – puis repris par l'analyse des problèmes publics, en particulier sous l'angle des politiques du risque (santé publique, déchets nucléaires, OGM, etc.) et d'enjeux touchant aux choix

technologiques, à leur plus ou moins grande réversibilité, aux incertitudes liées à leurs effets et à leur mise en délibération.

C'est une des propositions de ce travail que d'étudier les processus de mise en visibilité des enjeux de la 'justice prédictive' que déploient les acteurs, sachant que tous ne se mobilisent pas de la même manière. La controverse porte sur le caractère plus ou moins désirable du développement de la 'justice prédictive', c'est-à-dire sur ses enjeux techniques, juridiques, éthiques, économiques, politiques.

Une série de questions se posent en amont et en aval de ces dispositifs. Comment récupérer et garantir la qualité des données qui alimentent ces algorithmes ? Qui peut y avoir accès ; sous quels formats ? Comment les algorithmes sont-ils paramétrés ? Pour réaliser quelles opérations et activités ? À destination de quels publics (magistrats, avocats ?) et en requérant quel niveau de spécialisation ? Une régulation publique de ces dispositifs est-elle prévue, quels en sont les acteurs, vecteurs ? À quelle conception de la jurisprudence ces algorithmes sont-ils arrimés (Serverin, 1993) ?

Les acteurs de la controverse débattent de ces sujets ; testent la solidité des arguments et la validité des démonstrations ; évaluent la plausibilité des projections et scénarii existants, explorent les angles morts. Ces questions ne sont jamais strictement juridiques ou éthiques ou techniques ou politiques ou économiques, et la manière dont il y est répondu, non plus. Au fil de la controverse s'élaborent des agencements complexes d'éléments hétérogènes cimentés entre eux. Entrer dans la controverse, c'est rejoindre le « point de fusion » (Callon, 1981) pour saisir cet agglomérat qui, lorsque la question sera refroidie, sera passé de l'état liquide à solide, expurgé de certains éléments, orienté dans une certaine direction aux dépens d'autres, plus ou moins durablement écartées. Les controverses sociotechniques sont d'autant plus intéressantes que les objets technologiques sur lesquels elles portent ne sont pas encore validés, réifiés : « les négociations sont multiples, la nature des choix encore discutable, les acteurs impliqués nombreux et variés, les exclusions non définitives » (Callon, 1981). Elles donnent accès aux possibles non advenus, aux alternatives écartées.

Les controverses peuvent être plus ou moins confinées. C'est l'enjeu même des acteurs qui y prennent part que de s'efforcer d'ouvrir ou de restreindre l'espace social de la controverse. Les dynamiques de « publicisation et de confinement de la controverse » (Lemieux, 2007 : 200-201) sont intéressantes à étudier de même que celles de montée en généralité au sens de l'élargissement du périmètre des enjeux. En l'occurrence, le terme est lui-même au cœur du débat en France, bien plus que dans les pays anglophones où il est davantage question de *legal analytics* ou *predictive analytics*, expressions plus neutres. Celui de *predictive justice* est employé, mais beaucoup plus rarement qu'en France où il peut sembler avoir pris l'ascendant sur les autres — au moins pour un temps, nous y reviendrons plus loin. Il est en partie le résultat d'une rencontre mal nommée entre l'informatique et le droit. Dans les publications scientifiques d'informatique juridique (*law & informatics*) et d'analyse behavioriste des comportements des juges (*judicial behavior studies*) que prolongent les expérimentations d'apprentissage machine, le mot *predictive*, est utilisé pour désigner le niveau de prévisibilité que certains modèles peuvent atteindre. En français aussi,

l'adjectif prédictif est employé dans la littérature spécialisée de statistique et d'informatique pour décrire des modèles (régressions linéaires, arbres de décision). Plus encore, le Trésor de la langue française définit le verbe prédire comme « annoncer à l'avance un événement par inspiration surnaturelle, par voyance ou prémonition », mais également « par calcul, par raisonnement, par induction ». Donc la notion de prédiction comprend la dimension de modèles mathématiques. Mais à côté de cela, lorsque l'on écoute le bulletin météo pour savoir s'il est raisonnable de sortir sans parapluie, la voix à la radio ne nous parle pas de prédictions, mais des prévisions, même s'il s'agit de modèles mathématiques prédictifs. Ce qui véritablement fait problème, dans le contexte francophone, c'est d'accoler cet adjectif au nom de justice.

Nous verrons que les débats portent à la fois sur la pertinence du terme de 'justice prédictive', sur ce qu'elle désigne exactement ainsi que sur les effets réels qu'elle pourrait produire sur le droit et la justice. Ils sont constitutifs d'une controverse qui s'inscrit principalement dans le monde du droit et des juristes, mais déborde, bien que plus timidement, dans l'espace public et médiatique. Les questions qui méritent d'être posées sont les suivantes.

- D'où vient ce construit sémantique, où, comment, par qui est-il forgé ?
- Que recouvre-t-il ? À quels discours et représentations du jugement et de ce que c'est que juger ; à quelle conception des métiers d'avocat et de magistrat, à quelle vision des algorithmes et dispositifs sociotechniques est-il associé ?
- Comment se structure et s'organise cette controverse ? Autour de quels acteurs, de quelles alliances ? Quels en sont les temps forts ?
- La justice est marquée par une tension forte entre autonomie et contrôle qu'incarne le principe institutionnel de l'indépendance des juges. Dans quelle mesure cette expression, la façon dont elle est utilisée, mais aussi les arguments mobilisés dans la controverse rendent-ils compte — et selon quelles modalités — de cette tension ?
- Quels sont la portée et les effets de cette controverse sur la manière de penser la rencontre entre algorithmes et justice ? Cette expression parvient-elle à s'imposer, voire à devenir le label qui désigne la rencontre entre jugement juridictionnel et intelligence artificielle ? Devient-elle, ou pas, une catégorie de l'action publique ?

Analyser les controverses suppose de mener un méticuleux travail d'enquête de terrain, qui n'est pas complètement réalisé ici. En effet, ce manuscrit est le produit d'un travail en cours. Les résultats présentés demandent à être consolidés d'une part par des investigations complémentaires et d'autre part par l'exploitation de données d'enquête déjà collectées, mais pas utilisées dans le cadre du présent manuscrit.

3.3. L'analyse de dispositifs algorithmiques de traitement des décisions de justice

La 'justice prédictive' est donc un terme qui pose question. Elle ne constitue pas une catégorie d'analyse, mais un objet de recherche à déconstruire et reconstruire. Pour ce faire, je m'appuie sur le concept de dispositif, tel qu'il a été proposé par la sociologie de la traduction, à partir de la définition donnée par Michel Foucault. Il désigne un réseau qui relie les éléments d'un « ensemble résolument hétérogène, comportant des discours, des

institutions, des aménagements architecturaux, des décisions réglementaires, des lois, des mesures administratives, des énoncés scientifiques, des propositions philosophiques, morales, philanthropiques » (Foucault, 1994 [1977], 299²⁷). La notion de dispositif désigne des assemblages qui associent différents éléments hétérogènes entre eux et dont l'association confère une puissance et une résistance particulières à l'ensemble. Comme le soulignent Nicolas Dodier et Jeanine Barbot, à l'issue d'une analyse des usages de cette notion dans les sciences sociales, « l'hétérogénéité interne des dispositifs, c'est-à-dire la variété de leurs composantes au regard des grandes catégories d'existants, est une propriété décisive [des dispositifs] » (Dodier et Barbot, 2016 : 422).

Un dispositif sociotechnique est descriptible sous chacune de ses composantes, mais sa spécificité repose sur la combinaison de ces différentes dimensions. C'est une des caractéristiques essentielles d'un dispositif. Je propose de considérer la 'justice prédictive' comme un ensemble de dispositifs disparates qui associent plusieurs dimensions à examiner pour saisir en quoi ils sont neufs, en quoi ils recomposent la manière d'accéder à l'information juridique. Dans un précédent travail mené avec Christian Licoppe (2019), nous les avons qualifiés de « dispositifs algorithmiques de traitement et d'analyse des décisions de justice ».

Ils relèvent de la mise en relation d'éléments hétérogènes, humains et non humains, matériels et langagiers entre lesquels sont établis des points de passage. Toute une série d'objets et de technologies numériques (logiciels, serveurs informatiques, écrans de datavisualisation, etc.), de catégories et notions juridiques (la publicité d'une décision de justice par exemple), de normes techniques, de mots, de contextes d'activité (celui du cabinet d'un avocat ou d'un magistrat) sont associés. Leur encapsulage définit les contours et modalités de ce qui pourra être désigné comme la 'justice prédictive', sous réserve que les acteurs qui le portent parviennent à l'imposer au terme de la controverse. Cette notion met l'accent sur l'importance des discours, des mots et des récits qui sont produits, tout en tenant compte de la matérialité des objets et des aspects techniques, informatiques, etc. Or, dans la controverse, la pertinence du terme lui-même est interrogée et contestée.

La matière première de ces dispositifs est faite de décisions de justice ; certaines plateformes compilent également d'autres sources juridiques (textes de loi, commentaires de doctrine, etc.). L'enjeu déborde donc parfois sur l'information juridique plus largement. Par ailleurs, l'expression dispositifs algorithmiques de traitement et d'analyse des décisions de justice est longue et lourde. Pour des raisons de style et compte tenu de la froideur du sigle (Datadj), elle sera parfois remplacée par d'autres termes comme les algorithmes, les plateformes, les logiciels ou bien encore les dispositifs de 'justice prédictive'.

3.4. Justice pénale ou civile ?

À ce jour, en France et contrairement à ce qui se passe par exemple depuis les années 1980 aux États-Unis, les dispositifs algorithmiques étudiés ne portent quasiment que sur des décisions civiles puisque la Chancellerie ne dispose pas d'une base de données toute faite des décisions d'appel en matière pénale contrairement à la matière civile (Jurica). La

²⁷ Voir également le numéro thématique « Dispositifs » de la revue *Terrains et Travaux*, vol. 11, n° 2, 2006.

situation devrait évoluer rapidement d’abord parce que le cadre juridique sur l’*open data* des décisions de justice a été refondu et qu’un calendrier d’ouverture des décisions pénales d’appel et de première instance est désormais fixé (cf. *supra*). Ensuite, les enjeux économiques sont considérables, notamment pour les acteurs nationaux du marché de l’information juridique. Ils sont en effet engagés dans une concurrence internationale où ils affrontent les géants de l’information et de l’édition juridiques, qui ont accumulé depuis fort longtemps de l’expérience sur le traitement des données en matière pénale. Cela est d’autant plus crucial pour les *startup* qui n’ont pas (encore ?) été rachetées par l’un de ces grands groupes tels que [Brandt] (groupe néerlandais) ou [James] (groupe américain). Enfin, on observe déjà que, ponctuellement, des opérateurs nationaux développent des modules en matière pénale, à la demande de leurs clients²⁸.

Ce manuscrit porte donc principalement sur la justice civile même s’il comporte une ligne de fuite vers le pénal. Si jusque-là, les dispositifs algorithmiques étudiés concernent essentiellement des décisions de justice civile, les arguments de la controverse ne se limitent pas à l’évocation de cette matière. Au contraire, ils nourrissent la confusion. Les dispositifs algorithmiques de traitement et d’analyse des jugements et arrêts civils sont comparés, voire assimilés aux dispositifs existants en matière pénale. L’expression ‘justice prédictive’ est elle-même parfois utilisée pour faire référence aux dispositifs en vigueur principalement aux États-Unis pour le pénal comme le logiciel Compas, initiant un glissement vers la notion de justice actuarielle.

3.5. Deux enjeux de connaissance

Deux principaux enjeux de connaissance traversent ce travail. Le premier concerne la production et la trajectoire des réformes dans la justice ; le second porte sur la fabrique du jugement en justice.

3.5.1. Production, trajectoire des réformes et désingularisation de la justice ?

La justice est à la fois un secteur de l’action publique, une institution qui comporte des normes et des rôles ainsi qu’une organisation dans laquelle interagissent différents groupes professionnels. La force du droit, le caractère régalien de cette institution, son intense ritualisation et son fonctionnement en bureaucratie professionnelle, sont le terreau à partir duquel les mondes judiciaires défendent historiquement une identité et une exceptionnalité judiciaires (Commaille, 2000). Or, la justice n’échappe pas au mouvement de « gestionnarisation » (Vigour, 2019) qui touche les différentes institutions étatiques que sont l’école, l’enseignement supérieur et la recherche, la santé, en France comme au plan international (Hood, 1991 ; Pollitt et Boukaert, 2004). Ces réformes managériales « s’intéressent à la manière dont l’administration et le gouvernement remplissent leurs missions plus qu’au contenu de celles-ci. Elles reposent pour partie sur des rhétoriques, des savoirs et savoir-faire supposés universels et transposables à n’importe quel pays ou secteur d’action publique » (Vigour, 2019 : introduction). Une logique gestionnaire « caractérisée par l’orientation de l’action en fonction des coûts, de l’efficacité et de l’efficacité » (*ibid.*) a touché

²⁸ Source : entretiens réalisés les 17 février et 18 mai 2022, avec [Anajur].

la justice française à travers différents types de dispositifs : mise en place de tableaux de bord d'activité juridiction par juridiction, raisonnement en stocks et flux de dossiers, élaboration d'indicateurs de productivité et introduction de primes au rendement (Chelle, 2011), standardisation et certification des procédures via les normes type ISO, etc. Les technologies sont souvent au cœur de ce processus, en France comme ailleurs (Contini et Lanzara 2009 ; Fabri et Langbroek, 2000) et les justiciables eux, plutôt, à la marge.

Comment les réalisations et les débats relatifs aux dispositifs algorithmiques dans la justice s'inscrivent-ils dans ce contexte de transformation gestionnaire de ce secteur ? Le développement de la 'justice prédictive' relève-t-il du même type de logique ? Que peut-il nous apprendre sur la manière dont les réformes en matière de technologies sont introduites et conduites dans la justice ?

Traditionnellement, deux idéaux types de trajectoires d'innovation informatique et technologique coexistent dans la justice. Le premier repose sur de grands projets d'infrastructure informatique portés par les administrations centrales et leurs DSI. C'est le cas des premiers grands chantiers d'informatisation de la justice ; de la chaîne pénale Cassiopée (Féry, 2015). Le second procède d'innovations locales, spécifiques, portées par des magistrats et greffiers de terrain, parfois chefs de juridictions ou de greffes, qui les implantent là où ils sont en poste. Les « applicatifs locaux », des logiciels « maisons » de suivi des dossiers ont été conçus ainsi et le bilan, dans les années 2000, était celui d'une grande disparité des logiciels en fonctionnement dans les différentes juridictions françaises. Le même constat pouvait être dressé pour les trames informatisées de jugement. La notion de « carte changeante des innovations » proposée par Ackermann et Bastard (1993) vise à signifier que souvent, ces innovations se déplacent au gré des mutations de leurs promoteurs.

L'institutionnalisation de ces innovations locales peut donc s'avérer difficile sauf si elles parviennent d'une manière ou d'une autre à remonter jusqu'au niveau central pour inspirer des réformes, comme ce fut le cas pour la visioconférence utilisée dans le cadre d'audiences à distance : innovation hyper localisée devenue solution d'action publique largement promue et utilisée (Dumoulin et Licoppe, 2017). Dans ces deux modèles, la mise en œuvre du projet, c'est-à-dire la conception de l'architecture informatique, est confiée à des prestataires privés qui maîtrisent donc la traduction informatique du projet initial. Ils interviennent comme prestataires, dans le cadre d'une sous-traitance portant sur différents aspects (conception des systèmes techniques, programmation, équipements, formations), en fonction des cahiers des charges, dont le contenu, l'ampleur et les fonds engagés peuvent être très variables.

On peut se demander ce qu'il en est concernant le développement des dispositifs algorithmiques de traitement et d'analyse des décisions de justice. D'où vient l'expérimentation [PréviCompute] ? Comment est-elle née ? À l'initiative de quels acteurs, locaux et/ou centraux ; judiciaires et/ou administratifs. Qu'est-ce qui l'a rendue possible ? La première direction de l'investigation consiste à porter attention à l'historicité, à la genèse et à la trajectoire de la notion de 'justice prédictive' et aux liens qu'elle entretient avec l'expérimentation [PréviCompute].

Une des questions connexes qui émerge, même si elle reste probablement un peu sous-jacente, faute d'avoir pu encore lui donner sa pleine mesure, concerne le rôle des juristes dans l'explosion de la 'justice prédictive'. On sait en effet la puissance du verbe et le pouvoir d'un discours qui présente des transformations qu'il contribue à faire exister et advenir, à travers l'énoncé de prophéties autoréalisatrices, en particulier concernant le développement des technologies. On connaît également le rôle que les juristes académiques ont joué dans l'élaboration d'outils et de catégories analytiques pour penser les transformations juridiques et sociopolitiques, par exemple via la notion de *judicialization* (Commaille et Dumoulin, 2009). Comment les chercheurs en droit interviennent-ils dans les débats sur la 'justice prédictive' ? Dans quelle mesure la scène académique est-elle un lieu où se négocie la valeur de cette construction sémantique en émergence et quelles spécialités juridiques y sont plus sensibles (droit des affaires, droit de la famille, etc.) ? Quelles places respectives des juristes universitaires et des praticiens dans ce débat ?

3.5.2. *Décision, jugement et algorithmes : fabrique et processus du jugement en justice*

Un autre enjeu de connaissance de ce manuscrit a trait au jugement. Le mot est polysémique dans le langage courant. Il est utilisé par métonymie pour désigner l'ensemble des décisions juridictionnelles. D'un point de vue juridique, il ne s'applique, *stricto sensu*, qu'aux décisions de première instance. En appel et cassation, le terme précis est celui d'arrêt tandis que celui d'ordonnance est réservé aux décisions prononcées par des instances spécifiques comme le juge aux affaires familiales, le juge d'instruction, le parquet, etc. Le terme de « jugement » recouvre ensuite le processus de prise de décision (tout ou partie des séquences du *decision making*) et l'aboutissement de celui-ci (les mesures prises par une juridiction). Enfin, quand il est utilisé pour évoquer le résultat du processus juridictionnel, il s'applique à la matérialité de l'écrit, concrètement le document qui s'intitule « jugement du tribunal judiciaire de Rouen – 1^{re} Chambre » par exemple. En matière civile, ce texte est généralement composé de trois blocs qui sont juridiquement distincts : les moyens et prétentions des parties (ce qu'elles demandent, sur la base de quels fondements) ; les motifs de la décision (l'argumentaire qui fonde les mesures prises par la juridiction) et son dispositif (le contenu des mesures arrêtées). Le langage ordinaire confond souvent ces deux dernières notions, bien distinctes d'un point de vue doctrinal, mais reliées par une logique argumentative et stylistique.

Comprendre et analyser le développement de la notion de 'justice prédictive' et les abondants discours qu'elle suscite, donne l'occasion de mener une réflexion sur la manière dont les jugements sont pensés et décrits. Dans le fonctionnement des algorithmes, plusieurs réalités empiriques se côtoient. Les jugements dont il est question sont d'une part ceux qui fournissent des données à partir desquelles les algorithmes sont entraînés ou établissent des statistiques, des modèles, etc., et d'autre part ceux qui sont susceptibles d'être équipés par le dispositif algorithmique, dans le cadre d'expérimentations comme celle de [PréviCompute]. Il s'agit donc de jugements marqués par deux temporalités distinctes : certains sont saisis après coup, une fois advenus et d'autres en cours de fabrique. Mais puisque ces dispositifs sont aussi destinés aux avocats, juristes d'entreprise, sociétés d'assurance — en réalité leurs principaux

utilisateurs actuels — ils impliquent également des décisions non juridictionnelles (aller au contentieux ou favoriser un règlement amiable du litige).

Ce manuscrit remet en question la ou les conceptions des jugements que les discours sur la ‘justice prédictive’ contribuent à véhiculer en contexte juridique et judiciaire. Comment la question du jugement est-elle thématifiée ? Cette controverse n’est-elle pas également, et peut-être même avant tout, une dispute quant à la manière de concevoir l’activité de jugement, le rôle des magistrats et leur périmètre d’intervention ? Ne met-elle pas en exergue plusieurs options normatives : processus automatisable ou comportant une part d’indétermination à préserver ? Monopole des magistrats ou partageable avec d’autres acteurs, voire en partie susceptible d’être déléguée ? À quelles philosophies du jugement la ‘justice prédictive’ est-elle adossée ? Est-elle présentée comme une forme suprême de rationalisation du droit et/ou un processus qui contribue à rendre les décisions moins explicables ? Quels liens entretient-elle avec les enjeux de recherche d’efficacité des juridictions, de désencombrement de celles-ci ?

Historiquement, la rencontre entre l’informatique numérique et la justice est d’abord l’œuvre de marginaux sécants. Puis, la reprise en main au plan central (Dumoulin *et al.*, 2007) s’effectue via un référentiel gestionnaire qui endogénéise les outils informatiques pour mettre en place un projet de rationalisation et de réorganisation de la justice, centré sur l’efficacité. La chaîne pénale Cassiopée (Féry, 2015) comme le projet de chaîne civile Portalis, encore en cours, et d’autres grandes infrastructures d’information du ministère de la Justice²⁹ sont des entreprises à fort contenu technique et technologique dont l’enjeu de réorganisation et de remontée de l’information vers l’administration centrale est très clair. Construire ces grandes infrastructures est un des moyens de faire de la politique en la technicisant (*ibid.*). Le projet Cassiopée a consisté à concevoir un système centralisé conçu pour être interopérable tout au long de la chaîne pénale et assurer la circulation en temps réel des données généralement selon le schéma *bottom up*. La décision du projet intervient au niveau central, sans concertation des juridictions et avec l’objectif de disposer d’informations statistiques fiables, comparables et autorisant un pilotage de plus en plus centré sur une approche de la justice comme organisation à gérer, en matière de stocks, de flux, de délais, etc.

Le cas des audiences par visioconférence témoigne également de la prégnance que peuvent acquérir des considérations et enjeux gestionnaires, au point que des fondamentaux de l’audience (unité de lieu, de temps et d’action du drame judiciaire, droits de la défense, etc.) leur soient en partie subordonnés. Cette innovation organisationnelle, à partir d’une technologie de visiophonie, est née et s’est développée dans une perspective de résolution de problème, qui s’est ensuite doublée d’une logique managériale. L’enrôlement dans la logique gestionnaire s’est traduit par la formulation d’un programme de développement contraignant, quantifié et planifié, visant à évaluer la réalisation d’objectifs et de performances (Bezès, 2007 : 19), en l’occurrence la multiplication des comparutions à distance de personnes détenues, en lieu et place de comparutions corps présents pour

²⁹ (Jonkheere, 2013 sur l’informatisation des maisons de justice ; Larminat, 2013 sur l’informatisation des services d’insertion et de probation de l’administration pénitentiaire).

permettre une réduction du nombre des escortes judiciaires (Dumoulin et Licoppe, 2015). Dans ce processus, la logique proprement juridique et judiciaire autour de laquelle l'institution est structurée est passée au second plan (Dumoulin et Licoppe, 2016). L'introduction d'un dispositif sociotechnique tel que la vidéoconférence, parce qu'il crée une écologie fracturée et transforme la manière dont les interactions sont accomplies, a impliqué des négociations avec l'univers très institutionnalisé de la justice.

Qu'en est-il pour les dispositifs algorithmiques ici étudiés ? Comment agissent-ils ou sont-ils mobilisés dans cette tension entre justice comme organisation à gérer et justice comme institution porteuse d'un ensemble de valeurs et de normes spécifiques ? La question mérite d'autant plus d'être posée qu'elle concerne l'acte de juger, et donc le cœur de l'activité et de l'identité des magistrats (Cahiers de la justice, 2015). Comme l'audience publique, l'acte de juger condense la mission et le sens de l'exercice juridictionnel et incarne le principe cardinal de l'institution judiciaire qu'est l'indépendance à l'égard des pouvoirs, en particulier politique, mais aussi économique.

IV / Des terrains à géométrie et granulométrie variables

Le manuscrit repose sur une série de terrains dont la géométrie et la granulométrie sont variables. Toutefois, tous sont pensés pour servir l'objectif de recherche de ce mémoire d'HDR et administrer la preuve empirique de ce qui est avancé. Plusieurs corpus de différentes natures sont associés : articles de presse, *tweets*, archives et traces. À cela s'ajoutent, très marginalement, quelques observations et entretiens (réalisés, mais pas toujours mobilisés).

4.1. Articles de presse

Les articles de journaux sont un premier type de matériau collecté pour analyser le traitement de la 'justice prédictive' dans l'espace public. Il s'agit d'articles parus dans une série de titres de la presse française de 2010 à 2020 et comportant les mots « prédictif » ou « prédictive ». Sept journaux français ont été retenus pour la diversité de leurs positionnements, périmètres et formats : trois titres de la presse quotidienne nationale (*Le Figaro*, *Le Monde*, *Libération*), deux de la presse quotidienne régionale³⁰ (*Ouest-France*, *La Voix du Nord*) et deux hebdomadaires nationaux (*L'Obs* et *Le Point*). Pour chacun, les éditions papier et numérique ont été prises en compte. Sur la période 2010-2020, 1 622 articles ont été recueillis via l'interrogation de la base de données Europresse.

Au sein de ce corpus a ensuite été isolé un sous-corpus 'justice prédictive' qui comprend donc les seuls papiers comportant cette expression exacte³¹, au nombre de 122. Corpus et sous-corpus ont fait l'objet d'une analyse de sémantique lexicale visant à rendre

³⁰ Choisis parce qu'ils couvrent le ressort des deux cours d'appel françaises (Douai et Rennes) ayant participé à l'expérimentation du dispositif de 'justice prédictive' [PréviCompute].

³¹ Une relecture manuelle a permis de réintégrer dans le sous-corpus les quelques articles qui questionnent explicitement cette notion, mais qu'une recherche automatique avait écartés en raison de constructions grammaticales spécifiques (par exemple la justice qualifiée de prédictive). Ces articles ont été ajoutés au sous-corpus.

intelligible ce qui s'écrit sur la 'justice prédictive', en identifiant des univers de sens et la manière dont ils sont structurés, polarisés. Le logiciel d'analyse de données textuelles Iramuteq — version libre d'*Alceste* (méthode Reinert) — permet d'effectuer une analyse lexicométrique et sémiologique du corpus à partir d'une classification hiérarchique descendante puis ascendante. Autrement dit, « le logiciel fractionne de façon successive le texte et en extrait des classes de mots représentatives » (Delavigne, 2003) pour déterminer des univers de sens qui peuvent être ensuite mis en lien avec différentes variables (le lieu, la période de publication, etc.). L'analyse qualitative approfondie, via un logiciel CAQDAS du type de N-vivo, interviendra plus tard, dans l'ANR Just-IA³².

4.2. *Tweets*

Sur la même période de dix ans (2010-20), un corpus de *tweets* a également été constitué autour de la présence conjointe des termes « justice » et « prédictive » sur la plateforme Twitter. Il complète la gamme des discours produits sur la 'justice prédictive', tout en repérant les types d'acteurs, institutions, communautés et réseaux qui mobilisent le terme et le discutent. Les réseaux sociaux et les données qu'ils contiennent sont une source importante pour les sciences sociales (Boullier, 2015 ; Boyadjian, 2017). 6261 *tweets*, mentions et *retweets* ont été récupérés, répondant à différents mots-clés et hashtags. Une double collecte a été réalisée : synchrone par l'application DMI-TCAT lancée à partir de janvier 2020 pour couvrir l'année 2020 et asynchrone sur la période 2010-2020 à partir du script *Python GetOldTweets3*, qui consiste à scraper la page affichée lors de l'utilisation du moteur de recherche de Twitter. Ces données ont ensuite fait l'objet d'opérations de nettoyage, d'encodage et d'enrichissement, pour constituer une base de données, qui ne comprend que des *tweets* publics, publiés par des comptes eux-mêmes publics, en tenant compte des dispositions du RGPD³³.

La contribution de Florent Frasque, étudiant de M1 puis de M2, à Sciences Po Grenoble, a été décisive pour cette partie très technique que je n'aurais pu accomplir seule et pour laquelle il a bénéficié des conseils de Gilles Bastin, professeur de sociologie et d'Emmanuel Marty, maître de conférences en sciences de l'information et de la communication, dans l'environnement favorable de la chaire « Sociétés algorithmiques » de l'Institut MIAI Grenoble-Alpes. Il a complété la collecte d'articles de presse que j'avais commencée et a effectué seul la collecte, le nettoyage, l'enrichissement et la première analyse de la base de données de *tweets*³⁴.

Le choix de la plateforme Twitter — et non d'un autre réseau social numérique — est lié à ses caractéristiques telles que repérées par la littérature : les catégories

³² À ce stade, seule une analyse manuelle superficielle a été réalisée ; elle n'est pas véritablement restituée ici, seulement évoquée ponctuellement.

³³ Ces techniques ont été testées d'abord sur 200 *tweets*, grâce à un financement exploratoire du Data-Institute et de la chaire Société algorithmique MIAI-Université Grenoble Alpes puis réalisées en intégralité dans le cadre de l'ANR Just-IA.

³⁴ Je ne saurais assez le remercier de son travail si précieux d'assistant de recherche effectué dans le cadre d'un stage et de vacances financé par la chaire Société algorithmique, le Data Institute de Grenoble et l'ANR Just-IA.

socioprofessionnelles les plus favorisées y sont surreprésentées, elle est principalement tournée vers l'actualité politique, institutionnelle, médiatique et culturelle. Elle est utilisée comme un vecteur de communication (institutionnelle notamment), un moyen d'information et constitue un espace conversationnel (Cardon et Prieur, 2016 : 244-47) propre à l'émergence de controverses que « *l'usage possible des hashtags* » renforce en tant qu'ils fonctionnent « *comme balises d'un affrontement autour de la définition de l'objet de débat* » (Bottini et Julliard, 2017). Les *tweets* se prêtent bien à une analyse textuelle lexicométrique et sémiologique assistée par Iramuteq, mais aussi à une analyse de réseaux via les *retweets* et mentions. L'analyse de réseaux permet d'identifier les leaders d'opinion en la matière, les groupes qui se constituent, et donc de définir des communautés ; cette information est ensuite introduite comme une nouvelle variable. Après l'analyse du *tweet*, les corrélations entre classe de discours et appartenance à une communauté peuvent être mises en évidence (Smyrnaio et Ratinaud, 2013). Ces résultats peuvent être réintroduits dans une analyse de contenu et croisés avec elle, ce qui sera fait ultérieurement.

4.3. Archives et traces

Retracer la trajectoire de la controverse sur la 'justice prédictive' suppose de suivre les discours — et celles et ceux qui les portent — là où ils nous emmènent et d'en saisir les enchaînements. Comprendre comment l'expérimentation [PréviCompute] est née et ce qui l'a rendue possible implique là encore de démêler l'écheveau et suivre les fils, notamment en remontant le temps.

Un travail de reconstitution des séquences et des acteurs de cette histoire a donc été effectué à partir d'un ensemble hétérogène de traces : textes juridiques et travaux parlementaires, veille sur les textes juridiques et projets de réforme, en particulier autour du projet de décret sur l'anonymisation des décisions de justice, rapports officiels (Cadiet, Villani, Cnil, etc.), production écrite des acteurs institutionnels pertinents (Cepej, Dila, Etalab, Open Law, Juriconnexion, Cour de cassation, Conseil d'État, etc.) via leur communication publique (sites Internet) ou leurs listes de diffusion quand il est possible de s'y abonner (Juriconnexion), etc.

Si le grand angle était essentiel pour reconstituer l'ampleur du paysage, zoomer sur les discours et positionnements des concepteurs de dispositifs algorithmiques était également indispensable du fait de leur rôle dans la controverse sur la 'justice prédictive'. [PréviCompute] qui a donné lieu à une expérimentation a été analysé de près, de même qu'[Anajur]. Une analyse systématique de leurs sites Internet, newsletters, documentation papier, guides, etc. a été réalisée.

4.4. Entretiens semi-directifs et observations

Les entretiens et observations ethnographiques sont peu présents dans ce manuscrit. Six entretiens semi-directifs menés par Christian Licoppe autour de l'expérimentation [PréviCompute] informent la compréhension de ce qui se joue, mais ne sont pas utilisés en tant que tels dans le présent manuscrit pour ne pas brouiller les apports respectifs de ce travail et d'un article précédemment co-écrit (Licoppe et Dumoulin, 2019). Quelques entretiens avec une autre société qui propose des algorithmes de traitements des décisions de justice [Anajur]

ont été réalisés, mais n’ont pas nourri la démonstration ici présentée, en raison de contraintes liées à leur validation par les enquêtés — condition de l’accord pour accéder à certains terrains où les acteurs sont pour le moins frileux. Ils informent toutefois mes interprétations. Surtout, une campagne d’une trentaine d’entretiens auprès des *stakeholders* de cette histoire est prévue dans l’ANR Just-IA pour compléter le travail.

La dimension ethnographique sera pleinement développée dans le cadre de l’ANR Just-IA afin d’accéder aux usages concrets des dispositifs algorithmiques par les magistrats, avocats et autres acteurs du processus juridictionnel (greffiers et greffières, etc.).

Une enquête ethnographique est amorcée avec l’observation de la manifestation du 4^e Village de la *Legaltech* pendant deux jours, restituée dans le chapitre 3. Outre les conférences, les démonstrations et le *off* du salon, la visite des stands en place et les discussions avec les sociétés opératrices d’algorithmes de traitement des décisions de justice représentées, ont été riches. J’ai également recueilli de nombreux documents, plaquettes, affiches, etc. à cette occasion. Le protocole de recherche ici restitué est donc destiné à s’étoffer et se diversifier dans le cadre de l’ANR Just-IA.

4.5. Anonymiser oui, mais comment ?

Le sujet de la ‘justice prédictive’ est particulièrement chaud (même s’il a un peu refroidi ces derniers temps), ce qui rend la question de l’anonymisation cruciale. On considère généralement que l’anonymisation ne se justifie pas nécessairement pour tout ce qui est dans le domaine public. Ce n’est pas exactement la solution retenue ici, et il est important d’expliquer pourquoi les opérateurs de dispositifs algorithmiques de traitement et d’analyse des décisions de justice, qu’il s’agisse des entreprises ou de leurs salariés et dirigeants, sont mentionnés à travers des alias.

C’est un choix coûteux formellement. Mais c’est une nécessité pour limiter les risques d’interférences avec la suite, car ce manuscrit propose un « arrêt sur image » dans le cadre d’une recherche en cours qu’il est important de perturber le moins possible. De plus, l’anonymisation est un marqueur de mon positionnement et une façon de mettre en pratique ma revendication d’équidistance avec tous les acteurs de cette histoire. Vu l’actualité et la sensibilité de ces sujets, c’est selon moi la meilleure manière de concilier des perspectives et intérêts en partie divergents. En l’occurrence, je tiens à respecter les enjeux propres de mes enquêtés — dans un contexte commercial et concurrentiel en particulier — tout en conservant ma propre liberté d’analyse et en limitant les possibles « retours de flamme » (*i.e.* éviter de « griller » le terrain pour moi-même ou la communauté des chercheur.es, comme divers travaux précédents ont pu le faire).

Certes, ceux qui connaissent le sujet reconnaîtront facilement certains acteurs, notamment les protagonistes de cette histoire, plus faciles à identifier. Mais ce n’est pas le cas de tous. L’argument d’une anonymisation imparfaite, au sens où des réidentifications sont possibles, ne me semble pas rédhibitoire. Travailler en mode anonymisé peut aussi aider à ouvrir certains terrains (les plus confidentiels, auprès de *startup* encore peu connues et pas hostiles à la recherche en sciences humaines et sociales, et ce qu’elle peut leur apporter) pour la suite de cette recherche. De plus, mon horizon est également celui de la conformité avec le

RGPD ; en cas de problème, ce que la CNIL considère, c'est la bonne foi des chercheurs et le fait qu'ils ont fait à leur niveau ce qu'ils peuvent pour protéger les données des personnes. Je travaille surtout avec des données publiques, mais aussi avec des *tweets*, eux-mêmes publics. L'enjeu réside surtout du côté des observations (celles que j'ai faites pour ce manuscrit et les autres à venir) et plus encore des entretiens (que je vais recueillir dans l'ANR). Une stratégie de pseudonymisation conçue depuis le démarrage d'une recherche comme celle-ci me semble importante. C'est la raison pour laquelle dès l'article coécrit avec Christian Licoppe (2019), j'ai proposé de recourir à des alias ; même pour une expérimentation, dont les caractéristiques rendaient les acteurs assez vite identifiables pour qui connaissait le sujet ou souhaitait procéder à une réidentification.

Dans ce manuscrit, j'ai donc transformé tous les noms des sociétés étudiées (opérateurs d'algorithmes de traitement des décisions de justice), entreprises et acteurs directement concernés par des alias indiqués entre crochets, du type de [PréviCompute], y compris dans les références précises (titres d'articles de presse en ligne par exemple pour faire disparaître tout élément immédiatement identifiant). J'ai fait de même pour les articles de doctrine écrits par les leaders de *startup*, dont j'ai anonymisé les noms en bibliographie. Ils sont repérables car toujours entre crochets : par exemple [Philippe Sinclair]. En général, les structures publiques restent identifiées : Open Law, Etalab, etc. Certaines personnes également, lorsqu'elles sont publiques, car je présente leurs parcours à partir de données elles aussi publiques, accessibles à tout utilisateur d'Internet (et non tirées d'entretiens par exemple).

V / Annonce du plan : le contenu des chapitres

Ce manuscrit associe une revue de littérature sur la 'justice prédictive' et plus largement sur le thème droit, justice et algorithmes ; une proposition théorique d'ensemble de sociologie politique qui considère le jugement en justice comme une activité pratique équipée par une pluralité de dispositifs sociotechniques ; des investigations empiriques quantitatives et qualitatives déployées à partir de la collecte d'un corpus de sources, originales et diversifiées (traces et archives, *tweets* et articles de presse). Des premiers résultats de recherche sont restitués, en particulier l'analyse de la controverse sur la 'justice prédictive' et ce qu'elle révèle des logiques de transformation qui traversent actuellement la justice, en tant qu'institution, organisation et secteur d'action publique. La trajectoire que connaît l'expression 'justice prédictive' dans le débat public est décrite empiriquement avec l'identification d'un moment 2017 ; analysée et mise en lien avec l'expérimentation [Previcompute] et les intenses débats que celle-ci a suscités. Cette dernière est interprétée comme l'une des manifestations de l'émergence d'une nouvelle configuration de réforme administrative qui s'est mise en place depuis les années 2010 pour favoriser l'*open data* des décisions de justice et encourager l'innovation dans le droit. Les opérations développées par Open Law en lien avec Etalab et en particulier le programme entrepreneur d'intérêt général en offrent d'autres exemples.

Classiquement consacré à un état de l'art de la production académique, le premier chapitre rend compte de l'abondante littérature consacrée à la 'justice prédictive' et qui s'avère relever surtout du droit et des disciplines afférentes (philosophie du droit, théorie du droit) parfois en lien avec des mathématiques ou des sciences computationnelles. Ces publications *adoptent très majoritairement une posture normative et évaluative*, appréciant de manière générale et surplombante les atouts et limites des algorithmes de traitement des décisions de justice ainsi que leurs conséquences possibles. Un des effets de la multiplication des travaux de doctrine en un temps court est de consolider le terme de 'justice prédictive' qui fonctionne comme un label, alors même que son contenu, les significations et réalités qu'il désigne sont hétérogènes et peu partagées tant parmi les auteurs de doctrine, les praticiens (avocats, magistrats) que parmi les concepteurs d'algorithmes, qui tous — et parfois ensemble — s'expriment dans des espaces et sur des supports de science juridique (colloques, revues, ouvrages collectifs). Les formulations qui consacrent la notion, mais indiquent une prise de distance comme le fait d'utiliser des guillemets ou de mentionner la « justice dite prédictive », peu fréquentes en 2016-2017, deviennent plus courantes dans les publications des années postérieures. Comparativement, rares sont les analyses de sciences sociales, empiriquement nourries — et distancées des *stakeholders* — dans la littérature tant nationale qu'internationale. Quelques exceptions notables s'intéressent toutefois aux perceptions, discours, fabrique et usages de ces dispositifs, à partir d'approches de sociologie des usages, sociologie de l'innovation et sociologie du droit et de l'action publique. C'est dans le prolongement de ces travaux que la présente recherche s'inscrit, afin de faire progresser la connaissance sociologique sur la diffusion rapide de l'expression 'justice prédictive' et ce qu'elle recouvre.

Le second chapitre objective la percée de la 'justice prédictive' dans le débat public. À partir de l'analyse d'un vaste corpus d'articles de presse sur la période 2010-2020, il met en évidence la montée en flèche du terme prédictif.ve dans les discours relatifs à plusieurs secteurs : compétition électorale, santé, automobile, justice. Si les temporalités et connotations sont assez variées, deux transversalités discursives émergent : le caractère prédictif des approches par probabilités et modélisations qui dessinent des projections et des anticipations d'une part ; la délégation (ou le risque de) des activités classiquement accomplies par les humains à des algorithmes, machines ou robots d'autre part. La justice participe donc d'un tournant algorithmique (« prédictif ») qui la dépasse. Pour autant, la comparaison entre ces différents domaines révèle des spécificités dans la manière dont les discours sont produits concernant la justice : la place du monde académique et surtout juridique y est prépondérante de même que l'orientation vers une réflexion sur les enjeux sociétaux, notamment prospectifs. À partir de l'analyse comparée d'articles de presse et de *tweets* sur la seule expression de 'justice prédictive' cette fois, le chapitre confirme alors l'engouement remarquable autour de cette notion et documente sa trajectoire en forme de pic sur les années 2017-2018, stabilisée en 2019-2020 à un niveau supérieur à celui de 2015 dans les médias étudiés. Apparue pour la première fois en 2015 dans la presse et en novembre 2011 sur Twitter, la référence au caractère prédictif de la justice abrite une tension entre deux versants de la 'justice prédictive' : celui qui renvoie à des dispositifs de traitement et d'analyse

des décisions de justice d'une part et celui qui désigne davantage des dispositifs de contrôle et d'anticipation de la commission d'infractions et/ou de récidive dans le cadre de menaces notamment terroristes, et qui renvoie donc davantage à la dimension pénale de la justice d'autre part. En fonction des supports de communication, des locuteurs qui s'expriment, le périmètre de ce que cette expression recouvre varie considérablement.

Le troisième chapitre démontre que l'effet de label acquis par la 'justice prédictive' dans un temps court, tient au fait que cette construction sémantique est soutenue par une nouvelle configuration de réforme administrative. Porteuse de représentations, discours et alliances, en partie nouveaux, celle-ci s'est mise en place, en France, à partir de la seconde moitié des années 2010, en lien avec l'adoption de la loi pour une République numérique — laquelle proclame la mise à disposition publique et gratuite des décisions de justice, condition *sine qua non* du fonctionnement d'algorithmes d'apprentissage machine. Le chapitre revient sur l'enjeu de l'ouverture des décisions de justice. Il montre comment le répertoire de réforme de l'*open data* des données publiques, centré sur l'écosystème de la donnée, a été traduit dans le monde du droit et de la justice. Il identifie de nouveaux acteurs, parmi lesquels Open Law, porteurs de représentations et discours centrés sur le pouvoir transformateur des technologies numériques et les vertus d'une ouverture du droit. Les controverses qui ont lieu autour de la mise en œuvre de la loi dite Lemaire sont analysées comme l'affrontement de deux coalitions de cause, l'une particulariste et l'autre généraliste. La première, en position hégémonique, réunit les directions d'administration centrale du ministère de la Justice, la Cour de cassation, les éditeurs juridiques classiques, certains organes de représentation de la profession d'avocat. Elle milite pour contenir, réguler et encadrer l'*open data* des décisions de justice afin de le subordonner à des logiques spécifiquement juridiques et juridictionnelles. La seconde, en position de challengeur, rassemble Etalab, Open Law, le Village de la justice, et les entrants dans la justice que sont les *startup* et entreprises de *legaltech* ([Jurissimo], [PréviCompute], etc.) dont les fondateurs sont eux-mêmes plutôt des *outsiders* du droit. Elle défend une conception large de l'ouverture des données judiciaires, c'est-à-dire identique à celle des autres données publiques. Ces deux coalitions se distinguent quant à leurs visions de l'*open data* des décisions de justice, mais aussi en raison de systèmes de croyances contrastés.

Au croisement entre sociologie de l'innovation, sociologie de l'action publique et sociologie économique, le quatrième chapitre retrace la genèse, la mise en œuvre et les effets d'une expérimentation qualifiée de 'justice prédictive', issue de cette nouvelle configuration de réforme administrative et d'un contexte institutionnel hybride. Intervenue en 2017 dans les cours d'appel de Rennes et Douai, cette expérience a consisté à faire tester un dispositif algorithmique de traitement et d'analyse des décisions de justice, par quelques magistrats pour des affaires civiles. Cette opération est directement issue du répertoire d'action de l'*open data* porté par Open Law et Etalab. Elle incarne les valeurs et méthodes promues par la coalition des challengers, lesquels tentent de séduire et d'intéresser à leur dispositif, des acteurs installés du système judiciaire. C'est aussi une tentative de déplacer ce dispositif de la cible des avocats vers celle des magistrats. Si [PréviCompute] est une expérience d'ampleur limitée dont la trajectoire semble mener à l'impasse — certains acteurs ayant publiquement

mentionné son échec — les développements et effets de cet épisode sont marquants. L'expérimentation très localisée, mais largement publicisée, suscite et alimente une controverse ; cette épreuve marque un tournant, à la fois pour l'expérience locale, pour la coalition de cause généraliste, traversée par d'importantes divergences, pour la coalition de cause hégémonique qui contre-attaque, et pour la 'justice prédictive'. Elle joue un rôle central dans la cristallisation de ce terme, dans sa contestation et l'émergence de vocables alternatifs ainsi que dans l'ouverture d'un questionnement sur ce qu'est rendre justice, en premier lieu pour les magistrats et magistrates.

Le dernier temps de la réflexion est consacré à analyser la manière dont les jugements en justice sont appréhendés dans la controverse sur la 'justice prédictive'. À partir d'un parallèle entre le cas de l'expertise judiciaire et celui des dispositifs algorithmiques de traitement et d'analyse des décisions de justice est pointée la force de la représentation sociale du jugement comme pure décision du juge. Cet impensé décisionniste transparaît au fil des données recueillies dans le discours des concepteurs d'algorithmes, celui des médias, et les débats sur Twitter. Or, cette représentation souffre d'un décalage avec la réalité du fonctionnement de la justice, telle qu'elle est observée, décrite et analysée par la sociologie. Loin du mythe du juge seul, la pluralité des acteurs, des équipements et appuis matériels de l'action, invite à considérer le jugement au moins autant comme un processus et une activité de production en partie collaborative que comme le résultat de cette action.

Chapitre 1 : Les algorithmes dans la justice, de l'effervescence des expériences à l'analyse de leurs effets

L'effervescence autour de la notion de « justice prédictive » est récente et concerne surtout le monde des juristes, universitaires et praticiens, et ses marges interdisciplinaires. Si elle se développe brusquement dans la seconde moitié des années 2010 en France, c'est en réponse à la publication, dans des revues juridiques internationales ou sur des sites en ligne, d'articles relatant des expérimentations de dispositifs algorithmiques de traitement et d'analyse de grandes quantités de décisions de justice. Les auteurs de ces premiers articles — *i.e.* les équipes pluridisciplinaires ayant réalisé ces expérimentations — se positionnent comme des innovateurs, porteurs d'un discours de rupture, et prêts à contribuer à la transformation effective des systèmes de justice. Ils insistent sur la révolution que représentent ces dispositifs, nés du développement conjoint des capacités de calcul des ordinateurs, du *big data* et de l'intelligence artificielle, par l'apprentissage machine. Ils en soulignent les usages possibles par les juridictions et acteurs de justice et parfois même, les mettent à leur disposition, en les commercialisant par le biais de sociétés privées ou en créant des accès gratuits en ligne. Cet élan qui apparaît comme disruptif, accompagné de son lot d'applications pratiques, suscite d'abondantes réactions de juristes qui s'interrogent sur la « justice prédictive ».

La notion de « justice prédictive » est particulièrement polémique parce qu'elle semble indiquer que *l'on peut prédire les arrêts et jugements (et non pas progresser dans l'analyse de leur prévisibilité)*. Pourtant, il est clair que les dispositifs algorithmiques dont il est question ne sont jamais que des « machines à prévoir » (Cardon et Benbouzid, 2018). Par définition, l'avenir n'est pas écrit et donc jamais parfaitement prédictible. Mais des politiques publiques peuvent viser à ce que ces technologies qui proposent des modèles et des probabilités soient employées pour informer ou déterminer les décisions de justice. La question est alors celle du degré de contrainte à l'utilisation (une information parmi d'autres ou une information qui définit automatiquement la décision ?), de son périmètre (tout ou partie des affaires, selon quels critères ?) ainsi que celle des effets de ces politiques publiques sur les pratiques locales des acteurs (quels usages effectifs ?). Cet enjeu vient ajouter un niveau supplémentaire à la discussion et lui conférer une dimension presque tragique pour bon nombre de juristes, mais également de philosophes : est-on en train de créer les conditions de l'implantation d'une justice automatisée, en lieu et place d'une justice fondée sur le jugement humain – expression du partage par tous les membres d'une communauté de valeurs, d'intérêts, mais aussi d'une commune humanité ?

La controverse se développe principalement selon deux axes. Le premier est celui d'une discussion sur l'apport réel de ces dispositifs algorithmiques. En quoi sont-ils si radicalement nouveaux et en quoi tiennent-ils les promesses qu'ils avancent ? Les équipes pluridisciplinaires qui prétendent déclencher une révolution dans l'analyse, la modélisation et

l'amélioration de la prévisibilité des jugements et arrêts, ont intérêt à défendre, voire survaloriser la qualité et la fiabilité de leurs résultats. Sans se positionner de manière uniforme, elles partagent l'enjeu de convaincre leurs audiences du pouvoir des technologies computationnelles et de la pertinence de celles-ci pour transformer les systèmes judiciaires. Certaines d'entre elles jouent aussi leur place dans l'émergence d'un nouveau marché dont elles attendent des retombées économiques directes. Porteuses de valeurs relatives au changement dans le fonctionnement de la justice doublées d'intérêts explicites à ce que ce changement advienne, elles sont vues notamment par certains spécialistes d'informatique juridique comme des acteurs opportunistes, davantage motivés par la dimension technologique que par les enjeux de la justice, et dont les projets peuvent être dangereux pour le monde judiciaire.

C'est ainsi que le second axe de la controverse se développe autour d'une discussion des potentialités et risques ouverts par ces technologies appliquées à l'exercice de la justice et à la mise en œuvre du droit. Il s'agit là d'une approche typique de la doctrine, orientée vers l'analyse théorique, évaluative et prospective des « potentialités réelles et [...] risques éventuels » ([Desjardins], 2019), des « vices et vertus » (Jacquemin, 2019), des « avantages [...], mais aussi [des] risques » (Malabat, 2019 : 109), des espoirs et des dangers, « ce que l'on a à gagner [...] et à perdre » (Hyde, 2019 : 6). Car « les logiciels de justice prédictive [...] sont porteurs d'autant de promesses que de menaces, de progrès que de risques de régression » (Boré, 2018 : 2) que ces travaux cherchent à mettre en lumière. Cela est valable au plan international où « hypes, hopes and dreams » (Reiling, 2021) sont scrutés.

Quels sont les enjeux juridiques, anthropologiques, philosophiques, moraux et éthiques soulevés par la création et l'utilisation de dispositifs algorithmiques de traitement et d'analyse de décisions de justice dont une part se présente comme permettant de *prédire* ce que les juges décideront ? Cette discussion, fondamentale, profonde et prospective, devient très vite un sujet de préoccupation chez les juristes, relayé en particulier par des revues destinées aux praticiens ou spécialisées dans les technologies (Larrègue, 2019 : 10). D'intenses débats émergent également à l'occasion de conférences et colloques consacrés à la « justice prédictive » (Blanc et Mekki, 2019 ; Clavier, 2020 ; Hubin, Jacquemin, Michaux, 2019 ; Ordre des avocats au Conseil d'État et à la Cour de cassation, 2018 ; Jeuland, 2017)¹ et sont ensuite publiés². En France, dans un premier temps, ce sont surtout les spécialistes de technologies et d'informatique juridique qui s'expriment, puis, progressivement, chacun y va, qui d'une note, qui d'un article. La parution d'un appel à projets du GIP Mission de recherche Droit et Justice

¹ Par exemple, le colloque de la Cour d'appel de Paris « Justice prédictive : évolution, révolution ? » (27 mars 2017) ; la conférence de la société de législation comparée avec l'Ordre des avocats au Conseil d'État et à la Cour de cassation « L'administration de la justice 5.0 » (axe Code & Law 8 décembre 2017) ; colloque organisé par l'IRDA Université Sorbonne Paris Nord, la Cour de cassation et l'ENM (8 juin 2018).

² Blanc Nathalie et Mekki Mustapha (dir.) (2019), *Le juge et le numérique : un défi pour la justice du XXI^e siècle*. Paris, Dalloz, coll. « Thèmes et commentaires » ; Ordre des avocats au Conseil d'État et à la Cour de cassation (2018), *La justice prédictive*, Paris, Dalloz, coll. « Thèmes et commentaires ».

fléché sur le thème « Droit, justice et numérique »³ stimule la réalisation de recherches interdisciplinaires dont deux portent explicitement sur la « justice prédictive » (Godefroy, Lebaron et [Sinclair], 2019 ; Léonard et Moritz, 2020)⁴.

La focalisation de la controverse sur la « justice prédictive » s'explique aussi par le positionnement et la stratégie marketing des entreprises. La société qui a fourni le dispositif expérimenté par des juges dans deux cours d'appel françaises en 2017 a qualifié son produit d'outil de « justice prédictive » et a donné un écho remarquable à cette notion, y compris en publiant des articles dans des revues juridiques. L'arène scientifique est ainsi le lieu de l'encastrement d'une diversité d'acteurs, de discours et d'intérêts. Dans l'immense majorité des publications considérées, la posture épistémologique et ses enjeux éthiques ne sont pas abordés de manière frontale, s'agissant en particulier de la multipositionnalité et des intérêts croisés qu'elle peut générer. Bien au contraire, il peut arriver qu'un même auteur (ayant par exemple un pied dans l'enseignement supérieur et la recherche et l'autre dans une *startup*) ne se présente pas de la même manière suivant l'auditoire auquel il s'adresse.

La revue de littérature réalisée dans ce chapitre met en ordre cette nébuleuse, à travers le repérage de 'continents de littérature' qui se répondent séquentiellement et l'identification de lignes de débat structurantes. Elle couvre des disciplines variées même si les productions juridiques sont incontestablement les plus nombreuses. Le débat académique déborde constamment puisque de nombreux juristes praticiens contribuent aux activités de recherche et de publications. Un premier temps sera consacré à restituer le noyau de la controverse ouverte par les travaux européens et américains qui, au milieu des années 2010, ont annoncé des progrès considérables dans l'élaboration de modèles prédictifs des décisions de justice grâce à l'apprentissage machine (1). Nous reviendrons ensuite sur l'élargissement de la controverse. Par un effet boule de neige, celle-ci gagne des cercles plus vastes et des acteurs plus lointains, qui élaborent des réflexions prospectives sur les transformations que la « justice prédictive » pourrait provoquer dans le fonctionnement des systèmes juridiques et judiciaires (2). Force est de constater que rares sont les analyses empiriques des dispositifs algorithmiques existants, que ce soit sous l'angle de leur genèse, des discours auxquels ils donnent lieu, de leur fonctionnement concret ou bien encore des usages dont ils sont l'objet. Nous reviendrons sur celles qui, en sociologie et en science politique, dessinent une voie, à mes yeux, constructive (3).

I / Une controverse sur la performance des algorithmes dans la justice

Au tournant des années 2016-2017, des équipes pluridisciplinaires annoncent triomphalement des résultats marquants dans la « prédiction » de décisions de justice de

³ Mission de recherche droit et justice, 2016 ; reconduit en 2017, 2018 et 2019. <http://www.gip-recherche-justice.fr/wp-content/uploads/2013/06/2016-A.O-Droit-justice-et-num%C3%A9rique.pdf> [consulté le 20 novembre 2020].

⁴ La première a la spécificité d'associer des chercheurs de droit et sciences sociales et le cofondateur d'une *startup* du domaine.

différentes juridictions, et ce, par l'utilisation de l'apprentissage machine et le travail en langage naturel (1). Ces recherches, publiées dans des revues juridiques, suscitent d'intenses réactions de la part des académiques et des praticiens du droit, spécialistes du sujet. Ils questionnent la rupture technologique et la levée de verrous technologiques qu'elles prétendent opérer (2).

1.1. Des expériences disruptives concernant les décisions de justice ?

Ces expérimentations pluridisciplinaires, fondées sur des techniques d'apprentissage machine et de traitement du langage naturel, sont une nouveauté en ce qui concerne les décisions de la Cour européenne des droits de l'homme (CEDH) ou de certaines juridictions civiles françaises. En revanche, dans le cas de celles qui portent sur la Cour suprême des États-Unis, cela représente davantage une étape supplémentaire dans un processus enclenché depuis plusieurs décennies.

1.1.1. *Les décisions de la CEDH au prisme de l'apprentissage machine*

En 2016, est publiée une étude portant sur un échantillon de décisions de la CEDH menée par une équipe de trois informaticiens, spécialistes de *machine learning*, et d'un juriste (Aletras et al., 2016). Nikolaos Aletras est *senior lecturer*, spécialiste d'intelligence artificielle au Department of Computer Science, University College de Londres, mais il travaille également pour Amazon. Il a participé à différentes recherches qui touchent bien d'autres domaines que le seul droit, par exemple sur la parodie politique sur Twitter (Maronikolakis, Danae Sánchez, Preoțiuc-Pietro et Aletras, 2020). Vasileios Lampos, qui a aussi un profil d'informaticien, est en fonction dans la même institution. Daniel Preoțiuc-Pietro est spécialiste de traitement du langage naturel, de sciences sociales computationnelles et d'apprentissage machine, doublement affilié au Centre d'informatique et au Centre de psychologie positive de l'Université de Pennsylvanie à Philadelphie, à l'époque. Dimitrios Tsarapatsanis, quant à lui, est le juriste de l'équipe, spécialiste de théorie du droit, en poste à l'Université de Sheffield au moment où le papier est publié – il est désormais à l'Université de York. En réalité, la seule composition de l'équipe montre que le droit et la justice ne sont pas au cœur des préoccupations de ce collectif, mais un terrain à partir duquel tester, spécifier des techniques computationnelles utilisées dans d'autres domaines.

Ces auteurs s'intéressent à des décisions de la CEDH publiées dans la base de données de la Cour. Il s'agit de décisions déjà jugées, c'est-à-dire non seulement ayant fait l'objet d'une décision de la CEDH, mais ayant passé avec succès les différents filtres préalables mis en place dans les juridictions nationales et au sein de la Cour elle-même. Elles représentent donc une infime part des procédures qui lui sont adressées⁵. L'analyse a été effectuée à partir du texte des jugements publiés, lesquels s'y prêtent particulièrement bien puisqu'ils sont déjà

⁵ Environ 1 % puisque les auteurs indiquent que pour 2015, l'année étudiée, la Cour a rendu 891 jugements sur environ 83 700 procédures.

standardisés : structurés autour d'une trame type, définissant des paragraphes et des catégories d'informations, ce qui facilite la fouille automatique en texte intégral⁶.

L'équipe travaille à partir d'un ensemble de données composé d'affaires ayant trait aux articles 3, 6 et 8 de la CEDH. Le corpus constitué comporte 584 dossiers, tous rédigés en anglais, et représentant à égalité et en proportion du nombre de cas initial, les décisions de violation et de non-violation de l'article de la Convention en question. Deux cent cinquante cas sont relatifs à l'article 3 (qui interdit la torture et les traitements inhumains ou dégradants) ; quatre-vingts cas à l'article 6 (droit à un procès équitable) ; deux cent cinquante-quatre cas à l'article 8 (droit au respect de sa vie privée, de son domicile et de sa correspondance).

Les auteurs ont introduit dans la machine tout un ensemble de données dites d'apprentissage⁷, issues du texte des arrêts – expurgé toutefois des paragraphes où la décision finale de la Cour était mentionnée et ce, précisément parce que le but était de contrôler si l'algorithme permettait, à partir de l'analyse du texte du jugement, de prévoir la nature de la décision (violation ou pas de l'article en question). C'est donc une méthode *a posteriori* appliquée à des cas déjà jugés, où l'on peut vérifier si l'algorithme parvient aux mêmes décisions que la Cour, sachant que lui sont fournis : « les faits, la procédure précédente, les arguments des parties et les dispositions de droit positif pertinentes ». En l'occurrence, l'algorithme produit des décisions identiques à celles de la Cour dans 79 % des cas ce qui indique, selon N. Aletras *et al.*, qu'il existe bien une forme de cohérence, de stabilité et de logique des jugements humains. Mais il demeure aussi une disparité importante : dans un cas sur cinq, l'algorithme aboutit à une décision différente. Ce peut être parce que les processus de jugement relèvent en partie d'autres éléments que ceux qui ont été soumis à l'algorithme : préférences idéologiques ou personnalité des juges ; conditions particulières dans lesquelles la décision est prise ; anticipation des conséquences sociales, économiques, diplomatiques de l'arrêt, etc.

L'analyse du fonctionnement de ces modèles prédictifs met en évidence que le contexte factuel contenu dans la sous-section « Circonstances » de l'arrêt est la partie du texte la plus importante de l'affaire lorsqu'il s'agit de prévoir la décision de la CEDH. Les autres sections et sous-sections qui se réfèrent aux faits d'une affaire déterminée, à savoir « Procédure », « Droit applicable » et « Faits » sont un peu moins performantes, bien qu'elles restent systématiquement plus prédictives que le hasard. De cela, Aletras et ses collègues tirent la conclusion que l'algorithme a réussi plus facilement à prédire la solution en se fondant sur les données factuelles des dossiers que sur les données juridiques disponibles. Cela

⁶ L'algorithme dispose de l'information relative au statut du paragraphe : porte-t-il sur la procédure antérieure, sur les faits, etc.

⁷ N. Aletras et son équipe utilisent les caractéristiques du texte, c'est-à-dire les séquences de *n* caractères consécutifs (appelés *N*-grammes) et les clusters de mots proches dans les textes, pour former les classificateurs d'un ensemble de techniques d'apprentissage supervisé destinées à résoudre des problèmes de classification et de régression. Ils sont appelés SVM : *support vector machine* ou séparateurs à vaste marge. Ces modèles algorithmiques sont utilisés dans d'autres domaines.

validerait donc les thèses réalistes américaines⁸, selon lesquelles la prise de décision des juges repose davantage sur les caractéristiques factuelles du cas que sur les dispositions juridiques applicables (Aletras *et al.*, 2016 : 16).

Aletras et ses collaborateurs reconnaissent bien des limites à leur étude, mais la présentent toutefois comme « la première étude systématique de prévision de l'issue des affaires jugées par un grand tribunal international, réalisée par l'exploitation des informations textuelles disponibles » (2016 : 2). Ils affirment que, d'ores et déjà, leur algorithme peut être utile à l'amélioration du fonctionnement de la CEDH, en particulier pour lui permettre de mieux gérer l'afflux de demandes qu'elle reçoit. Les délais longs de la Cour pourraient être raccourcis par le recours à des méthodes prédictives qui faciliteraient la sélection des demandes reçues en fonction de la probabilité qu'elles contiennent des violations d'articles. Sur la base d'une fouille automatique par algorithme, celles-ci seraient dirigées en priorité vers un juge. D'autres utilisations sont également avancées par Aletras et ses collègues : aider les avocats à s'orienter dans les décisions de la CEDH, à comprendre les critères de décisions, doter les juges d'outils d'aide à la décision. D'autant plus que si les auteurs ont travaillé seulement à partir des textes des arrêts de la cour, ils avancent que leur approche pourrait être « utilisée de façon fructueuse pour prévoir les résultats à partir d'[...] autres types de textes » (4), comme les demandes introduites par les parties ou les conclusions des avocats⁹.

Nous verrons plus loin que le travail de cette équipe a représenté un point de focalisation important pour les auteurs internationaux et français. D'autres travaux du même type ont par la suite été publiés, par exemple sur les décisions de la CEDH relatives à l'article 9 (Medvedeva, Vols et Wieling, 2019). Ils sont toutefois bien moins connus : rares sont les auteurs qui les citent ou les discutent. Peut-être aussi parce qu'en quelques années, l'effet de nouveauté, voire de révolution est devenu beaucoup moins fort.

1.1.2. *La modélisation des décisions de la Cour suprême américaine*

Une autre équipe, réunie autour de Daniel Martin Katz, a développé, à la même époque, une approche du même type, à partir de décisions de la Cour suprême américaine (Katz, Bommarito et Blackman, 2017). Daniel Martin Katz est professeur de droit au Chicago-Kent College of Law ; il a cofondé avec Michael J. Bommarito, également juriste et consultant dans le secteur des technologies pour le droit, la société Lexpredict. Celle-ci assure des missions de consultance dans les domaines des *legal analytics*, *legal data science and training*, *risk management*, and *legal data strategy*. Josh Blackman est, quant à lui, professeur assistant à la faculté de droit du Texas du Sud à Houston. Leur équipe est beaucoup plus juridique que la précédente, mais encore plus typique d'un double positionnement, dans l'université et le secteur marchand, avec ici la création d'une *startup* destinée à la clientèle des *law firms*. La technique qu'ils utilisent est basée sur un autre type d'algorithme classificateur appelé

⁸ Le réalisme s'est développé en réaction au positivisme juridique. Il affirme l'importance de la description et de la compréhension de ce que font les acteurs chargés d'appliquer le droit. Voir *infra*.

⁹ Sous réserve, précisent-ils, qu'il y ait suffisamment de similarité entre les sections de jugements qu'ils ont analysées et celles des demandes ou conclusions.

algorithme des forêts d'arbres décisionnels (*random forest classifier*)¹⁰, qui est présenté comme particulièrement sophistiqué¹¹.

Le travail de Katz et son équipe s'inscrit dans une longue tradition de modélisation des décisions de la Cour suprême américaine, depuis les travaux fondateurs de Kort, Loevinger (1963) dans les années 1950-1960¹². La science jurimétrique proposée par Lee Lovinger, juriste américain qui fut juge à la Cour suprême du Minnesota, consistait en « une application de la méthode scientifique à l'étude du droit », avec pour objet « le calcul de la prédictibilité des décisions judiciaires » (Loevinger, 1963 cité par Zambrano, 2016 : 1). La science juridique est intéressée de longue date à comprendre ce qui fonde les décisions des juges en particulier ceux de la Cour suprême américaine, sur laquelle de très nombreux travaux ont été menés, à commencer par ceux de Lovinger lui-même. Pour Zambrano, ce qui est en jeu, surtout pour les systèmes de *common law*, c'est la prétention et la capacité de la science juridique à développer un discours scientifique sur le droit et à permettre une intelligibilité de la manière dont le juge exerce son pouvoir de créateur de norme.

Au carrefour du droit et de la science politique américaine, à partir de différents types de méthodes statistiques, un sous champ s'est constitué, celui des *Judicial Behavior Studies* (Tate, 1983 ; Howard et Randazzo, 2017 ; Epstein et Lindquist, 2017). Leur enjeu est d'identifier les facteurs qui pèsent sur les décisions prises par les membres des juridictions, en particulier au sein de la Cour suprême. La question des préférences idéologiques et politiques des juges, de la manière dont ils les font prévaloir dans leurs comportements, c'est-à-dire leurs prises de décision et leurs votes y est centrale¹³. Déclinaison du domaine plus large des *judicial politics*, ce sous-champ est traversé par différentes théories : le choix rationnel¹⁴, le modèle

¹⁰ Pour une explication du fonctionnement : YIU Tony, « Understanding Random Forest » [en ligne], mis en ligne en juin 2019, <https://towardsdatascience.com/understanding-random-forest-58381e0602d2> [consulté le 7 septembre 2020].

¹¹ Celui-ci fonctionne à partir d'arbres de décision individuels qui sont combinés pour former une forêt aléatoire – sur le modèle de la forêt végétale dont les arbres individuellement différents se protègent les uns les autres. Les forêts aléatoires permettent par la combinaison de la multitude et l'absence de liens entre les arbres de minorer les erreurs par l'effet de nombre, ce qui permet d'obtenir des modèles prédictifs robustes. Katz et ses collègues (2017 : 7) expliquent qu'ils ont entraîné l'algorithme avec 125 arbres sur la période 1791-1816, chaque arbre représentant un juge. « À chaque nouveau mandat, en l'absence de changement dans la composition de la juridiction, cinq arbres supplémentaires ont été formés et ajoutés à la forêt du mandat précédent. » Ce n'est que lorsqu'un membre change (décès, etc.) qu'une nouvelle forêt est créée.

¹² En France, Guillaume Zambrano (2015), maître de conférences en droit privé spécialiste d'IA, coauteur de publications consacrées aux modèles d'analyse statistique des décisions judiciaires (Zambrano et al., 2019 & 2020) et responsable d'un DU intitulé « Justice prédictive et *legaltech* », est un des premiers à rappeler que le projet de rendre plus prévisibles les décisions judiciaires par analyse des décisions passées n'est pas neuf, surtout outre-Atlantique, et qu'il constitue même un pan entier de la science juridique américaine.

¹³ (Nagel, 1963 ; Schubert, 1963 ; Segal, 1984 ; Segal & Spaeth, 2002).

¹⁴ Les économistes du droit étudient aussi les facteurs qui pèsent sur les décisions juridictionnelles, et ont mis en avant un « effet juge » bien connu (Ferey, 2018 : 79 ; Sunstein, 2001). On note également des approches de psychologie « *behavioriste* » qui tentent de moduler les approches *Law & Economics* fondées principalement sur la théorie du choix limité.

attitudinal (Segal et Spaeth, 1993), le modèle stratégique (Epstein et Knight, 1998) ou bien encore celui de l'audience des juges (Baum, 2017 ; 2009). Dans ces travaux, il s'agit le plus souvent d'expliquer *ex post* les décisions produites par la plus haute juridiction des États-Unis, à partir de modèles mathématiques. Les caractéristiques individuelles, liées au profil du juge, à la manière dont il interagit avec les autres membres de la Cour sont particulièrement testées.

S'il existe donc une tradition d'analyse prédictive des décisions de la Cour suprême des États-Unis – qui se présente comme telle –, les algorithmes d'apprentissage supervisé n'ont toutefois été utilisés que depuis le début des années 2000 (Martin *et al.*, 2004 ; Ruger *et al.*, 2004). Les tentatives récentes de construire des modèles de prédiction des décisions de la Cour suprême sont basées sur des périodes-tests dont les caractéristiques sont bien spécifiques, ce qui rend leurs résultats non généralisables au-delà de la période étudiée. Katz et son équipe revendiquent, quant à eux, de construire un modèle « général » pertinent pour « des données *out-of-sample* [c'est-à-dire non présentes dans l'échantillon initial] et applicable à l'ensemble du passé et de l'avenir de la Cour, et pas seulement à un mandat spécifique. » (Katz, Bommarito et Blackman, 2017 : 2) Pour ce faire, ils ont utilisé les bases de données de la Cour suprême (SCDB Legacy et Modern¹⁵), lesquelles contiennent l'intégralité des décisions intervenues entre 1791 et 2015 (compris), soit 240 000 votes pour 28 000 décisions. Chaque affaire comporte jusqu'à 240 variables chronologiques, relatives à son contexte, aux membres de la Cour qui ont participé au délibéré et à la décision elle-même.

Leur modèle prédictif atteint une probabilité de 70,2 % au niveau de l'issue des affaires et 71,9 % au niveau du vote de chacun des juges individuellement considéré. En comparant leur modèle avec un modèle de régression dans lequel les paramètres sont tous équivalents à 0 (appelé le modèle nul), ils ont dépassé de 5 % les résultats de celui-ci, pour les arrêts allant de 1915 à 2015. Les auteurs estiment que « Ces améliorations de la modélisation devraient intéresser, entre autres, les observateurs des tribunaux, les plaideurs, les citoyens et les marchés. En effet, en ce qui concerne les marchés, étant donné que les décisions judiciaires peuvent avoir un impact sur les sociétés cotées en bourse [...] *même des gains modestes en matière de prédiction peuvent produire des récompenses financières importantes.* » (Katz *et al.*, 2017 : 14) Ils affirment encore que leurs « résultats représentent une avancée importante pour la science de la prédiction juridique quantitative et laissent présager toute une série d'autres applications potentielles. » (*ibid.*) En d'autres termes, cela augure selon eux d'un progrès important susceptible d'être traduit dans la vie du droit. L'espace académique est un lieu de valorisation de ces recherches dont les enjeux dépassent de loin le monde universitaire et sont explicitement liés à des valorisations dans le monde de la pratique et sous des formes marchandes.

1.1.3. Les analyses algorithmiques des décisions de justice françaises disponibles

Le même mécanisme est observable en France, toujours à la même période, où se font également connaître des dispositifs qui revendiquent d'apporter une innovation de rupture dans le domaine juridique et judiciaire, en ce qui concerne l'analyse des décisions de justice

¹⁵ <http://scdb.wustl.edu/>

et à partir de techniques d'apprentissage machine. [Morgan Forestier], fondateur du site [Legalplus], publie en 2016, sur le blog du Village de la justice, un article portant sur une expérience menée sur les arrêts des Cours d'appel administratives en annulation des obligations de quitter le territoire (OQTF), prononcées en première instance. Il revendique haut et fort le caractère pionnier de cette réalisation : « C'est à notre connaissance la première fois que des algorithmes ayant une forme de compréhension du sens des mots et des phrases sont appliqués à la jurisprudence française. » ([Forestier], 2016) Ils sont utilisés pour « extraire [automatiquement] entre autres la qualité du demandeur et du défendeur, la nature du dispositif ou le thème de la décision. [...] Les résultats sont mis à disposition pour être interrogés par les juristes en quelques secondes. La précision de nos résultats se situe entre 90 % et 99 % en fonction du champ extrait. » (*ibid.*)

L'article entend démontrer l'intérêt de ce dispositif pour analyser de manière comparée les décisions des différentes formations de jugement sur le territoire français. En l'occurrence, il révèle une disparité considérable, de l'ordre de 40 points, stable à l'échelle de la période étudiée (2012 à 2015) entre les présidents de chambre qui valident le plus et ceux qui valident le moins les demandes d'annulation des OQTF. Parmi les 20 % qui rejettent le plus les demandes d'annulation des OQTF, il sélectionne et compare les trois présidents qui rejettent le plus ou le moins ces demandes lorsqu'elles sont formées par des particuliers contre l'administration. Il s'attache alors à étudier pour les trois mêmes présidents les taux de rejet quand c'est, au contraire, l'administration qui procède au recours en annulation. Ceux qui rejettent le plus les demandes des particuliers sont globalement aussi ceux qui rejettent le moins les demandes formulées par l'administration, ce que M. [Forestier] interprète comme attestant de l'existence de profils de juges « rebelles » et d'autres « plus soumis » à l'administration. L'enjeu est ici d'utiliser des algorithmes pour extraire automatiquement des données des décisions de justice (c'est ici qu'intervient l'apprentissage machine), ce qui permet ensuite des traitements statistiques classiques pour rendre lisibles des régularités dans la fabrique des décisions (ici à l'échelle de chaque président de formation de jugement). Cette expérience a été discutée notamment sur le plan de son éthique puisque les magistrats administratifs y sont nommément cités et que l'accès à cet algorithme est complètement ouvert (il est mis gratuitement à disposition du public).

Presque concomitamment, un groupe de quatre chercheurs en poste dans des universités roumaine, anglaise et allemande publie une recherche relative aux décisions de la Cour de cassation française (Sulea *et al.*, 2017). L'investigation repose sur les mêmes techniques de classification textuelle que N. Aletras (2016) avec un apprentissage machine supervisé et entraîné sur les caractéristiques lexicales des décisions de la juridiction suprême. Le corpus d'arrêts, allant des années 1800 à nos jours, comprend 126 000 décisions, chacune contenant une description de l'affaire et quatre types d'étiquettes différentes qui renvoient aux métadonnées de la décision : l'instance qui a eu à s'occuper de l'affaire (chambre civile, commerciale, assemblée plénière), la date de l'arrêt, l'arrêt lui-même et une liste de références d'articles et de lois cités dans la description. Certaines de ces informations sont toutefois masquées, en fonction de la question posée. Par exemple lorsqu'il s'agit de déterminer la nature de la décision adoptée par la Cour (cassation, rejet, irrecevabilité, QPC,

etc.), l'information relative au dispositif de l'arrêt a été masquée dans la description du cas pour se rapprocher le plus possible d'une situation prospective dans laquelle l'issue de l'affaire ne serait pas connue. Les auteurs avancent des résultats de 96 % dans la prévision d'une décision pour un cas, 90 % dans celle du domaine juridique auquel se rapporte le cas et donc de la Chambre concernée ; 75,9 % dans l'estimation de la période à laquelle une décision a été rendue. Cette recherche, mise en ligne sur différents sites spécialisés dans l'IA profonde, est passée inaperçue ; elle n'est guère mentionnée par la littérature nationale et internationale (hormis G'sell, 2020b).

D'autres expériences, comportant un volet commercial explicite, sont également présentées dans des revues juridiques ([Monta] et [Sinclair], 2017) ainsi que dans des colloques et les ouvrages collectifs qui en résultent¹⁶, y compris par leurs propres fondateurs, très proactifs dans l'arène scientifique. Au plan national, la *startup* [PréviCompute] capte une grande partie de l'attention, car l'expérimentation auprès de juges fournit une opportunité de visibilité qu'elle exploite massivement. [Anajur] qui apparaît comme sa principale concurrente développe également une politique de communication volontariste. Toutes deux sont régulièrement citées et prises en exemple dans les débats sur la 'justice prédictive'. Nous aurons l'occasion de revenir longuement sur elles, dans le chapitre 4.

La littérature scientifique sur les dispositifs algorithmiques de traitement et d'analyse des décisions de justice est donc traversée par un enjeu de reconnaissance des acteurs porteurs d'innovations liées à l'intelligence artificielle. Ils affirment qu'une révolution est en cours, revendiquent une posture disruptive qui précisément va être discutée par celles et ceux qui, spécialistes de longue date de la rencontre entre droit et informatique, s'interrogent sur ces arrivants, sur leur brutale percée, sur leur intérêt réel pour la justice et sur ce que ces 'cavaliers seuls' apportent véritablement.

1.2. Ces modèles prédictifs sont-ils si neufs et performants ?

Le courant *Law & Informatics* existe au plan national et international depuis les années 1960. Il s'est progressivement institutionnalisé comme en témoigne la création d'associations professionnelles comme l'*American Association for Artificial Intelligence* ou l'*International Association for Artificial Intelligence and Law*, de fondations telle la *Foundation for Legal Knowledge Systems* (JURIX), de conférences internationales comme l'*International Conference on Artificial Intelligence and Law* ou bien encore de revues spécialisées telles que *Jurimetrics*, *Artificial Intelligence and Law*, *European Journal of Law and Technology*, etc. Aujourd'hui, il est de plus en plus question de *computational legal studies* (Ashley, 2017), parfois traduit par l'expression « droit et mathématiques » (Garapon et Lassègue, 2018 : 104 ; G'sell, 2018 : 48).

Les spécialistes de ce vaste domaine s'interrogent sur le caractère vraiment neuf et puissant des résultats avancés par les équipes d'Aletras et de Katz. En mai 2020, dans une adresse à la 7^e conférence internationale sur l'intelligence artificielle et le droit, Bart Verheij, le président de l'association qui organise cette conférence, lui-même spécialiste de

¹⁶ ([Monta] & [Sinclair], 2018 ; [Monta], 2018 ; [Chevalet], 2018a, 2018b).

modélisation de l'argumentation juridique à l'Université de Groningen, aborde les différentes dimensions de la recherche sur IA et droit, dont il pointe les acquis et les enjeux. Il alerte son auditoire sur le fait que les recherches actuelles en matière d'IA feraient bien de se poser la question de ce qu'est « une bonne IA » et qu'à cet égard, ils devraient se tourner vers l'expertise accumulée par la communauté qui laboure depuis plusieurs décennies le champ IA et droit. Celle-ci a emmagasiné une importante expertise face aux problèmes particulièrement complexes qu'elle a eu à affronter compte tenu de son objet, le droit, et qui touchent à la fois au raisonnement, à la connaissance, à l'apprentissage et au langage.

Pour lui, les travaux des équipes d'Aletras d'une part et de Katz d'autre part ne franchissent pas une étape marquante ni ne lèvent les verrous technologiques bien identifiés par la communauté scientifique. Au contraire, leurs résultats témoignent de la difficulté d'atteindre une forte prévisibilité en matière de décisions de justice, comme il le dit dans cette longue citation : « Quelques exemples récents montrent à quel point la prédiction judiciaire est difficile, y compris lorsqu'on utilise de sérieuses techniques d'apprentissage automatique. Katz *et al.* (2017) affirment que leur machine de prévision des décisions de la Cour suprême des États-Unis pourrait atteindre une précision de 70 %. Cela représente une légère amélioration par rapport à la base de référence du résultat majoritaire historique (qui tend à toujours confirmer une décision antérieure) qui est de 60 %, et encore plus légère en comparaison avec le résultat majoritaire sur 10 ans qui est de 67 %. Le système a basé ses prévisions sur des caractéristiques telles que l'identité du juge, le mois, la juridiction d'origine et la question abordée, de sorte que des résultats modestes ne sont pas surprenants. Dans une autre étude, Aletras et ses collègues (2016) ont étudié les affaires de la Cour européenne des droits de l'homme. Ils ont utilisé des n-grammes et des sujets comme point de départ de leur formation, et ont utilisé un ensemble de données préparées pour établir une base de référence plus précise de 50 % en faisant des suppositions aléatoires. Ils ont atteint une précision de 79 % en utilisant l'ensemble du texte, et ont noté qu'en n'utilisant que la partie où les circonstances factuelles sont décrites, une précision de 73 % est déjà atteinte. *En prenant naïvement les ratios de 70 sur 60 et de 79 sur 50, on constate que les facteurs d'amélioration de 1,2 et de 1,6 sont des résultats de recherche pertinents, mais pratiquement modestes.* Et surtout, ces systèmes se concentrent uniquement sur les résultats, *sans dire quoi que ce soit sur la manière d'arriver à un résultat ni sur les raisons pour lesquelles un résultat est justifié ou non.* » (Verheij, 2020 : 196)

Bart Verheij relativise ainsi cruellement l'apport de ces travaux du point de vue de la spécificité de la matière juridique et opère également une critique sévère de leur effet boîte noire pour finalement mettre l'accent sur le long chemin qu'il reste à parcourir. Il n'est pas le seul à accueillir avec tiédeur les résultats des équipes d'Aletras et de Katz. Plusieurs auteurs passent au crible les méthodes utilisées et les résultats obtenus par l'équipe d'Aletras. Le fait que cette recherche porte sur les décisions d'une juridiction européenne, pour laquelle il n'existe pas la même tradition d'analyse des décisions que pour les juridictions américaines, aide certainement à comprendre la focalisation sur cette recherche. Celle très comparable, qui, en 2017, a porté sur les décisions de la Cour de cassation française n'a pas du tout connu le même écho.

1.2.1. La critique ciblée des travaux de l'équipe de N. Aletras

L'expérimentation de l'équipe d'Aletras a suscité d'importantes réactions dans le monde académique autour des méthodes utilisées, de la nature des décisions collectées, de la validité et la généralité des résultats produits (Pasquale et Cashwell, 2018 ; Barraud, 2017 ; [Vial], 2018). Sur le plan international, Franck Pasquale, juriste, auteur du très remarqué *The Black Box Society* consacre un long article coécrit avec Glyn Cashwell (2018) à discuter le papier publié par Aletras et ses collègues. En France dès 2017, cette étude est examinée de manière détaillée par un petit cercle d'auteurs dont Boris Barraud, enseignant vacataire en droit du numérique, spécialiste de l'IA, et [Thomas Vial], magistrat chargé des questions d'IA au sein de la Commission pour l'efficacité de la justice (Cepej) du Conseil de l'Europe.

Plusieurs points de critiques sont soulevés. L'argument central de Pasquale et Cashwell porte sur le caractère excessivement positiviste des présupposés à partir desquels cette expérimentation est réalisée. Recourir à des modèles conçus pour des sciences de la nature dans des contextes judiciaires relève d'une théorie du droit comportementaliste, reposant sur un modèle de psychologie cognitive¹⁷ dans lequel les processus mentaux des juges sont réduits à la transformation opaque d'*inputs* en *outputs*, par le biais de boîtes noires, ce qui est éloigné de la réalité empirique.

La performance du dispositif est elle-même discutée. Boris Barraud (2017) relativise les 29 % d'écart entre les décisions de la Cour et celles de l'algorithme en estimant que, comme il s'agit d'une décision binaire (reconnaître la violation/la rejeter), en réalité l'algorithme a d'emblée une chance sur deux de décider à l'instar de la Cour, donc 50 %. En obtenant un score de 29 %, l'algorithme est en réalité beaucoup plus loin des décisions de la Cour que ne le disent les auteurs de l'expérience. Ce constat n'est guère contestable.

De surcroît, les résultats de l'étude seraient faussés puisque l'algorithme travaille à prédire des décisions déjà produites par la Cour européenne des droits de l'homme, à partir du texte même de ces décisions ; et non à partir de documents complètement étrangers comme le sont les demandes ou les conclusions des avocats (Pasquale et Cashwell, 2018). Cette étude ne 'prédit' pas, elle analyse « un matériel en grande partie issu du raisonnement du juge que cette étude était pourtant censée devancer » ([Vial], 2018). Elle informe donc davantage « sur la pratique professionnelle des juristes du greffe de la Cour (qui utilisent parfois des paragraphes prérédigés en fonction du traitement et de l'issue donnés à l'affaire) que sur la réflexion juridique ayant conduit à la décision ». De ce fait, elle ne reproduit pas à proprement parler le raisonnement de la Cour et ne peut permettre d'anticiper sa décision sur le seul fondement « du récit brut d'un futur requérant devant la Cour de Strasbourg » (*ibid.*).

De plus, en enlevant du corpus de décisions utilisé pour nourrir l'algorithme un certain nombre de demandes qui ont été rejetées parce qu'elles ne répondaient pas aux "critères de recevabilité" (essentiellement de nature procédurale), les auteurs risquent d'avoir évacué d'emblée une grande partie des cas pour lesquels des points de droit ont été marquants dans

¹⁷ Thèse qualifiée de skinnerienne comme le psychologue controversé Skinner en a défendu.

la décision – en l’occurrence du fait d’irrecevabilité ou de radiation, alors même qu’elles représentent une immense proportion des cas. Aussi est-il pour le moins excessif de considérer que ces résultats constituent une confirmation des théories réalistes du droit (Pasquale et Cashwell, 2018 : 70) ; ils devraient peut-être même être interprétés inversement, comme leur réfutation (Barraud, 2017).

Travailler à partir de corrélations entre des mots considérés indépendamment de leurs significations comme le fait l’équipe d’Aletras fait également courir le risque d’être victime des corrélations fallacieuses (*spurious correlations*) qui ne manquent pas d’apparaître dès lors que l’on confronte des ensembles de données suffisamment importants. Ces corrélations sont dénuées de toute signification autre que statistique. Pasquale et Cashwell donnent l’exemple de celle, très forte, qui existe entre le taux de divorce dans l’État du Maine et la consommation individuelle de margarine entre 2000 et 2009¹⁸. De manière comparable, si l’on considère le corpus de décisions comme un « sac de mots », et que l’on introduit suffisamment de variations, certaines relations se dégageront, sans qu’elles soient nécessairement significatives autrement que sur le plan de la corrélation statistique. Les modèles et calculs mathématiques utilisés par Aletras et ses collègues sont certes sophistiqués, il n’en reste pas moins qu’ils ne tiennent pas compte du sens des mots, lequel est capital pour définir des droits. Or certains mots peuvent être logiquement associés à des violations de l’article 8 (par exemple enterrement, attaque et mort), mais d’autres sont beaucoup plus polysémiques tel celui de « résultat ».

Comme le soulignait déjà Bart Verheij, la démarche ne fournit pas « d’explication motivée [...] quant à la raison pour laquelle de tels mots devraient être prédictifs en ce qui concerne le corpus, sans parler de leur prédictivité pour la jurisprudence future ». Pasquale et Cashwell concluent que si les analyses prédictives peuvent avoir une certaine utilité, ce n’est certainement pas celle de décider des affaires qui doivent être prioritaires ou même de remplacer les juges. En revanche, si elles révèlent des dysfonctionnements dans les processus décisionnels comme l’irruption de facteurs indésirables — une plus grande sévérité des juges avant qu’après déjeuner¹⁹ —, ce dévoilement, précieux, autorise le développement de parades, pour y remédier (par le grignotage régulier des juges, suggèrent-ils avec humour !). Ils alertent en conséquence sur le risque que ces dispositifs, malgré leur opacité et leurs limites, soient appropriés tels quels par des juridictions débordées, ne disposant pas de suffisamment de moyens humains pour faire face aux demandes.

Par ailleurs, si ce type d’outil devait être utilisé par les juridictions, il serait alors de la responsabilité de celles-ci de s’assurer que toutes les parties y ont également accès. De fait,

¹⁸ Les auteurs renvoient au site de Tyler Vigen, étudiant en criminologie à l’École de droit d’Harvard. Il a conçu un programme informatique afin de déceler des corrélations statistiques dans d’immenses corpus de données et s’est amusé à publier les plus saugrenues. <http://www.tylervigen.com/spurious-correlations>

¹⁹ Cela fait référence aux résultats d’une étude (Danziger *et al.*, 2011) dans laquelle les auteurs démontrent statistiquement que la sévérité de juges israéliens varie en fonction de l’heure à laquelle le dossier arrive en audience, plus ou moins proche de la pause-déjeuner. Une autre recherche a relevé les erreurs et simplifications outrancières de cette étude reposant sur un modèle à une seule variable (Weinshall-Margel & Shapard, 2011). Voir (Jean, 2021).

un énorme déséquilibre existe déjà, entre les petites structures de quelques avocats et les vastes organisations juridiques aux moyens considérables. Il pourrait s'agir là d'une nouvelle inégalité creusant encore le fossé qui existe déjà entre joueurs réguliers et occasionnels du jeu judiciaire (Galanter, 2013 [1974]). Dans la discussion générale sur la « justice prédictive », cet argument de « l'inégalité des armes dans le procès » est parfois mis en avant (Gamet, 2019 : 119). Il est vraisemblable que sur ce point les effets dépendent surtout de la politique de positionnement des opérateurs. En réalité, ils semblent plutôt s'orienter vers des propositions commerciales plus accessibles pour les petits cabinets d'avocats que ne l'étaient celles des éditeurs juridiques traditionnels, jusque-là en situation de quasi-monopole sur l'information juridique.

1.2.2. Un héritage à mettre en valeur

D'une certaine manière, la controverse scientifique qui s'ouvre autour des travaux de l'équipe de N. Aletras, et au-delà, vise à la déconstruction d'un discours dont la dimension d'autopromotion est très forte. Dans le contexte de l'effervescence soudaine des publications et expériences d'IA, leur côté « hype » et « vendeur » (Barthe, 2021[2017] : § 1) ne doit pas éblouir. C'est ce que rappelle à sa manière Emmanuel Barthe. Documentaliste juridique, il a fait carrière dans différentes firmes et à la Cour de cassation où il a contribué à mettre en place l'intranet et le site web. Il est aussi formateur *freelance* dans le domaine des données juridiques. Il publie une très longue note en ligne en 2017²⁰, dans laquelle il entreprend de « démystifier les fonctionnalités et performances des applications présentées comme des IA en droit » (Barthe, 2021[2017] : chapeau de l'article). Il passe en revue les différentes propositions techniques et commerciales disponibles au plan national et international, et les situe du point de vue de leurs performances, en les restituant dans une perspective de long terme.

Il pointe que le terme d'IA comme celui de *legaltech* invisibilise les éditeurs qui, depuis les années 1990, sont implantés dans le secteur de l'informatique juridique à travers différentes offres : « Juris-Data des Éditeurs techniques, futures Éditions du JurisClasseur sur Minitel, Lexis (de Lamy) sur terminal dédié, cédéroms Juridiques Lamy » (*ibid.*). L'informatique juridique des années 1960 et 1970 s'est principalement développée à partir de systèmes experts et de formes peu élaborées de traitement automatique du langage, « du TAL à l'ancienne, avec ses limites, mais bien rodé, notamment dans les cartouches de langage d'Expert System (ex-Temis), utilisées par plusieurs grands éditeurs juridiques », y compris D et Légifrance. Considérées comme une « IA "de papa" », ces technologies équipent des dispositifs tels que « Jurisprudence chiffrée » conçu par l'éditeur Francis Lefebvre à la fin des années 2000, mais aussi des systèmes actuels. « À l'exception totale de l'ex-[Legalplus] et partielle de [Anajur], les applications de justice prédictive [actuellement commercialisées en France] utilisent essentiellement, pour leur moteur de recherche et l'analyse des décisions, des systèmes fondés sur les chaînes de caractères et les expressions régulières (regex), dont

²⁰ Régulièrement actualisée depuis et dont une version abrégée a été publiée dans la revue *La Semaine juridique* sous le titre « Les outils de l'intelligence artificielle pour le droit français » (8 avril 2019).

le précurseur en France sont les cartouches [de langage] de la société Expert System, ex-Temis. » (Barthe, 2021[2017])

Autrement dit, technologies sont en grande partie beaucoup moins révolutionnaires que le discours marketing ne le laisse penser. Ce sont principalement des moteurs de recherche en langage naturel qui associent des filtres permettant d'effectuer des recherches sur des décisions de justice en fonction de certains critères, associés à des analyses statistiques du contentieux (sauf [Anajur] qui utilise des modélisations mathématiques). Le discours actuel sur l'IA tend à occulter ce type de technologies au profit de celles plus en pointe comme les *legal data analytics* qui reposent sur des logiciels conçus spécifiquement pour traiter d'immenses quantités de données. Par le recul historique et l'analyse fine des technologies utilisées, ce que décrit Emmanuel Barthe c'est donc une sorte de double mystification effectuée par le langage de la *legaltech* et de l'IA : la revendication de s'appuyer sur les formes les plus innovantes et prometteuses de l'IA (réseaux de neurones, etc.) tout en pratiquant des technologies d'IA anciennes sur lesquelles une sorte de *hold up* est opéré, aux dépens de leurs initiateurs. Les *startup* sont au premier plan de ce qui ressemble fort à un habile tour de passe-passe où d'importants profits matériels et symboliques sont en jeu.

Il est frappant en effet de constater que l'approche par systèmes experts (Bourcier, 1995) semble avoir disparu comme si elle n'avait pas existé ; les acquis de l'informatique documentaire tellement acquis qu'ils en sont oubliés (Bories, 2003) ; et les pionniers de l'informatique juridique et documentaire quasi gommés de la courte histoire de la « justice prédictive » qui a commencé à être écrite. Or, les moteurs de recherche actuellement intégrés dans les applications de « justice prédictive » reposent sur des méthodes de sélection, formalisation, analyse et croisement des données contenues dans les décisions de justice, méthodes qui ont été élaborées dans le cadre des premiers travaux d'informatique juridique (Serverin, 2018 : 24-30 ; Catala, 1998 ; Bing, 2010). En France comme au plan international, l'informatique a été utilisée par les juristes pour construire des bases de données de jurisprudence dans lesquelles effectuer des recherches automatisées, analyser différents types de corpus de textes juridiques, former au droit (*legal education*) et diffuser l'information juridique, y compris vers le grand public (Erdelez et O'Hare, 1997 ; Paliwala, 2010). Décorticage et analyse systématiques des décisions par la constitution de fichiers de métadonnées, la rédaction d'abstracts, l'élaboration d'index et de thesaurus ; constitutions de banques de données multisources comprenant des corpus de textes juridiques, des décisions contentieuses, plus ou moins enrichies (doctrine, etc.). Des années 1960 aux années 2000, l'informatique a été appliquée aux sources juridiques et a donné lieu à de premières réalisations, qui sont le support à partir duquel se développent les technologies actuelles (Serverin, 2018 : 24). La critique consiste donc aussi, et très simplement, à rappeler l'histoire longue dans laquelle les *legaltechs* et *startup* s'inscrivent, bien que la plupart ne la mentionnent guère, tant le discours marketing autour de la nouveauté radicale est dominant. C'est d'ailleurs dans cette perspective que se place le séminaire E-juris coordonné par Isabelle Sayn (2019a) et réalisé en 2018-2019. Il se propose justement d'explorer des perspectives ouvertes par les algorithmes en développement, en faisant le lien avec les précédentes conceptualisations et réalisations. Centré sur les « potentialités des bases de données

décisionnelles », il implique des chercheurs et chercheuses français investis de longue date sur ces sujets (Danièle Bourcier, Antoine Jeammaud, Jean-Claude Ray, Isabelle Sayn, Évelyne Serverin), des juristes d'une nouvelle génération (Marianne Cottin, Olivier Leclerc...) ; des praticiens (magistrats judiciaires et administratifs dans les juridictions et à la sous-direction de la statistique et des études du ministère de la Justice, à la CEPEJ, un avocat) et des opérateurs privés, acteurs historiques des banques de décisions de justice (LexisNexis et Lefebvre-Sarrut). Ce séminaire replace les dispositifs algorithmiques de traitement et d'analyse des décisions de justice dans un questionnement plus large et inscrit dans les acteurs, opérateurs et structures historiques de l'informatique juridique.

1.2.3. La preuve par l'exemple

Si les auteurs évoqués jusque-là relativisent le discours des opérateurs de dispositifs algorithmiques qui se présentent comme radicalement révolutionnaires, les auteurs d'une recherche de type économétrique suivent une autre voie : recenser les difficultés concrètes qui doivent être surmontées pour produire des résultats prédictifs robustes, en reproduisant ce que les opérateurs de dispositifs algorithmiques prétendent faire (Jeandidier *et al.*, 2020)²¹.

Membres du Bureau d'économie théorique et appliquée de l'Université de Lorraine, Bruno Jeandidier, Jean-Claude Ray et Julie Mansuy s'intéressent à la fixation des montants de la prestation compensatoire dans les cas de divorce²². Ils disposent d'une base de données composée de 5 453 décisions de divorce qu'ils se proposent d'analyser selon les « questions de prédiction » et « types d'offre » avancés par « les sites des quatre ou cinq principales *Legal Tech* de "justice prédictive" française » (*ibid.*, p. 2). Il s'agit en particulier « d'estimer une probabilité (ou un pourcentage) de succès dans une procédure judiciaire [et] une valeur issue de la décision (par exemple un montant d'indemnité) » (*ibid.*, p. 2). Les auteurs interrogent leur base de données en procédant de façon traditionnelle, c'est-à-dire sans recours à des techniques de fouille de données automatisée ni d'intelligence artificielle. Leur ambition est de mettre en évidence les obstacles rencontrés « avec comme perspective de poser la question de savoir si ces techniques associées au *big data* seraient susceptibles de dépasser de telles limites (ou si, au contraire, elles risquent d'être confrontées aux mêmes obstacles) » (*ibid.*, p. 2).

Nombreuses sont les restrictions et complications que révèle cette étude : certaines informations ne sont pas disponibles telles quelles dans les décisions, elles sont à construire soit par la réunion de plusieurs items (par exemple pour la prestation compensatoire elle-même, la combinaison d'une soulte et d'une rente mensuelle), soit par la recherche d'informations hors du jugement lui-même. En effet, même si l'argumentaire ne les

²¹ Il s'agit d'une recherche menée dans le cadre de l'atelier de recherche « E-juris » financée par la Maison des sciences de l'homme de l'Université de Lyon et avec le support financier du Pôle scientifique SJPEG de l'Université de Lorraine (projet « Économie et Justice : de nouveaux défis méthodologiques »). Le financement E-juris a également fait émerger le séminaire animé par Isabelle Sayn, dont il a été question *supra*.

²² Il faut préciser que Jean-Claude Ray a participé à une recherche antérieure sur les décisions des juges en matière de pensions alimentaires : BOURREAU-DUBOIS Cécile, DORIAT-DUBAN Myriam, RAY Jean-Claude (2006), « Caractéristiques du juge et décisions en matière de pensions alimentaires : Une étude à partir de données expérimentales », *Revue économique*, vol. 57, p. 563-572.

mentionne pas, des déterminants importants peuvent agir, comme l'existence ou pas d'un accord entre les parties. Seule une connaissance préalable du contentieux du divorce permet de savoir que cette distinction est capitale pour expliquer et donc prévoir les décisions des magistrats en matière de prestation compensatoire. Par ailleurs, l'enquête conclut à un taux d'erreur important et au fait qu'en ce qui concerne les prestations compensatoires, « L'ampleur de ces erreurs (en euros comme en proportion) est telle qu'il serait très risqué de recourir à ces estimations pour effectuer de la justice prédictive à destination des couples en instance de divorce. En effet, dans un cas sur quatre, l'erreur de prédiction excède la moitié du montant observé de la PC. » La critique, qui procède cette fois par l'exemple, est donc extrêmement cinglante à l'égard de ce que les dispositifs algorithmiques de traitement et d'analyse des décisions de justice prétendent faire. Cela rejoint d'ailleurs les limites qu'ont pointées les magistrats qui ont testé [PréviCompute] (Licoppe et Dumoulin, 2019).

Depuis une diversité de points de vue, de disciplines et de méthodes, émerge donc la mise en question de l'originalité réelle et de la solidité des résultats avancés par les acteurs qui prétendent révolutionner la compréhension et la gestion des décisions de justice dans différents contextes institutionnels. Le fait que certaines propositions s'inscrivent d'emblée dans une stratégie commerciale est un des éléments qui suscite des réactions hostiles, y compris de la part d'acteurs juridiques eux-mêmes impliqués dans la transformation du droit.

Ainsi dans un article consacré à la « justice prédictive », Emmanuel Jeuland (2017), professeur de droit à Paris 1, défend une autre conception du rapport entre droit et technologies, laquelle ne se place ni en concurrence ni en alternative des approches prédictives. Il explique en effet que ce qui l'intéresse ce n'est pas le prédictif, mais le potentiel, c'est-à-dire la capacité à cultiver une part d'aléa juridique au sens de créativité.²³ Par exemple, « réécrire un grand arrêt en lui donnant le sens inverse de sa position », ce qui a été tenté à partir du célèbre arrêt Blanco du Tribunal des conflits²⁴. Outre la dimension de jeu sur la langue juridique et ses formes, dans une perspective de pataphysique, ce projet intellectuel et politique s'exprime à travers le modèle économique non marchand dont il est inséparable puisqu'il s'agit d'un atelier ou bien, est-il précisé, d'« une *startup* de *legaltech* à but non lucratif créée avec des doctorants en 2013 ».

Si un ensemble de publications orientées vers la restitution d'expériences d'utilisation de l'apprentissage machine appliqué à des décisions de justice émerge autour de la seconde moitié des années 2010, il suscite donc des réactions critiques immédiates qui ne portent

²³ Il développe cette ambition à travers l'Oudropo, « un espace d'inventivité juridique et de théorie du droit » dont l'ambition est de « générer du droit [au sens de "norme, acte, lien, prérogative ou personne juridique"] à l'aide de contraintes que les membres se fixent à eux-mêmes » (Jeuland, 2017 : 11). L'Oudropo, OUvroir de DROit Potentiel, s'inscrit dans la mouvance de l'Oulipo, l'ouvroir de littérature potentielle créé en 1960 par Le Lionnais et Queneau. <https://www.oudropo.com/2017/12/19/publication-oudropo/>, consulté le 11 décembre 2020.

²⁴ Lequel a posé le principe selon lequel la responsabilité, qui peut incomber à l'État, pour les dommages causés aux particuliers par le fait des personnes qu'il emploie dans le service public, ne relève pas du Code civil, mais de dispositions spéciales, et que par conséquent l'autorité administrative est la seule compétente en la matière (justice administrative). <https://www.oudropo.com/portfolio-item/reecriture-darret/>, consulté le 11 décembre 2020.

toutefois pas sur les technologies en tant que telles. La montée en puissance de cette critique s'explique davantage par la manière dont les expérimentations sont présentées et positionnées par celles et ceux qui les mettent en place. La controverse naît en réaction à la prétention de révolutionner la compréhension et la pratique de la production des décisions de justice de différentes juridictions par ces technologies ; prétention avancée par des acteurs relativement éloignés du monde juridico-judiciaire lui-même, perçus comme des outsiders généralistes n'adhérant pas forcément à ses valeurs et enfin, motivés, pour une partie d'entre eux, par la perspective de construction d'un marché et de ses retombées économiques. Les réactions des spécialistes d'informatique juridique ou de « droit et intelligence artificielle » inaugurent une mise en question de ces dispositifs, à partir d'arguments précis liés à la manière dont ils sont conçus ou en référence à l'histoire du domaine dans lequel ils s'inscrivent.

II / Une controverse autour des effets possibles des algorithmes sur le système juridique et judiciaire français

La communauté des juristes et des autres spécialistes du droit s'engouffre dans cette controverse et la diffracte. L'intelligence artificielle appliquée aux arrêts et jugements et plus particulièrement la « justice prédictive » devient le lieu d'interrogations sur la possible « révolution » (Garapon, 2017 ; Zambrano, 2015 : 7) qu'ils induisent pour la justice et le droit. Le terme même de « justice prédictive », en même temps qu'il est repris comme en écho, est interrogé et progressivement soumis à la critique et avec lui, la place de l'aléa dans la décision juridictionnelle : évitable ou non ? Souhaitable ou source d'injustice ? (1) Au-delà de l'expression elle-même et des débats philosophiques sur la nature de l'acte de juger, de nombreuses publications s'attachent à repérer et évaluer les avantages et risques des dispositifs algorithmiques de traitement et d'analyse des décisions de justice (Biard *et al.*, 2021). Plusieurs lignes de débat apparaissent. Ces dispositifs sont-ils les vecteurs d'une plus grande sécurité juridique et/ou d'une standardisation accrue des décisions, réduisant l'autonomie professionnelle des juges (2) ? Vont-ils conduire à une déjudiciarisation bienvenue des conflits ou bien participer d'une remise en cause du principe fondamental du droit d'accès au juge pour toutes et tous (3) ? Peut-on espérer d'eux qu'ils libèrent les avocats, en les délestant de tâches répétitives, ou bien craindre qu'ils signent la disparition d'une partie d'entre eux, au profit de machines et de robots (4) ? Quelles transformations profondes vont-ils entraîner sur le système juridique et judiciaire (5) et enfin, comment les réguler (6) ?

2.1. Peut-on prédire le jugement ?

Le débat autour du terme « justice prédictive » émerge progressivement dans la littérature (1). Justice 'prédictive' et prévisible sont progressivement distinguées (2). Le débat se focalise alors sur la nature même de la justice : est-elle encore humaine si elle est dépourvue de tout aléa (3) ?

2.1.1. *La « justice prédictive » : un néologisme malheureux*

Dans un premier temps, l'expression « justice prédictive » n'est pas interrogée et plus encore elle est reprise telle quelle²⁵, parfois en pointant le fait que plusieurs appellations sont possibles (Barraud, 2017). Apparemment, cela ne dure pas. Il faudrait bien sûr opérer un comptage systématique pour le vérifier. Mais, pour prendre un exemple, le contraste de ces premiers articles publiés en 2017 avec l'ouvrage *La justice prédictive* édité par l'Ordre des avocats en 2018 est très net. Sur la quinzaine de contributions, quatre auteurs seulement utilisent le terme « justice prédictive » sans aucune marque de distance (ni guillemet, ni remarque sur son caractère approprié ou pas). Sept questionnent la pertinence de la notion de manière cosmétique ou ponctuelle – et parfois varient au long de leur propre texte (Péré, 2018). Quatre seulement nourrissent une critique plus consistante. Comme s'il était devenu inévitable pour se faire comprendre d'utiliser cette expression (plusieurs estiment d'ailleurs qu'elle est désormais instituée), mais que reconnaître ses limites était aussi devenu, en quelque sorte, un passage obligé.

Plusieurs manières de prendre de la distance coexistent. Les guillemets en sont la manifestation minimale et restent parfois assez formels. L'affirmation du caractère inadapté, peu « approprié » (Molinié, 2018 : 64), voire « trompeur » (Amrani-Mekki, 2018 : 39), de ce néologisme en est une autre, un peu plus marquée qui s'exprime dans les formules suivantes : « la justice dite prédictive » (Jeuland, 2019 : 194), la « justice (dite) prédictive » (Barthe, 2021[2017] : § 7 du résumé), ou « la justice qui n'a de prédictive que le nom » (Chaduteau, 2018 : 74-75), « qui porte assez mal son nom » (Jacquemin, 2019 : 121), dont l'adjectif « semble être celui qui convient le moins bien » (Hyde, 2019 : 2), ou bien encore « oxymore » (Mouillard, 2018 : 81) tellement contre-nature que « la justice prédictive, à proprement parler, n'existe pas » (Gaudemet, 2018 : 111).

Le développement de raffinements terminologiques (Barraud, 2017 : 7 ; Hyde, 2019 : 2), les efforts définitionnels²⁶, la proposition de substituts en sont d'autres, plus fortes encore. « Justice prédite par des algorithmes » et non « justice qui prédit » (Lebreton-Derrien, 2018 : 4), « justice algorithmique » (Barraud, 2017), « justice potentielle » (Jeuland, 2017 : 11), « justice quantitative » ([Monta] et [Sinclair], 2017), « justice digitale » (Garapon et Lassègue, 2018), « modes algorithmiques d'analyse des décisions (MAAD) » (Godefroy, 2019a : 38), « justice probabiliste » (Guyomar, 2018 : 97), « justice prévisionnelle²⁷ », « justice numérique » (*Les Cahiers de la justice*, 2019), « jurimétrie » et « justice simulative » (Degos, 2020), « analyse juridique prédictive » (Zambrano, 2015), « analyse mathématique et algorithmique des décisions de justice » (Barthe, 2021[2017] : § 7, point 2), « système algorithmique d'aide à la décision (SAAD) » (Desmoulin-Canselier et Le Metayer, 2020), sont

²⁵ ([Strippoli], 2017 ; Dondero, 2017 ; [Moline], 2017 ; Garapon, 2017).

²⁶ (Barthe, 2021[2017] ; Malabat, 2019 ; Louvaris, 2018 : 25-26 ; Lebreton-Derrien, 2018 : 4-5).

²⁷ Appel à recherches GIP Mission de recherche droit et justice, 2016 ; reconduit en 2017, 2018 et 2019. <http://www.gip-recherche-justice.fr/wp-content/uploads/2013/06/2016-A.O-Droit-justice-et-num%C3%A9rique.pdf> [consulté le 20 novembre 2020].

autant de propositions qui visent à se démarquer d'une expression qui sous-entend l'abandon de toute forme d'indétermination du jugement.

2.1.2. « Justice prédictive » ou prévisible ?

Un des principaux points d'achoppement concerne en effet la distinction entre prédire et prévoir (Barthe, 2021[2017] ; [Vial], 2018 ; Marin, 2018 ; Brigant, 2018 : 240). « La prédiction est l'acte d'annoncer (*prae*, avant – *dictare*, dire) par avance des événements futurs par inspiration surnaturelle. La prévision, quant à elle, résulte de l'observation (*visere*, voir) d'un ensemble de données, afin d'envisager une situation future. », affirme [T. Vial] (repris par Hyde, 2019 : 2). Le rôle de la traduction malheureuse de l'anglais vers le français est pointé par plusieurs auteurs ([Vial], 2018 ; Barthe, 2021[2017]).

Cela dit, même en français, il est d'usage dans les domaines scientifiques d'évoquer la « capacité prédictive d'une théorie » et de renvoyer ainsi à « sa valeur scientifique » (Rottier, 2018 : 190). Comme souligné en introduction de ce chapitre, le problème réside donc plutôt dans le fait d'accoler l'adjectif prédictif au mot justice pour en faire une expression nouvelle dont le sens est alors vacillant. Qualifier la justice de « prédictive » et non de « prévisible » reviendrait donc à simplifier à l'extrême le processus de fabrique du jugement, à négliger la pluralité de facteurs plus ou moins bien identifiés dont il est le produit ([Desjardins], 2019 : 270 ; [Vial], 2018 : 6 ; Sayn, 2019c) et sacrifier finalement à ce « lent glissement dans l'esprit collectif [qui] nous porte à croire que des machines, dénuées d'affects, seront un jour mieux à même de 'fiabiliser' l'acte de juger. » Et c'est bien ce qui est en question : dans quelle mesure les jugements peuvent-ils comporter une part d'incertitude, d'aléa qui, non seulement n'est pas anormale, mais est l'expression même d'une justice humaine ?

L'aléa judiciaire est souvent présenté comme « une propriété ontologique du Droit » (Zambrano, 2015 : 7), intrinsèque au jugement, en tant que celui-ci est une interprétation (Dondero, 2017 ; [Vial], 2018). L'œuvre de justice est fondamentalement « œuvre humaine » (Buart-Ménard, 2019 : 270) ; elle comporte une part d'« imprévisibilité nécessaire » (Rottier, 2018 : 189) et s'adresse à ce qu'il y a d'humain dans l'homme (Garapon, 2017 : 9 ; Garapon et Lassègue, 2018 : 134-135). L'intime conviction, qui est au cœur de l'exercice de la justice pénale, comporte en effet une « irréductible humanité » (Brigant, 2018 : 250), une « dimension psycho-affective », une « part indispensable d'humanité » dont certains assument la conséquence logique, à savoir reconnaître « la part d'erreur qui l'accompagne » ([Strippoli], 2017). Ils sont toutefois rares, ceux qui concèdent que cet argument de l'humanité de la justice implique d'accepter comme inévitable la part d'erreur propre au jugement humain. C'est un point aveugle de la discussion sur la « justice prédictive » puisque d'un autre côté, les promoteurs de l'utilisation de dispositifs algorithmiques ne vont pas jusqu'à les présenter comme un antidote aux erreurs, mais seulement comme un moyen de sécuriser les pratiques professionnelles des avocats et des juges. La notion de prévisibilité et l'adjectif 'prévisible' se dégagent progressivement comme des vocables acceptables pour la justice. Ils permettent de penser comme non antinomiques la rationalisation des décisions autour de régularités existantes (et qu'il faudrait peut-être intensifier), et la part d'indétermination et d'aléa de la décision que la reconnaissance de l'*imperium* du juge implique. En effet, pour les juristes et praticiens de justice, le principe de l'indépendance des juges à l'égard des pouvoirs

institués, constitue à la fois une règle du jeu posée et reconnue par les textes juridiques et une valeur cardinale constitutive de l'identité des magistrats, et plus largement de la justice en tant qu'institution.

2.2. Sécurité juridique accrue ou standardisation des décisions ?

L'argument d'une meilleure prévisibilité des décisions de justice pour les justiciables et leurs conseils est récurrent dans les études qui s'efforcent d'évaluer les atouts et les risques du recours aux algorithmes de traitement et d'analyse de grandes quantités de décisions de justice²⁸. Il l'est aussi dans les discours de certains acteurs des *legaltechs* qui les développent et les commercialisent (Katz, 2013 ; [Chevalet], 2018a, 2018b).

2.2.1. La sécurité juridique, un argument positiviste aux effets puissants

L'enjeu de prévisibilité est relié à la notion de sécurité juridique, principe général qui est devenu une règle de droit dans le système judiciaire français (Tulkens, 1990). Au tournant des années 2000 en France, face à ce qu'elles dénonçaient comme le fléau de l'inflation législative, les plus hautes juridictions (Conseil d'État, Cour de cassation, Conseil constitutionnel) se sont positionnées comme des acteurs de la sécurisation du système juridique, légitimant l'exercice d'un contrôle accru des juges sur la production législative et gagnant ainsi un rôle sociopolitique étendu (Vanneuille, 2008). L'enjeu de la sécurité juridique a donc été historiquement une manière pour les juristes d'asseoir une fonction de gouvernement du droit par les autorités juridictionnelles et les juges, face aux parlementaires et aux exécutifs. L'argument de la sécurité juridique est aujourd'hui repris pour défendre l'utilisation par les professionnels du droit d'outils algorithmiques de nature à fiabiliser les conseils des avocats et les prises de décision des juges.

2.2.2. Sécuriser comment ?

C'est le cœur du métier d'avocat et de conseiller juridique, que d'évaluer un cas, à partir de la connaissance d'un type de contentieux et de la jurisprudence nationale ou locale, et de le comparer à d'autres plus ou moins similaires. Pouvoir accéder à des analyses portant sur des milliers ou des dizaines de milliers d'affaires, les comparer à celle qu'un client vient lui soumettre, devrait permettre à l'avocat d'être plus solide dans son diagnostic et ses anticipations, concernant le contentieux lui-même, les voies de résolution du litige ou les perspectives de négociation (Chabert, 2019 : 2-3). Le recours à des algorithmes fiabiliserait, systématiserait voire vérifierait les régularités que les professionnels connaissent déjà (Chaisemartin, 2018). Là où ils procèdent de manière immédiate, intuitive, probablement partielle et orientée (le souvenir d'une affaire emblématique ou récente), les algorithmes établissent des calculs systématiques, sur un grand nombre de décisions. L'analyse du cas apporté par le client s'en trouve objectivée et le discours envers lui, plus argumenté. Les avocats, les directions juridiques de grands groupes en particulier dans le secteur de l'assurance (Kestenare, 2018 : 271), ont donc tout intérêt à se doter de ces outils, pour affiner leur expertise juridique. L'enjeu est aussi économique : rendre plus tangible le contenu de leurs prestations, par exemple sous la forme d'un dossier de statistiques résumant l'analyse

²⁸ ([Strippoli], 2017 : 334 ; Dondero, 2017 : 532 ; Bruguès-Reix & Pacquetet, 2018).

du cas au regard d'affaires comparables, et ainsi rendre plus acceptables les honoraires demandés. En ce qui concerne les juges, ces outils encadreraient les excessives disparités qui existent d'un juge à l'autre, d'une juridiction à une autre et qui sont présentées comme les marqueurs d'une absence d'égalité de traitement des justiciables. C'est évidemment tout le sens de la démonstration de [Morgan Forestier] avec [Legalplus] et l'étude sur les décisions des juges administratifs en matière d'OQTF.

2.2.3. *Les conditions d'une sécurité juridique accrue : exactitude et transparence*

Les algorithmes disponibles ne sauraient représenter une réelle plus-value en matière de sécurité juridique qu'à plusieurs conditions. Qu'ils soient effectivement performants, c'est-à-dire, d'une part, qu'ils ne fournissent pas de résultats erronés et, d'autre part, qu'ils livrent des statistiques pertinentes pour ceux qui les utilisent, autrement dit qui répondent aux questions précises qu'ils se posent, dans le contexte d'action qui est le leur (Chabert, 2019 : 2-3 ; Rottier, 2018 : 190-91). À cet égard, les bases de données utilisées déterminent le champ à l'intérieur duquel les résultats sont valables. Elles doivent donc être actualisées en temps réel pour intégrer les très nombreux changements du droit positif, au risque, sinon, de conduire à des erreurs, par exemple si l'adoption d'une nouvelle loi ou le changement de régime juridique intervenu, rendent caduc tout ou partie de la jurisprudence précédemment accumulée (Rottier, 2018 : 192). La validité des résultats repose en effet sur la qualité de ce travail, certes automatisable en partie, mais qui pour le reste doit être supervisé par du personnel qualifié en droit, pour lequel il convient donc de mettre les moyens nécessaires. L'interprétation juste des décisions de justice requiert une réelle technicité juridique et une méthode de dépouillement qui garantit des fausses interprétations et contresens. Dans un secteur où l'avantage compétitif réside pour beaucoup dans l'avance prise sur les concurrents, les *legaltechs* risquent d'être tentées de lésiner sur les moyens à mettre dans cette partie de leur activité — au profit de la logique marketing — pour valoriser rapidement des résultats et se positionner sur le marché (Dondero, 2017 : 6).

De plus, encore faut-il, pour s'assurer de la qualité de ces outils, qu'ils soient transparents (Sayn, 2019b : 3 ; Mouriesse, 2018), c'est-à-dire, dans le cas des techniques d'apprentissage machine, que leurs utilisateurs, et plus largement le public, aient « accès au code source de ces algorithmes, mais aussi [...] aux bases de données qui les ont entraînés ainsi qu'aux critères de sélection utilisés pour construire les bases d'apprentissage » (Bourcier et De Filippi, 2018 : 527). La remarque vaut également pour les bases de données à partir desquelles sont élaborées des statistiques et modélisations plus classiques. Il s'agit de permettre l'intelligence des résultats produits et l'exercice d'un regard critique sur eux. Ce dernier semble d'autant plus indispensable que des biais sont possibles et même avérés, pour tous les types d'algorithmes.

2.2.4. *De la sécurité juridique à la performativité des algorithmes sur les juges*

Le risque que les biais de discrimination raciale déjà présents dans un système judiciaire soient renforcés par les algorithmes est mentionné par de nombreux auteurs²⁹ qui

²⁹ (*inter al.* Déziel, 2018 : 256-259 ; Garapon, 2017 : 8 ; Garapon & Lassègue, 2018 ; Defferrard & Papineau, 2017 ; Brigant, 2018 : 243-244, G'sell, 2018 : 52).

se réfèrent généralement à l'étude publiée par Propublica (Angqin *et al.*, 2016) à propos des biais algorithmiques du logiciel Compas d'analyse des risques de récidive de personnes condamnées, utilisé dans certains États américains. Ils rejoignent alors le débat sur les effets pervers, notamment en matière d'amplification des discriminations, des échelles actuarielles utilisées pour la mesure des risques de récidive au stade pré et postsentenciel (Hildebrandt, 2018 ; G'sell, 2018) et plus largement encore de la littérature sur l'*evidence-based sentencing* (Starr, 2014). Les rapports de la CNIL (2017 : 31) et du Conseil de l'Europe (2017 : 29) font une large part à cet aspect.

La contribution des algorithmes à une plus grande prévisibilité des décisions de justice se pose donc de manière singulièrement plus aigüe si ce sont les magistrats, en charge de rendre des décisions, qui ont accès à ces outils et les utilisent. Quels peuvent être les effets d'une confrontation d'un magistrat aux nombreux jugements et arrêts de ses collègues ? Ne devient-il pas plus difficile de résister face à cette norme rendue visible et publique ? L'effet performatif que peuvent avoir ces algorithmes est souligné par de très nombreux auteurs³⁰. Est pointé le risque d'un comportement « moutonnier » des magistrats qui, pour de multiples raisons dont la pression à l'efficacité, pourraient se contenter de suivre leurs pairs. On peut imaginer en effet qu'il devient plus difficile d'aller à contre-courant des autres juges, de la norme qui se dessine surtout lorsque celle-ci est exprimée sous forme mathématique – par exemple à travers des pourcentages (Garapon, 2017) et des outils de visualisation (histogrammes, etc.). Plane alors le risque que la prévision devienne « prophétie autoréalisatrice » ([Desjardins], 2019 : 274 ; [Vial], 2018 : § 19) puisque les décisions prévisibles, statistiquement pertinentes, seraient les décisions prises. Subrepticement et avec leur consentement, les usages des algorithmes pourraient empiéter sur la liberté d'appréciation des juges. Ils abdiqueraient en partie le pouvoir officiel d'interpréter les textes qui leur est pourtant confié par le législateur, et dont la Cour de cassation rappelle qu'il doit s'exercer en faisant valoir la primauté de la loi sur les outils techniques d'aide à la décision que sont les barèmes³¹ ([Vial], 2018). Le risque est, pour le magistrat lui-même, celui « d'une perte de contrôle sur le processus de décision » (Clément, 2019 : 6), à laquelle sont particulièrement vigilants, dans leurs écrits, les magistrats judiciaires et administratifs de formation (Garapon) ou en exercice ([Desjardins], [Strippoli], Clément, Gimabias, [Vial]). Par le fait, le pouvoir des concepteurs d'algorithmes en deviendrait considérable et la dépendance qu'entreprendraient les magistrats à l'égard de ces acteurs économiques privés serait de nature à compromettre leur indépendance, pourtant prévue par les textes qui régissent leur statut (Cholet, 2018 : 231).

En outre, l'existence même de cette photographie des décisions de justice ne fournirait-elle pas les clefs d'un contrôle externe facilité ([Desjardins] et [Bourg], 2017) ? Cette interrogation naît notamment de l'expérimentation de [Legalplus], scénario-catastrophe bien

³⁰ (*inter al.* Amrani-Mekki, 2018 ; [Desjardins], 2019 ; Garapon, 2017 ; Garapon & Lassègue, 2018 ; Jacquemin, 2018 ; [Vial], 2018 ; Lebreton-Derrien, 2018 : 11 ; Hildebrandt, 2018 ; Sayn, 2019b).

³¹ [T.Vial] cite en particulier un arrêt de la Cour de cassation relatif au cas des barèmes de pensions alimentaires dans lequel la plus haute juridiction réaffirme la primauté de la loi sur les outils techniques : Cass. 1^{re} civ., 23 octobre 2013, n° 12-25.301 : JurisData n° 2013-023208.

réel du point de vue des magistrats : la dénonciation de pratiques individuelles rendues visibles et objectivées par un outil d'intelligence artificielle. Ce dispositif algorithmique a fait émerger des métriques de l'activité par président de chambre, point d'appui d'un discours sur l'absence d'objectivité ou d'impartialité dans les décisions et chez celles et ceux, nominativement identifiés, qui les avaient prises. Pour Yannick [Vial] (2018 : § 33), magistrat détaché au Conseil de l'Europe, « le fait est que les modèles prédictifs sont de véritables radiographies de la production des juges avec les risques de contrôle que cela comporte ». La mise en calcul des décisions conduit à un risque avéré de contrôle accru, direct ou indirect, sur les productions des juges, mais aussi sur la gestion de leurs carrières ou le fonctionnement des juridictions, si ces dispositifs étaient utilisés par les hiérarchies judiciaires. Comme nous le verrons dans le chapitre 3, ce précédent a joué un rôle important dans la mobilisation d'acteurs judiciaires pour l'adoption d'un article de loi interdisant certains usages des données identifiantes de magistrats³².

2.2.5. De l'harmonisation à la standardisation

La perspective de réduire les écarts entre décisions de justice, et donc de minimiser l'incertitude (sécurité juridique) ainsi que les disparités (sources d'injustices), peut apparaître séduisante. L'aléa judiciaire n'est alors pas pensé comme une nécessaire adaptation du droit et des juges aux cas toujours singuliers qu'ils rencontrent, mais comme l'expression d'une inexplicable variabilité dans l'appréhension des situations individuelles par les juridictions ([Monta] et [Sinclair], 2017). La logique de réduction d'écarts considérés comme pas toujours justifiés s'exprime alors à travers la recherche d'une harmonisation des décisions de justice, présentée comme bénéfique et souhaitable tant pour la cohérence du système judiciaire que pour ses destinataires, les justiciables. Peu de magistrats y sont favorables, à la notable exception de [Valère Desjardins]. Il a fait sa carrière à la fois au ministère de la Justice (sous-direction de la performance et des méthodes de la direction des services judiciaires) et à la Cour de cassation où il est devenu responsable du service des études et de la documentation. Il s'est beaucoup investi dans le développement de projets liés à l'*open data* des décisions de justice³³. Il défend l'idée qu'un rapprochement des décisions de justice serait positif, à condition qu'il conduise à une harmonisation et non à une standardisation. « La principale [perspective d'évolution particulièrement attendue] réside dans l'effet d'harmonisation qui renforcerait la prévisibilité de la décision de justice et, par voie de conséquence, la sécurité juridique. Il s'agirait là, répétons-le, de favoriser une saine convergence des jurisprudences, et non une standardisation irraisonnée, par le moyen d'un dialogue "horizontal" des juges renouvelé grâce aux avancées de la technique. » ([Desjardins], 2019 : 275) La « saine

³² L'article 33, al. 4 de la loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice, interdit que « les données d'identité des magistrats et des membres du greffe [fassent] l'objet d'une réutilisation ayant pour objet ou pour effet d'évaluer, d'analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées ». Il reste à savoir si cette disposition s'appliquera aussi aux organisations auxquelles magistrats et greffiers appartiennent, dans le cadre de leurs fonctions de gestion de l'activité et des ressources humaines...

³³ Sources : Légifrance et site de la CEPEJ, <https://www.coe.int/fr/web/cepej/biography-eloi-buat-menard> [consulté le 9 octobre 2021].

convergence » est ici distinguée d'une « standardisation irraisonnée », perçue comme dangereuse.

Le choix du mot harmonisation et son opposition à celui de standardisation sont bien plus qu'un effet de style ou une subtilité linguistique. On les retrouve d'ailleurs chez un magistrat du ressort de la cour d'appel de Douai, [Simon Place], — favorable aux outils permettant de contribuer à la convergence des décisions — et qui a testé le dispositif algorithmique [PréviCompute]³⁴. Ils créent un espace pour penser le rapprochement des décisions de justice sans leur uniformisation ainsi que la recherche de cohérence entre les juges sans la remise en cause de leur pouvoir d'interprétation ou de leur *impérium*. L'harmonisation procède d'une logique acceptable juridiquement et judiciairement : harmoniser le droit ou les manières de faire y compris pour la justice, c'est faire ce qui a été entrepris, par exemple au niveau européen, au fil de la progression de la construction européenne (Vauchez, 2013). Le terme ouvre une perspective tolérable, notamment de dialogue des juges. Plus encore, il dessine un espace de rencontre possible entre la logique proprement juridico-judiciaire d'autonomie professionnelle des magistrats (appuyée sur le principe d'indépendance) et d'autres logiques, administrative et gestionnaire de recherche de comparabilité des productions juridictionnelles entre elles. Il ouvre aussi à une possible rencontre avec des enjeux citoyens de recherche d'égalité de traitement de cas comparables et de pédagogie des décisions.

Mais de la prévisibilité et l'harmonisation à la standardisation, il n'y a qu'un pas. En réalité, le développement des dispositifs algorithmiques de traitement et d'analyse des décisions de justice peut être un aiguillon pour avancer plus vite dans la convergence des pratiques juridictionnelles, par exemple à travers la systématisation de trames ou de formulaires préconstruits de jugement. En effet, les dispositifs algorithmiques seront d'autant plus performants et compétitifs ([Strippoli], 2017) que la matière première qu'ils analysent sera déjà organisée et normée, par exemple avec un rubricage précis des différentes parties du jugement ou le recours à des vocables stabilisés, ce qui est loin d'être le cas actuellement dans toutes les affaires et toutes les juridictions. Certains types de litiges, par leurs caractéristiques, seraient standardisables plus facilement que d'autres.

Parmi ceux qui embrayent sur le développement de dispositifs algorithmiques figurent donc des praticiens qui, comme [Simon Place], se sont déjà positionnés et investis activement dans la recherche d'une plus grande cohérence des décisions de justice, en travaillant à la formulation de référentiels (local puis reconnu au plan national et enseigné à l'ENM) en matière de dommages corporels ([Place], 2014). D'autres qui ont à cœur d'adapter le droit aux enjeux des acteurs économiques se positionnent du côté d'une recherche de rationalité accrue. C'est le cas de [Thibault Strippoli], magistrat, conseiller à la cour d'appel de Paris qui

³⁴ Il estime ainsi que l'algorithme en question « permet d'être sécurisé dans la décision et d'être sûr que d'autres ont pu avoir des appréciations comparables. *C'est une harmonisation plus qu'une uniformisation* », article édité par C. Mimaut et L. Carrive, mis en ligne le 17 mai 2017, https://www.francetvinfo.fr/societe/justice/justice-predictive-le-big-data-frappe-a-la-porte-des-tribunaux-et-des-cabinets-d-avocats_2196972.html [consulté le 13 octobre 2021].

s'est intéressé dès 2017 aux enjeux de l'IA dans la justice. Vice-président de Présage, un *think tank* portant sur les domaines de l'économie, du droit et de la justice, il est aussi directeur de la collection « Macro droit - Micro droit » que soutient cet institut en partenariat avec les éditions Larcier et Bruylant et dans laquelle [T.Vial] (2020) a fait paraître un ouvrage sur l'IA³⁵. [T.Strippoli] estime que, pour des contentieux massifs, déjà très standardisés et pour lesquels il existe parfois déjà des barèmes, les dispositifs algorithmiques iraient plus loin vers une forme d'automatisation des décisions. Cela permettrait de gagner en efficacité et d'orienter les ressources judiciaires limitées vers les contentieux à forts enjeux sociétaux. Il exclut toutefois la matière pénale par principe.

La discussion dépasse donc les seuls dispositifs informatiques, numériques ou algorithmiques ; elle s'active et se réactive depuis plusieurs décennies à l'occasion de tous les dispositifs orientés vers la mise en cohérence ou le rapprochement des décisions de justice (Licoppe et Dumoulin, 2019 ; Sayn, 2014 et 2020) : nomenclatures établissant des catégories de références (par exemple pour les préjudices en matière de dommage corporel), référentiels et barèmes fournissant des modes de calcul et des fourchettes de montants d'indemnités (de licenciement, pension alimentaire, préjudices corporels, etc.). Que ce soit en matière civile (Cousin, 2017) ou pénale (Gautron, 2014), l'enjeu de ces outils est de gagner en efficacité dans la fabrique des décisions, dans une logique de gestionnarisation des services publics comme la justice, mais aussi de rationalisation des processus et des produits, afin de rendre les décisions plus homogènes. Cela fait directement écho aux débats que nous avons suivis lors de la mise en place de l'informatique et d'Internet dans les juridictions. Les trames de saisie des jugements développées par quelques magistrats et greffiers dans les années 1990 poursuivaient le même but — rationaliser pour gagner du temps lors de la rédaction et progresser en cohérence — et se heurtaient aux mêmes réactions hostiles de leur milieu, inquiet des formes d'optimisation qui encadrent et minent peu à peu la liberté d'appréciation des juges (Dumoulin et al., 2007). Car si certains praticiens sont demandeurs d'une plus grande cohérence et agissent en ce sens, d'autres mettent en avant les risques d'une remise en cause du pouvoir des juges, de plus en plus cadrés dans leurs décisions d'une part et de plus en plus concurrencés par des plateformes privées de règlement en ligne des conflits d'autre part, dans un contexte où les réformes de justice visent à encourager la déjudiciarisation.

2.3. Déjudiciarisation des conflits vs accès à la justice

Dès lors que les décisions sont plus prévisibles, il devient non seulement plus facile d'en automatiser certaines, mais aussi de stimuler les alternatives au procès et les règlements amiables, sources d'économies de moyens dans la gestion des ressources limitées de la justice (Rottier, 2018 : 191). Les algorithmes de « justice prédictive » seraient « un facteur essentiel de dynamisation des modes alternatifs de règlement des litiges, en permettant aux parties de transiger sur la base d'une anticipation crédible et solide de leurs perspectives de succès. »

³⁵ La collection d'ouvrages s'adresse aux acteurs institutionnels et économiques pour « leur permettre d'agir dans un environnement juridique et économique en mutation », notamment dans un contexte de droit globalisé où normes nationales, européennes et internationales s'interpénètrent.

([Desjardins], 2019 : 275) D'où il découlerait « un évitement vertueux du procès [au bénéfice] d'une concentration des moyens de la justice sur les situations méritant une plus grande attention, notamment, en ce qu'elles ne se rattachent pas d'évidence à un courant jurisprudentiel établi. » (*ibid.*) La lecture positive de l'automatisation et de l'accroissement de la part des affaires traitées sur un mode amiable, repose sur l'idée selon laquelle le temps gagné sera réinjecté dans les cas plus complexes et/ou dans la relation d'écoute et de proximité avec les justiciables ([Vial], 2018 ; [Strippoli], 2017). Ce type d'argumentation qui repose sur une mécanique de vases communicants n'est pas nouvelle ; elle est régulièrement invoquée, par exemple dans le cadre du développement de la médiation pénale. Or elle raisonne à nombre d'affaires constant, ce qui est théorique. La lecture plus circonspecte consiste à pointer les dangers pour la justice, en tant que valeur et qu'institution, d'une organisation dans laquelle le juge ne connaîtrait plus que des questions hautement controversées ou des situations nouvelles, pour lesquelles il « ne trancherait que pour lever les cas incertains décidés en dehors de lui » (Cholet, 2018 : 235).

En outre, un débat se dessine qui met en avant le droit à l'accès au juge et à la justice et le risque qui consisterait, en pratique, à priver les justiciables de cette possibilité. En effet, depuis plusieurs décennies, différentes réformes ont favorisé le règlement non contentieux des litiges et cherché des voies parallèles pour délester les juridictions et les magistrats débordés (Cadiet, 2017 : 522). Les dernières en date prévoient, pour les litiges d'un montant inférieur à 4000 €, l'obligation pour les parties de tenter un règlement amiable avant de pouvoir entamer une procédure contentieuse. La médiation préalable obligatoire trouve, à travers le développement des algorithmes d'analyse de décisions de justice, un appui technique précieux. Ils apportent des informations bienvenues lorsqu'il s'agit de mener une négociation entre les parties, et peuvent « faciliter l'émergence des transactions souhaitées », comme l'indique le directeur d'un groupe privé d'assurances à l'issue de l'expérimentation de logiciels de « justice prédictive » (Kestenare, 2018 : 276). C'est d'ailleurs ce qui ressort de la discussion autour de la nouvelle plateforme de médiation du Conseil national des Barreaux (Garenne et Muller, 2019 : 6). Les algorithmes pourraient, *a fortiori* s'ils étaient ensuite associés à des plateformes de règlement en ligne des litiges (*online dispute resolution* [ODR]), opérer comme des « filtres privés d'accès au juge » (Louvaris, 2018 : 33). Le recours au juge pourrait être dissuadé sur la base de données issues d'algorithmes, pas entièrement fiables, comportant peut-être des biais et conçus par des entreprises privées, mues davantage par leur intérêt économique que par l'intérêt général (Sauvé, 2018 : 10). Le risque est donc que les dispositifs algorithmiques participent d'un système dans lequel des solutions de substitution au procès sont non seulement proposées, mais conçues de telle manière qu'elles détournent les justiciables de leur droit fondamental à l'accès au juge (Amrani-Mekki, 2018). À terme, les *legaltechs* qui proposent ces services et/ou des règlements en ligne des litiges³⁶ pourraient réaliser des missions de service public facturées aux justiciables qui disposent pourtant d'un droit d'accès au juge étatique. Emmanuel Jeuland note qu'un déséquilibre important est déjà en train de se créer entre différents types d'utilisateurs : alors que pour les litiges liés à la consommation, « les services ODR privés payants sont obligatoires » pour les

³⁶ Par exemple [Demanderjustice.fr].

justiciables (souvent des particuliers), en revanche, « pour les injonctions de payer ce sont les professionnels non payés [commerçants, etc.] qui vont avoir un titre exécutoire sans audience qu'ils pourront utiliser sans frais alors que le débiteur sera, lui, dépourvu de moyens de défense. » (Jeuland, 2018 : 196) La controverse déborde donc l'enjeu technologique des dispositifs pour questionner celui des modèles de justice en concurrence et de la place du marché. L'irruption d'acteurs économiques privés, porteurs d'innovations en matière d'aide à la décision des avocats et des juges, interroge le secteur de la justice, ses acteurs et pratiques, quant à l'émergence de formes de marchandisation des activités de règlement des conflits.

2.4. Libération vs disparition des professionnels du droit

L'hypothèse de la lente disparition des professionnels du conseil juridique que sont les avocats suscite nombre d'inquiétudes parmi les juristes et génère des réflexions sur une comment revoir l'organisation des professions juridiques (Cassar, 2020). D'autant plus que pour certains, il s'agirait plutôt de leur remplacement rapide et brutal par des robots, automates ou machines nés de l'IA (Barthe, 2021[2017] : introduction). C'est en tout cas la prévision de Richard et Daniel Susskind (2015) dans leur ouvrage intitulé *L'avenir des professions libérales : Comment les nouvelles technologies vont transformer le travail des experts humains* et de toute une série d'analyses qui au plan international vont dans le même sens (Fenwick et Vermeulen, 2019). L'emballement autour de la mise en danger de la profession est notable et provoque comme une onde de choc.

L'interprétation de ce qui est en train de se jouer n'est toutefois pas stabilisée tant l'exercice de prospective laisse entrevoir une diversité de scénarii (Katz, 2013). En France, Louis Boré, avocat à la Cour et au Conseil et président de l'Ordre, en propose deux. Une histoire qui finit bien : les avocats tirent parti de ces outils, se délestent de ce qui peut être automatisé et se consacrent davantage aux tâches plus intéressantes et en particulier à l'écoute de leur client ; et une histoire qui finit mal dans laquelle cette révolution technologique n'oblige pas les humains à se former, à évoluer, comme les précédentes révolutions schumpéteriennes, mais supprime l'humain et le remplace par des automatisations (2018 : 4). Le spectre de la disparition de certains métiers et de la transformation d'autres plane (Gamet, 2019 : 119). Les jeunes générations de juristes risquent d'être particulièrement touchées, voire sacrifiées dans ce processus qui pourrait avoir pour effet de supprimer les emplois occupés par les assistants qui effectuent des tâches automatisables (recherche de documentation juridique, rédaction de contrats simples, etc.) au sein des cabinets et firmes juridiques. D'autant qu'à l'inverse, le nouveau marché juridique en construction pourrait leur échapper, ce qui les empêcherait d'en tirer parti. D'ores et déjà, les *startup* qui interviennent dans le secteur du droit et de la justice sont davantage créées par des étudiants en commerce que par des étudiants en droit. Si cela se confirmait, les juristes pourraient être les grands perdants de la nouvelle chaîne de valeur créée, en faveur de profils d'ingénieurs et de commerciaux (Croze, 2017 : 175).

Le courant dit *Law & Economics* consacré à l'analyse économique du droit (Deffains et Ferey, 2010 ; Kirat, 2012) s'intéresse précisément aux recompositions du marché des services juridiques entraînés par le développement de l'IA (Deffains, 2019). Et il ne dit pas forcément

autre chose. D'une part, l'IA peut contribuer à élargir le marché en faisant émerger de nouveaux segments tout en provoquant la baisse des prix des prestations juridiques pour les consommateurs de droit (Fortuit et Hamidou, 2017 : 27), mais dans le cadre d'une évaluation des prix et de la qualité des services toujours délicate (Kern, 2017). D'autre part, les technologies de l'IA risquent de conduire à la disparition de certains types de tâches répétitives, désormais automatisées, au profit d'une plus grande efficacité des professionnels du droit qui seront parvenus à investir pour s'équiper. La chaîne de valeur des prestations juridiques devrait s'en trouver modifiée de même que le marché lui-même avec un fort risque de concentration entre quelques acteurs dominants, éventuellement de nouveaux entrants mieux adaptés pour répondre à ce nouveau contexte (Deffains, 2019 : 46). Les perspectives d'intégration de dispositifs algorithmiques d'analyse des décisions de justice à des plateformes de justice en ligne dessinent ce qui pourrait être une transformation importante de l'économie même de l'exercice de la justice extrajudiciaire (Biard, 2019 : 180).

2.5. Transformations profondes du système juridique et du rôle du droit

Outre les professionnels du droit et le marché juridique, c'est peut-être la logique d'ensemble du système juridique français qui serait menacée d'une ou plusieurs évolutions radicales.

2.5.1. Vers un système de précédents ?

Traditionnellement légicentrique, celui-ci repose sur la prééminence du droit légiféré que le juge met en œuvre. La jurisprudence contribue certes à affiner les conditions d'application de la loi, mais elle n'agit pas comme un précédent obligatoire, sauf en ce qu'elle émane de l'autorité supérieure qu'est la Cour de cassation. « Elle constitue un "précédent d'interprétation", et non un "précédent de cas", comme dans les pays de *common law* » (Serverin, 2018 : 23-24). L'application de l'informatique documentaire aux sources juridiques a mis à l'épreuve cette conception verticale de la jurisprudence, sans parvenir à la transformer. Les bases de données ont été organisées pour préserver ce modèle hiérarchique, et les Cours suprêmes, Cour de cassation et Conseil d'État, forts de leurs prérogatives ont joué un rôle crucial en la matière. Le principe de sélectivité des décisions indexées dans ces bases, la non-intégration des jugements de première instance en témoignent de même que la gestion en direct de Juri-Ca par les services de la Cour de cassation. Or, la loi Lemaire de 2016 en alignant le régime des décisions de justice sur celui de l'ensemble des décisions publiques a créé un accès ouvert et gratuit à toutes les décisions du fond, de quelque juridiction qu'elles proviennent, première instance, appel ou juridictions suprêmes. Comme l'analyse très bien É. Serverin (2018 : 43), ce changement profond, consistant à injecter « une masse de décisions non sélectionnées dans un système jurisprudentiel historiquement sélectif », n'a pas donné lieu à une théorie et suscite bien des interrogations sur ce qui pourra en découler.

Dès lors, deux évolutions sont possibles : soit vers un système de précédents où chaque jugement comporte une puissance normative, quelle que soit la juridiction dont il émane ; soit vers un renforcement du système actuel dans lequel la fonction d'interprétation des juridictions sera confortée, « avec l'objectif plus ou moins assumé de contrôler et de réduire les divergences » (*ibid.*, p. 43). La plupart des observateurs juristes se focalisent sur la

première option, celle d'un renforcement du rôle de la jurisprudence signant l'évolution voire la dénaturation du système français, entraîné dans une logique de *common law* (Croze, 2017 ; Garapon, 2017 ; Epineuse et Garapon, 2018). Dans un système qui ne connaît pas le *stare decisis*, ni le précédent au sens de l'obligation de juger de la même manière un cas déjà jugé, comme c'est le cas en *common law*, la mise à disposition de décisions juridictionnelles produites par différents niveaux de juridiction peut participer à une transformation importante des mécanismes concrets de production du droit, à travers la montée en puissance de ce qui est qualifié de « jurisprudence horizontale » (Cottin, 2019 : 24).

Avocat, Hervé Croze (2017 : 175), ajoute que si des outils d'aide à la décision comme les dispositifs algorithmiques de traitement et d'analyse des décisions de justice existent, et bien que le juge reste maître de sa décision en principe, « il n'empêche que si un juge prétend s'écarter de la décision "normale", il semblerait logique qu'il s'en explique au vu des circonstances particulières. On reconnaît là le modèle anglo-saxon du précédent et la technique de la distinction qui est, pour le moment, étranger au système de *Civil law*. » Antoine Garapon le rejoint, regrettant que les opérations algorithmiques reposent sur « un savoir en lui-même conservateur qui donne un grand rôle à la jurisprudence aux dépens du législateur » (Garapon, 2017). D'autant plus que le risque est double en réalité : la remise en cause du système de droit écrit, par une mobilisation accrue et qualitativement différente des précédents d'une part, et l'appauvrissement de la faculté d'innovation propre au système juridique lui-même (Hildebrandt, 2018) d'autre part, puisque le système sera autoalimenté « avec des décisions identiques » ([Vial], 2018 : § 33). La logique algorithmique et statistique pourrait donc entraîner des effets majeurs sur la production et la standardisation du droit, considérés à un niveau global ([Vial], 2018 ; [Desjardins], 2019).

2.5.2. Vers une factualisation du droit ?

L'équilibre ou le déséquilibre entre fait et droit dans la production des décisions de justice est également en cause et participe de ce débat. Nous avons vu plus haut que l'équipe de N. Aletras a d'emblée positionné les résultats de ses travaux comme une confirmation des thèses réalistes américaines. La discussion par B. Barraud (2017) a justement contesté cette conclusion. De manière plus générale, la « justice prédictive » suscite des débats qui mobilisent les théories réalistes américaines et/ou scandinaves, soit parce que les auteurs s'inscrivent à proprement parler dans le champ de la théorie du droit (Barraud, 2017, Zambrano, 2015), laquelle s'est questionnée de longue date sur le rôle du juge et du jugement dans la fabrique du droit ; soit parce qu'ils abordent de manière plus secondaire — et parfois superficielle — certains auteurs, principes et débats de la théorie réaliste. Au fil de la littérature, une double interrogation est développée : dans quelle mesure les dispositifs algorithmiques actuellement commercialisés donnent-ils raison aux théories réalistes ou permettent-ils de vérifier leurs affirmations (Aletras et al., 2016 ; Barraud, 2017) ? Dans quelle mesure sont-ils les vecteurs par lesquels s'opérationnalise le passage vers une conception réaliste du jugement (Garapon, 2017 ; Croze, 2017) ?

Le réalisme s'est développé en réaction au positivisme juridique et en affirmant l'importance de la description de ce que font les acteurs chargés d'appliquer le droit. La vulgate réaliste consiste à affirmer le triomphe du fait sur le droit et à considérer que les

dispositifs algorithmiques incarnent celui-ci. « Au fond, la justice prédictive se rattache à une théorie réaliste du droit où tout est fait, tout est au même niveau : le passé d'un juge, la position majoritaire des autres juges, la décision de première instance et celle de la Cour de cassation. La règle issue de la loi ou l'information donnée dans une circulaire ou une recommandation ont la même valeur. » (Jeuland, 2017 : 10) Pour Zambrano (2015 : 7), le *big data* et les algorithmes doivent permettre d'enfin mettre en œuvre le programme des réalistes américains, celui d'Holmes en particulier, en refondant la science juridique sur un substrat empirique, et ce, à travers une « révolution méthodologique ».

La réflexion en termes de « factualisation du droit » contribue à cette discussion dans le cadre francophone. Elle a été proposée par Hervé Croze (2017) dans une note de doctrine de janvier 2017 et abondamment reprise par la suite³⁷. Elle repose sur le constat selon lequel « les statistiques relèvent de l'empire du fait. C'est l'une des manifestations de l'importance des éléments factuels dans la réflexion, la production et l'activité juridique. Une sorte de revanche du fait qui grignote l'empire du droit » estime H. Croze (2017 : 174). Cette « revanche du fait » est observable à travers une diversité d'indices qui témoignent de la relativisation de la place des énoncés juridiques dans la fabrique des jugements et arrêts. Le développement de la « justice prédictive » n'en serait pas la première manifestation. Depuis longtemps, certains contentieux hautement techniques confèrent une place prédominante aux faits dans les décisions de justice. C'est le cas des brevets d'invention dans le domaine de la propriété intellectuelle ou du contentieux de la concurrence « où les avis de l'Autorité de la concurrence sont essentiellement d'ordre économique », c'est-à-dire qu'ils procèdent d'une analyse de l'intérêt et des conséquences de la décision sur un plan économique (pour le marché, les entreprises, etc.) davantage que d'une analyse strictement juridique de la situation.

Mais de fait, les dispositifs algorithmiques de traitement et d'analyse des décisions de justice semblent accentuer considérablement ce mouvement déjà en marche de « factualisation du droit ». Ils mettent sur le même plan tous les éléments qui sont également transformés en données informatiques, « qu'il s'agisse de données juridiques ou des données factuelles à chaque espèce » (Cholet, 2018 : 233). Avec cette « factualisation », c'est un changement de paradigme profond qui serait instauré, non seulement dans la logique interne du système juridique français, mais plus largement dans le rôle de régulation normative qui lui est habituellement conféré, puisque l'énoncé juridique n'aurait plus de force normative supérieure aux autres énoncés contenus dans les jugements et arrêts. Se produirait un nivellement qui araserait le surcroît de normativité que les juristes reconnaissent au droit dans l'élaboration des décisions de justice.

Le sujet est d'autant plus préoccupant que les résultats statistiques obtenus par les algorithmes pourraient se révéler faux juridiquement puisqu'ils « sont totalement étrangers à la question de la conformité en droit de telle ou telle solution, dès lors que ces outils procèdent à des calculs de probabilités et ne peuvent pas discriminer entre les raisonnements juridiquement valides et ceux qui ne le sont pas. " Le risque existerait donc que ces calculs

³⁷ (Jeuland 2017 ; 2019 ; Garapon, 2017 ; Garapon & Lassègue, 2018 ; Cholet, 2018 : 233).

produisent des solutions logiques sur le plan des corrélations et probabilités par rapport aux données sur lesquels ils auraient été entraînés, mais inexacts juridiquement. Il est donc ici question non seulement d'un risque de changement du statut de la jurisprudence dans le système juridique français ; d'une relativisation de la portée des énoncés juridiques au profit des faits ; d'une sclérose du système juridique par la reproduction de décisions passées, mais plus encore d'un risque de duplication de solutions fausses juridiquement. Quid des erreurs de droit provoquées par des statistiques et inférences erronées ? « Tout le droit devient un fait et réciproquement n'importe quel fait, légitime ou non, devient normatif. Ce postulat nécessaire à la méthode prédictive est extrêmement révélateur de la lente substitution du registre cognitif au registre normatif à laquelle on assiste aujourd'hui. » (Garapon, 2017) [T.Vial] comme Antoine Garapon, tous deux magistrats, pointent les risques d'un glissement vers un gouvernement par les nombres, au détriment des principes juridiques en vigueur, et même en les outrepassant.

2.5.3. *Les algorithmes, moyens de rationaliser le droit ou de le désymboliser ?*

Un débat d'un niveau encore plus méta se profile alors. Les nouvelles possibilités ouvertes par l'analyse et le traitement algorithmiques de la masse des décisions de justice sont-elles le moyen de rationaliser davantage le droit ou bien travaillent-elles à un affaiblissement de sa puissance symbolique ?

Guillaume Zambrano (2015) lui-même impliqué dans la transformation du droit par les technologies, célèbre les promesses du *big data* combiné aux algorithmes pour la science du droit. Pour lui, ils sont l'occasion que la doctrine juridique opère enfin un tournant empirique et développe un volet jurisprudentiel dont l'objectif serait la prévision fiable des décisions judiciaires. Il assimile ainsi méthode scientifique et analyse prédictive, à partir d'une conception radicalement positiviste de la science : « Le but ultime de toute méthode scientifique réside dans l'aptitude à effectuer une prédiction fiable d'événements futurs sur la base d'une connaissance expérimentale. [...] Une approche scientifique du Droit — qui ne se réduit pas à une théorie pure des normes — doit nécessairement inclure l'étude de l'application effective des normes par les tribunaux, et partant proposer des prédictions jurisprudentielles vérifiables. » (Zambrano, 2015, p.2-3) Il associe foi dans le positivisme scientifique et dans le programme réaliste. Pour lui, la notion de « prédiction juridique » telle que formulée par Holmes³⁸ doit être pleinement investie par les nouveaux moyens scientifiques à la disposition des juristes et profiter du « renouvellement méthodologique » en cours. « Il est à présent possible de procéder à l'analyse statistique exhaustive des décisions rendues par les tribunaux, pour former des prédictions précises et quantifiables sur ce qui sera décidé à l'avenir dans tel ou tel type de cas ». (*Ibid.* p.7) Il appelle à un « changement de paradigme, au profit d'une épistémologie probabiliste décrivant le système juridique en termes statistiques d'application de telle ou telle règle à tel ou tel cas. » Les dispositifs algorithmiques sont des alliés de la « grandeur » du droit dans la mesure où ils sont (et avec

³⁸ Il s'agit du célèbre aphorisme d'Oliver Holmes (1897 : 460-61) selon lequel le droit n'est rien d'autre que la prévision de ce que décideront les juges (ma traduction de : « The prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law »).

eux, la « nouvelle génération de chercheurs en Droit ») « la prothèse technologique qui permettra de restaurer la prévisibilité obscurcie du Droit. » (Zambrano, 2015, p. 7) On retrouve ici l'idée de la rationalisation qui, par le truchement des algorithmes et du *big data*, poursuit les ambitions de grandeur, de clarté et de prévisibilité d'un « droit-référence » (Commaille, 2015). Les algorithmes permettraient donc d'aller plus loin dans le processus inachevé de rationalisation formelle du droit (Coutu, 2018 : 72-74), sur le plan de sa structure formelle et logique interne, mais surtout sur celui de sa validité empirique, c'est-à-dire de sa capacité à apparaître stable, prévisible et donc à être intégré dans les calculs et comportements des acteurs.

Or, à rebours de cette interprétation, il peut être considéré que ce qui est à l'œuvre ne relève pas d'une étape de plus dans la rationalisation du droit, mais au contraire d'un processus de déclin de la capacité de celui-ci à opérer la médiation institutionnelle et symbolique qu'il effectue classiquement. Pour Garapon (2017) et Garapon et Lassègue (2018), le droit et le numérique entretiennent une relation de concurrence autour du pouvoir de symbolisation en place dans les sociétés : « le savoir algorithmique [désymbolise dans la mesure où il] fait ressortir la fragile humanité des juges, les limites de la justice 'juridique' et l'artificialité des fictions pourtant nécessaires à la vie du droit ; [il resymbolise] en leur substituant la certitude scientifique. » (Garapon, 2017) Cela est particulièrement clair à travers l'argument — évoqué *supra* — selon lequel les outils prédictifs aideraient à pallier la trop grande variabilité des décisions produites par les juges et à les rendre plus cohérentes entre elles. Des auteurs comme Emmanuel Jeuland (2017 ; 2019), Hervé Croze (2017) ou Stéphanie Lacour et Daniela Piana (2019) font écho à l'idée selon laquelle les technologies numériques parce qu'elles dématérialisent et fragmentent le processus et l'espace judiciaires le désymbolisent et le déritualisent. Or la médiation technologique n'est pas en soi porteuse de sens et demande à être investie de significations relatives au lien social, au projet sociopolitique dans lequel elles sont inscrites.

2.6. Quelle régulation ? *Soft law*, autorégulation ou réglementation ?

Si de nombreux travaux s'interrogent sur les potentialités et les risques de l'utilisation d'algorithmes de traitement de bases de décisions de justice, la réflexion aboutit généralement sur le caractère crucial de la régulation des innovations et pratiques émergentes³⁹, en continuité avec les enjeux de régulation (Ganascia, 2019) et notamment de responsabilité posés par les développements de l'intelligence artificielle dans différents secteurs (Barfield et Pagallo, 2018 : 144-280). Mais si tous les acteurs et auteurs reconnaissent l'intérêt d'une telle régulation, toutes et tous ne la conçoivent pas de la même manière. Assez logiquement, les concepteurs de dispositifs algorithmiques mettent en avant la capacité d'autorégulation du secteur par ses entreprises, à travers des chartes et codes de bonne conduite ([Chevalet], 2018a). De fait, les acteurs de la transformation du droit par le numérique ont commencé à s'autoréguler par la mise en place de tels instruments comme la charte éthique proposée par les associations Open law et l'Association pour le développement

³⁹ ([Desjardins], 2019 ; Desmoulins-Canselier & Le Métayer, 2020 ; Garapon, 2017 ; G'sell, 2020a ; Louvaris, 2018 ; [Vial] 2018 et 2020 ; Vermeys, 2020 ; Amoroso & Tamburrini, 2021).

de l'informatique juridique (ADIJ) sur lesquelles on reviendra dans le chapitre 3. Le Comité d'éthique et la charte adoptés par la société [PréviCompute] en sont un autre exemple ([Chevalet], 2018a : 88). Nées d'une démarche d'autorégulation, ces initiatives n'instaurent toutefois ni processus de contrôle ni capacité de contrainte externe sur ceux qui les mettent en place. Plusieurs auteurs (Épineuse et Garapon, 2018 : 19 ; [Vial], 2018 et 2020), expriment leur circonspection à l'égard de ces pratiques d'autorégulation compte tenu de leur portée limitée d'une part et des enjeux éthiques et financiers considérables d'autre part. Le risque est bien celui de succomber à « l'illusion actuelle [...] de l'éthique comme un paravent qui renvoie à un mélange entre morale et intérêt d'entreprises se voulant “socialement responsables” ». Dans ce cadre, l'autorégulation ne risque-t-elle pas de conduire à un système de normes privées fonctionnant en circuit fermé, car totalement autocentré ? Les principes et les chartes éthiques adoptés comme guides de référence pour l'intelligence artificielle et le *big data* n'ont-ils pas eu pour conséquence jusqu'à présent de dissuader les autorités de légiférer alors que le marché de la transformation numérique est encore un marché naissant ? » (Deffains, 2018 : 278)

D'autres formes de régulation sont également possibles, faisant appel aux instruments classiques de la puissance publique, à l'échelle nationale ou européenne. La loi française qui interdit d'utiliser les données identifiantes des magistrats en est un exemple. Les magistrats investis sur le sujet de l'IA, comme c'est le cas de [T.Vial], publient articles et ouvrages (2020) pour défendre cette option. Il s'agit d'aller plus loin encore que les dispositifs existants — essentiellement la Charte éthique européenne d'utilisation de l'intelligence artificielle dans les systèmes judiciaires et leur environnement (Cepej) — ne le prévoient. Les pistes évoquées sont un régime juridique des algorithmes comme le recommande la CNIL, ou la création d'une autorité administrative indépendante (AAI) chargée de certifier ou labelliser les algorithmes en fonction d'un certain cahier des charges (Godefroy, 2019a : 46-52 ; Biard, 2019 : 188). Un embryon de régulation de ce type a été mis en place à travers la possibilité pour les plateformes de résolution en ligne des litiges (ODR) de recourir à une certification⁴⁰. Cette dernière est prévue par la loi sans être rendue obligatoire. De plus, cette mesure ne concerne pas les opérateurs qui proposent seulement des algorithmes d'aide à la décision, mais ceux qui commercialisent des solutions intégrées, associant outils d'analyse du contentieux et de règlement amiable en ligne. Parmi les acteurs publics, l'enjeu consiste à naviguer entre deux impératifs et à les rendre compatibles : s'assurer de la qualité de ce que produisent les algorithmes d'analyse des décisions de justice, tout en n'introduisant pas de barrières réglementaires susceptibles de défavoriser les *startup* françaises dans ce marché en émergence (Biard, 2019 ; [Bourg], 2018).

La voie de la corégulation, c'est-à-dire la coopération des *legaltechs* avec les barreaux et les instances ordinales des professions du droit « dans le respect des statuts professionnels d'ordre public et du droit de la concurrence » (Louvaris, 2018 : 38), s'inscrit dans ce cadre. De fait, les instances de régulation de la profession d'avocat s'en sont saisies. Le Conseil national des barreaux (CNB) qui représente l'ensemble de la profession s'est positionné d'abord à

⁴⁰ Instauré par la loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme de la justice.

travers la commande auprès d'une société de conseil en « transformation digitale⁴¹ » d'une étude comparée des principaux systèmes algorithmiques de traitement et d'analyse des décisions de justice sur le marché national. À la suite de la remise de ce rapport⁴², en octobre 2020, quatorze préconisations ont été adoptées qui affirment la volonté du CNB⁴³ de contribuer à la régulation de ce nouveau marché et des professionnels qui se positionnent dessus.

En somme, les publications portant sur le développement des dispositifs algorithmiques de traitement et d'analyse des décisions de justice reposent principalement sur le repérage de potentialités et de risques identifiés à partir d'exemples ponctuels, souvent inspirés par le cas d'outils de profilage et d'aide à la décision en vigueur aux États-Unis. Comme on l'a vu, une part non négligeable de cette littérature est produite par des acteurs engagés : concepteurs d'algorithmes et créateurs de *startup* d'une part, mais aussi magistrats et avocats impliqués dans la régulation de ces dispositifs d'autre part. L'ensemble contribue à dessiner un paysage où dominent les discours axiologiques et normatifs, reprenant en boucle des arguments qui restent assez généraux.

III / Les algorithmes à l'épreuve du terrain

De fait, il convient de reconnaître que très peu de recherches fondées sur des analyses systématiques et empiriquement nourries ont à ce jour étudié les algorithmes de traitement des décisions de justice. Parmi les quelques travaux repérés, deux orientations distinctes apparaissent toutefois. La première est dirigée vers l'étude des perceptions et discours sur les algorithmes de traitement des décisions de justice (1). La seconde se focalise sur la fabrique et les usages des dispositifs algorithmiques (2)⁴⁴.

3.1. Que disent concepteurs et magistrats des dispositifs algorithmiques ?

Une des rares recherches empiriquement documentées a été produite par une équipe pluridisciplinaire (droit, sociologie et mathématique) dans le cadre d'un financement par le GIP Mission de recherche Droit et Justice (1). Ses résultats, bien qu'exploratoires et inégaux, ont dégagé des pistes heuristiques en commençant à examiner « l'écosystème de l'IA dans la justice » ainsi que les perceptions et discours produits (2).

⁴¹ Sopra Steria Next, <https://www.soprasteria.fr/services/conseil> [consulté le 25 février 2021].

⁴² <http://lemondedudroit.com/decryptages/72373-cnb-preconisations-actions-legaltech-jurimetrie.html> [consulté le 25 février 2021].

⁴³ <https://www.cnb.avocat.fr/fr/actualites/preconisations-dactions-pour-les-legaltechs-du-domaine-de-la-jurimetrie> [consulté le 25 février 2021].

⁴⁴ Il faut ajouter la thèse en cours de Camille Girard-Chanudet intitulée « La fabrique sociale de l'intelligence artificielle : concevoir et mettre en œuvre une 'justice prédictive', sous la direction de Nicolas Dodier et Valérie Beaudoin. Elle propose l'analyse empirique de la fabrique d'un algorithme de pseudonymisation des décisions de justice au sein de la Cour de cassation et repose notamment sur une enquête ethnographique très fine menée en 2021 au sein du pôle *open data* de cette cour (Girard-Chanudet, 2021).

3.1.1. Une équipe pluridisciplinaire et une logique de recherche-action

L'appel d'offres « Droit, justice et numérique » du GIP, reconduit chaque année de 2017 à 2019, incite à s'emparer de cette thématique large — parmi lesquelles la *blockchain*⁴⁵ et — dans une perspective interdisciplinaire et de recherche-action⁴⁶. C'est dans le cadre de la réponse à l'édition 2017 de cet appel à projets qu'en 2019, une équipe pluridisciplinaire a livré un rapport dont le titre indique d'emblée la visée : « Comment le numérique transforme le droit et la justice vers de nouveaux usages et un bouleversement de la prise de décision. Anticiper les évolutions pour les accompagner et les maîtriser ». L'équipe comprend Lémy Godefroy, maîtresse de conférences en droit privé à l'Université de Nice Sophia Antipolis, Frédéric Lebaron, professeur de sociologie à l'ENS Paris-Saclay et [Philippe Sinclair], titulaire d'un doctorat en informatique, directeur de recherche à l'INRIA, et cofondateur de la *startup* [Anajur]. La collaboration a été entreprise dès 2015-2016 à l'initiative de [Philippe Sinclair] (Lebaron, 2019 : 112-113) et s'est ensuite greffée à un autre projet porté par L. Godefroy. Un élève normalien au département de sciences sociales de l'ENS Paris-Saclay, Tom Balteau, par ailleurs issu d'un cursus d'ingénieur, a contribué à la réalisation du volet sociologique de la recherche.

En dehors du rapide préambule qui vaut introduction générale au rapport, chacune des quatre parties, rédigée par un auteur, est indépendante des autres. La première signée par [P.Sinclair], porte sur l'analyse spécialisée et comparée des approches par statistiques et par modélisation appliquées aux décisions de justice ; elle s'attache à situer [Anajur] par rapport aux autres dispositifs algorithmiques et à présenter le type de résultats qui peuvent être obtenus suivant les domaines de droit concernés. La seconde partie émet des recommandations relatives à un encadrement juridique des modes algorithmiques d'analyse des décisions (MAAD). La troisième s'intéresse à « l'appropriation des MAAD » par les acteurs du secteur, en particulier les magistrats chefs de juridictions. Ces deux parties sont rédigées par Lémy Godefroy, juriste. Enfin, la quatrième partie, proprement sociologique, analyse la manière dont le champ juridique est reconfiguré par l'arrivée d'algorithmes prédictifs portés par de nouveaux acteurs « en général plutôt “subalternes” [...], voire, s'agissant des dirigeants de *startup*, clairement périphériques » (Lebaron, 2019 : 115). Frédéric Lebaron en est l'auteur, avec la collaboration de Tom Balteau.

⁴⁵ En 2020, un autre appel centré sur la *blockchain* est lancé.

⁴⁶ On peut lire en effet qu'« il paraît essentiel de ne pas uniquement se limiter à la réflexion sur le numérique dans son rapport au droit et à la justice, mais de soutenir des initiatives techniques dans ce domaine. En lien avec des spécialistes de l'outil numérique (développeurs, etc.) et des professionnels du droit et de la justice, les chercheurs pourraient proposer des outils (bases de données, MOOC, outils d'aide au partage d'information, etc.) novateurs en matière de droit et de justice (par exemple outils de traductions juridiques spécifiques, bases de données ayant pour but de faciliter les travaux juridiques, de rationaliser les lois en déterminant leurs incohérences éventuelles ou de mesurer leur impact, etc.). » Appel d'offres *Droit, justice et numérique*, Mission de recherche droit et justice, 2016, p. 4. Appel reconduit en 2017, 2018 et 2019. <http://www.gip-recherche-justice.fr/wp-content/uploads/2013/06/2016-A.O-Droit-justice-et-num%C3%A9rique.pdf> consulté le 20.11.2020.

Si la structure du rapport semble indiquer un travail d'écriture compartimenté entre les auteurs, la description des réunions et concertations régulières de l'équipe suggère des effets transversaux possibles : élaboration d'une réflexivité partagée et peut-être aussi esquisse de rapprochement dans certaines manières de penser le sujet. [P.Sinclair] démarre sa partie sur le fait qu'« il est important de comprendre que le concept de justice prédictive est non seulement vide de sens, mais, de plus, dangereux. » ([Sinclair], 2019 : 7) ; les deux parties rédigées par Lémy Godefroy (2019a ; 2019b) s'appuient et donnent corps à une notion alternative, celle de MAAD. Par ailleurs, il n'est pas inutile de remarquer que la notion initialement utilisée dans le questionnaire adressé aux magistrats⁴⁷ est un troisième terme, celui de « justice quantitative », expression déjà croisée sous la plume des fondateurs de [Anajur] en lieu et place de celle de « justice prédictive »⁴⁸ ([Monta] et [Sinclair], 2017 : 500).

3.1.2. Analyse des parcours de concepteurs d'algorithmes

La partie rédigée par Frédéric Lebaron (avec Tom Balteau) se distingue des deux précédentes dans la mesure où elle ne mobilise pas la notion de MAAD. Quant à celle de « justice prédictive », elle la met à distance à travers un paragraphe consacré à l'analyse des enjeux de cette notion à partir de ce qu'en disent les acteurs interrogés (119-122). Elle en fait donc un objet sociologique en tant que tel, ce qui est à notre connaissance une première. Elle conclut à la « forte hétérogénéité » des points de vue à l'intérieur même du petit monde des concepteurs d'algorithmes et souligne : « ni les finalités ni les enjeux des algorithmes juridiques ne sont véritablement partagés par les acteurs rencontrés » (122).

À partir de quelques entretiens⁴⁹, les caractéristiques des trajectoires sociales et professionnelles des concepteurs d'algorithmes sont repérées. Il s'agit surtout de quelques individus aux parcours très atypiques combinant compétences juridiques et mathématico-informatiques, avec une volonté de se lancer dans l'entreprise innovante. C'est un petit nombre d'acteurs jeunes, occupant des positions subalternes dans le champ juridique voire étant des outsiders venus d'espaces périphériques, disposant d'une forte visibilité sociale et d'importants soutiens institutionnels (115). Ils tirent parti du secteur de l'innovation, déjà bien structuré en France. Dans leurs discours transparaissent les différences de conception de la nature même des algorithmes d'analyse des décisions de justice et du rôle qu'ils peuvent jouer dans l'évolution du droit, le système de règlement des litiges et la fabrique de la décision juridictionnelle. Il apparaît que « ni le contenu, ni les finalités, ni les enjeux des algorithmes juridiques ne sont véritablement partagés » (Lebaron, 2019 : 121). Ces concepteurs

⁴⁷ « Un changement de terminologie a paru nécessaire depuis l'énoncé des questions adressées aux magistrats », explique-t-elle (Godefroy, 2019b : 88), reprenant le terme de MAAD proposé par [P. Sinclair].

⁴⁸ « La justice prédictive, que nous préférons pour notre part nommer justice quantitative » ; « Ne parlons plus de justice prédictive, mais de justice quantitative, car l'aléa est inhérent au jugement » ([Monta] & [Sinclair], 2017 : 500).

⁴⁹ S'il est précisé qu'il n'a pas été possible d'enregistrer tous les entretiens réalisés, le nombre exact de ceux-ci n'est pas indiqué. En s'appuyant sur les extraits de citations et descriptions de parcours restitués en partie dans les citations, il est vraisemblable que seuls deux ou trois aient été enregistrés. Les non enregistrés, quant à eux, ne sont pas du tout mentionnés. Comme l'indique F. Lebaron, il s'agit bien de quelques illustrations qui permettent toutefois de donner un début de fondement empirique à des hypothèses qui restent à approfondir.

d'algorithmes, créateurs d'entreprises innovantes, se rejoignent en revanche dans une description très critique du fonctionnement actuel de la justice, dans ses aspects institutionnels, organisationnels et professionnels. Leur volonté réformatrice ne s'ancre toutefois pas dans une vision univoque du rôle du droit dans la société ni de la capacité de changement des outils algorithmiques dans le monde judiciaire.

Ces premières conclusions sont prolongées par les résultats d'une autre étude, également réalisée dans le cadre de l'appel d'offres du GIP (Léonard et Moritz, 2020) et qui montre la profonde disparité interne du sous-champ « droit et mathématiques », en train d'émerger à l'intérieur du champ juridique. À partir d'un petit corpus d'entretiens auprès de fondateurs de *startup*, elle isole deux figures idéales typiques de *startupeur* : celle du concepteur-entrepreneur, marqué par la démarche scientifique, se distingue de celle de l'entrepreneur de l'innovation, plus tourné vers le marketing. La focale se resserre ensuite sur les différents groupes qui se constituent et s'affrontent à partir d'alliances entre acteurs et segments spécifiques à l'intérieur du monde du droit et de ses professionnels. Autour de la promotion de la notion de « justice prédictive », les auteurs identifient « une première alliance entre les “entrepreneurs de l'innovation”, le barreau d'affaires et certaines fractions de certains “établissements d'enseignement supérieur d'élite”, lesquels valorisent les cursus universitaires à vocation pratique mêlant droit et économie » (Léonard et Moritz, 2020 : 10). Ce groupe correspond à l'alliance formée autour de la *startup* [PréviCompute]. Ensuite, ils isolent « une fraction particulière parmi les entrepreneurs, laquelle se caractérise par la revendication plus ou moins explicite d'une démarche de “disruption” jouant sur la défiance à la fois vis-à-vis de l'État et vis-à-vis des entreprises, ainsi que par des liens noués d'abord avec d'autres acteurs de l'entrepreneuriat et des nouvelles technologies. » (*ibid.*) Ce second pôle se constitue autour de [Jurissimo]. Enfin, ils relèvent « une alliance implicite entre les “entrepreneurs-concepteurs”, le “barreau classique”, une fraction de la magistrature et du monde universitaire. » (*ibid.*) ; cette coalition prend forme autour de [Anajur].

La sociologie des producteurs d'algorithmes, et la manière dont ils constituent des réseaux plus ou moins ancrés dans le champ juridique ou plus ou moins en rupture avec lui, commence donc à être mieux connue.

3.1.3. Les chefs de juridiction sont-ils tellement favorables aux algorithmes ?

Les attitudes des magistrats à l'égard des algorithmes de traitement et d'analyse des décisions de justice, quant à elles, sont à peine dévoilées. Frédéric Lebaron et Tom Balteau s'y sont intéressés à partir de quelques entretiens, mais ils ont surtout réutilisé les réponses au questionnaire administré par Lémy Godefroy dont il a été question plus haut (2019b : 114).

Il convient de s'attarder sur ce questionnaire et le résultat tiré par L. Godefroy selon lequel une grande majorité des magistrats (85 %) est favorable à l'utilisation d'outils algorithmiques (2019b : 91). Si les conclusions font continûment référence aux « magistrats », en réalité elles ne concernent que ceux du siège⁵⁰ et encore, parmi ceux-ci, seuls celles et ceux

⁵⁰ Ce qui est plutôt logique, compte tenu du fait que les outils évoqués ne traitent à ce jour que des décisions civiles, dans lesquelles le Parquet est généralement peu présent.

qui sont en position hiérarchique dans l'institution ont été sollicités : le premier président de la Cour de cassation, les premiers présidents des Cours d'appel et les présidents de TGI (à l'époque). L'enquête de L. Godefroy repose sur les réponses de 67 magistrats chefs de cour (sur 197, ce qui représente un taux de réponse d'environ un tiers : 34,01 %). Autant dire que ce qui est présenté comme une étude sur l'« appropriation des modes algorithmiques d'analyse des décisions (MAAD) par les magistrats » (titre de la troisième partie, p. 85) est en réalité une étude des perceptions, avis et attentes des seuls responsables de juridictions.

Encore faut-il préciser que la question à laquelle ils ont répondu est la suivante : « Êtes-vous favorable aux outils de justice quantitative ? Pourquoi ? », suivie de trois autres questions qui incitent à explorer des aspects négatifs puis positifs, et enfin la dimension d'aide à la décision de ces outils. Une telle question en ouverture du questionnaire, incite plutôt à une réponse positive. De plus, celui-ci est accompagné d'un document de quatre pages (p. 147-150) intitulé « Problématiques soulevées par les outils de justice quantitative », dont le statut exact n'est pas très clair, mais qui place les enquêtés dans un certain état d'esprit, voire leur impose une problématique. Sur le plan méthodologique, l'enquête comporte donc des fragilités. Ce qui intéresse l'autrice, c'est de recueillir les points de vue des magistrats sur une certaine manière d'envisager les MAAD, et sur des prospectives qu'elle a elle-même dessinées dans la deuxième partie du rapport⁵¹. Cette recherche vise un objectif d'ingénierie sociale dans une perspective d'acceptabilité sociale de ces dispositifs. Cette approche n'est pas sans évoquer la sociologie législative chère à Jean Carbonnier (Commaille, 2007) dans laquelle les savoirs et instruments de l'investigation sociologique sont mis au service d'une démarche de rénovation du droit, en accord avec la société dans laquelle il s'inscrit. L'intérêt en est, outre de créer des passerelles et dialogues entre disciplines — singulièrement ici entre science juridique et sociologie —, de contribuer à une production de connaissances directement utiles pour l'action et ancrées dans une quête de pertinence sociale. D'une certaine manière, la contribution de L. Godefroy peut être vue comme le prolongement empirique de questionnements plus généraux rencontrés plus haut dans ce chapitre. Les différents débats autour de la performativité des dispositifs algorithmiques, des risques de biais, des effets sur le système juridique dans son ensemble et en particulier de la place des précédents, y sont abordés frontalement. L'enjeu est de saisir ce que les magistrats en pensent. En l'occurrence, il s'avère, comme le souligne Barthe (2021 [2017]), qu'ils sont hostiles au traitement automatique des dossiers, mais pas à un traitement de masse. Ils sont aussi très exigeants quant aux performances des dispositifs algorithmiques.

L'analyse lexicométrique des réponses des magistrats à l'aide du logiciel Alceste conduit à deux conclusions. D'abord, les registres de discours des chefs de cour interrogés s'organisent principalement autour des bénéfices et risques potentiels des algorithmes

⁵¹ Cette orientation est partagée par les financeurs du projet, dont elle précise qu'ils ont validé le questionnaire administré (Godefroy, 2019b : 85). D'ailleurs, les questions comportent à de multiples reprises des exemples de réponse que l'on retrouvera dans ses recommandations (par exemple sur l'obligation de motivation pour éviter les effets performatifs des algorithmes) (*ibid.*, p. 81 et 89).

d'analyse des décisions de justice. En cela, ils ne diffèrent guère de la manière, évaluative et prospective, dont la littérature juridique aborde le sujet.

Ensuite, cette analyse — contrairement à l'approche statistique des mêmes discours — semble mener à un constat beaucoup plus nuancé concernant les attitudes des responsables de juridiction à l'égard des algorithmes de traitement de décisions de justice. Ils « apparaissent relativement divisés et ambivalents sur le sujet » (136) et sur « *l'utilisation possible des algorithmes pour améliorer le fonctionnement de la Justice* ». Ces premiers éléments, très lacunaires, appellent de plus amples investigations.

3.2. Les algorithmes à travers leurs usages

Une autre voie d'analyse sociologique est ouverte par les travaux d'Angèle Christin, professeure assistante au département de communication de l'Université de Stanford aux États-Unis, également rattachée à la faculté de sociologie du programme science, technologie et société. Après avoir réalisé une première recherche sur les comparutions immédiates dans le cadre de juridictions françaises (Christin, 2008), elle s'est intéressée aux effets de la quantification et des dispositifs algorithmiques sur les pratiques des journalistes, sous l'angle des mesures d'audience des articles mis en ligne. Ponctuellement, elle a tenté des comparaisons entre journalistes et magistrats des deux côtés de l'Atlantique. Un de ses résultats majeurs porte sur l'importance à accorder aux usages pour saisir les effets réels des technologies de calcul et de quantification. Elle a mis en évidence plusieurs stratégies d'évitement qu'elle a empiriquement documentées.

3.2.1. Des usages pluriels qui créent un découplage entre prescriptions et pratiques

Si les mondes de la justice et du journalisme diffèrent par bien des aspects, ils ont en commun d'être organisés autour de l'expertise de professionnels, mais aussi de vivre « un moment algorithmique comparable » (Christin, 2017). Ces deux mondes sont confrontés à la montée en puissance d'une logique quantifiée introduite au cœur de leurs métiers et appuyée sur des outils algorithmiques. Si elle note l'omniprésence des chiffres dans les activités quotidiennes des juges et procureurs, dans la façon dont ils décrivent leurs activités, elle constate que cette logique de la mesure a toutefois des effets plus forts chez les journalistes dont l'activité est de type marchand (« *for-profit* ») et s'inscrit dans un marché concurrentiel. Le *big data* est également beaucoup plus sophistiqué et omniprésent dans le monde de l'information que dans celui de la justice.

Mais dans tous les cas, journalisme ou justice, il y a loin de la coupe aux lèvres, nous dit A. Christin. Ce n'est pas parce que des discours, consignes et outils sont promus qu'ils sont immédiatement et mécaniquement utilisés, en conformité avec les significations, objectifs, voire vertus, qui leur sont attribués de manière dominante, et notamment par la hiérarchie de l'organisation — *a fortiori* dans des univers marqués par l'autonomie traditionnelle de leurs professionnels. Si les discours produits par les instances de pilotage et les acteurs en position hiérarchique valorisent et encouragent le recours aux outils algorithmiques, à l'échelle microsociologique de l'organisation, les individus ne s'y conforment pas nécessairement. Il existe donc un important décalage entre les discours des managers sur l'utilisation des algorithmes et l'usage qui en est réellement fait. Si, dans le journalisme, les managers

souhaitent que la référence aux statistiques soit systématique de la part de tous les journalistes, certains n’y font même pas attention. Les observations menées révèlent que les journalistes ne consultent pas les logiciels qui analysent les audiences de leurs productions. Ils sont pourtant sensibles à la dimension d’audience — comme le montre l’envoi par mail de statistiques sur l’article qui a le plus marché la veille par exemple — mais ils ne s’intéressent pas en profondeur aux statistiques proposées par les technologies algorithmiques. De leur côté, les administrateurs des cours pénales sont enthousiastes à propos des avantages fournis par ces nouveaux algorithmes, achetant même régulièrement de nouvelles technologies. Cependant, la plupart des juges n’utilisent pas ces outils pour la préparation des jugements, ils continuent de travailler sur des documents papier.

Cet écart, ce « découplage » entre l’encadrement hiérarchique et les acteurs de la production fait bien entendu écho aux acquis de la sociologie des organisations, à travers les dichotomies entre travail prescrit et réel, régulation de contrôle et autonome (Reynaud, 1988). Il s’appuie aussi sur une sociologie des usages classique dans l’étude des technologies (Proulx, 2015). Même si elles ne sont pas citées, ces références représentent le socle sur lequel s’appuie l’analyse des algorithmes sous l’angle des pratiques et usages qu’ils suscitent chez leurs utilisateurs.

3.2.2. Critiquer ouvertement, traîner des pieds et ruser

Trois techniques de contournement (« *buffering strategies* ») visant à « minimiser l’impact des algorithmes sur le travail quotidien » (Christin, 2017 : 9-10) sont repérées. La première, la critique ouverte (« *open critique* ») peut s’accompagner d’une résistance aux systèmes technologiques de quantification, au sens d’un refus d’utilisation. Journalistes et magistrats trouvent dans leur *ethos* professionnel des ressources pour entreprendre la critique de la légitimité des démarches d’évaluation par la mesure. La « bonne justice » comme le « bon journalisme » ne peuvent pas s’apprécier quantitativement ; ils supposent un temps long, ce qui explique de faibles performances en nombre de cas traités pour les magistrats étudiés ou de clics pour les journalistes dont les articles sont diffusés sur des plateformes numériques (Christin, 2016). Les deux autres techniques sont moins exposées et moins exposantes pour ceux qui les pratiquent. Le « *foot-dragging* » (traîner des pieds) consiste à ignorer les outils de quantification, en n’y recourant pas dans sa pratique — volontairement, mais discrètement. Quant au « *gaming* » (ruser), il repose sur une utilisation détournée des règles ou des chiffres afin de maintenir ses pratiques professionnelles inchangées. Chez les journalistes, il s’agit par exemple de titrer son article d’une manière « choc », cherchant à provoquer des clics, au prix d’un décalage entre le contenu et le titre du papier. Dans la justice, entrer seulement certaines variables dans l’algorithme afin d’obtenir le score voulu, selon une rationalité inversée, procède de la même logique.

C’est à partir d’une enquête par observations et entretiens dans plusieurs juridictions pénales des États-Unis, qu’A. Christin (2018) a mis en évidence ces techniques de distanciation pratiquées par des juges et procureurs à l’égard de Compas, un logiciel fourni par leur juridiction et dont l’utilisation était encouragée par les chefs de cour. Pour chaque personne, l’algorithme évalue le risque de récidive, en fonction d’une centaine de critères, tirés de données antérieures elles-mêmes construites à partir des décisions que les juges ont prises

pour des cas similaires. Compas fournit un score global de probabilité de récidive sur une échelle de 1 à 10, destiné à aider la prise de décision des juges ou commissions en matière de détention provisoire ou de libération conditionnelle (Harcourt, 2007 ; Sapir, 2008).

Elle montre qu'au sein des juridictions pénales étudiées, certains juges ou procureurs sont réticents à utiliser ce logiciel et développent les stratégies évoquées plus haut pour s'y soustraire. Certes, le cas qu'elle étudie est un dispositif dit de justice actuarielle et non un algorithme de traitement des décisions de justice en matière civile⁵². Dans tous les cas, quel que soit le type d'algorithme étudié, l'intérêt d'une sociologie ouverte aux usages demeure intact lorsque l'on aborde les technologies les plus récentes. On ne comprendra subtilement les effets de celles-ci qu'en s'intéressant non seulement à ce qu'elles peuvent faire *a priori* (ou ce que leurs concepteurs et promoteurs disent de leurs performances), mais aussi à la manière dont elles sont effectivement utilisées, mobilisées ou pas, selon des modalités variables, par les systèmes d'acteurs en présence.

Sur la relative indétermination des technologies, A. Christin va même plus loin en évoquant la mise en place de systèmes numériques de gestion des dossiers (*paperless case-management system*). Cette transformation importante pose la question de l'évolution des habitudes de travail des acteurs judiciaires, du fonctionnement concret des juridictions, mais surtout celle des usages de masses considérables de données recueillies. Ces systèmes complexes, de grande ampleur permettent d'effectuer une grande diversité de tâches et d'assumer différentes fonctions au cœur du système judiciaire : programmation et gestion des audiences en ligne, mise en état dématérialisée des affaires, accès directs aux dossiers des accusés et condamnés pour les acteurs qui interviennent au stade de l'avant-procès comme de la probation. Dès lors que ces systèmes enregistrent, compilent et stockent des quantités massives de données dans un même lieu, la question de leur exploitation va se poser. A. Christin suggère qu'il convient d'être vigilant sur ce que les juridictions commencent à faire et feront à partir de ces données. Car là encore, une pluralité d'usages est disponible. Elles peuvent permettre de construire des « *predictive analytics unit* » en s'appuyant également sur les données fournies par les procureurs et la police, et ce, afin de contribuer à prévoir et donc à prévenir les futurs crimes et délits, dans une optique orientée principalement vers la protection de la société. D'autres juridictions ont également commencé à développer une approche réflexive pour comprendre le fonctionnement de ces technologies, leurs effets systémiques, etc. Différents objectifs sont visés : réformer les pans les moins efficaces de l'organisation, mieux identifier comment agir en vue d'une plus grande justice sociale, limiter l'arbitraire (*discretion*) des procureurs dans l'orientation des affaires entre négociation et procès ; repérer les peines favorables à une meilleure réinsertion des ex-détenus dans la société, etc. La nature même des transformations que pourraient alimenter les outils algorithmiques et numériques n'est donc ni tracée ni même partagée *a priori* parmi les utilisateurs.

⁵² Nous reviendrons plus loin sur les enjeux de labellisation de la « justice prédictive », par contraste avec la justice actuarielle.

Que sont exactement les algorithmes de traitement et que font-ils concrètement aux pratiques des utilisateurs ? Il est heuristique de se pencher sur ces infrastructures de calcul et sur la manière dont elles sont déployées, d'autant que, quel que soit le discours des concepteurs sur le pouvoir prédictif des algorithmes (Benbouzid et Cardon, 2018 ; Cardon, 2019), tout peut être remis en jeu par les pratiques, raisonnements et décisions, selon qu'à un extrême, les acteurs délèguent les décisions au système, ou au contraire choisissent d'ignorer les apports de celui-ci. Les premiers retours d'usages (Licoppe et Dumoulin, 2019) à partir de l'analyse d'une expérimentation de [PréviCompute] auprès de magistrats des cours d'appel de Rennes et de Douai indiquent que les utilisateurs judiciaires de ces dispositifs algorithmiques sont déjà équipés de ressources variées : textes, codes, bases de données type Jurica, référentiels, barèmes, forums de discussion, etc. (Sayn *et al.*, 2014), produits en interne. Ils suggèrent donc de saisir les effets des dispositifs algorithmiques comme un déplacement de manières de faire déjà équipées, plutôt que les poser d'emblée en rupture radicale.

En outre, considérer ces logiciels et plateformes en ligne non comme des formes de « justice prédictive », mais bien comme des « dispositifs de traitement et d'analyse des décisions de justice » (Licoppe et Dumoulin, 2019) met l'accent sur ce qu'ils incorporent, à savoir des procédures informatisées qui « permettent de transformer une entrée en un résultat », processus qui passe par une série d'opérations et de calculs, et ce, afin de produire des métriques et des analyses (Christin, 2020). La question des usages, si elle n'est pas incluse dans la présente recherche, a commencé à être défrichée. Elle continuera de l'être dans le cadre du projet ANR Just-IA sous l'angle des pratiques des différents professionnels du droit (avocats, directions juridiques de grandes entreprises dont des sociétés d'assurance, magistrats), mais aussi à travers une thèse en cours⁵³ plus spécifiquement centrée sur les appropriations des dispositifs algorithmiques par les avocats, mais aussi par la profession, à travers ses organes de représentation et de régulation.

Conclusion :

À l'issue de cet état de l'art, l'effervescence et la profusion de discours sur la « justice prédictive » ne peuvent que frapper. Ce premier chapitre a rendu compte de l'abondante littérature de droit et de philosophie et de théorie du droit consacrée à la 'justice prédictive'. J'ai montré combien ces publications *adoptent très majoritairement une posture normative et évaluative*, appréciant de manière générale et surplombante les atouts et limites des algorithmes de traitement des décisions de justice et de leurs effets possibles. C'est en termes de potentialités et de risques pour le système juridique de droit écrit, pour la justice comme activité régaliennne et pour l'avenir des professionnels que sont les avocats et magistrats, qu'est envisagé le développement de ces dispositifs. L'espace académique est ainsi saturé de publications qui mêlent les registres analytique, évaluatif, prescriptif et prospectif.

⁵³ Héloïse Eloi-Hammer, « Justice et algorithmes : vers une transformation du métier d'avocat ? », Thèse de sociologie sous la codirection de Dominique Cardon et Laurence Dumoulin.

Un des signes de l'intensification du débat autour de la 'justice prédictive' est le fait que progressivement à partir de 2017, le cercle de celles et ceux qui y participent s'élargit. Non seulement les publications se multiplient, mais la discussion gagne des juristes non spécialistes et se diffracte dans toutes les différentes branches du droit. Chaque auteur s'exprime à partir du prisme qui est le sien, et en particulier de la matière juridique dont il est spécialiste : droit de la responsabilité civile (Jacquemin, 2019), procédure civile (Amarani-Mekki, 2018 ; Cholet, 2018), droit social (Gamet, 2019), droit constitutionnel (Gaye-Palette, 2020) et même droit substantiel pénal (Malabat, 2019 ; Brigant, 2018), alors que dans le cas français les décisions pénales sont exclues des dispositifs algorithmiques. Les auteurs convergent vers certains topos, qui associent auteurs, thèmes, exemples étrangers récurrents. Des spécialistes d'un type de procédure ou d'un domaine (Valérie Malabat pour la procédure pénale ; Natalie Fricero pour la procédure civile), consacrent un ou plusieurs textes à la « justice prédictive », comme elles le font à la réforme des juridictions ou à d'autres sujets de l'actualité juridique et judiciaire. L'enjeu est celui des effets sur le droit, la justice et leurs professionnels.

Un des effets de la multiplication des travaux de doctrine en un temps court est de consolider le terme de 'justice prédictive' qui fonctionne alors comme un mot-clef, un label, alors même que son contenu, les significations et réalités qu'il désigne sont hétérogènes et peu partagées tant parmi les auteurs, qu'ils soient académiques, praticiens (avocats, magistrats) ou concepteurs de dispositifs algorithmiques de traitement et d'analyse des décisions de justice. Les formulations qui consacrent la notion mais indiquent une prise de distance (comme « la justice dite prédictive ») sont rares en 2016-17. Elles deviennent plus courantes dans les actes de colloques et écrits ultérieurs.

Si la science juridique et les juristes contribuent clairement à la montée en puissance de la notion de 'justice prédictive' et à sa mise en débat, les dispositifs d'analyse et de traitement de grandes quantités de décisions de justice apparaissent aussi comme parties prenantes de la construction d'un espace d'interdisciplinarité entre droit, informatique et sciences sociales comme en témoignent les équipes qui mettent en place des expérimentations (Aletras *et al.* ; Katz *et al.* ; [Monta] et [Sinclair], 2017), celles qui produisent des savoirs orientés vers l'action (Godefroy, [Sinclair] et Lebaron, 2019) et les quelques duos d'auteurs juriste/spécialiste d'informatique (Garapon et Lassègue, 2018, Desmoulins-Canselier et Le Métayer, 2020).

Pour autant et comparativement, il convient surtout de souligner le contraste entre l'abondance des analyses juridiques et la rareté des études sociologiques reposant sur des recherches empiriques. Très peu nombreuses sont les analyses de sciences sociales, empiriquement nourries et distanciées des *stakeholders*, dans la littérature tant internationale que nationale. Quelques exceptions notables s'intéressent toutefois aux perceptions, discours, fabrique et usages de ces dispositifs, à partir d'approches de sociologie des usages, sociologie de l'innovation et sociologie du droit et de l'action publique. C'est dans le prolongement de ces travaux que la présente recherche s'inscrit. Considérer la 'justice prédictive' comme une construction technique, sociale et politique, c'est s'attacher à objectiver la percée de l'expression, à interroger les conditions de genèse de cette catégorie

et de la mise à l'agenda de ce nouveau problème public, ce qui est l'objet des chapitres suivants.

Chapitre 2 : La ‘justice prédictive’, controverse autour d’une notion floue

L’expression ‘justice prédictive’ fait l’objet d’un débat qui se déploie sur la scène académique, à l’occasion de colloques, publications d’ouvrages, articles et dossiers de revues, principalement juridiques et auxquels prennent part surtout des juristes. L’effervescence autour de cette notion semble toutefois particulièrement récente. Une première analyse bibliométrique, réalisée à partir de quelques plateformes bibliographiques¹, a montré qu’entre 2012 et 2018, 166 articles juridiques francophones sont publiés. Quasiment les trois quarts d’entre eux (118) le sont en 2017 (Larrègue, 2019 : 10). Cette année apparaît comme charnière sur la période. La controverse se développe principalement dans l’espace académique juridique. Elle déborde toutefois de cet espace spécialisé, que ce soit à travers la presse généraliste qui s’en fait écho ou bien via l’activité de ses protagonistes (fondateurs de *startup* dans le domaine, institutions organisatrices de colloques, etc.) sur les réseaux sociaux.

Ce chapitre vise à poursuivre le voyage en ‘justice prédictive’ en s’attachant à objectiver la trajectoire de l’expression, à circonscrire et qualifier son essor. À partir de quand l’adjectif prédictif est-il accolé à la justice pour former ce nouveau construit sémantique ? Qui le reprend et qui s’en dissocie ? L’enjeu est ici de mieux cerner la genèse et les usages du terme, à partir d’une approche empirique.

Il s’agit en particulier d’objectiver son déploiement sur une décennie, c’est-à-dire sur une période de moyen terme ; de tester l’hypothèse selon laquelle cette construction sémantique serait apparue dans la seconde moitié des années 2010, et plus probablement autour de 2017 comme les premiers résultats de J. Larrègue incitent à le penser. Cette hypothèse est probable, au vu de la profusion de publications scientifiques observée dans le premier chapitre à ce moment-là. De plus, l’expérimentation par deux cours d’appel françaises de [PréviCompute], particulièrement médiatisée à l’époque, se déroule en cette même année 2017. L’objectif est donc de s’intéresser aux périodes et séquences qui scandent la trajectoire de ce terme ; aux connotations et liens qui sont rattachés à cette innovation sémantique ; aux espaces et types d’acteurs qui débattent de ‘justice prédictive’.

L’arène et les acteurs médiatiques jouent un rôle considérable dans les processus de cadrage (Neveu, 2015) de même que les réseaux sociaux (Neuman *et al.*, 2014) d’où l’intérêt d’étudier la manière dont ils mentionnent, reprennent, discutent cette notion. Saisir les dynamiques à l’œuvre dans l’arène médiatique est une manière d’accéder à une dimension concrète et empiriquement opérationnalisable de l’espace public. Or, la justice n’est pas la seule, au même moment, à être qualifiée de ‘prédictive’ : la santé par exemple l’est aussi. Il semble donc judicieux d’ouvrir la focale pour se tourner vers les différents secteurs traversés

¹ Les bases de données sont pour les sciences sociales : Cairn et Web of Science et pour le droit : Dalloz, Lamyline, Lexis Nexis, Lextenso. Il semblerait donc que l’auteur ait sélectionné *a priori* une majorité de plateformes bibliographiques spécialisées dans le droit, mais l’article publié est court et donne très peu de détails méthodologiques sur la manière dont l’investigation a été menée.

par ce mouvement. Pour quelles autres innovations, quels autres secteurs la notion de prédictif.ve est-elle mobilisée concomitamment ? Avec quelles significations et/ou connotations ? De la comparaison peuvent émerger avec plus de netteté les traits distinctifs de la ‘justice prédictive’ et la façon dont la rencontre entre justice et dispositifs algorithmiques est cadrée.

Poser ces questions simples est indispensable dans une démarche sensible aux effets de construction sociale des catégories. Une nouvelle notion livre une clef de lecture renouvelée du monde social, un cadrage qui reconfigure la description du monde et son entendement, à partir duquel faire émerger de nouveaux problèmes publics et contribuer à leur mise à l’agenda.

Afin d’identifier les discours produits sur la ‘justice prédictive’ et les resituer dans un contexte plus large, plusieurs ensembles de données ont été constitués et analysés sur la période 2010-2020. Ils ont été rapidement présentés en introduction. Un premier corpus réunit les articles contenant les mots « prédictif » ou « prédictive » parus dans les éditions papier et numérique de sept journaux français (*Le Figaro*, *Le Monde*, *Libération*, *Ouest-France*, *La Voix du Nord*, *L’Obs* et *Le Point*). Entre 2010 et 2020, 1 622 articles comprenant cet adjectif ont été publiés. À partir de la base de données Europresse, le texte intégral des articles recueillis a été mis en forme et importé dans le logiciel Iramuteq pour réaliser une analyse des caractéristiques de ce corpus de presse et de la manière dont s’y distribuent des univers de discours identifiés comme des classes. Concrètement, le logiciel procède au découpage des articles en segments de textes de longueur identique (classification hiérarchique descendante) puis procède à des regroupements (classification hiérarchique ascendante) à partir de la plus ou moins grande proximité des formes textuelles utilisées.

L’intérêt de cette approche lexicale et lexicométrique est de construire des classes de discours qui renvoient à des univers de sens, interprétables à la fois dans leur structure, leur poids respectif et leurs significations. En l’occurrence, l’analyse de ce premier corpus dessine quatre grands univers de sens qui se dégagent transversalement de l’ensemble des articles. Ils correspondent à quatre domaines : la santé, l’automobile, les élections et la justice. C’est de cette analyse que sont tirés les premiers enseignements sur la ‘justice prédictive’ présentés dans ce chapitre.

Au sein de ce corpus, est ensuite isolé un sous-corpus ‘justice prédictive’ qui comprend donc les seuls articles comportant cette expression exacte¹. Celui-ci réunit 122 articles et a fait l’objet de la même analyse, toujours par Iramuteq. En outre, compte tenu de l’importance acquise par les réseaux sociaux, *a fortiori* sur un sujet touchant aux technologies, un corpus de messages diffusés sur Twitter et comportant l’expression ‘justice prédictive’ a été constitué². Le choix de cette plateforme numérique est lié à ses caractéristiques, au fait qu’elle est principalement tournée vers l’actualité. Elle est utilisée comme un vecteur de

¹ Une relecture manuelle a permis de repérer les quelques articles qui remettent explicitement en question cette notion, mais qu’une recherche automatique avait écartés en raison de constructions grammaticales spécifiques (par exemple la justice qualifiée de prédictive). Ces articles ont été ajoutés au sous-corpus.

² En collaboration avec Florent Frasque, comme indiqué en introduction.

communication, institutionnelle notamment et comme un moyen d'information. Elle fonctionne aussi comme un espace conversationnel (Cardon et Prieur, 2016 : 244-247) propre à l'émergence de controverses que « l'usage possible des hashtags » renforce en tant qu'elles agissent « comme balises d'un affrontement autour de la définition de l'objet de débat » (Bottini et Julliard, 2017). Les messages sont des « productions écrites auto-publiées » (Boyadjian, 2016) issues de comptes individuels ou institutionnels.

Sur la période 2010-2020, 6 261 *tweets* ont été recueillis, à partir de requêtes automatiques¹ en croisant les mots-clefs « justice » et « prédictive ». Le corpus obtenu a été codé et enrichi de variables, pour constituer une base de données. Celle-ci comporte des informations comme la date de publication du *tweet*, l'existence ou pas de *retweets* et mentions, etc. qui sont utilisées comme des variables pour décrire le corpus. Il en va de même avec la profession des personnes twittant depuis des comptes personnels. Les informations disponibles ont été codées, par type de profession juridique (magistrats, avocats, juristes universitaires, autres professionnels du droit), par type de discipline exercée (universitaires en droit vs dans d'autres disciplines ; étudiants en droit vs dans d'autres disciplines) et par grandes catégories professionnelles (personnel politique, métiers de l'information, métiers de l'information juridique, métiers liés à la data vs autres professions) ; enfin par types de structures et d'institutions au sein desquelles ces métiers sont exercés.

Les *tweets* ont alors donné lieu à une analyse textuelle lexicométrique et sémiologique assistée par Iramuteq : qu'est-ce qui est dit sur la 'justice prédictive', en quels termes, quelles sont les classes de discours présentes et leurs axes de polarisation ? Les classes obtenues ont été réintroduites comme variables explicatives associées au contenu du *tweet* pour étudier l'évolution des types de discours sur la période. D'autres recherches sont envisagées dans le cadre de l'ANR Just-IA. En particulier, une analyse de réseaux via les *retweets* permettra d'identifier qui sont les leaders d'opinion, quelles communautés se constituent, etc. Ces résultats peuvent ensuite être injectés dans une analyse de contenu et croisés avec elle.

Ce chapitre ne prétend évidemment pas faire le tour du sujet ni, par exemple, établir avec certitude l'origine et la trajectoire de l'expression 'justice prédictive' dans l'espace francophone. Les données recueillies reposent sur des corpus élaborés selon une base nationale et francophone, avec un nombre limité de titres de la presse généraliste (n=7) et la prise en compte d'un seul réseau social (Twitter), qui pourrait être utilement complété par d'autres (par exemple LinkedIn). En outre, en raison des contraintes de volume, les corpus ne sont pas parfaitement symétriques entre eux, puisque la collection d'articles de presse « prédictif.ve » n'a pas son équivalent pour les *tweets*². À ce stade, l'ambition est d'offrir une première vue d'ensemble. Il n'en demeure pas moins que la démarche, parce qu'elle est

¹ Réalisées via l'ouverture d'un compte spécifique auprès de Twitter.

² Le nombre limité d'articles de presse (moins de 2 000 au total) a permis de constituer un corpus large relatif aux mots-clefs « prédictif », « prédictive » et à l'expression « justice prédictive », et donc de replacer cette dernière (qui n'est présente que dans 122 articles donc moins de 1/10^e du corpus) dans l'ensemble de ce qui est qualifié de prédictif/prédictive. En revanche, vu la masse de *tweets* publiés, le corpus a été délimité de manière à appréhender la seule notion de 'justice prédictive' (rassemblant à elle seule plus de 6 000 *tweets*).

empiriquement fondée et repose sur des méthodologies transparentes et contrôlées, peut contribuer à l'objectivation recherchée.

Ces premières investigations livrent plusieurs enseignements, sur la manière dont émerge la question de la 'justice prédictive', sa chronologie et ses liens avec un contexte plus général. En l'occurrence, la justice est un secteur parmi d'autres à propos duquel l'adjectif « prédictif.ve » est mobilisé. La comparaison avec les autres domaines — réalisée à partir du corpus d'articles de presse — révèle quelques traits distinctifs (1). L'analyse chronologique — à partir des articles de presse et des *tweets* — établit quant à elle que l'expression 'justice prédictive' connaît une courbe en forme de montagne russe, en quatre phases sur la deuxième moitié de la décennie 2000 (2). Un zoom sur le corpus de *tweets* met en évidence la façon dont la controverse se structure sur Twitter à travers l'identification des classes de discours, des arènes concernées, au fil de différentes phases identifiées (3).

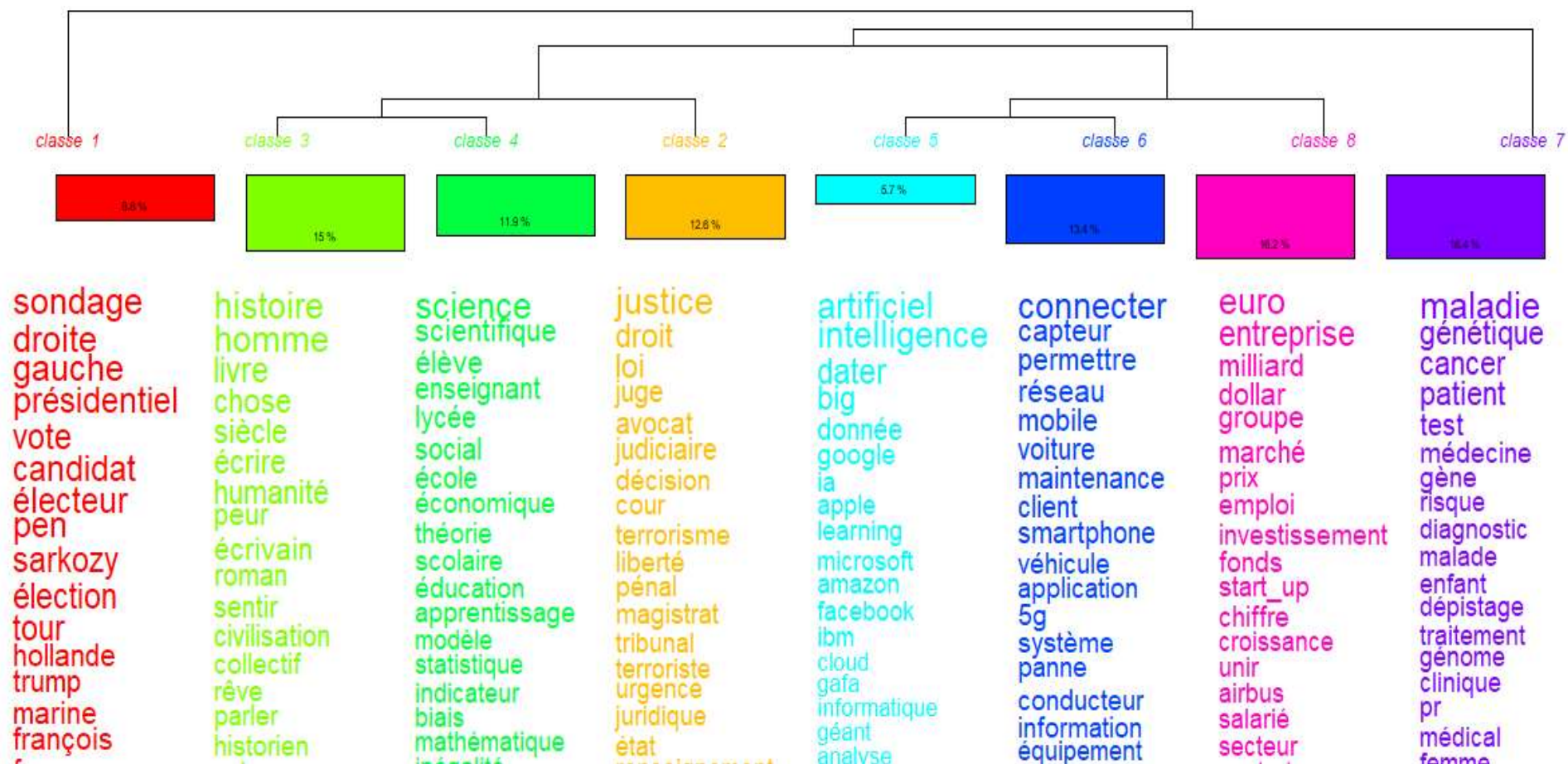
I / Une justice a-technologique confrontée à la modernité ?

Je présente d'abord l'ensemble de la classification réalisée par la méthode Reinert (1). J'en tire plusieurs enseignements : le double sens pénal et civil de la 'justice prédictive' d'abord (2), la double référence à l'anticipation et à la délégation des décisions judiciaires ensuite (3), l'intensité des débats prospectifs dans le milieu académique enfin (4).

1.1 Variations autour du tournant prédictif

L'analyse lexicométrique sur l'ensemble des articles de presse comportant les formes « prédictif » ou « prédictive » et apparentées (n=1 622) fait apparaître huit classes de discours, plus ou moins proches les unes des autres. Le graphique suivant offre une première représentation visuelle de ces univers de sens. Il met en évidence la succession de bipartitions dont procèdent les huit classes de discours isolées. Pour chacune d'elle, le poids qu'elle représente au sein de l'ensemble des segments de texte classés est affiché (en pourcentage), de même que les mots les plus caractéristiques de chaque classe.

Graphique 2.1. Dendrogramme représentant le résultat de la classification de Reinert. Corpus d'articles de presse « prédictif.ve » (2010-2020)

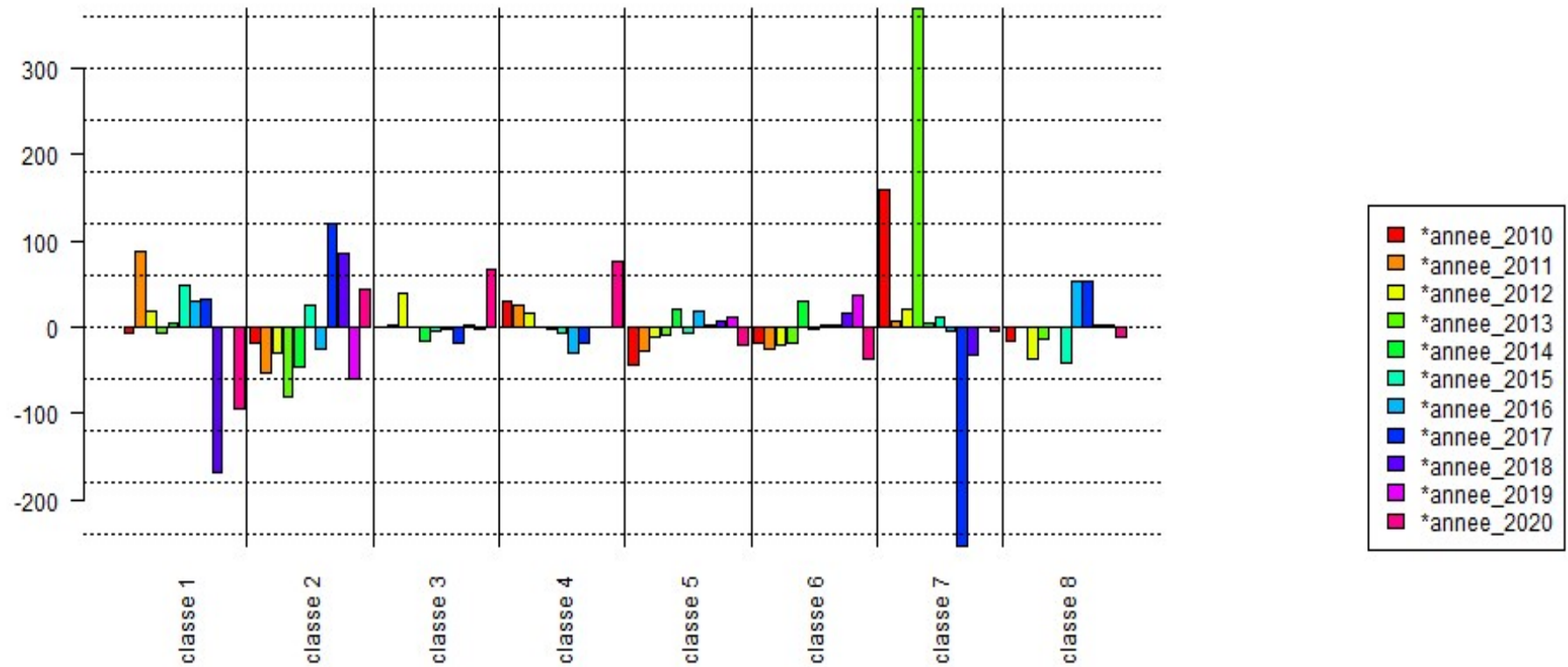


L'interprétation de ce dendrogramme conduit à isoler quatre secteurs distincts pour lesquels le prédictif fait sens : la compétition électorale, la santé, l'automobile et la justice. Il se lit comme un arbre aux ramifications successives. On observe deux premières bipartitions aboutissant pour chacune d'entre elles à deux classes très spécifiques (1 et 7) et donc éloignées des autres. L'univers sémantique de la première est tourné vers la compétition politique et les sondages (classe 1), celui de la seconde vers le domaine médical et la médecine prédictive (classe 7). Le système composé par les classes restantes est scindé en deux groupes de trois. Le premier contient une classe portant sur la justice (classe 2), qui est elle-même reliée au monde de la recherche et de l'enseignement (classe 4) et à la réflexion quant à l'état et l'avenir de la société (classe 3). Quant au second groupe, il comporte une classe consacrée au secteur automobile et notamment à la voiture autonome (classe 6). Elle est, quant à elle, associée à l'évocation d'innovations technologiques (classe 5) et aux perspectives économiques offertes par ces innovations (classe 8).

La saillance dans le temps des techniques et enjeux qualifiés de « prédictifs » n'est pas équivalente pour les différents secteurs. Le graphique suivant montre les variations d'intensité des classes de discours par année. On y constate déjà que dans la première moitié de la période analysée, le prédictif ne renvoie que rarement au domaine de la justice¹. En revanche, l'année 2017 est significativement surreprésentée pour la classe associée à la justice (classe 2) alors qu'elle est très sous-représentée pour celle associée à la santé (classe 7) par exemple. Cela apporte un premier élément de vérification de l'hypothèse d'une année 2017 marquante pour la 'justice prédictive'.

¹ On verra plus loin que ceci converge avec les résultats obtenus sur Twitter puisque seulement sept *tweets* traitent de la 'justice prédictive' entre janvier 2010 et mars 2015.

Graphique 2.2. Distribution de la variable année au sein de chacune des classes. Corpus d'articles de presse (n=1 622), 2010-2020



termes génériques se référant à la *justice* : *droit* et *loi* sont les mots les plus significatifs de la classe. Les termes *judiciaire*, *cour*, *décision*, *juridique* et *tribunal* complètent le registre lexical dédié strictement au monde judiciaire. Une série de termes relatifs à la menace terroriste est également présente : *terrorisme*, *pénal*, *liberté*, *terroriste* et *Etat d'urgence*. Cette proximité dans la même classe de discours reste à interpréter.

Plusieurs ensembles lexicaux se distinguent du nuage de mots (colonne centrale). Le premier porte sur la *criminalité* à travers des termes *délinquant*, *cambriolage*, *acte*, *commettre*, *crime*, *infraction*, *criminel*, *délinquance*, *terrorisme*, *détenu*, *délit*, *attentat*, *terroriste* et *violence*. Agissant comme une réponse à ce champ lexical de l'insécurité, se dégage un second registre lié à la *surveillance* avec les mots *fichier*, *exception*, *sécurité*, *antiterroriste*, *contrôle*, *renseignement*, *régime*, *police*, *résidence*, *État*, *urgence*, *intérieur*, *sûreté*, *assignation*, *sécuritaire*, *dangerosité*, *pénal* et *perquisition*.

Sont également citées de nombreuses institutions judiciaires telles que la *Cour de cassation*, les *cours d'appel*, le *Conseil constitutionnel*, les *bâtonniers*, le *barreau*, le *procureur* et, de manière plus générale, les *instances* et *juridictions*. Le registre lexical du procès et du jugement est également présent par l'intermédiaire des mots *peine*, *prison*, *procès*, *procédure*, *indemnité*, *condamner*, *juger*, *jugement*, *litiges* et *contentieux civils*. Enfin, certains termes renvoient aux dispositifs algorithmiques d'analyse des décisions de justice, de manière plus ou moins directe (*jurisprudence*, *décision*, *CNIL*, *algorithme*, *indemnité*). Les destinataires de la justice (*citoyens*, *justiciables*) sont également mentionnés.

Je peux avancer plusieurs conclusions quant à la façon dont justice et dimension prédictive sont associées dans cette classe de discours. D'abord, les termes relatifs aux dispositifs algorithmiques de traitement des décisions de justice civile sont présents, mais ils n'en constituent pas le cœur — comme le montrent aussi bien les mots les plus significatifs de la classe que le nuage de mots. Ensuite, et c'est complémentaire, apparaît un autre registre lexical qui porte sur la criminalité, le terrorisme, le contrôle, la surveillance. Il suggère que la dimension 'prédictive' de la justice pourrait aussi être associée dans la presse à des dispositifs de type actuariels, dans un contexte sécuritaire accru, ce qui résonne avec l'état de l'art réalisé dans le précédent chapitre. L'on y a vu que la discussion mêle souvent les algorithmes d'analyse des décisions de justice avec ceux d'évaluation des risques utilisés en matière pénale, comme *Compas*.

La lecture des cinq segments de texte les plus caractéristiques de la classe conforte l'idée d'une double connotation de la 'justice prédictive' : associée d'une part à des dispositifs d'aide à la décision (« juges virtuels ») dans les segments 2, 3 et 5 ci-dessous ; et d'autre part à la surveillance et à l'état d'urgence instaurés à la suite des attentats de 2015 dans les segments 1 et 4 (« loi sur l'état d'urgence », « déséquilibre » entre « défense des libertés et de l'ordre public »).

Encadré 2.1. Les 5 segments de texte les plus caractéristiques de la classe 2

**** *annee_2017 *AnneeMois_2017-10 *Date_2017-10-31 *Journal_Le_Monde_WEB *predictif_OUI *predictive_NON *justicepredictive_NON *sourcesconfondues_Le_Monde *typepresse_PQN *support_WEB *couleurpol_Le_Monde *antepost9oct_POST *phases_3 *953_21563

score : 11309.25

pour paul cassia professeur de **droit** à l'université de paris i panthéon sorbonne la **loi** de l'état d'urgence a déséquilibré la balance que le **juge** doit opérer entre la **défense** des **libertés** et celle de l'**ordre public** mais les **juridictions administratives** et **judiciaires** ont **aggravé** ce déséquilibre

**** *annee_2018 *AnneeMois_2018-03 *Date_2018-03-19 *Journal_Ouest_France_WEB *predictif_NON *predictive_OUI *justicepredictive_OUI *sourcesconfondues_Ouest_France *typepresse_PQR *support_WEB *couleurpol_PQR *antepost9oct_POST *phases_4 *1072_24153

score : 10268.67

l'aléa en la **matière** doit être **préservé** marylise couraud **magistrats** et **avocats** émettent des doutes **justice** prédictive **juges** virtuels **aide** à la **décision** on peut leur donner le **nom** qu'on veut les différents **logiciels** qui tambourinent à la porte du monde **judiciaire** occupent les **esprits**

**** *annee_2018 *AnneeMois_2018-03 *Date_2018-03-20 *Journal_Ouest_France_PAPIER *predictif_NON *predictive_OUI *justicepredictive_OUI *sourcesconfondues_Ouest_France *typepresse_PQR *support_PAPIER *couleurpol_PQR *antepost9oct_POST *phases_4 *1072_24269

score : 10080.90

magistrats et **avocats** émettent des doutes **justice** prédictive **juges** virtuels **aide** à la **décision** on peut leur donner le **nom** qu'on veut les différents **logiciels** qui tambourinent à la porte du monde **judiciaire** occupent les **esprits**

**** *annee_2015 *AnneeMois_2015-11 *Date_2015-11-26 *Journal_Le_Monde_WEB *predictif_NON *predictive_OUI *justicepredictive_OUI *sourcesconfondues_Le_Monde *typepresse_PQN *support_WEB *couleurpol_Le_Monde *antepost9oct_ANTE *phases_2 *528_11423

score : 9658.54

tant que la **loi semble** respectueuse pour les **nationaux** les écarts du **droit** ne paraissent pas dramatiques le problème avec la **guerre** contre le **terrorisme** c'est qu'elle n'aura sans doute pas de **fin antoine garapon secrétaire général** de l'institut des hautes études sur la **justice**

**** *annee_2017 *AnneeMois_2017-02 *Date_2017-02-15 *Journal_Le_Monde_WEB *predictif_OUI *predictive_OUI *justicepredictive_OUI *sourcesconfondues_Le_Monde *typepresse_PQN *support_WEB *couleurpol_Le_Monde *antepost9oct_ANTE *phases_3 *793_17717

score : 9372.19

plusieurs start_up tricolores surfent déjà sur l'**open data juridique** la **loi** pour une **république numérique impose** la diffusion de toutes les **décisions judiciaires** et tentent d'investir le champ de la **justice** prédictive **case law analytics** prédictive

1.2.2. Un double sens inégalement réparti sur la période

L'analyse manuelle des articles mentionnant la 'justice prédictive' sur la période 2010-2017 dans seulement deux journaux (*Le Monde* et *La Voix du Nord*) montre que cette interprétation d'un double sens est robuste. Les premières occurrences du terme apparaissent en 2014 dans le journal *Le Monde*, (et *LeMonde.fr*) dans un article relatant l'avis du contrôleur général des lieux de privation de liberté, Jean-Marie Delarue à l'époque, portant sur la mise en œuvre de la loi de 2008 sur la rétention de sûreté¹. Celle-ci autorise la rétention de personnes condamnées après qu'elles ont purgé leur peine de prison, lorsqu'elles ont commis des atteintes aux personnes très graves et dans la mesure où elles présentent une dangerosité particulière se traduisant par un risque très élevé de récidive. Évidemment, cela ne veut pas dire que le terme est composé à ce moment-là, mais qu'à partir de 2014, il acquiert une visibilité.

Au printemps 2015, il réapparaît cette fois dans la bouche de la juriste Mireille Delmas-Marty, à l'occasion d'un entretien portant sur les dérives sécuritaires dans les réponses au terrorisme (en France comme aux États-Unis). Est pointé le risque du passage d'« une justice pénale à une justice prédictive », centrée sur la protection de la société par l'enfermement de personnes en raison de leur dangerosité. Ce terme est utilisé par M. Delmas-Marty et repris en intertitre de l'interview. À l'automne de la même année, il émerge à nouveau, là encore dans un *verbatim* de cette figure du droit et de la défense des droits humains où elle fait le lien entre la loi sur la rétention de sûreté de 2008 et les mesures prises pour la lutte contre le terrorisme. Elle estime qu'un pas a été franchi avec cette loi : « La rétention de sûreté instaure donc une justice prédictive. Elle repose sur une logique de suspicion — fondée sur des pronostics — et non plus sur une logique d'accusation — fondée sur des preuves. C'est grave,

¹ Franck JOHANNÈS, « Le fondement juridique de la rétention de sûreté jugé "hasardeux et incertain" », *Le Monde*, 27 février 2014.

car il est impossible de prédire tous les comportements humains.¹ » En 2016, c'est à nouveau en lien avec la menace terroriste qui pèse sur les démocraties et en référence à Mireille Delmas-Marty que la notion de 'justice prédictive' apparaît dans les colonnes du *Monde*.

On note donc que l'expression apparue dans le corpus en 2014 est ponctuellement présente les deux années suivantes, à propos de la justice pénale, dans le cadre de la mise sous tension des libertés publiques par des mesures de sécurité, notamment dans un contexte terroriste. Le terme apparaît dans la bouche d'une juriste et d'un ancien conseiller d'État, tous deux spécialistes de questions juridiques, dans le cadre d'articles réalisés par les journalistes du service Justice du *Monde* (Anne Chemin, Franck Johannès). L'expression 'justice prédictive' leur sert à désigner et qualifier les risques encourus du fait des nouvelles législations adoptées (qu'il s'agisse de la loi de 2008 sur la rétention de sûreté ou celles sur l'état d'urgence).

L'année 2017 représente ensuite une césure dans la mesure où la fréquence de l'expression s'accroît considérablement — avec un effet relatif de pic sur lequel on va revenir — et où le sens et les connotations se diversifient. Pour la seule année 2017 en effet, toujours dans le seul journal *Le Monde*, 11 articles comprennent le terme 'justice prédictive' (contre 2 en 2016).

Une partie d'entre eux continue à traiter de la justice pénale et en particulier dans le cas d'affaires de terrorisme islamiste. La 'justice prédictive' est alors un label critique destiné à stigmatiser des pratiques pénales qui visent des groupes plus que des individus et fragilisent les principes fondamentaux de l'État de droit. Une autre partie assimile la 'justice prédictive' à l'utilisation d'algorithmes reposant sur l'analyse automatisée de grandes quantités de décisions de jurisprudence. Un dernier ensemble porte enfin précisément sur l'expérimentation [PréviCompute] (qu'elle soit explicitement nommée ou pas). Ces derniers articles se retrouvent essentiellement dans *La Voix du Nord* (aucun article jusqu'en 2016 ; 8 en 2017).

Il faudrait poursuivre l'analyse qualitative des articles sur les années 2018 à 2020 pour les deux titres — et la réaliser pour toute la période sur les autres. À ce stade, on peut faire l'hypothèse que le mouvement amorcé en 2017 se poursuit avec le maintien d'une partie des articles qui traitent de l'utilisation d'algorithmes de traitement des décisions de justice et/ou de l'expérimentation [PréviCompute]. On peut se demander si l'utilisation de l'expression 'justice prédictive' comme label critique pour dénoncer des mesures et pratiques de type actuariel disparaît complètement ou continue à exister — et, dans ce dernier cas, si elle est reliée ou pas à la question des algorithmes d'aide à la décision pour les avocats et les juges. Dans tous les cas, l'hétérogénéité de la notion est ici avérée, tout comme se révèle sa progressive assimilation à des dispositifs algorithmiques de traitement et d'analyse des décisions de justice.

1.3. Anticipation et délégation à des machines

Ce qui fait lien entre les différentes classes de discours permet également de mieux saisir ce qu'exprime le « prédictif » de 'justice prédictive'. Transversalement aux quatre

¹ Anne CHEMIN, « État d'urgence : le débat piégé », *Le Monde*, 28 novembre 2015.

secteurs (santé, automobile, élections, justice), il renvoie à deux dimensions : les approches par probabilités et modélisations qui dessinent des projections et des anticipations d'une part, la délégation des activités classiquement accomplies par les humains à des algorithmes, machines ou robots d'autre part.

La première concerne, dans le cas de la santé, *l'évaluation et l'anticipation* de risques liés aux maladies, principalement les cancers, via des tests génétiques. Le discours attaché à la « médecine prédictive » est appuyé sur la logique d'identification de critères d'appartenance à un groupe — dont les membres sont porteurs de certaines caractéristiques qualifiées de facteurs de risque. L'enjeu est d'anticiper la survenance de ces risques médicaux. La même logique a cours pour la 'justice prédictive' lorsqu'il s'agit d'évaluer la dangerosité des personnes, selon une logique actuarielle et probabiliste.

L'anticipation est aussi celle de résultats pour lesquels il existe une incertitude. Dans le domaine de la compétition électorale, la dimension prédictive s'exprime à travers les *sondages* — terme le plus significatif de la classe — dans la mesure où ils cherchent à prévoir le résultat des *élections*. Le parallèle avec la justice est limpide : le *sondage* réduit l'élection au *résultat* du *vote*, comme la 'justice prédictive' réduit le jugement à une décision. Dans un cas comme dans l'autre, l'accent est mis sur le résultat final. Quant à l'anticipation de ce résultat, elle se fonde sur des *intentions* de *vote* des électeurs et électrices ; comme les dispositifs de justice actuarielle appliquent des mesures de prévention à des personnes dont il est considéré qu'elles pourraient commettre un délit ou en auraient l'intention.

La seconde direction est relative au transfert d'activités usuellement accomplies par des humains (conduire, juger) vers des robots et machines. Le prédictif se focalise dans le secteur de l'automobile sur le cas des voitures autonomes ou connectées. Plus qu'une dimension prédictive au sens strict, la voiture autonome renvoie à l'usage des *big data* et de l'intelligence artificielle pour déléguer des tâches jusqu'alors réservées aux humains à des robots. L'analyse des cooccurrences du mot *autonome* au sein de la classe 6, outre qu'elle confirme que ce terme est bien associé en majorité aux mots *voiture* et *véhicule*, met en évidence que la perspective d'une *voiture autonome* revient à confier la *conduite* à un *robot*, ce qui fait écho à certains discours sur la 'justice prédictive' et renvoie à la controverse autour de la déshumanisation et de l'éventuelle prise de pouvoir des machines sur les hommes (confier la fonction de juger à un robot). Les interrogations sur ce vers quoi la 'justice prédictive' entraîne l'institution judiciaire sont d'autant plus vives qu'elle est associée au fait de déposséder les humains de l'exercice d'une fonction essentielle à la régulation des conflits et de la société, au profit de machines.

1.4. De quoi l'avenir sera-t-il fait? Stupeurs et tremblements...

Considérer de près le système de classe constitué autour de la justice (classes 2-3-4) met en lumière quelques spécificités des discours sur la 'justice prédictive' et plus généralement sur le monde judiciaire. L'ampleur des questionnements prospectifs et la force de l'inquiétude qui les traverse sont notables.

Encadré 2.2. Les cinq segments de texte les plus caractéristiques de la classe 3. Corpus d'articles de presse, mot-clef « prédictif.ve » 2010-2020

**** *annee_2010 *AnneeMois_2010-09 *Date_2010-09-09 *Journal_Le_Point_WEB *predictif_NON *predictive_OUI *justicepredictive_NON *sourcesconfondues_Le_Point *typepresse_HEBDO *support_WEB *couleurpol_DROITE *antepost9oct_ANTE *phases_1 *47 712

score : 1654.65

sans **prétendre** trancher ou apporter une **vérité** à une matière | **histoire** qui n'est pas la sienne **philippe** forest fait ici **oeuvre** d'**écrivain** dans un **roman** magnifique pour finalement **raconter** quelque **chose** qui **ressemble** beaucoup à l'**âme** française

**** *annee_2011 *AnneeMois_2011-09 *Date_2011-09-09 *Journal_Le_Monde_PAPIER *predictif_NON *predictive_OUI *justicepredictive_NON *sourcesconfondues_Le_Monde *typepresse_PQN *support_PAPIER *couleurpol_Le_Monde *antepost9oct_ANTE *phases_1 *85 1498

score : 1652.19

vous êtes **écrivain** vous **sentez** les **choses** avant tout le **monde** vous **inventez** une **histoire** plus **vraie** que nature celle d'un nantais au faite de sa gloire dont le **destin** bascule soudain à cause d'un invraisemblable **scandale sexuel**

**** *annee_2018 *AnneeMois_2018-02 *Date_2018-02-12 *Journal_Le_Figaro_WEB *predictif_NON *predictive_OUI *justicepredictive_NON *sourcesconfondues_Le_Figaro *typepresse_PQN *support_WEB *couleurpol_DROITE *antepost9oct_POST *phases_3 *1034 23393

score : 1628.54

à l'un des plus grands essais de la **fin** du **siècle** dernier tout d'abord les découvreurs bouquins **robert** laffont de daniel boorstin qui nous **livre** une **histoire** enlevée et passionnante de la plus grande épopée de l'**homme**

**** *annee_2010 *AnneeMois_2010-09 *Date_2010-09-09 *Journal_Le_Point_WEB *predictif_NON *predictive_OUI *justicepredictive_NON *sourcesconfondues_Le_Point *typepresse_HEBDO *support_WEB *couleurpol_DROITE *antepost9oct_ANTE *phases_1 *47 706

score : 1564.29

juché sur l'épaule paternelle c'est avant tout l'**histoire** d'un **siècle** que l'**écrivain** raconte dans ce **livre** quand l'aviation était une **religion** et incarnait l'optimisme du **xxe siècle** à **travers** l'épopée de mermoz

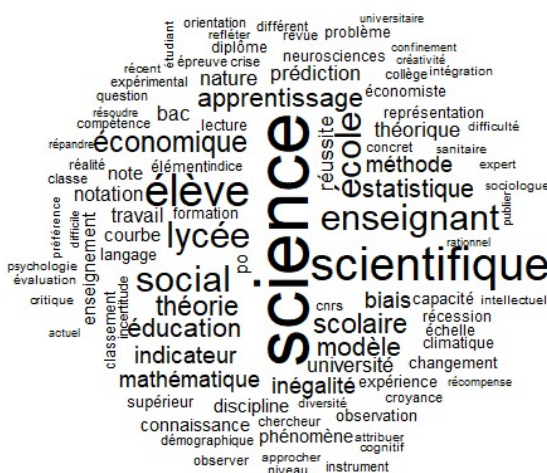
**** *annee_2020 *AnneeMois_2020-05 *Date_2020-05-05 *Journal_Le_Figaro_WEB *predictif_NON *predictive_OUI *justicepredictive_NON *sourcesconfondues_Le_Figaro *typepresse_PQN *support_WEB *couleurpol_DROITE *antepost9oct_POST *phases_4 *1543 36105

score : 1495.10

ce qui n'**empêche** pas l'**historien philippe** fabry de porter en lui cette tournure d'**esprit** qu'il développe brillamment dans ses **livres** dans son **histoire** du **siècle** à **venir** il **expose** un modèle cyclique de deux millénaires

La classe 4, quant à elle, est construite autour du monde scientifique et académique, de l'enseignement et de la recherche.

Graphique 2.7. Nuage représentant les 100 mots les plus significatifs de la classe 4



Parmi les mots les plus significatifs et dans le nuage de mots, se distingue un registre relatif au contexte scolaire et à l'apprentissage (*éducation, école, lycée, élève, bac, langues, diplôme, enseignant, étudiant, enseignement, langage, évaluation, collège et université*). L'université, en tant que lieu de production de savoir, participe aussi du registre lexical de la recherche scientifique qui s'exprime au fil des mots *universitaire, chercheur, méthode, expérimental, connaissance, observation, revue, CNRS*. Diverses disciplines sont mentionnées : *sociologie, psychologie, mathématique, neurosciences, économiste*. Cet univers de discours,

centré sur l'enseignement et la recherche, semble indiquer que le prédictif devient un sujet d'intérêt pour les enseignants, chercheurs ou essayistes.

La diversité des disciplines évoquées suggère une certaine transversalité, laquelle est confortée par les cinq segments de texte les plus caractéristiques de la classe. Ils laissent entrevoir une dimension méthodologique et épistémologique.

Encadré 2.3. Les cinq segments de texte les plus caractéristiques de la classe 4

**** *annee_2019 *AnneeMois_2019-10 *Date_2019-10-03 *Journal_Le_Point_WEB *predictif_NON *predictive_OUI *justicepredictive_OUI *sourcesconfondues_Le_Point *typepresse_HEBDO *support_WEB *couleurpol_DROITE *antepost9oct_POST *phases_4 *1431 33109
score : 2462.06

et pourtant avant la seconde guerre mondiale on ne s'intéressait pas à cette question alors qu'après 1945 les revues de sciences sociales et politiques se focalisent de plus en plus sur l'approche rationnelle voire scientifique du politique

**** *annee_2012 *AnneeMois_2012-03 *Date_2012-03-22 *Journal_Le_Monde_PAPIER *predictif_OUI *predictive_NON *justicepredictive_NON *sourcesconfondues_Le_Monde *typepresse_PQN *support_PAPIER *couleurpol_Le_Monde *antepost9oct_ANTE *phases_1 *120 2241
score : 2456.82

la formation des enseignants questions à christophe michaud maître de conférences en sciences à l'université de nantes enseignant en sciences de l'éducation à nantes christophe michaud est spécialiste de l'évaluation des politiques de sélection des étudiants

**** *annee_2019 *AnneeMois_2019-10 *Date_2019-10-17 *Journal_Le_Monde_WEB *predictif_NON *predictive_OUI *justicepredictive_NON *sourcesconfondues_Le_Monde *typepresse_PQN *support_WEB *couleurpol_Le_Monde *antepost9oct_POST *phases_4 *1441 33478
score : 2398.40

l'approche expérimentale qu'ils ont promue dite expériences contrôlées repose sur des méthodes statistiques qui étaient déjà connues des sciences sociales et largement mobilisées en médecine mais peu utilisées par les économistes

**** *annee_2019 *AnneeMois_2019-10 *Date_2019-10-04 *Journal_L_Obs_WEB *predictif_OUI *predictive_NON *justicepredictive_NON *sourcesconfondues_L_Obs *typepresse_HEBDO *support_WEB *couleurpol_GAUCHE *antepost9oct_POST *phases_4 *1433 33148
score : 2323.97

il entend désormais chasser les fraudeurs sur les réseaux sociaux et les places de marché amazon etc gregory lewkowicz professeur de droit à l'université libre de bruxelles enseignant à assas hec paris et sciences po

**** *annee_2014 *AnneeMois_2014-02 *Date_2014-02-21 *Journal_Le_Monde_WEB *predictif_NON *predictive_OUI *justicepredictive_NON *sourcesconfondues_Le_Monde *typepresse_PQN *support_WEB *couleurpol_Le_Monde *antepost9oct_ANTE *phases_1 *248 4843
score : 2320.55

n'ayant jamais terminé de thèse en psychologie ou en sciences de l'éducation il n'était pas officiellement considéré comme chercheur pourtant dans les débats qui traversent la communauté des chercheurs sur les premiers apprentissages scolaires ses travaux comptent

L'analyse de sémantique lexicale par classification du corpus d'articles de presse étudié montre donc que la justice est un des domaines pour lesquels la question prédictive se pose, surtout à partir de la seconde moitié des années 2010, en lien avec des univers de discours relatifs au monde académique et à travers des interrogations sur la société et son avenir ; et ce, davantage comparativement aux élections (sondages), à l'automobile (voiture autonome), et à la santé (médecine prédictive).

1.1.2. Un avenir inquiétant... pour l'institution et ses professionnels ?

Compte tenu de l'orientation de la classe 3, il apparaît logique qu'elle soit extrêmement proche de la classe 4 (monde académique) puisque les scientifiques, chercheurs, enseignants et autres intellectuels sont celles et ceux qui, fonctionnellement et par leur positionnement, réfléchissent et s'expriment sur l'état du monde et soulèvent des enjeux prospectifs. Quant au lien avec la classe 2, dédiée à la justice, cela incite à conclure que les discours sur la 'justice prédictive' sont structurellement plus connectés aux espaces académiques et suscitent plus d'interrogations philosophiques que la santé ou les sondages. La 'justice prédictive' est traitée à travers des réflexions prospectives larges, et ce, bien plus

que d'autres secteurs qui devraient aussi susciter d'importantes questions de ce type, comme la santé par exemple.

Cette mise en doute de l'avenir m'apparaît être le signe, en creux, d'une représentation dominante de la justice comme insuffisamment modernisée voire a-technologique ; une institution aux normes, rituels et processus, d'abord et avant tout, traditionnels, voire passéistes. L'interprétation fait sens compte tenu du fait que cette représentation est une antienne du débat public en général. En outre, elle fait écho à une partie des sentiments qu'expriment les citoyens et citoyennes quand ils sont interrogés sur leurs perceptions de la justice et de son fonctionnement (Dumoulin et Vigour, 2019). Ils décrivent un monde de traditions, un peu figé autour de symboliques parfois vues comme décalées, en rupture avec les mondes vécus ordinaires. Ce type de cadrage autour d'une justice peu ouverte aux technologies est de nature à favoriser une forme de réticence et/ou de peur à l'égard des innovations surtout si elles sont présentées comme disruptives et relevant d'un saut dans l'inconnu. Les opérateurs d'algorithmes ont d'ailleurs intérêt à cultiver cette représentation et à contribuer à ce *framing* pour mieux asseoir l'idée d'une révolution liée à l'intelligence artificielle.

1.1.3. *Formes savantes, acteurs et espaces académiques au premier plan*

L'unité lexicale de la classe 4 se forme autour du monde scientifique et académique et elle est davantage reliée à la justice qu'aux autres domaines. On peut imaginer que cela traduit la place spécifique qu'occupent le savoir académique et les formes savantes (colloque, livres, etc.) dans la réflexion sur la 'justice prédictive' et plus largement sur la justice comme institution et secteur d'action publique. D'autant plus que les professionnels du droit sont particulièrement multi-positionnés et souvent à la charnière entre plusieurs mondes. Les praticiens que sont les avocats et les magistrats gardent souvent un pied dans l'université, à travers des enseignements, la participation à des colloques ou la publication d'articles dans des revues juridiques. Le monde de la justice est de surcroît peuplé de professionnels du droit pour assurer l'administration et pour concevoir les politiques judiciaires, et leurs réformes (Vauchez et Willemez, 2007 ; Vigour, 2008).

Si les articles de presse sont écrits par des journalistes, les *tweets* émanent pour les deux tiers de comptes personnels (4089 sur 6261). Pour 65 % de ces comptes personnels, nous sommes parvenus à identifier la profession des utilisateurs et utilisatrices, ce qui permet de montrer que si les avocats et avocates sont de loin le groupe professionnel le plus actif de notre corpus avec un tiers des *tweets* pour lesquels nous connaissons la profession de l'utilisateur à l'origine du message (793 *tweets* soit 29,52 %) ; les universitaires juristes sont finalement bien représentés avec 13% des messages postés (350 *tweets* soit 13,03 %). D'autant que parmi le tiers restant des *tweets* (2172 sur 6261), émis par des comptes appartenant à des structures, les institutions académiques sont également très présentes : universités, facultés, écoles, masters et centres de recherche twittent sur la 'justice prédictive'¹.

¹ Pour les comptes personnels, les magistrats et magistrates sont très marginalement présents (74 *tweets* soit 2,76 %), et les autres professionnels du droit représentent aussi une portion congrue (8,45 % soit 227 *tweets*). Il

Les professionnels du droit et de la justice sont donc massivement présents, mais parmi eux avocats et juristes académiques sont particulièrement actifs. Ils relaient en particulier des messages relatifs à des manifestations scientifiques comme des colloques.

Voyons à présent quelle est la dynamique de développement des discours sur la 'justice prédictive', à partir d'analyses croisées du sous-corpus d'articles de presse (n=122) et du corpus de *tweets* (n=6 261) qui, de 2010 à 2020, comportent ce terme.

II / La trajectoire ascendante de l'expression 'justice prédictive' au milieu des années 2010

L'analyse comparée des corpus d'articles de presse et de *tweets* confirme d'abord que la 'justice prédictive' est une construction sémantique récemment apparue, ponctuellement repérable à partir de 2010, mais plus régulièrement mobilisée à partir du milieu de la décennie, avec une explosion en 2017 (1). L'analyse approfondie du corpus de *tweets* sur un plan chronologique et à différentes échelles (année, mois, jour) révèle ensuite une courbe structurée en quatre séquences, elles-mêmes marquées par d'importantes variations. L'étude des pics d'intensité et de la façon dont ils se forment, renseigne sur la dynamique de diffusion de l'expression 'justice prédictive' sur Twitter, compte tenu des spécificités de production et de circulation de l'information sur ce réseau social et des pratiques de ses utilisateurs (2).

1.2. Des trajectoires comparables dans la presse et les *tweets* : le pic de 2017

Les volumes d'articles de presse et de *tweets* diffèrent considérablement, une centaine pour les uns, plusieurs milliers pour les autres. Toutefois, l'émergence de l'expression 'justice prédictive' est concomitante dans les deux corpus : deux premières occurrences dans la presse¹ en 2010 et la première dans un *tweet*² fin 2011. Les trajectoires sont ensuite similaires : apparition ponctuelle avant 2015, percée à partir du milieu des années 2010, avec

faut ajouter que 13% proviennent des métiers de l'information juridique (371 *tweets* soit 13,81 %) et 5% de professionnels de l'information généraliste (134 *tweets* soit 4,99 %) ; 5% d'universitaires non juristes (141 *tweets* soit 5,25 %) ; 2,35 % des messages (63 *tweets*) sont postés par des utilisateurs exerçant un métier en lien avec la *data*. Seulement 39 *tweets* (1,45 %) le sont par le personnel politique. Quant aux *tweets* des étudiants, 68 émanent d'apprentis juristes (2,53 %) et 22 d'étudiants d'autres disciplines (0,82 %). Pour les *tweets* provenant de comptes de structures ; On y retrouve également la présence des avocats — via des cabinets ou instances de représentation de la profession comme les barreaux, leurs incubateurs ou le Conseil national des barreaux. Des acteurs économiques du monde du droit y participent aussi : de grands groupes d'édition juridique, des *legaltechs* dont les sociétés qui proposent des dispositifs algorithmiques. Les médias généralistes et spécialisés tweetent également ; le Village de la justice poste à lui seul 125 *tweets*. En revanche, les acteurs institutionnels comme le ministère de la Justice (avec seulement 2 *tweets*) la Cour de cassation (5) et plus largement les juridictions, à l'image des magistrats et magistrates, sont quasiment absents.

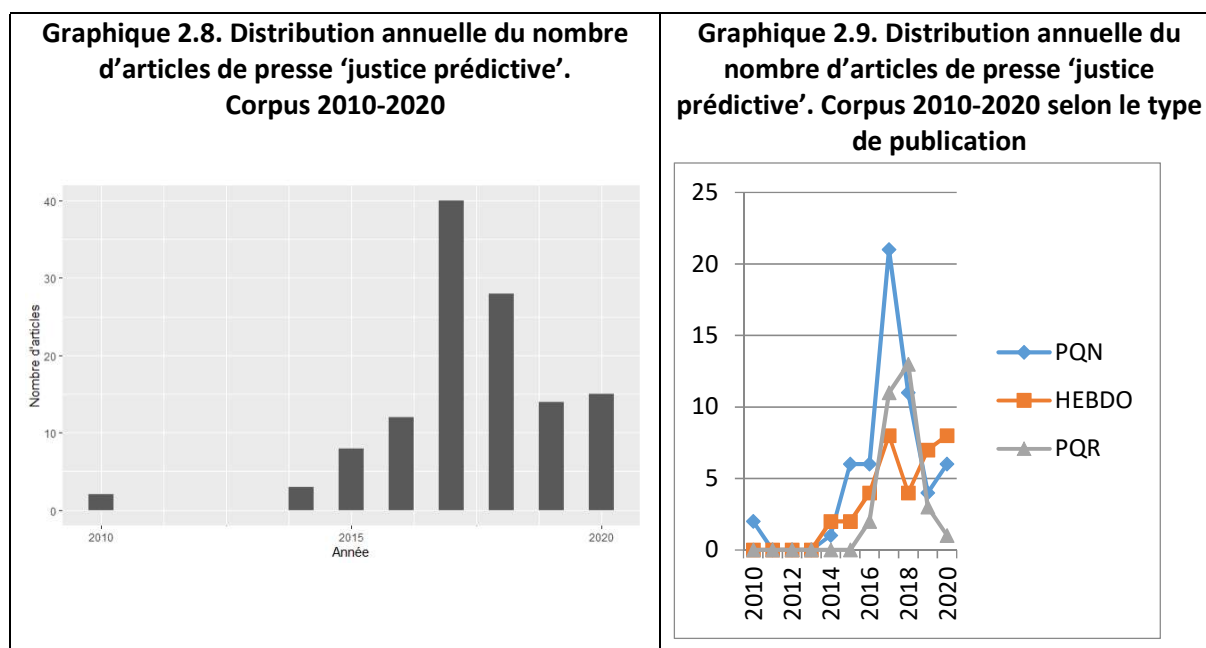
¹ « Un logiciel d'aide à la décision pour la justice », *Le Monde.fr*, 22 avril 2010 ; Sonya Faure, « Violences conjugales : les cris étouffés des victimes », *Libération*, 22 juin 2010, p. 8.

² « Fantôme d'une justice prédictive et hyper répressive : un excellent billet ! », 22 novembre 2011, 12:13. Le *tweet* répercute un post de *Rue 89* lequel reprenait lui-même un article publié par *Le Nouvel Observateur* : <https://www.nouvelobs.com/rue89/rue89-nos-vies-connectees/20111122.RUE5871/le-meurtre-du-chambon-sur-lignon-et-le-fantome-minority-report.html>

un effet de pic brutal en 2017 puis lente décade et stabilisation à un niveau plus élevé qu'auparavant.

1.2.1. Une courbe en forme de montagne russe dans la presse

Le graphique sur la distribution annuelle du nombre d'articles de presse (cf. ci-dessous) dessine une courbe dont l'allure générale est celle d'une montagne russe où la montée est plus rapide et marquée que la descente. Le pic est clairement repérable en 2017 de même que le contraste entre la première et la seconde moitié de la décennie. C'est en effet à partir de 2014 qu'intervient une augmentation progressive et continue du nombre d'articles traitant de la 'justice prédictive'. 2018 représente un recul par rapport à 2017, mais avec un nombre qui reste élevé. La chute se poursuit en 2019 pour sembler se stabiliser en 2020 à un niveau de publication plus important que celui connu avant le pic de 2017. De 2014 à 2020, l'évolution est considérable puisque l'on est passé de quelques articles sporadiques (3 en 2014) à plusieurs dizaines chaque année : 40 en 2017, 28 en 2018, encore 19 en 2019 et 20 en 2020. Les années 2017 et 2018 réunies représentent un peu plus de la moitié des articles publiés sur la décennie (68 sur 122). Cette évolution touche de la même manière les différents supports de publications, presse quotidienne et hebdomadaire, nationale (PQN) et régionale (PQR) comme le montre l'autre graphique. La chute, plus accentuée dans la PQR, est probablement liée au fait que les sujets journalistiques relatifs à la 'justice prédictive' sont indexés sur les actualités locales qu'elle a suscitées : expérimentation de [PréviCompute] à Douai et Rennes, manifestations scientifiques et judiciaires (journées d'études)¹.

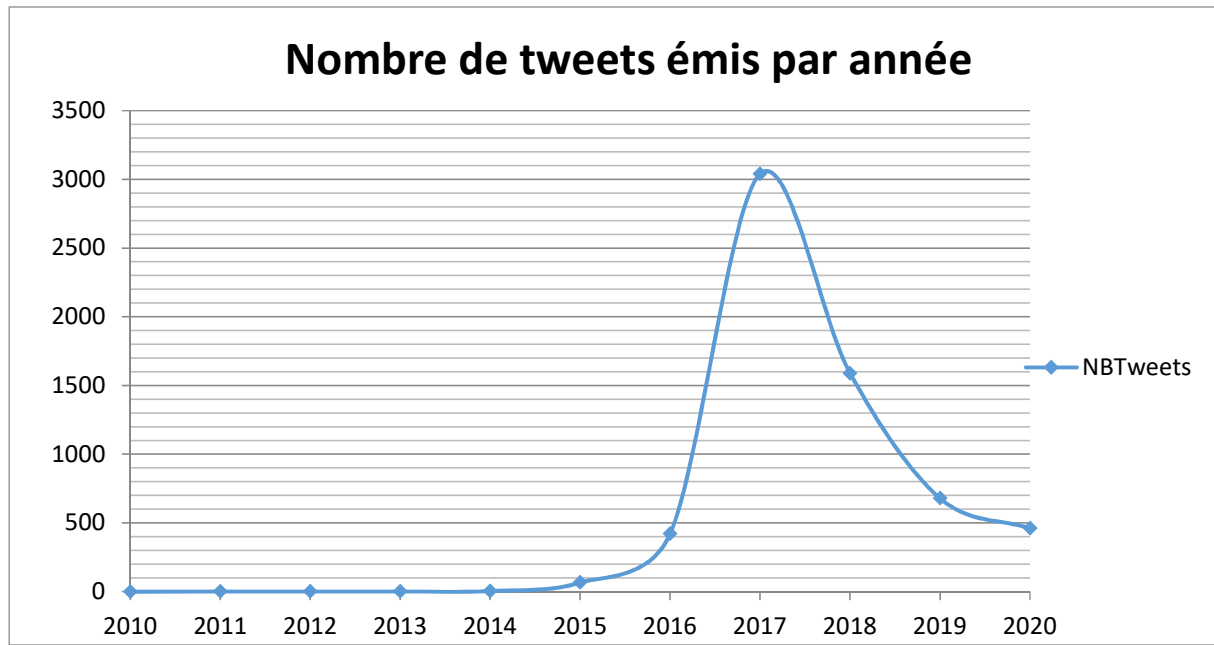


¹ Le sous-corpus de 122 articles est ainsi structuré : 38 sont issus du journal *Le Monde*, 27 proviennent du *Point*, 17 articles de *Ouest-France*, 13 de *La Voix du Nord*, 10 du *Figaro*, 9 de *Libération* et enfin seulement 8 de *L'Obs*. On observe donc un investissement différentiel selon les titres de presse sur le sujet de la 'justice prédictive'. *Le Monde* et *Le Point* sont les journaux en pointe quant à cette question ; il faut toutefois pondérer cela par le fait que certains sont des quotidiens et d'autres des hebdomadaires, donc l'espace disponible pour traiter les différents sujets n'est pas le même. Cela ne fait que renforcer la surreprésentation de l'hebdomadaire *Le Point*, particulièrement actif via son site web.

1.2.2. Le pic de 2017 sur Twitter

Même si les ordres de grandeur diffèrent, l'agrégation des *tweets* au niveau annuel dessine une courbe très comparable dans le temps, avec un pic encore plus marqué en 2017.

Graphique 2.10. Nombre de tweets émis par année sur la période 2010-2020



Seulement 7 *tweets* sont collectés de 2010 à 2014. De fait, la 'justice prédictive' pénètre timidement sur Twitter en 2015. Suit alors une phase ascendante, lente en 2015 puis très rapide en 2016, durant laquelle l'intensité de la controverse augmente jusqu'à culminer à un moment d'effervescence en 2017. Quasiment la moitié du corpus total (3 040 *tweets* sur 6 261) correspond à la seule année 2017 : l'effet de pic est plus accentué que pour les articles de presse. S'amorce ensuite une phase descendante jusqu'en 2019, où l'activité se stabilise à un niveau plus élevé qu'avant 2017. Les années 2018 et 2019 représentent respectivement la deuxième et la troisième année pour lesquelles le plus de *tweets* relatifs à la 'justice prédictive' sont collectés. De 2019 à 2020, la baisse continue pour se stabiliser autour d'un peu moins de 500 *tweets* sur l'année.

L'année 2017 est donc clairement le principal pic d'intensité repéré. Ce que nous avons pressenti et que Larrègue (2019) a rapidement montré pour la littérature en droit et en sciences sociales est donc confirmé plus généralement, pour le débat appréhendé tant à travers la presse généraliste que sur la plateforme Twitter. Cela s'explique probablement en partie parce que les publications scientifiques relatives à la 'justice prédictive' se multiplient en 2017, que les colloques se succèdent et que cette actualité académique alimente les sujets journalistiques et les échanges de *tweets*. C'est également au cours de cette année 2017 que les cours d'appel de Douai et Rennes expérimentent pour la première fois en France un logiciel [PreviCompute] qui est qualifié par ses promoteurs de 'justice prédictive'. Ce logiciel a

également été testé par des cabinets d’avocats appartenant au barreau lillois à l’initiative du bâtonnier [Sébastien Rouve], vraisemblablement au même moment. Le pic d’intensité repéré dans la presse et sur Twitter témoigne donc de ce moment particulier où la ‘justice prédictive’ est au centre de l’attention des professionnels du droit et où les discours mobilisent effectivement cette expression pour en parler. Il y a bien un moment 2017.

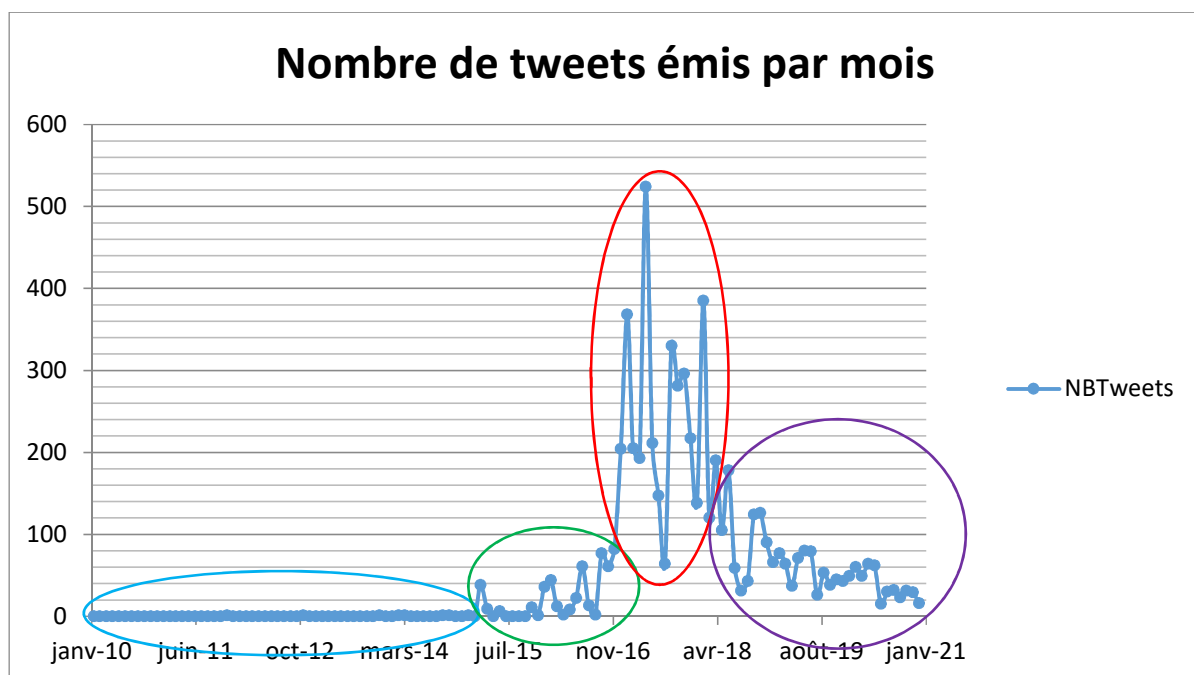
1.3. Analyse de séquences et de pics d’intensité

L’activité sur le réseau social Twitter est fortement marquée par son immédiateté et sa viralité, ce qui explique que le pic de 2017 soit quantitativement plus marqué sur cette plateforme que dans la presse. Les séquences identifiées doivent donc être scrutées à une échelle temporelle plus fine, mensuelle et même quotidienne.

1.3.1. Quatre phases et de fortes oscillations liées à la viralité sur Twitter

Afin d’affiner les premiers résultats obtenus, le nombre de tweets postés pendant chaque mois de la période étudiée a été calculé. Le graphique ci-dessous, qui en est issu, est plus contrasté, avec des oscillations importantes et des pics d’intensité plus courts, qui mettent en évidence quatre séquences distinctes.

Graphique 2.11. Nombre de tweets émis par mois sur la période 2010-2020



Après une première phase de latence qui s’étire jusqu’en mars 2015, au cours de laquelle la ‘justice prédictive’ n’est pas, ou extrêmement peu, présente sur Twitter, s’en ouvre une autre qui dure jusqu’à la fin de l’année 2016 (en vert), pendant laquelle l’activité discursive liée à la ‘justice prédictive’, marginale, augmente par petits à-coups. La troisième séquence, celle du bouillonnement, ne se limite pas à 2017 : elle se prolonge en réalité au

moins jusqu'en février 2018 (en rouge). Elle est, elle aussi, caractérisée par de fortes variations d'intensité. Enfin, durant la dernière phase, l'intensité décline tout en se maintenant à un niveau supérieur à celui précédant 2017 (en violet). Le vocable de 'justice prédictive' demeure présent sur Twitter, au rythme de quelques *tweets* par jour, environ une centaine par mois jusqu'en août 2019 et moins d'une cinquantaine par mois ensuite.

Sur cette plateforme, l'information circule vite, de manière très condensée sur de courtes périodes. Cela suggère de resserrer encore l'échelle temporelle et de retenir comme unité d'analyse, non plus le mois, mais le jour. À l'intérieur des quatre séquences identifiées, des pics d'intensité sont repérables pour lesquels l'analyse à l'échelle journalière a été réalisée, et synthétisée en colonne 3 du tableau présenté un peu plus loin (cf. tableau 2.1). La place manque ici pour en restituer le détail. En revanche, leur accumulation permet de mettre en évidence plusieurs résultats d'ensemble.

1.3.2. Deux modèles à l'origine des pics d'intensité sur Twitter

Concernant l'origine des pics d'intensité, deux modèles se dessinent qui renvoient à deux situations typiques distinctes : la parution d'un énoncé de type médiatique, ensuite relayé par les utilisateurs de Twitter et le commentaire en direct d'un événement en cours sur le réseau social.

1.3.2.1. Le relais d'un énoncé médiatique

Le premier cas correspond au relais d'un énoncé médiatique. Celui-ci est d'abord publié en dehors de la plateforme avant que plusieurs abonnés ne le partagent sur Twitter. Ce renvoi est opéré notamment par l'ajout, au sein du *tweet*, d'un lien vers cet énoncé médiatique (Rieder et Smyrnaio, 2012). L'activité sur Twitter dépend ici « *de l'agenda médiatique* » (Olivesi et Hubé, 2016, p. 9), entendu au sens large, c'est-à-dire dépassant les seuls médias de référence. Si ces derniers bénéficient d'une diffusion de plus grande ampleur, on relève dans le corpus des pics provoqués par des publications très diverses, mêlant de nombreux supports (presse écrite et Web, reportages radiophoniques ou télévisuels, mais aussi blogs). L'agenda médiatique n'est plus sous le contrôle exclusif des *gatekeepers* traditionnels de l'espace public (Cardon, 2019, chap. 3 et 4). Il est également déterminé par les publications d'amateurs éclairés, d'universitaires, de collectifs d'avocats à travers des blogs. Les deux pics d'intensité qui surviennent au cours du mois de juillet 2016 en sont symptomatiques. Le premier (11 juillet) provient de *tweets* relayant la parution d'un article sur le blog juridique *Le Petit Juriste*¹. Le nombre de *tweets* postés augmente car ce billet de blog est largement partagé. Le second (21-25 juillet) s'explique de la même manière par le renvoi vers un article paru sur le site *Le Village de la justice* à propos de l'arrivée de nombreuses *startup* sur le marché des innovations technologiques pour le droit². Ce modèle de construction d'un pic d'intensité se caractérise par sa viralité liée à la rapidité de réaction

¹ Clémentine Anno, « La justice prédictive : nouvel horizon juridique », *Le Petit Juriste*, 11 juillet 2016, consulté le 15 mai 2020, <https://www.lepetitjuriste.fr/justice-predictive-nouvel-horizon-juridique/>

² Laurine Tavitian, « Justice prédictive : où en est-on ? », *village-justice.com*, 21 juillet 2016, consulté le 15 mai 2020, <https://www.village-justice.com/articles/Justice-predictive-est,22683.html>

des utilisateurs. À plusieurs reprises, la hausse d'activité sur Twitter liée au relais d'un article de presse ou de blog est intervenue le jour même de la publication de celui-ci.

1.3.2.2. Le commentaire direct d'un événement en cours

Le second cas typique de hausse d'activité sur Twitter correspond au commentaire spontané et direct d'un événement. Lorsque les usagers de la plateforme assistent à une manifestation sur la 'justice prédictive', de type colloque ou séminaire, ils la commentent instantanément, et quelquefois les organisateurs prévoient des *live-tweets*, pratique consistant à relayer en direct ce qui se passe pendant l'événement. Twitter est alors un « *baromètre de la vie [...] réelle* » (Écormier-Nocca et Louis-Sidois, 2019 : 490). La distance temporelle entre le déroulement de l'événement et son commentaire est considérablement réduite par rapport à d'autres moyens d'information disponibles pour le monde juridique, en particulier les bulletins d'informations spécialisés de la presse écrite. *La Gazette du palais* ou *La Semaine juridique* assurent traditionnellement une fonction d'information sur l'actualité mais pas selon la même périodicité. Les *tweets* qui répondent à ce schéma ne contiennent généralement pas de liens, mais se caractérisent plutôt par l'inscription du *hashtag* relatif à l'événement commenté. C'est le cas du pic d'intensité repéré le 23 mai 2017. Il est le fruit de nombreux commentaires spontanés produits lors ou à propos du colloque sur la justice 'prédictive' organisé ce jour-là à la Cour d'appel de Paris.

1.3.2.3. Agenda médiatique et commentaires d'événements dans les pics de *tweets* sur la 'justice prédictive'

Ces deux schémas ne s'excluent pas, ils peuvent cohabiter concomitamment sur la plateforme et engendrer un seul et même pic. Un unique événement peut aussi provoquer deux pics d'intensité, légèrement espacés dans le temps, qui suivent chacun l'un des deux modèles. Cela apparaît nettement lorsque l'on récapitule les différents types de pics d'intensité rencontrés sur la période, comme le montre la colonne 5 dans le tableau ci-dessous.

Tableau 2.1. La 'justice prédictive' sur Twitter, de 2010 à 2020 : Deux modèles à l'origine des pics d'intensité

Phases	Mois clés	Pics d'intensité	Nombre de tweets	Modèle typique
1 : Janvier 2010 à février 2015, la justice prédictive absente de Twitter	-	-	-	-
2 : Mars 2015 à décembre 2016, des entrées timides et ponctuelles sur la plateforme	Mars 2015 (n=38)	26-31 mars 2015	38	1
	Janvier et février 2016 (n=36 et n=44)	28-29 janvier 2016	33	1
		19-24 février 2016	24	1
	Juillet 2016 (n=61)	11 juillet 2016	7	1
		21-25 juillet 2016	38	1
	Octobre, novembre et décembre 2016 (n=77, n=61 et n=82)	5-11 octobre 2016	31	1
		2-3 novembre 2016	19	1
		7 décembre 2016	12	1
		27-28 décembre 2016	35	1
	3 : Janvier 2017 à Février 2018, un moment d'effervescence	Janvier 2017 (n=204)	9 janvier 2017	16
20 janvier 2017			19	1
27 janvier 2017			32	2
30-31 janvier 2017			29	1 + 2
Février 2017 (n=368)		2 février 2017	23	1
		10-11 février 2017	49	1
		15-16 février 2017	78	1 + 2
		24 février 2017	54	1
Mai 2017 (n=524)		5-6 mai 2017	42	1
		12 mai 2017	25	1
		19 mai 2017	262	1 + 2
		23 mai 2017	19	2
Juillet, août 2017 (n=147 et n=64)		4-5 juillet 2017	29	1 + 2
		31 juillet 2017	40	1
Septembre, octobre et novembre 2017 (n=330, n=281 et n=296)		9-10 septembre 2017	46	1
		17-18 septembre 2017	96	1
		25 septembre 2017	24	2
		10 octobre 2017	26	2
		19-20 octobre 2017	81	2
		2-3 novembre 2017	67	1
	24 novembre 2017	33	1	
Février 2018 (n=385)	12 février 2018	170	2	
4 : Mars 2018 à décembre 2020, une baisse d'intensité	Juin 2018 (n=178)	15 juin 2018	19	2
		18 juin 2018	38	2

Le premier schéma, dans lequel l'activité sur Twitter dépend de l'agenda médiatique, est le seul modèle présent lors de la phase d'entrée du sujet sur la plateforme. La 'justice prédictive' pénètre donc ce réseau social en relayant des contenus médiatiques, spécialisés ou généralistes.

Lors de la phase d'effervescence, les deux modèles cohabitent, parfois au sein d'un même pic d'intensité. De plus, un événement peut engendrer, le jour de son organisation, un pic lié aux commentaires en direct des usagers puis, quelques jours plus tard, en provoquer un second du fait cette fois-ci de publications médiatiques. Les Assises du droit et de la compétitivité 2017 sont caractéristiques de cet agencement. Elles génèrent un premier surcroît de *tweets* le jour où elles se tiennent (le 27 janvier 2017), lié aux commentaires en direct. Un second est mesuré les 30 et 31 janvier, en partie causé par la publication et la diffusion d'un compte-rendu de cette manifestation¹. Quant au pic d'intensité majeure du 19 mai 2017 (262 *tweets* pour cette seule journée, soit le nombre le plus élevé par jour sur toute la décennie étudiée), il regroupe des *tweets* relatifs à la diffusion d'un reportage radiophonique dédié à la 'justice prédictive'², et d'autres concernant un colloque sur le même thème organisé le même jour à l'Université catholique de Lille. Celui-ci provoque de nombreuses réactions des différents participants sur Twitter et fait l'objet d'un *live-tweet*.

Enfin, lorsque se conclut la phase d'effervescence en février 2018, les pics d'intensité ne sont plus induits par des énoncés médiatiques. Seul le second modèle engendre encore des hausses significatives de l'activité relative à la 'justice prédictive' sur Twitter. Le sujet, qui reste d'actualité, continue de susciter des réactions, via les arènes universitaire et politique. Le recul de l'effervescence proprement médiatique semble toutefois caractériser la phase de déclin de l'intensité des échanges sur Twitter. Plus précisément, les deux modèles cohabitent tout au long de la phase d'effervescence, mais on pourrait dire que ce second modèle est déjà plus fréquent dès l'automne 2017, en réalité dès l'annonce publique de l'échec de l'expérimentation de [PreviCompute], le 10 octobre 2017 — moment clef sur lequel reviendra le chapitre 4. Il devient dominant ensuite. Il convient à présent de se pencher sur la nature des discours produits au fil de cette trajectoire.

¹ Stéphanie Pourtau, « La justice prédictive : progrès ou menace ? », *actualitesdudroit.fr*, 30 janvier 2017, consulté le 15 mai 2020, <https://www.actualitesdudroit.fr/browse/vie-des-professions-juridiques-et-judiciaires/magistrat/5198/la-justice-predictive-progres-ou-menace>

² Lorélie Carrive, « Justice prédictive : le "big data" frappe à la porte des tribunaux et des cabinets d'avocats », *France Info*, 19 mai 2017, consulté le 18 mai 2020, https://www.francetvinfo.fr/societe/justice/justice-predictive-le-big-data-frappe-a-la-porte-des-tribunaux-et-des-cabinets-d-avocats_2196972.html

III / La 'justice prédictive', une innovation algorithmique qui pose question

Les statistiques lexicales réalisées via le logiciel Iramuteq¹ dessinent tout d'abord le portrait de la 'justice prédictive' comme la rencontre entre le droit, les innovations technologiques et l'université confirmant les premières conclusions déjà esquissées plus haut (1). L'exploration de sémantique lexicale selon la méthode Reinert² identifie plusieurs registres de discours (9 classes) qui mettent en évidence les espaces, acteurs et enjeux des débats sur la 'justice prédictive' (2). L'analyse des métadonnées liées à chaque *tweet*, croisée avec les classes de discours, permet de décrypter cette controverse, telle qu'elle se développe sur Twitter et à laquelle contribuent activement certains acteurs et arènes (3). Dans l'ensemble, les *tweets* associent la 'justice prédictive' aux dispositifs algorithmiques de traitement et d'analyse des décisions de justice en matière civile. La connotation actuarielle et pénale n'est toutefois pas absente. La controverse se construit aussi autour de cette tension entre types de dispositifs, ce que la comparaison des supports de communication et des locuteurs contribue à objectiver (4).

1.4. La rencontre du droit, des innovations et de l'université

Dans l'analyse des statistiques lexicales des différents corpus, la 'justice prédictive' apparaît au carrefour de trois champs sémantiques principaux : le droit et la justice, les innovations numériques et les activités universitaires.

Le corpus de *tweets* apparaît très tourné vers le monde du droit (842 occurrences du substantif et 348 de l'adjectif *juridique*), comparativement à celui des articles de presse plus orienté vers celui de la justice (le mot *judiciaire* y est parmi les plus fréquents)³.

¹ Les lecteurs et lectrices ont bien compris que sur l'expression et les mots-clefs « justice prédictive », nous disposons de deux corpus, un d'articles de presse (n=122) et un de *tweets* (n=6 261). Nous avons réalisé séparément les analyses de classification puisque dans le cas des articles de presse, la classification est opérée à partir de segments de 40 occurrences et que dans celui des *tweets*, vu la brièveté des messages, c'est chaque *tweet* qui est classé. Faute de pouvoir restituer l'étude détaillée des deux corpus ici, les lignes qui suivent porteront principalement sur l'analyse réalisée à partir des *tweets*, mise en perspective, lorsque cela est pertinent, avec les résultats obtenus sur le corpus de presse. Une comparaison synthétique entre *tweets* et articles de presse, du point de vue de la sémantique lexicale et des univers de discours, est toutefois proposée en fin de partie.

² La classification est généralement réalisée sur des segments de textes de 40 occurrences. C'est ce que nous avons réalisé pour le corpus d'articles de presse. Mais la taille d'un *tweet* (140 puis 280 signes) ne se prête pas à une telle segmentation, c'est pourquoi nous procédons ici à une classification dans laquelle c'est le *tweet* qui est retenu comme unité textuelle à classer. Nous cherchons, en réalisant plusieurs essais, à obtenir des classes suffisamment fines pour que l'interprétation soit intéressante et suffisamment consistante pour être significatives statistiquement, et ce, tout en veillant à ce que le plus grand nombre de *tweets* soient classés. Le meilleur compromis nous semble atteint lorsque nous demandons 62 classes à la fin de la phase descendante. Le nombre de classes finales est alors égal à 9, et 72,58 % des *tweets* sont bien classés. S'il est généralement préconisé de veiller à ce que plus de 80 % de segments, ou textes, soient classés, cela semble difficile à atteindre ici. La difficulté sera contournée en travaillant ensuite les données sous NVivo ; pour être assuré d'avoir bien pris en compte la totalité du corpus de *tweets* et pas seulement les 72,6 % classés par Iramuteq.

³ Même en considérant que le terme « *justice* » est logiquement surreprésenté puisque comme celui de « *prédictive* », il a servi à la constitution du corpus. En réalité, outre ces deux termes, ce sont les différentes écritures combinant ces deux mots, telles que « *justiceprédictive* » et « *justiceprédicitive* » qui ont été retenues. Le terme « *prédictive* » est ici lemmatisé sous la forme « *prédicatif* ».

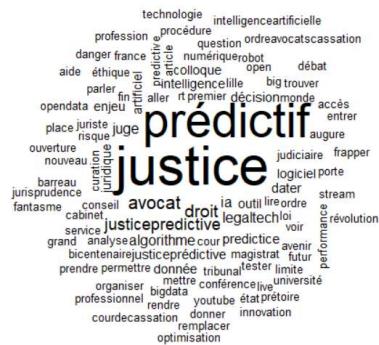
Tableau 2.2. Les 20 termes les plus fréquents au sein du corpus de tweets 2010-20

Forme	Freq. ↓
justice	5995
prédictif	5092
avocat	939
droit	842
justiceprédictive	771
algorithme	540
legaltech	535
ia	514
juge	449
justiceprédictive	404
juridique	348
predictive	340
dater	337
décision	326
outil	302
intelligence	289
artificiel	277
donnée	242
logiciel	224
colloque	212

Tableau 2.3. Les 20 termes les plus fréquents au sein du corpus de presse 2010-20

Forme	Freq. ↓
justice	601
droit	461
avocat	420
juge	307
décision	290
loi	280
judiciaire	233
aller	220
état	215
risque	197
algorithme	193
donnée	187
permettre	179
mettre	170
france	168
outil	166
premier	150
mesure	149
juridique	146
magistrat	146

Graphique 2.12. Nuage des 100 mots les plus fréquents du corpus de tweets 2010-20



Graphique 2.13. Nuage des 100 mots les plus fréquents du corpus de presse 2010-20



Dans les deux cas, les *avocats* sont particulièrement centraux, beaucoup plus que les *juges*, même si la différence est moins marquée dans la presse. Sur Twitter, la ‘justice prédictive’ implique des discours beaucoup plus centrés sur les avocats que sur les juges, et ce, même si l’on prend en compte les occurrences des formes magistrat (154) et magistrature (5 en cumulant les différentes orthographe). À l’intérieur de cette dernière catégorie, les parquets sont marginaux (*procureur*, 9 occurrences ; *procureur général*, 3) ce qui indique soit que le débat se focalise sur les enjeux en matière civile soit qu’il exclut la part considérable des décisions effectuées par ces magistrats qui depuis les vingt dernières années ont acquis un rôle de plus en plus déterminant dans la définition des peines (Bossan et Leturmy, 2019 ; Mouhanna et Bastard, 2010 : 49 ; Vigour, 2018 : chap.5). Dans les nuages de mots, le champ lexical propre au monde du droit et de la justice s’étend avec la présence des termes *procédure*, *tribunal*, *juriste*, *barreau*, *cabinet* ou encore *Ordre des avocats*.

Outre le champ sémantique du droit, celui des innovations technologiques est également très présent avec les termes *legaltech* (535), *algorithme* (540) et *intelligence*

artificielle (289+277 = 566) ou encore *ia* (514 auxquels il faudrait ajouter les 78 occurrences de *intelligenceartificielle* qui n'apparaissent pas dans le tableau). Cela est d'autant plus remarquable que ce vocabulaire est loin d'être aussi présent dans le corpus de presse. Le mot *legaltech* est par exemple absent des 100 formes actives les plus fréquentes et n'y compte que quelques occurrences. Sans grande surprise, la dimension technologique s'avère donc beaucoup plus marquée au sein du corpus de *tweets*. *A contrario*, la *donnée* ou *data*¹ (mais aussi *bigdata* et *opendata*) reste un enjeu central du fonctionnement des innovations algorithmiques, quels que soient les corpus. Elle est présente dans le corpus de presse général « prédictif.ve » (forme active la plus présente avec 2 501 occurrences) et le sous-corpus 'justice prédictive' (donnée = 187 occurrences) ainsi que dans les *tweets* (dater pour data = 317 occurrences).

La dimension universitaire, fréquemment présente sur Twitter au travers des mots *colloque* (20^e forme la plus présente avec 212 occurrences), *université*, *conférence* est presque introuvable dans la presse (où cette forme est absente des 100 formes actives les plus fréquentes et n'y compte que quelques occurrences). Cela s'explique par ce qui a été établi *supra* à propos de la construction des pics d'intensité sur Twitter et des répercussions que des manifestations scientifiques comme les colloques peuvent provoquer via des *live-tweets* et des commentaires en différé. Le terme *outil* est également très courant dans le corpus de *tweets* (15^e forme la plus citée avec 302 occurrences) comme dans la presse (16^e forme la plus citée avec 166 occurrences). Il renvoie au discours sur les fonctions et les usages des innovations de 'justice prédictive' et met l'accent sur leur dimension instrumentale. On peut imaginer que cette manière de qualifier les dispositifs algorithmiques est affirmée, mais aussi discutée : sont-ils seulement des outils ? Que permettent-ils de faire ? Les transformations en cours et leurs enjeux pour l'avenir occupent l'arrière-plan de la réflexion. Les mots *futur*, *remplacer*, *révolution*, *avenir*, *augure* témoignent d'un univers orienté vers la prospective, répondant à la question « de quoi l'avenir sera-t-il fait ? ».

Enfin, du côté de la presse, la dimension institutionnelle (*loi*, *droit*, *État*, *France*) semble davantage valorisée que dans les *tweets*. Le mot *risque* en est aussi plus structurant. Inversement, certains termes très fréquents dans le corpus de presse tels que *police*, *crime*, *criminel*, *pénal*, *attentat*, *terrorisme* sont absents des *tweets*. La dimension civile est donc *a priori* plus marquée dans les débats sur Twitter ; le pénal, davantage dans les articles de presse, ce qui confirme le double ancrage civil/pénal de la 'justice prédictive'.

Ces statistiques rapides livrent un premier aperçu des différents univers lexicaux mobilisés au sujet de la 'justice prédictive' : sur Twitter, le discours associe principalement le monde du droit, à travers les avocats, celui des technologies et innovations numériques portées par des *startup* comme [PréviCompute] et celui de l'activité académique. Autre résultat qui sera confirmé plus loin et sur lequel la fin du chapitre reviendra (3.4.) : les usages de l'expression 'justice prédictive' renvoient surtout aux dispositifs algorithmiques d'analyse

¹ Le terme « *data* » a été lemmatisé sous la forme « *dater* ». Les 317 occurrences du mot correspondent à cette situation.

des décisions de justice civiles tandis que les articles de presse laissent une place plus grande aux dispositifs utilisables en matière pénale.

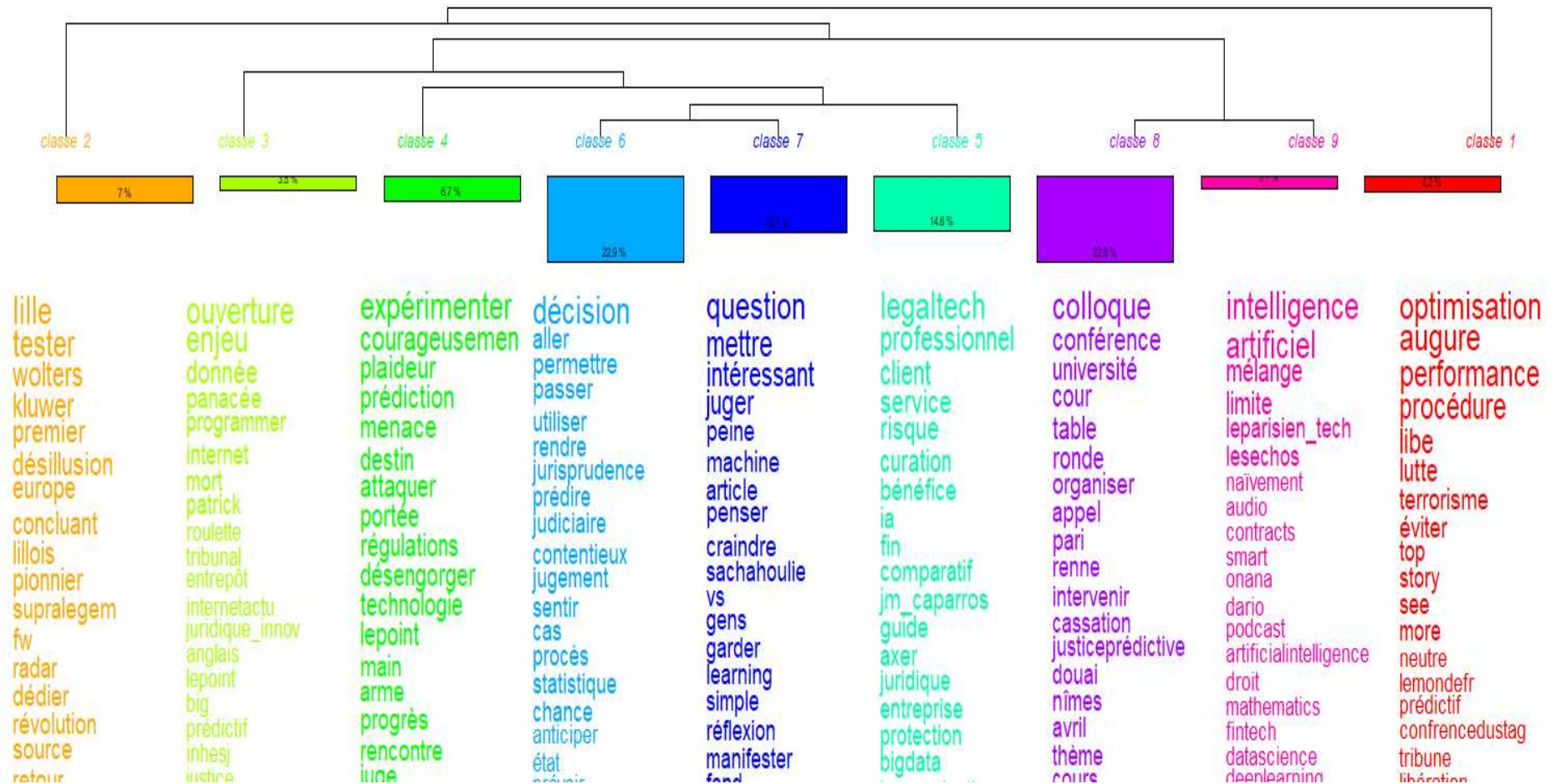
1.5. La controverse sur les algorithmes d'analyse des décisions de justice civiles

L'analyse de classification des *tweets* identifie une structuration du discours en neuf classes dont l'architecture globale sera d'abord présentée (1). Sur Twitter, la 'justice prédictive' est directement associée à l'expérimentation [PréviCompute] par les magistrats des cours d'appel de Douai et Rennes (2). Si l'effervescence discursive est alimentée par la réflexion et les activités universitaires (classe 8 avec presque 23 % des *tweets*) (3), elle prend surtout la forme d'un débat autour des potentialités et limites de la 'justice prédictive' (presque 40 % des *tweets* classés) (4). La controverse est toutefois en grande partie le reflet du discours et de l'actualité médiatiques (un quart des *tweets* classés) (5). Enfin, l'activité discursive sur la 'justice prédictive' rattache celle-ci à un ensemble de transformations, activement orchestrées par les acteurs de la *legaltech*, ce dont témoigne la classe 5, qui représente environ 14 % des *tweets* classés (6).

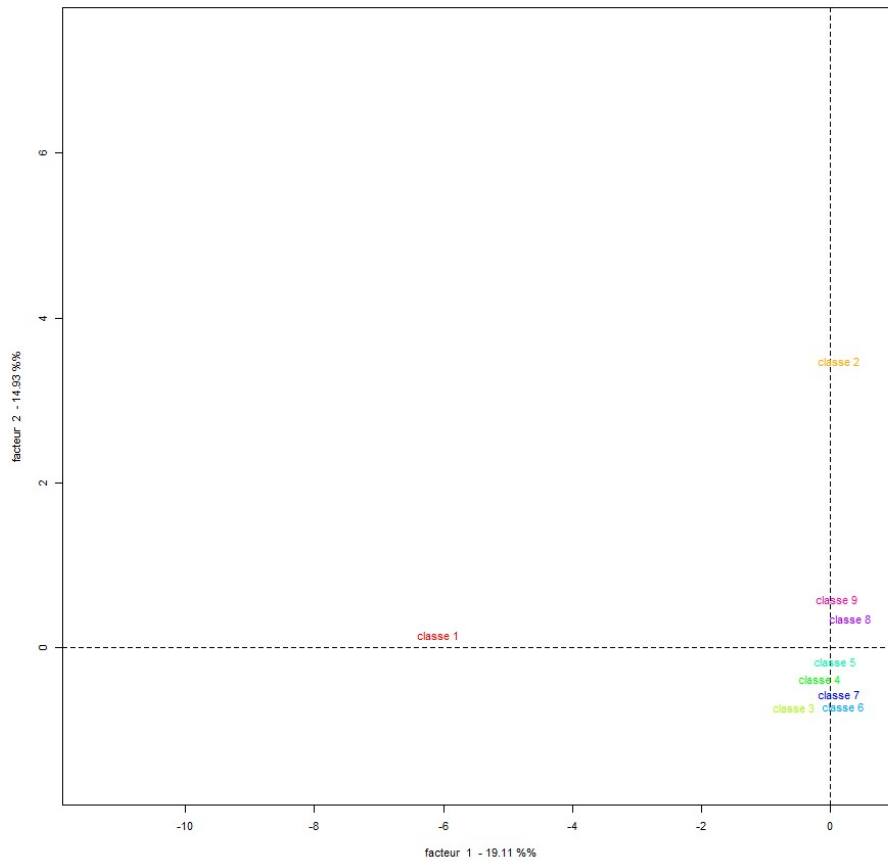
1.5.1. Structuration générale des discours

Neuf classes de poids très inégal apparaissent à l'issue de la classification puisque cinq d'entre elles regroupent moins de 7 % des *tweets* chacune (classes 1, 2, 3, 4 et 9) et totalisent environ 25 % de l'ensemble, tandis que deux autres représentent à elles seules 45 % (classes 6 et 8) ; 30 % pour les deux dernières (classes 5 et 7). L'analyse des bipartitions successives sur le dendrogramme ci-dessous et sur la projection factorielle des classes identifiées (AFC) permet de visualiser les situations respectives et de remarquer que les classes 3 à 9, sans être imbriquées, sont proches entre elles, à la différence des classes 1 et 2.

Graphique 2.14. Dendrogramme représentant le résultat de la classification de Reinert



Graphique 2.15. Projection factorielle des classes identifiées. Corpus 2010-2020



La classe 1, très réduite, trouve sa cohérence dans la référence à trois articles médiatiques. Quant à la classe 2, elle fait sens autour de l'expérimentation de [PréviCompute].

1.5.2. La 'justice prédictive' à l'épreuve : l'expérimentation [PréviCompute]

L'expérimentation [PréviCompute] est saisie, dans les discours médiatiques, comme une épreuve pour la 'justice prédictive'. C'est le sens des classes de discours 2 et 4, toutes les deux constituées de *tweets* relayant des articles de presse.

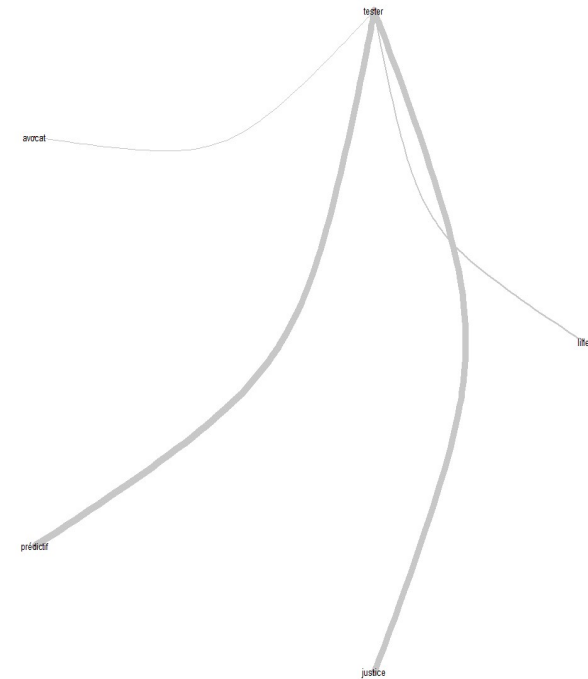
Graphique 2.16. Les mots les plus significatifs de la classe 2



Graphique 2.17. Nuage de mots représentant les 100 mots les plus significatifs de la classe 2



Graphique 2.18. Analyse de graphe du mot tester



La première porte sur le test de [PréviCompute] (racheté par le géant de l'édition juridique qu'est [Brandt]) par des *avocats* du barreau de Lille, comme le montrent le nuage de mots et les *tweets* les plus caractéristiques. Dans cette classe, certains discours liés à l'évaluation de l'expérimentation du même logiciel dans la cour d'appel locale (Douai) semblent également se mêler (*désillusion, concluant*).

Encadré 2.4. Les 5 tweets les plus caractéristiques de la classe 2¹

« Lille est le **premier barreau à tester la justice prédictive en Europe #avocat** »
 « Lille est le **premier barreau à tester la justice prédictive en Europe @annonce_legales #avocat** »
 « Droit : **Pionniers en Europe, les avocats lillois testent la justice prédictive #Lille** »
 « Lille est le **premier barreau à tester la justice prédictive en Europe** »
 « #curation #legaltech Lille est le **premier barreau à tester la justice prédictive en Europe** »

La classe 4 porte plus directement sur l'expérimentation par les magistrats, laquelle est le point de départ d'une réflexion sur ses effets, positifs ou négatifs, comme le montrent les titres des articles relayés : « Les technologies de justice prédictive sont-elles une menace pour la société ? », « La justice prédictive, progrès ou menace ? », ou encore « La justice prédictive, l'outil qui donne des armes aux plaideurs ». Cela est confirmé par les *tweets* les plus caractéristiques.

Encadré 2.5. Les 5 tweets les plus caractéristiques de la classe 4

« #IA #Droit #Magistrat #Justicepredictive La **France bannit les IA de prédiction des comportements des juges. Les technologies de justice prédictive sont-elles une menace pour la société ?** via @developpez »
 « La **France bannit les IA de prédiction des comportements des juges. Les technologies de justice prédictive sont-elles une menace pour la société ?** via @developpez » (x4)
 « La **France bannit les IA de prédiction des comportements des juges. Les technologies de justice prédictive sont-elles une menace pour la société ?** » (x7)
 « #curation #legaltech La **France bannit les IA de prédiction des comportements des juges. Les technologies de justice prédictive sont-elles une menace pour la société ?** »
 « Les **juges expérimentent courageusement la justice prédictive** » (x23)

Ces deux classes peuvent être considérées comme principalement médiatiques au sens où elles relaient massivement des articles de presse, que le travail concomitant sur les différents corpus a souvent permis de repérer — sans qu'une recherche de rapprochement systématiquement ait été entreprise.

¹ Nous retranscrivons ici les *tweets* sous leur forme au moment de l'introduction des données sous Iramuteq, c'est-à-dire après suppression des liens hypertextes contenus en leur sein. Les mots en gras et en rouge correspondent aux mots significatifs de la classe.

Tableau 2.4. Récapitulatif des énoncés médiatiques dont les mots sont présents dans les mots les significatifs des classes 2 et 4

Classe 2 : articles médiatiques	Classe 4 : articles médiatiques
<ol style="list-style-type: none"> 1. « Lille est le premier barreau à tester la justice prédictive¹ » 2. « Pionniers en Europe, les avocats lillois testent “la justice prédictive” »² 	<ol style="list-style-type: none"> 1. « Les juges expérimentent courageusement la justice prédictive » 2. « Justice prédictive : l’outil qui donne des armes aux plaideurs³ » 3. « La justice prédictive, progrès ou menace ? » 4. « La France bannit les IA de prédiction des comportements des juges. Les technologies de justice prédictive sont-elles une menace pour la société ? »

1.5.3. L’effervescence universitaire autour de la ‘justice prédictive’

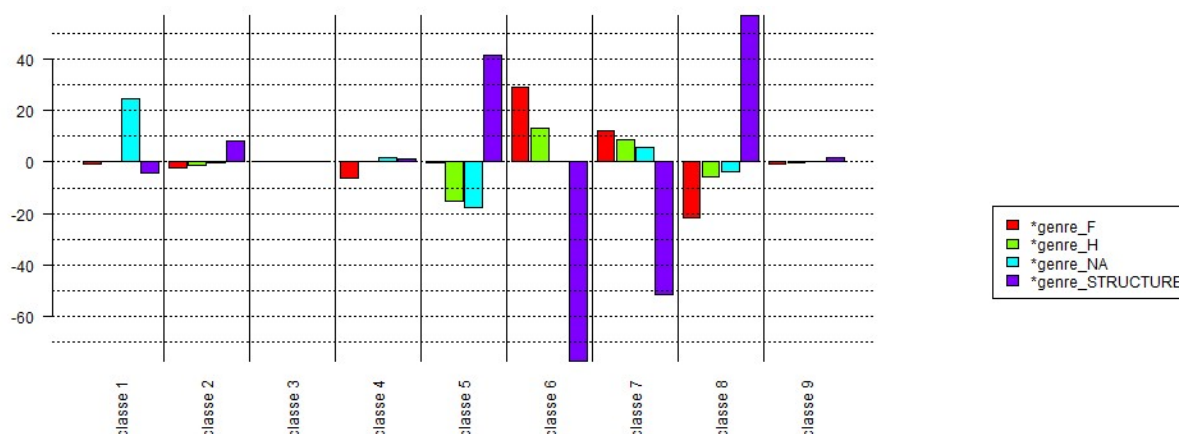
Les *tweets* font référence à la réflexion sur la ‘justice prédictive’ telle qu’elle est proposée par les universitaires via des formes établies du débat intellectuel : il est question de *colloques, tables rondes, conférences, ateliers* et *séminaires*, ce qui explique la présence de termes associés à ces activités comme *programme, inscription, intervention, intervenant* et bien sûr *université*. Plusieurs noms d’intervenants comme *Antoine Garapon*, secrétaire général de l’Institut des hautes études sur la justice (*Ihej*), mais aussi *Alain Benssoussan*, [*Thomas Vial*] sont repérables. Cette classe de discours recouvre une partie très importante du corpus étudié à savoir presque un quart des *tweets* classés (22,8 % soit 1037 *tweets*), ce qui témoigne d’une véritable effervescence du monde juridique des praticiens et universitaires autour de cette question.

¹ [Victor Chevalet], « Lille est le premier barreau à tester la justice prédictive », LinkedIn, 24 janvier 2017, consulté le 26 mai 2020, [https://www.linkedin.com/pulse/lille-est-le-premier-barreau-%C3%A0-tester-la-justice-\[victor-chevalet\]](https://www.linkedin.com/pulse/lille-est-le-premier-barreau-%C3%A0-tester-la-justice-[victor-chevalet])

² Chantal David, « Pionniers en Europe, les avocats lillois testent “la justice prédictive” », *La Voix du Nord*, 20 janvier 2017, consulté le 26 mai 2020, <https://www.lavoixdunord.fr/106005/article/2017-01-20/pionniers-en-europe-les-avocats-lillois-testent-la-justice-predictive?>

³ Laurence Neuer, « Justice prédictive : l’outil qui donne des armes aux plaideurs », *Le Point*, 27 décembre 2016, consulté le 26 mai 2020, https://www.lepoint.fr/editos-du-point/laurence-neuer/justice-predictive-l-outil-qui-donne-des-armes-aux-plaideurs-27-12-2016-2093180_56.php#

Graphique 2.24. Distribution de la variable genre au sein de chacune des classes¹



Les tweets les plus caractéristiques de la classe confirment cette activité d'organisation de manifestations savantes sur la 'justice prédictive', portée par les mondes du droit et de la justice, en particulier universités et juridictions étroitement associées.

Encadré 2.6. Les 5 tweets les plus caractéristiques de la classe 8

« À l'occasion de son bicentenaire, l'ordre des avocats au Conseil d'État et à la Cour de cassation organise en partenariat avec l'Université Paris-Dauphine un colloque sur la justice prédictive, 12 février 2018 »

« J'ai couvert ce vendredi le colloque "la justice prédictive : risques et avenir d'une justice virtuelle" à Laval, organisé par Le Mans Université, le barreau de Laval et la Cour d'appel d'Angers. Un évènement instructif qui a fait salle comble. Photos en ligne très bientôt »

« La justice prédictive. Actes du colloque du 12/02/2018 de l'Ordre des avocats au Conseil d'État et à la Cour de cassation en partenariat avec l'Université Paris-Dauphine #psl #algo #avocat #bigdata #digital #droit #ia #juge #justicepredictive #opendata »

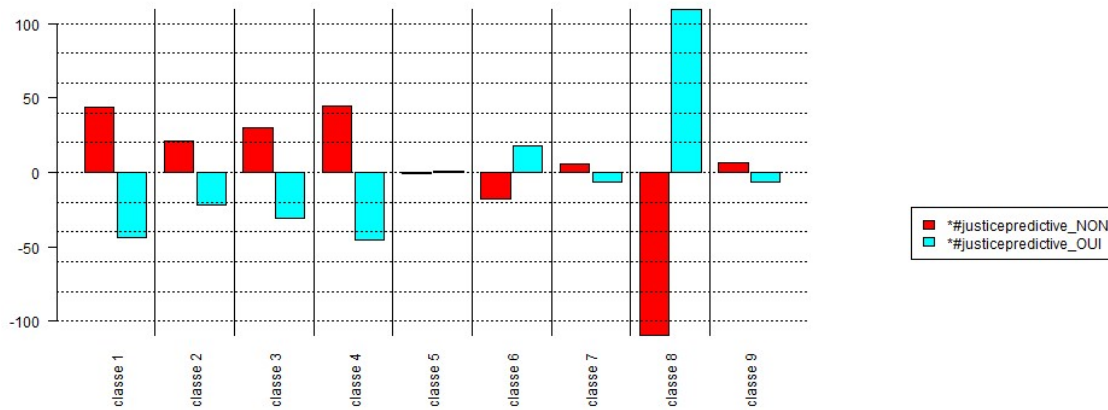
« On parle au colloque de la Cour de cassation de l'expérimentation par les Cours d'Appel de Rennes et de Douai d'un très médiatique logiciel prétendument de justice prédictive #opendata »

« Co-organisé avec la Cour d'Appel de Douai et le barreau de Lille, colloque sur la justice prédictive le 19 mai »

L'analyse de la distribution du hashtag #justicepredictive montre que celui-ci est très significativement surreprésenté au sein de cette classe de discours. Il est vraisemblable qu'il soit utilisé pour annoncer des colloques sur ce sujet et confirme que, dans l'espace académique, il devient agrégateur.

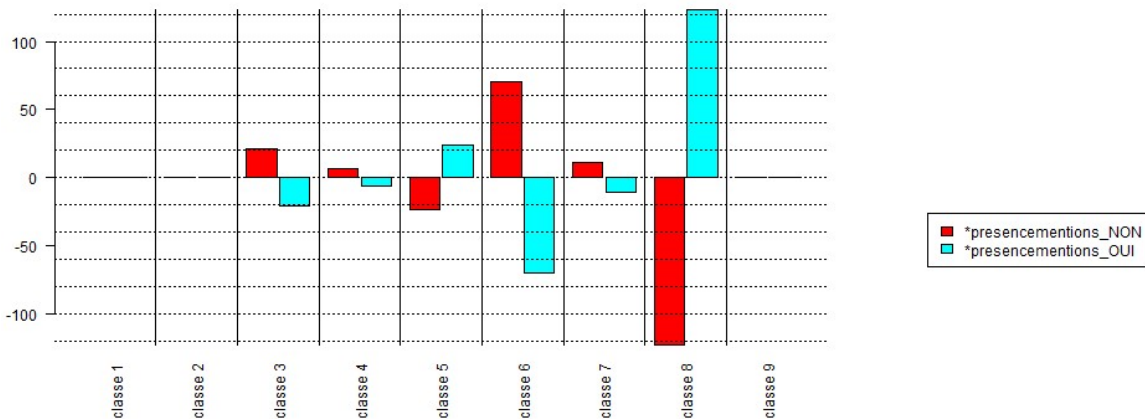
¹ N. B. : les résultats des classes 3, 4, 9 ne sont pas statistiquement significatifs.

Graphique 2.25. Distribution de la variable #justicepredictive au sein de chacune des classes



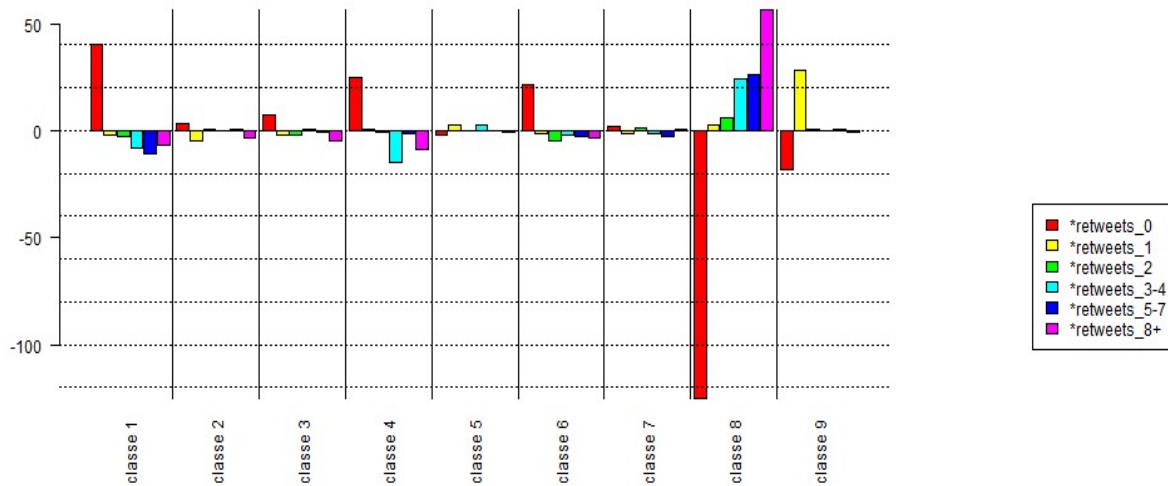
Les *tweets* qui comportent une mention sont surreprésentés dans cette classe, ce qui est logique puisque c'est le moyen de citer un intervenant.

Graphique 2.26. Distribution de la variable presencementions au sein de chacune des classes



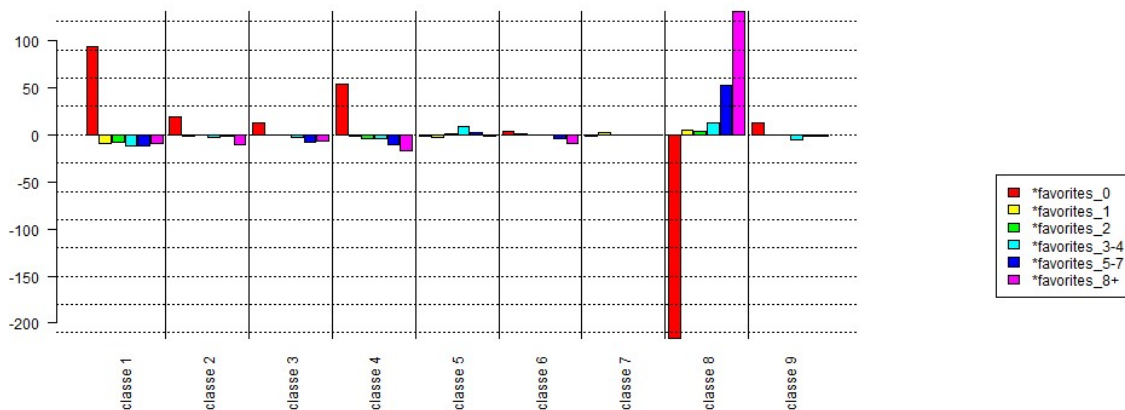
C'est la classe qui génère le plus de *retweets* : l'annonce et le commentaire de l'activité de type universitaire suscitent des réactions visant à faire circuler le plus largement possible l'information sur les manifestations organisées.

Graphique 2.27. Distribution de la variable retweets au sein de chacune des classes



C'est aussi la classe qui génère le plus de *likes*, forme numérique privilégiée de manifestation publique d'une appréciation positive du contenu des *tweets*. Dans le graphique ci-dessous, la modalité « plus de 8 likes » de la variable est nettement surreprésentée.

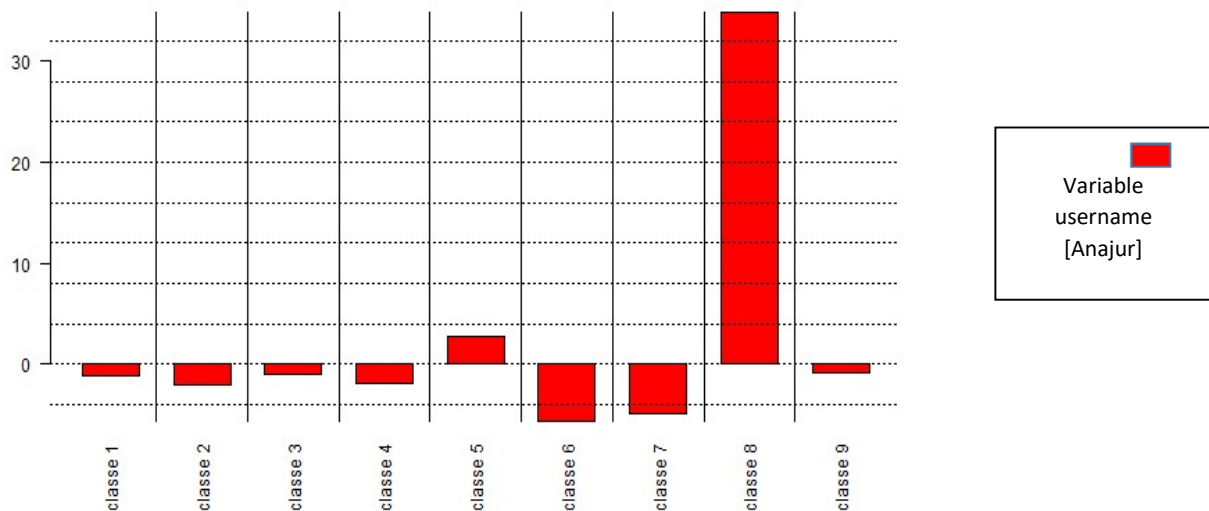
Graphique 2.28. Distribution de la variable favorites au sein de chacune des classes



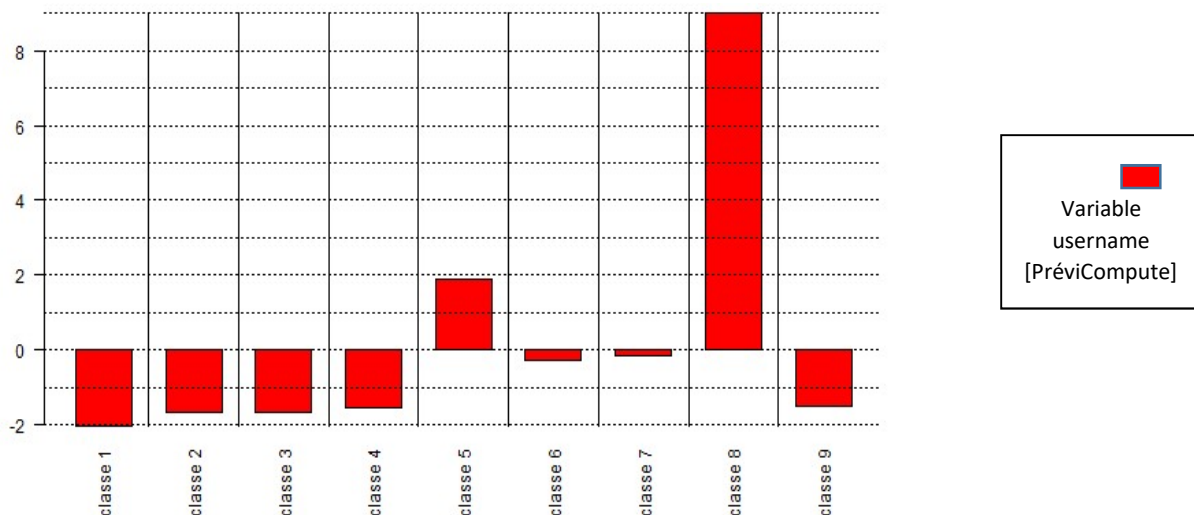
Enfin, parmi les *tweets* postés par des comptes de structures, ceux de [PréviCompute] et [Anajur] sont surreprésentés au sein de cette classe, ce qui suggère que ces acteurs promeuvent les événements universitaires liés à la 'justice prédictive'¹.

¹ Les autres résultats ne sont pas statistiquement significatifs.

Graphique 2.29. La distribution des tweets émis par [Anajur] au sein des classes



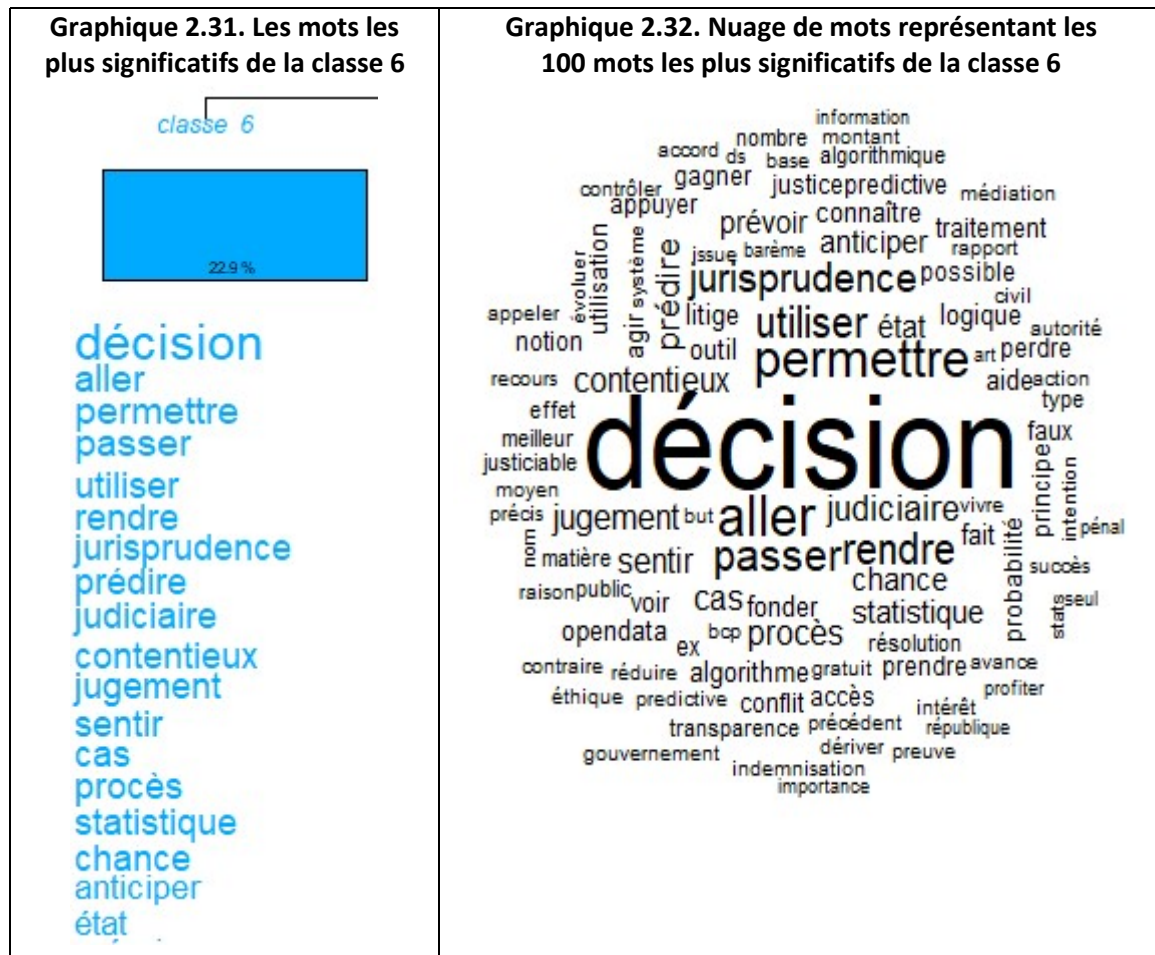
Graphique 2.30. La distribution des tweets émis par [PréviCompute] au sein des classes



1.5.4. La 'justice prédictive' en débat : potentialités au civil et limites au pénal

Une autre partie très importante des échanges de *tweets* qui mentionnent la 'justice prédictive' s'inscrit dans le registre de la controverse sur les potentialités et limites de celle-ci, comme le montrent deux classes (6 et 7) qui représentent ensemble quasiment deux cinquièmes des *tweets* classés (38 %). Le discours n'est pas le reflet d'articles de presse ou de blogs, mais l'expression d'un questionnement sur la 'justice prédictive' par les auteurs et autrices des messages postés. Ces deux classes fonctionnent doublement en miroir. La première (classe 6) est plus centrée sur les dispositifs algorithmiques en matière civile et évalue plutôt positivement l'aide à la décision qu'ils peuvent apporter. La seconde (classe 7) porte davantage sur les dispositifs actuariels et le versant pénal, envisagés de manière plus critique.

La classe 6 regroupe des discours liés à l'utilité et aux potentialités de la 'justice prédictive'. Elle est perçue comme un dispositif qui *va permettre d'anticiper le jugement*, de *prédire la décision*, de connaître ses *chances* lors d'un *procès*, la probabilité *statistique* pour le justiciable de gagner un *contentieux* à partir de l'étude de la *jurisprudence*. La 'justice prédictive' est vue comme une innovation plutôt profitable, même si cela est assez implicite et si des questionnements, par exemple sur les biais possibles, sont présents dans les messages les plus caractéristiques de la classe.



Encadré 2.7. Les 5 tweets les plus caractéristiques de la classe 6

« *L'open data permettra la mise en place d'une justice prédictive, ce qui devra contribuer à une meilleure transparence et prévisibilité de la justice, renforcera la cohérence de la jurisprudence, rendra la décision judiciaire plus objective en la privant de biais et d'erreurs* »

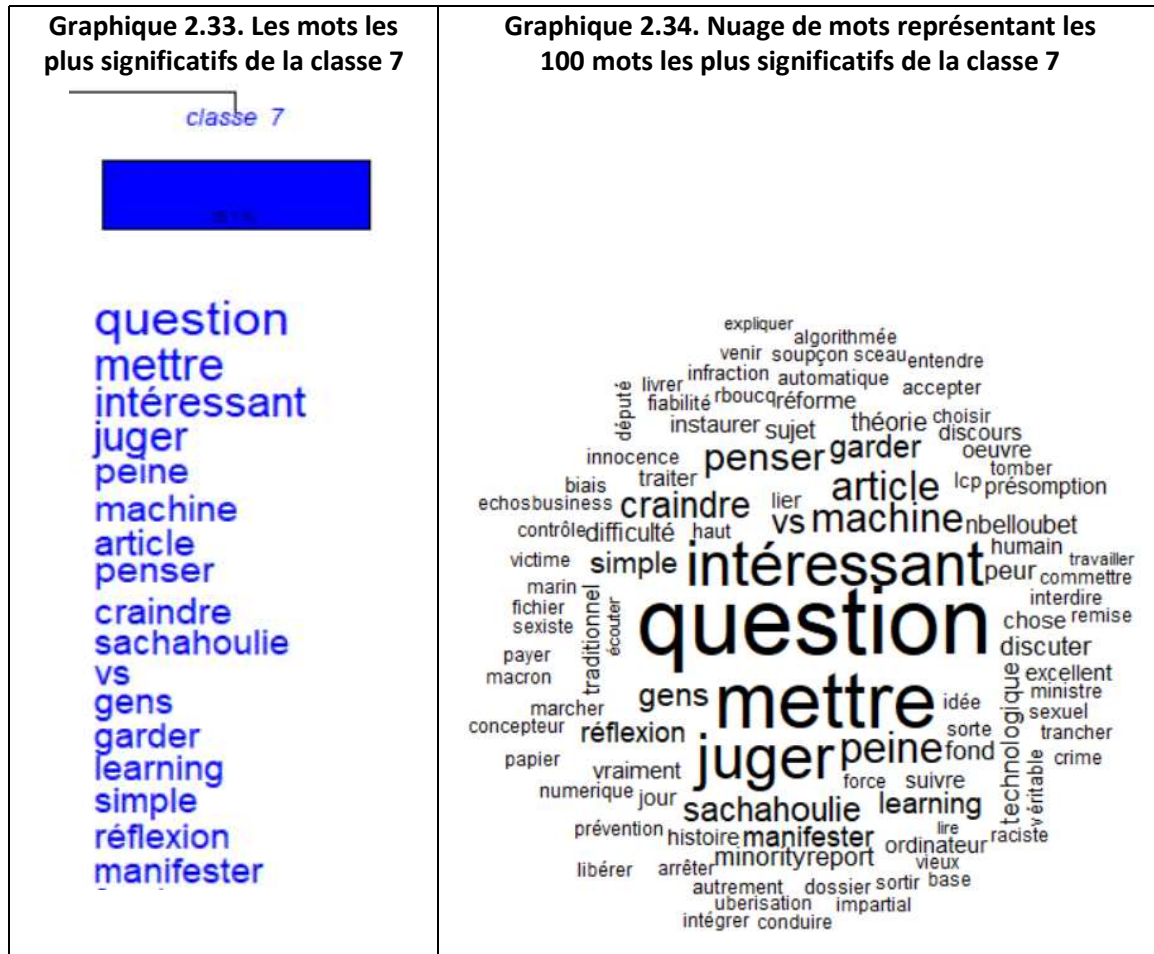
« *#JusticePredictive Le logiciel permet de rechercher les décisions qui vont dans un seul sens* »

« *Et les codes sources des logiciels dits de justice prédictive (en fait des stats sur les décisions de justice)? Vu leur effet prochain prévisible sur les jugements ou les assurances, on va faire comment pour connaître leurs biais ?* »

« @xberne @paulafortea @etalab Et les **codes** sources des **logiciels** dits de justice prédictive (en fait des **stats** sur les **décisions** de justice) ? **Vu** leur **effet** prochain **prévisible** sur les **jugements** ou les assurances, on **va** faire comment pour **connaître** leurs **biais** ? #avocat #avoc »

« Thibault nous **parle** #JusticePredictive. La justice prédictive **va** à **partir** de **bases** de **données** de **décisions** de justice **donner** le % de **chance** de remporter ou non un **procès** »

La classe 7 est centrée sur l'évaluation critique de la 'justice prédictive' que traduisent les termes *question*, *réflexion* et le verbe *penser*.



D'autres verbes (*craindre*, *garder*) expriment une certaine réserve dans la manière dont sont traités les sujets en débat, comme le *machine learning* par exemple. Des thématiques propres au pénal, au risque actuariel comme *présomption d'innocence*, *souçon* émergent du nuage de mots et des *tweets* les plus représentatifs. Ces derniers mettent en scène le risque que le film *Minority Report* de Steven Spielberg devienne réalité avec l'anticipation de la commission des actes délinquants. L'enjeu de l'individualisation des peines est également mentionné. Cette septième classe regroupe donc des discours qui abordent différents points de débats qui animent la controverse relative à la 'justice prédictive'. Les propos sont nuancés, formulés davantage comme des interrogations que comme des opinions tranchées, mais marquent une certaine réserve, plus prononcée que dans les autres classes.

Encadré 2.8. Les 5 tweets les plus caractéristiques de la classe 7

« La **justice** face à la révolution open data : ouverts à cette **évolution technologique** de **fond**, les **magistrats craignent** une **remise en question** de l'**individualisation** des **peines** et **mettent en garde** contre le fantasme d'une **justice prédictive** #podcast »

« @kanikpascal @gmutemuni @kiskia4 @avoatscd **Belle idée**, c'est déjà le cas aux **USA** avec les technologies de **justice prédictive**, la France s'y **met** peu à peu, en Afrique on y encore loin. **Lire** cet **article intéressant** sur la **question** »

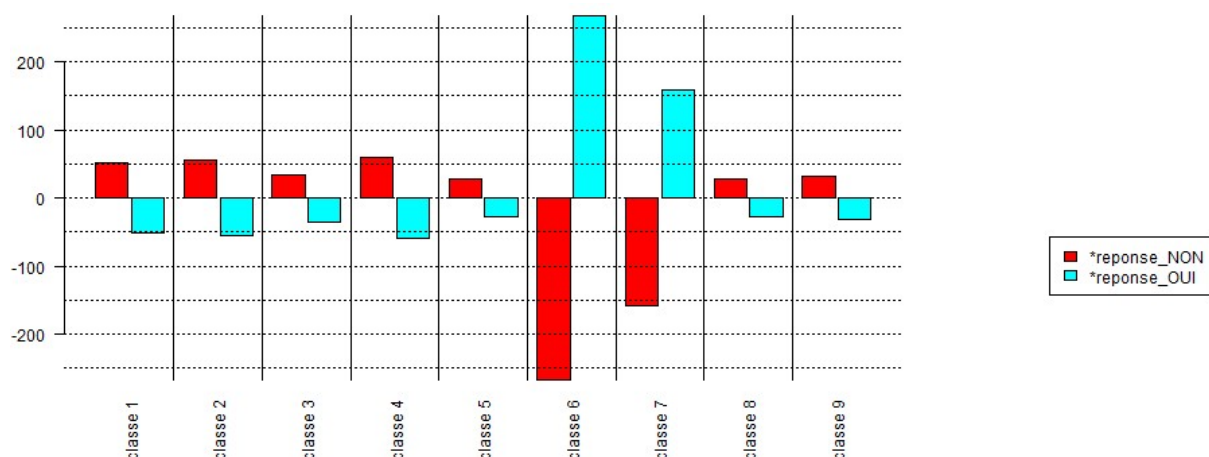
« Nicole **Belloubet** a globalement **soutenu** la **proposition** tout en **mettant** en avant un impérieux **besoin** d'équilibre, pas **question** d'**instaurer** une quelconque **forme** de **justice prédictive** sur la **base** d'un **simple soupçon** »

« Pour une **mise en perspective** de la **question** de la **justice prédictive** avec la **réforme** de la **motivation** des décisions, petite auto promo pour mon **article** de 2018 au @jcp_g #legaltech »

« Je **viens d'apprendre** qu'il **existe** une **justice prédictive** en France, **instauré** par le **code pénal**. Selon la disposition en **question**, il est prévu de **juger** par pour un **acte**, mais le **soupçon** qu'un **acte** pourrait être **commis** @oxotadji »

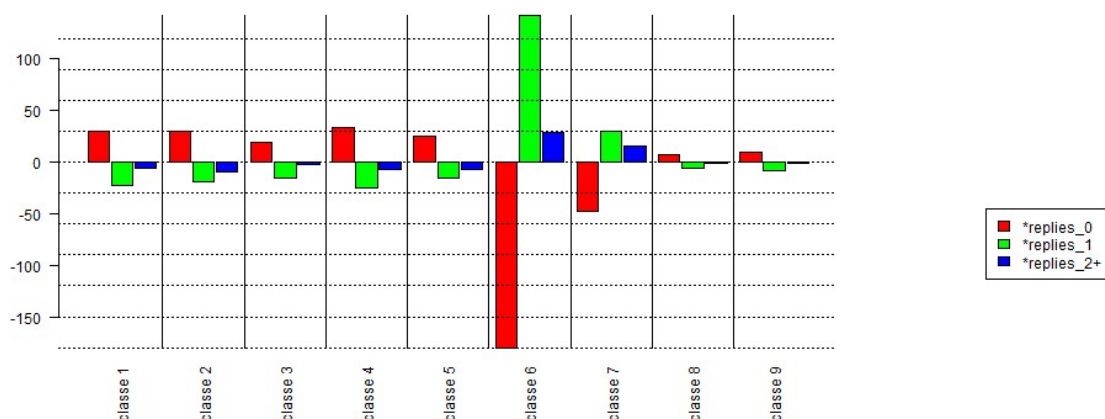
Ces deux classes de discours partagent la forme débat entre utilisateurs de la plateforme puisqu'elles comportent, plus que d'autres, des chaînes de *tweets* (des messages qui se répondent). Elles se caractérisent par un mode conversationnel : les usagers se répondent entre eux pour débattre et échanger des opinions. Dans le graphique ci-dessous, les *tweets* qui sont des réponses à d'autres sont surreprésentés de manière significative au sein des classes 6 et 7.

Graphique 2.35. Distribution de la variable réponse au sein de chacune des classes



De plus, les *tweets* qui engendrent le plus de réponses (au moins une et plus de deux) sont également surreprésentés de manière significative au sein de ces classes.

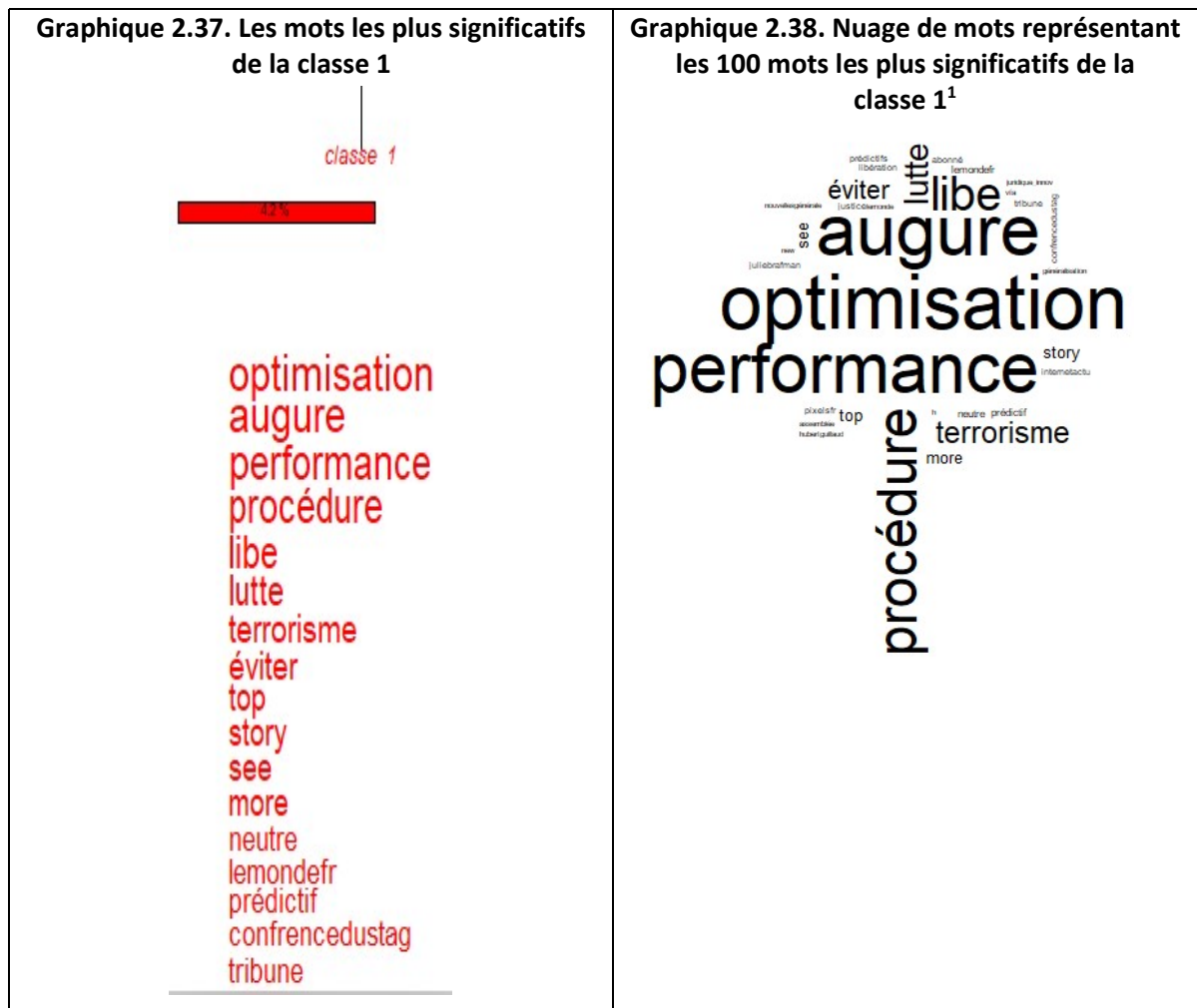
Graphique 2.36. Distribution de la variable replies au sein de chacune des classes



1.5.5. La 'justice prédictive', objet du discours médiatique relayé sur Twitter

Si la 'justice prédictive' est l'objet de débat sur Twitter, c'est aussi parce que les utilisateurs de la plateforme relaient le débat qui existe dans l'arène médiatique classique ou sur d'autres réseaux sociaux (comme LinkedIn). Les classes 2 et 4 portant sur l'expérimentation (cf. *supra*) sont constituées d'articles de presse et de blogs. C'est également le cas des classes 1, 3 et 9 dont il va être question maintenant. Elles correspondent à un faible nombre de *tweets* (un peu plus de 10 % de l'ensemble) et portent sur quelques articles (nous en avons identifié trois, cf. encadré). En outre, sur le plan statistique, peu de formes sont caractéristiques de ces classes, d'où des nuages de mots moins riches. Tout ceci converge pour empêcher d'aller très loin dans l'interprétation.

La première classe incarne le double sens, civil et pénal, de la 'justice prédictive'. Elle porte d'un côté sur les dispositifs de traitement et d'analyse des jugements et arrêts. Leur rôle dans l'*optimisation* et la *performance* dans l'aide à la décision en matière civile est discuté. D'un autre côté, elle traite des dispositifs de justice actuarielle dans le contexte de la *lutte* contre le *terrorisme* et du pénal.



Encadré 2.9. Les articles médiatiques de la classe 1

1. « Justice prédictive, l’augure des procédures ?² »
2. « La justice prédictive : entre performance et optimisation³ »
3. « Dans la lutte contre le terrorisme, évitons une justice prédictive⁴ »

¹ Remarquons que cette classe de discours compte moins de 100 termes statistiquement significatifs.

² Julie Brafman, « Justice prédictive, l’augure des procédures », *Libération*, 23 février 2017, consulté le 18 mai 2020, https://www.liberation.fr/france/2017/02/23/justice-predictive-l-augure-des-procedures_1550628

³ Hubert Guillaud, « La justice prédictive (3/3) : entre performance et optimisation », *lemonde.fr/blog/internetactu*, 17 septembre 2017, consulté le 18 mai 2020, <https://www.lemonde.fr/blog/internetactu/2017/09/17/la-justice-predictive-33-entre-performance-et-optimisation/>

⁴ Collectif, « Dans la lutte contre le terrorisme, “évitons une justice prédictive” », *Le Monde*, 31 juillet 2017, consulté le 18 mai 2020, https://www.lemonde.fr/idees/article/2017/07/31/dans-la-lutte-contre-le-terrorisme-evitons-une-justice-predictive_5166775_3232.html

Encadré 2.10. Les 5 tweets les plus caractéristiques de la classe 1

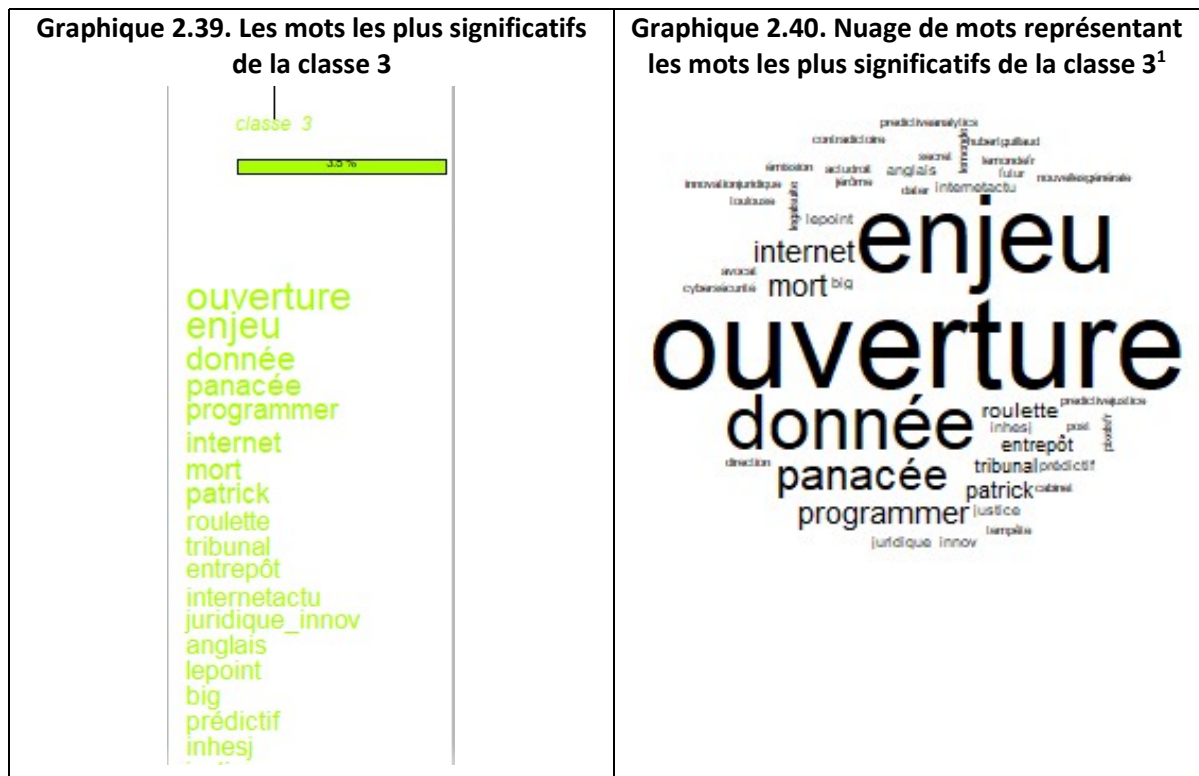
« **Top story** : la **justice prédictive** (3/3), entre **performance** et **optimisation** in **see more** » (x2)
« **Justice Prédictive** : l'**augure** des **procédures**, via **@libe @juliebrafman** »
« **Justice Prédictive** : l'**augure** des **procédures**, **@juliebrafman** via **@libe** »
« **Justice Prédictive** : l'**augure** des **procédures**, via **@libe** » (x2)
« **Justice Prédictive** : l'**augure** des **procédures**, via **@libe** #prédiction #algorithme #bigdata #droit »

Les mots les plus caractéristiques de la troisième classe renvoient à un billet de blog hébergé sur le site internet du journal *Le Monde* signé par Hubert Guillaud, intitulé « Justice prédictive : l'enjeu de l'ouverture des données¹ ». Ce billet revient sur le colloque dédié à la 'justice prédictive' qui s'est déroulé fin juin 2017 et reprend notamment les propos d'un magistrat, conseiller référendaire à la Cour de cassation, adjoint au directeur du service de documentation, des études et du rapport². L'analyse porte sur le mouvement d'*open data* des décisions de justice, les enjeux de définition de la 'justice prédictive' et des appellations concurrentes ainsi que sur les risques et les avantages des dispositifs algorithmiques. Dès sa publication, ce billet a été relayé de manière importante sur Twitter : il a été partagé de manière virale tout au long du mois de septembre 2017. En réalité, il fait partie d'une trilogie : deux autres billets lui succéderont³, qui circulent eux aussi abondamment. Cette classe de discours ne regroupe toutefois que 159 messages, soit 3,5 % des *tweets* classés.

¹ Hubert Guillaud, « La justice prédictive (1/3) : l'enjeu de l'ouverture des données », [lemonde.fr/blog/internetactu](https://www.lemonde.fr/blog/internetactu), 9 septembre 2017, consulté le 18 mai 2020, <https://www.lemonde.fr/blog/internetactu/2017/09/09/la-justice-predictive-13-lenjeu-de-louverture-des-donnees/>

² Information tirée de [Valère Desjardins] (2019), « La justice dite "prédictive" : prérequis, risques et attentes – l'expérience française », *Les Cahiers de la Justice*, n° 2, p. 269-276.

³ Hubert Guillaud, « La justice prédictive (2/3) : prédictions et régulations », [lemonde.fr/blog/internetactu](https://www.lemonde.fr/blog/internetactu), 13 septembre 2017, consulté le 18 mai 2020, <https://www.lemonde.fr/blog/internetactu/2017/09/13/la-justice-predictive-23-predictions-et-regulations/>; Hubert Guillaud, 17 septembre 2017, « La justice prédictive (3/3) : entre performance et optimisation », [lemonde.fr/blog/internetactu](https://www.lemonde.fr/blog/internetactu), consulté le 18 mai 2020, <https://www.lemonde.fr/blog/internetactu/2017/09/17/la-justice-predictive-33-entre-performance-et-optimisation/>



Encadré 2.11. L'article viral au cœur de la classe 3

1. « Justice prédictive : l'enjeu de l'ouverture des données² »

Encadré 2.12. Les 5 tweets les plus caractéristiques de la classe 3

« La **justice prédictive** (1/3) : L'**enjeu** de l'**ouverture** des **données** @internetactu blogs lemonde.fr »

« Rt @pixelsfr : la **justice prédictive** (1/3) : L'**enjeu** de l'**ouverture** des **données** @internetactu »

« La **justice prédictive** (1/3) : L'**enjeu** de l'**ouverture** des **données** par @hubertguillaud @internetactu »

« La **justice prédictive** (1/3) : L'**enjeu** de l'**ouverture** des **données** @internetactu » (x2)

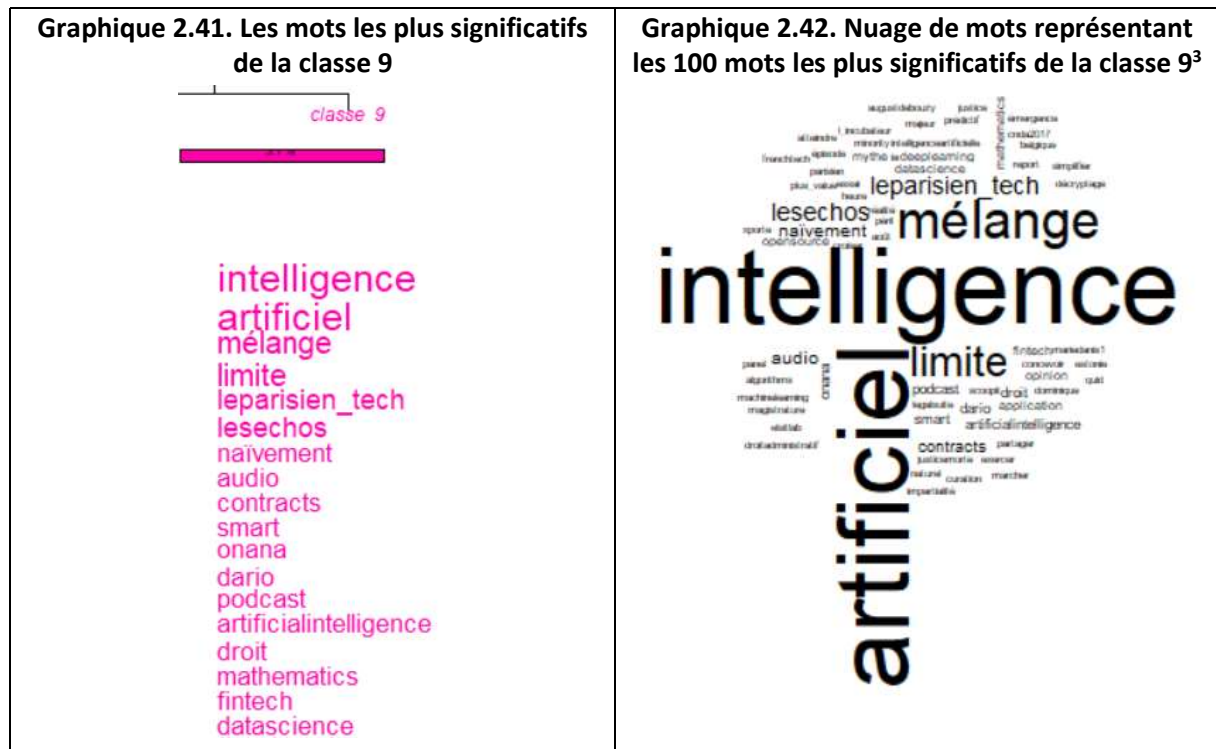
« #curation #legaltech, la **justice prédictive** (1/3) : L'**enjeu** de l'**ouverture** des **données** @internetactu »

Quant à la classe 9, elle regroupe deux articles de presse, qui mettent l'accent sur la 'justice prédictive' comme étant liée à l'intelligence artificielle : « La justice prédictive,

¹ Il y a moins de 100 termes statistiquement significatifs rattachés à cette classe.

² Hubert Guillaud, « La justice prédictive (1/3) : l'enjeu de l'ouverture des données », lemonde.fr/blog/internetactu, 9 septembre 2017, consulté le 18 mai 2020, <https://www.lemonde.fr/blog/internetactu/2017/09/09/la-justice-predictive-13-lenjeu-de-louverture-des-donnees/>

mélange de droit et d'intelligence artificielle¹ » et « Intelligence artificielle : les limites de la justice prédictive² ».



Encadré 2.13. Les 5 tweets les plus caractéristiques de la classe 3

« La **justice prédictive, mélange** de **droit** et d'**intelligence artificielle** @leparisien_tech #legaltech #ia »

« La **justice prédictive, mélange** de **droit** et d'**intelligence artificielle** #IA via @leparisien_tech »

« La **justice prédictive, mélange** de **droit** et d'**intelligence artificielle** via @leparisien_tech » (x7)

« La #justiceprédictive, **mélange** de **droit** et d'**intelligence artificielle** via @leparisien_tech #lawtech #data #IA »

« **Intelligence artificielle** : selon @lesechos, la **justice prédictive montre** ses **limites**. Tant mieux ! @scoopit »

¹ Anonyme, « La justice prédictive, mélange de droit et d'intelligence artificielle », *Le Parisien*, 1^{er} février 2017, consulté le 18 mai 2020, <http://www.leparisien.fr/high-tech/la-justice-predictive-melange-de-droit-et-d-intelligence-artificielle-01-02-2017-6645696.php>

² Philippe Rozec et Louise Thiebaut, « Intelligence artificielle : les limites de la justice prédictive », *Les Échos*, 2 novembre 2017, consulté le 18 mai 2020.

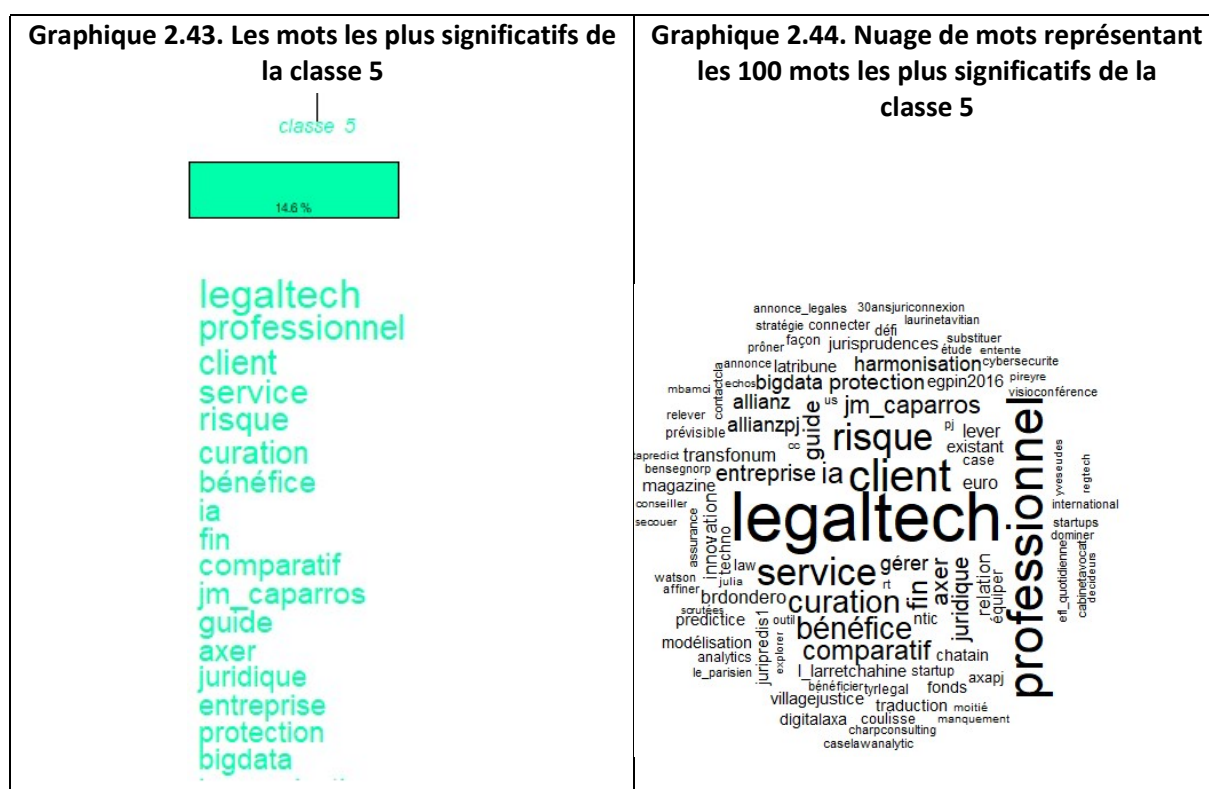
³ Il y a moins de 100 termes statistiquement significatifs rattachés à cette classe.

Encadré 2.14. Les articles de presse de la classe 3

1. « La justice prédictive, mélange de droit et d'intelligence artificielle¹ »
2. « Intelligence artificielle : les limites de la justice prédictive² »

1.5.6. La 'justice prédictive', révolution numérique portée par les legaltechs et enjeux commerciaux

La 'justice prédictive' apparaît dans de très nombreux tweets comme une innovation qui s'inscrit dans la révolution numérique du droit, portée par les acteurs de la *legaltech*, au premier rang desquels les *startups* et le réseau sur lequel elle s'appuie, engagée dans un monde marchand et un rapport à la clientèle. C'est le sens de la classe 5 qui représente à elle seule 664 tweets, soit 14,61 % de l'ensemble. Elle est marquée par l'abondance de mots, souvent des anglicismes, liés à l'innovation, aux technologies numériques et algorithmiques et au travail de la donnée avec les termes *legaltech*, *ia*, *bigdata*, *curation*, *startup*, *NTIC*. Est également présente une dimension économique (*service*, *client*, *jm_caparros*, *bénéfice*, *axer*, *entreprise*, *euro*).



Dans le nuage de mots figurent également les acteurs de cette « révolution numérique » : *Village de la justice*, *Juriconnexion*, *[Anajur]*, *[PréviCompute]*, *[Prédidroit]*,

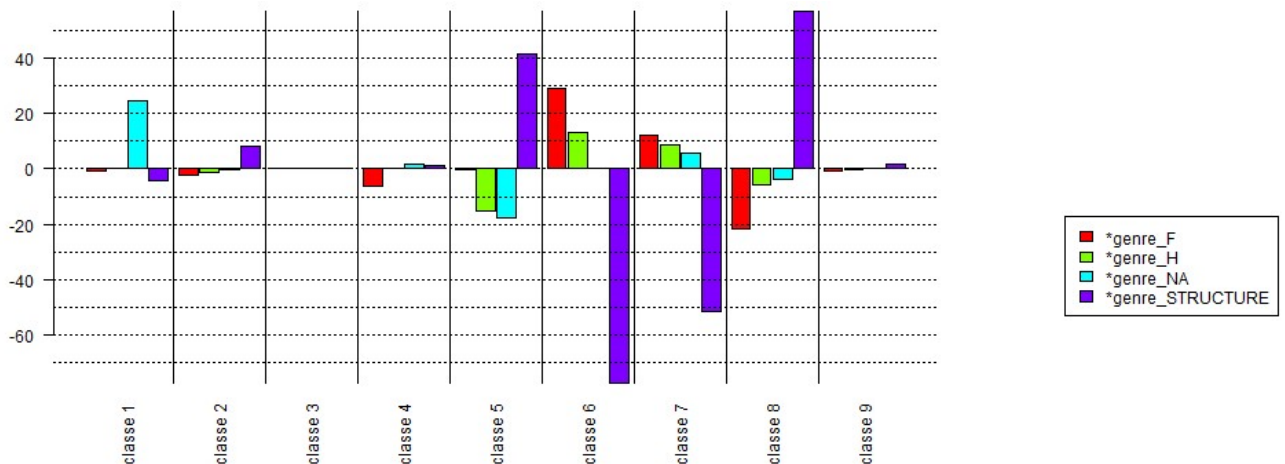
¹ Anonyme, « La justice prédictive, mélange de droit et d'intelligence artificielle », *Le Parisien*, 1 février 2017, consulté le 26 mai 2020.

² Philippe Rozec et Louise Thiebaut, « Intelligence artificielle : les limites de la justice prédictive », *Les Échos*, 2 novembre 2017, consulté le 26 mai 2020.

[Juri+], [Victor Chevaleret]¹, ou les compagnies d'assurance Axapj² et Allianzpj³. Ces technologies numériques semblent perçues en majorité au sein de cette classe comme des *services*, des *solutions* en direction des *clients*.

Dans cette classe, la variable relative à qui détient le compte à partir duquel le *tweet* est posté, montre une surreprésentation des comptes de structures (par distinction avec les comptes personnels d'hommes ou de femmes). Cela s'explique par la présence d'un grand nombre de *tweets* postés par des comptes officiels de *startup*, institutions et organisations promouvant l'innovation technologique dans le droit.

Graphique 2.45. Distribution de la variable genre au sein de chacune des classes⁴



Les *tweets* les plus caractéristiques de la classe mettent l'accent sur un acteur central du monde de la 'justice prédictive', à savoir l'entreprise [Anajur]. Celle-ci semble donc s'être imposée au cœur des discours traitant de la 'justice prédictive' comme une innovation prenant part à un mouvement de révolution numérique au potentiel économique très fort. Les *tweets* se rapprochent d'annonces de type publicitaire sur la proposition commerciale d'[Anajur] et de l'actualité économique liée à la levée de fonds de cette *startup*.

¹ Le compte Twitter de [Victor Chevaleret] est @[v_chevaleret]

² *Digitalaxa* est le nom du programme de développement numérique de la compagnie d'assurance et *axapj* le service de protection juridique. Axa a conclu un partenariat avec [PréviCompute].

³ Allianzpj est l'assurance de protection juridique proposée par la compagnie d'assurance allemande Allianz.

⁴ N. B. : les résultats des classes 3, 4, 9 ne sont pas statistiquement significatifs.

Encadré 2.15. Les 5 tweets les plus caractéristiques de la classe 5

« **[Anajur]** au **service** de la **quantification** du **risque juridique** #IA #legaltech #justicepredictive »

« **RT** kbratings : @**[anajur]** qui **développe** et propose aux **professionnels** du **droit** des **outils** d'**anticipation** des **aléas juridiques** lève **2 millions d'euros** pour **financer** son **développement** à l'**international** #justicepredictive #legaltech #regtech #fren »

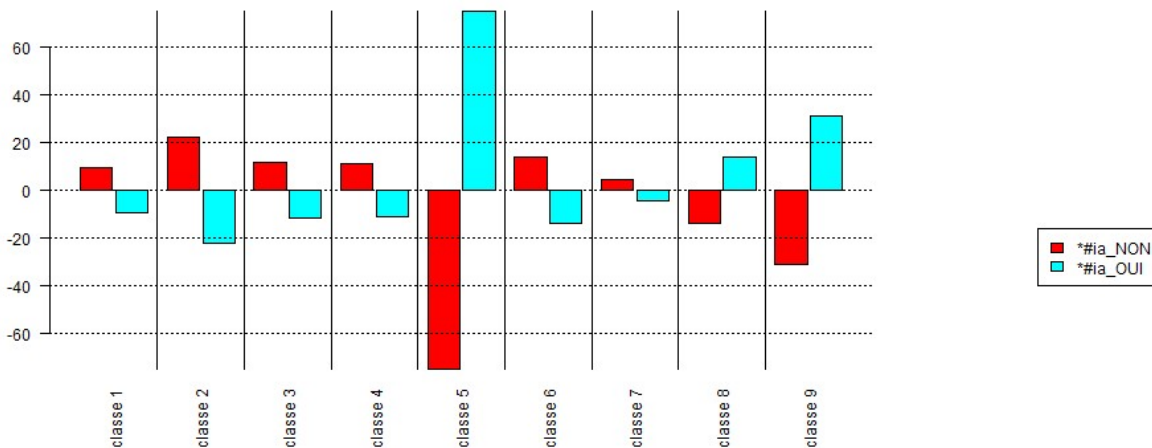
« @**[anajur]** propose aux **professionnels** du **droit** des **solutions** de **quantification** du **risque** qui se fondent sur une collaboration entre juristes et **mathématiciens** #1000startups #justiceprédicative #legaltech #regtech #IA #INRIA »

« @**contact[anajur]** Les **mathématiques** au **service** de la prise de décision **juridique**. À **lire** aussi pour mieux comprendre : Justice prédictive, vers une **analyse** très **fine** du **risque juridique** #legaltech »

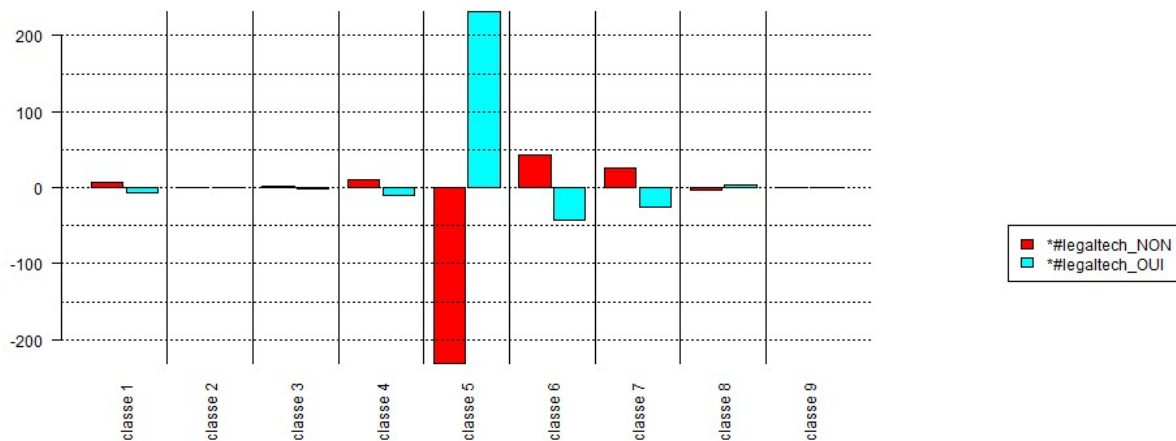
« @**[anajur]** qui **développe** et propose aux **professionnels** du **droit** des **outils** d'**anticipation** des **aléas juridiques** lève **2 millions d'euros** pour **financer** son **développement** à l'**international** #justicepredictive #legaltech #regtech #frenchtech »

Par ailleurs, l'étude des *hashtags* montre que #ia et #legaltech sont très fortement surreprésentés au sein de cette classe de discours ; ils y jouent le rôle de balise. Le terme *legaltech* semble être le néologisme caractérisant l'innovation dans le monde du droit derrière lequel se rassemblent discours et acteurs. Si, en général, le *hashtag* est un procédé qui « assemble sans rassembler » (Cardon, 2019, p. 244), dans le cas ici étudié, il apparaît que certains d'entre eux peuvent fonctionner comme des points de rassemblement de discours similaires. Cela est particulièrement net pour la classe de discours relative à la révolution numérique puisqu'elle est fortement corrélée aux *hashtags* #ia et #legaltech.

Graphique 2.46. Distribution de la variable #ia au sein de chacune des classes



Graphique 2.47. Distribution de la variable #legaltech au sein de chacune des classes



En somme, l'analyse de classification du corpus de *tweets* sur la 'justice prédictive' fait apparaître et objective la place de plusieurs types d'acteurs dominants dans l'émergence et la diffusion de l'expression 'justice prédictive' parmi lesquels les *startup* de la *legaltech*. L'expérimentation de [PréviCompute] dans deux cours d'appel et au sein du barreau lillois apparaît comme une source importante d'activité discursive sur Twitter au sens où elle offre une prise à partir de laquelle se développe la réflexion sur les enjeux des dispositifs algorithmiques, leurs potentialités et leurs limites. Le milieu juridique et judiciaire contribue directement à ces réflexions, dans le prolongement de cette expérimentation. Les échanges de *tweets* sont alimentés, presque à hauteur d'un quart, par l'annonce d'événements consacrés à la 'justice prédictive', où universitaires et praticiens réfléchissent sur ce phénomène en émergence. Il ne semble toutefois pas que ce soit des événements grand public : le débat reste au contraire assez confiné parmi les spécialistes. Jouant un effet relais pour des articles, Twitter apparaît comme une caisse de résonance des réflexions et débats parus dans la presse spécialisée ou plus généraliste (les grands quotidiens d'information) (pour 5 classes sur 9, représentant environ un quart des messages postés). Les scènes médiatique d'une part et académique d'autre part contribuent donc fortement aux échanges de *tweets* relatifs à la 'justice prédictive'. Elles participent à la circulation et à la consolidation de cette expression sur la plateforme et plus largement dans l'espace public. Enfin, le *hashtag* #justiceprédictive acquiert un pouvoir d'étiquetage d'un ensemble de débats et de manifestations. On pourrait considérer que cela est un effet lié à la sélection du sous-corpus de *tweets* autour de 'justice prédictive'. Il est vrai que nous n'avons pas effectué la recherche systématique à partir de mots-clefs concurrents. Mais à l'intérieur de notre corpus, nous avons recherché l'existence d'autres *hashtags*, correspondant à des formules repérées dans la littérature (cf. précédent chapitre) : #justicenumérique, #justicealgorithmique, #justicequantitative, #justicedigitale. Or, ils sont quasi absents des échanges sur Twitter : trois messages comportent #justicedigitale, un seul, #justicenumérique. Les autres *hashtags* sont absents du corpus de *tweets*¹. Or, si ces expressions avaient le vent en poupe, elles pourraient apparaître comme des *hashtags* complémentaires de celui de 'justice prédictive', comme c'est

¹ Pour être tout à fait précise, deux *tweets* comportent le *hashtag* #quantitative et quatre, #algorithmique.

le cas pour *#ia* ou *#legaltech*. Le pouvoir de désignation et d'agrégation du terme 'justice prédictive' et de l'étiquette numérique *#justiceprédictive*, leurs usages et leur circulation sont donc établis assez solidement, en dépit de cette réserve.

1.6. La dynamique de la controverse, à travers le croisement des phases et des classes de discours

Pour avancer dans la compréhension de la structuration — y compris temporelle — de la controverse, il est intéressant de vérifier la manière dont les classes de discours identifiées se combinent avec les quatre phases chronologiques repérées *supra*. Pour ce faire, chacun des *tweets* a été associé au numéro de la classe de discours à laquelle il correspond¹. Pour chaque pic d'intensité, le nombre de *tweets* appartenant à chacune des classes a été comptabilisé afin de relever la classe de discours dominante².

¹ Les *tweets* non classés par le logiciel Iramuteq sont regroupés au sein de la classe 0.

² Comme la classe 0 n'a pas de cohérence interne, lorsque celle-ci est la plus fréquente au sein d'un pic d'intensité, la seconde classe la plus présente est également reportée entre parenthèses.

**Tableau 2.5. La 'justice prédictive' sur Twitter, de 2010 à 2020 :
pics d'intensité et classes de discours**

Mois clés	Pics d'intensité	Nombre de tweets	Modèle typique	Classe de discours principale
-	-	-	-	-
Mars 2015 (n=38)	26-31 mars 2015	38	1	0
Janvier et février 2016 (n=36 et n=44)	28-29 janvier 2016	33	1	0 (3)
	19-24 février 2016	24	1	0 (6)
Juillet 2016 (n=61)	11 juillet 2016	7	1	7
	21-25 juillet 2016	38	1	0 (6)
Octobre, novembre et décembre 2016 (n=77, n=61 et n=82)	5-11 octobre 2016	31	1	6
	2-3 novembre 2016	19	1	0 (6)
	7 décembre 2016	12	1	3
	27-28 décembre 2016	35	1	9
Janvier 2017 (n=204)	9 janvier 2017	16	1	9
	20 janvier 2017	19	1	9
	27 janvier 2017	32	2	1
	30-31 janvier 2017	29	1+2	9
Février 2017 (n=368)	2 février 2017	23	1	9
	10-11 février 2017	49	1	0 (9)
	15-16 février 2017	78	1+2	9
	24 février 2017	54	1	0 (3)
Mai 2017 (n=524)	5-6 mai 2017	42	1	0 (3)
	12 mai 2017	25	1	0 (9)
	19 mai 2017	262	1+2	3
	23 mai 2017	19	2	3
Juillet, août 2017 (n=147 et n=64)	4-5 juillet 2017	29	1+2	0 (1)
	31 juillet 2017	40	1	0 (3)
Septembre, octobre et novembre 2017 (n=330, n=281 et n=296)	9-10 septembre 2017	46	1	8
	17-18 septembre 2017	96	1	0 (8)
	25 septembre 2017	24	2	0 (6)
	10 octobre 2017	26	2	2
	19-20 octobre 2017	81	2	3
	2-3 novembre 2017	67	1	0 (9)
	24 novembre 2017	33	1	9
Février 2018 (n=385)	12 février 2018	170	2	0 (3)
Juin 2018 (n=178)	15 juin 2018	19	2	3
	18 juin 2018	38	2	3

La mise en évidence des classes de discours les plus utilisées au sein de chacun des pics d'intensité vient confirmer le raisonnement par phases. Les pics de la deuxième phase, celle de l'entrée timide et ponctuelle de la 'justice prédictive' sur Twitter, contiennent plusieurs messages relevant de la classe 6, tournée vers le débat sur les potentialités de la 'justice prédictive'. L'autre classe de débat, centrée sur l'évaluation critique sous ses aspects actuariels et pénaux (classe 7), connaît son seul pic d'intensité ici. La classe 3 qui porte sur le billet de blog viral relatif à l'ouverture des données est présente dans deux pics d'intensité de cette phase. Enfin, la classe 9 qui met l'accent sur la 'justice prédictive' comme liée à l'intelligence artificielle (IA) apparaît dans le dernier pic, fin décembre 2016.

Au cours de la troisième phase, celle de l'effervescence, de nombreux pics se dessinent au sein desquels cette même classe lexicale, qui fait la part belle à l'IA, est la plus présente. Les billets de blog d'Hubert Guillaud sur l'ouverture des données (classe 3) partagés de manière virale pendant tout le mois de septembre 2017 continuent à engendrer des pics. Durant cette troisième phase, d'autres classes de discours causent parfois une activité virale. Le pic d'intensité survenu à l'occasion de *la Nuit de la Legaltech* (25 septembre 2017) comprend principalement des *tweets* qui ont trait à la classe de discours 6 tournée vers les potentialités des dispositifs algorithmiques. Celui du 10 octobre 2017, au lendemain de l'annonce de l'échec de l'expérimentation de [PréviCompute], comprend majoritairement des *tweets* relevant de la classe 2, celle qui est précisément associée à cette expérimentation. La classe 8, correspondant à l'annonce et au commentaire d'activités académiques, est celle qui est la plus présente les 9-10 et 17-18 septembre 2017.

Enfin, les deux pics d'intensité repérés au cours de la quatrième phase de baisse d'intensité relèvent tous deux principalement de l'utilisation de la classe 3 et donc de la question de l'ouverture des données, via le billet de blog d'H. Guillaud qui continue de circuler intensément.

Par ailleurs, deux classes de discours ne sont jamais responsables d'un pic d'intensité. Il s'agit de la classe 4 qui porte sur l'appréciation des conséquences, positives ou négatives, de l'expérimentation [PréviCompute] et de la classe 5 consacrée aux enjeux commerciaux et économiques des innovations numériques dans le droit. Elles se déroulent sur un temps plus long et ne provoquent donc jamais d'activité virale. C'est une des caractéristiques de l'algorithme de Twitter que de privilégier la viralité dans la sélection des sujets les plus « tendance ». Certains sujets de fond ne sont jamais en *trending topics* car ils ne provoquent pas une activité suffisamment rapprochée dans le temps, bien qu'ils soient massivement présents sur la plateforme (Gillespie, 2012).

Enfin, les modèles typiques précédemment établis sont renforcés par cette analyse. Pour 21 pics sur 33, les classes de discours identifiées comme médiatiques (1, 2, 3, 4 et 9) lors de l'analyse sous Iramuteq correspondent au premier modèle, au sein duquel Twitter dépend « de l'agenda médiatique » (Olivesi et Hubé, 2016, p. 9) ou aux deux modèles ensemble. Lorsque les pics d'intensité observés répondent au second modèle, selon lequel Twitter est un « baromètre de la vie [...] réelle » (Écormier-Nocca et Louis-Sidois, 2019, p. 490), une plus

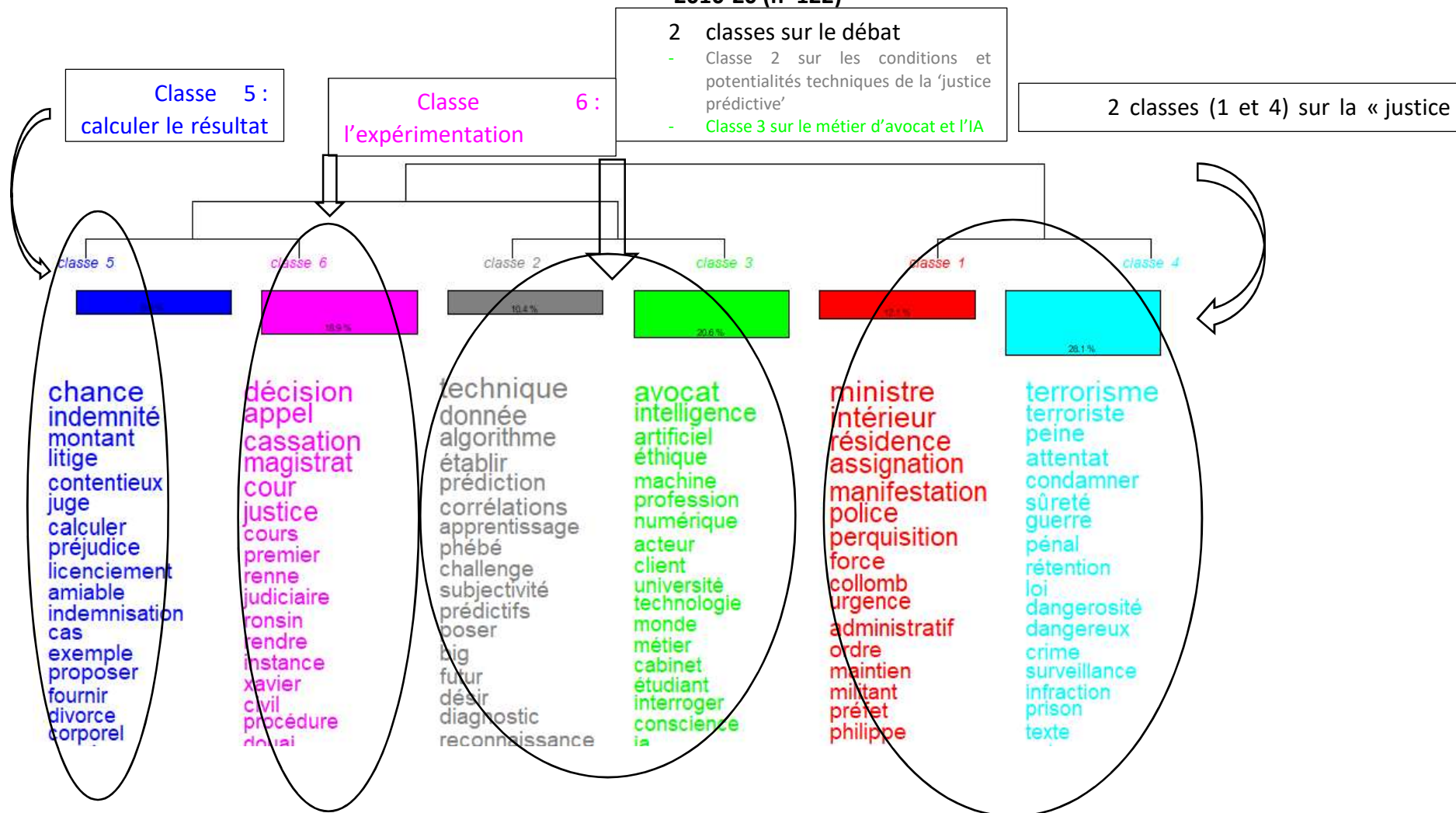
grande diversité est observable, avec notamment les classes de discours non médiatiques (2, 6, 8).

1.7. Comparaison avec l'analyse lexicale des articles de presse

La même analyse de classification a été menée sur le sous-corpus d'articles de presse consacré à la 'justice prédictive'. Elle n'a pas été restituée ici faute de place, mais nous en reproduisons le dendrogramme synthétique. Sans entrer dans le détail d'une comparaison fastidieuse, on retiendra qu'elle met également en évidence la mise en débat de la 'justice prédictive' dans deux directions : les transformations du marché juridique et du métier d'avocat, liées aux innovations de l'IA (classe 3) et les conditions techniques et les potentialités de développement de la 'justice prédictive' (classe 2). En revanche, la dimension universitaire n'apparaît pas comme saillante, compte tenu probablement du fait que les articles de presse n'ont, dans l'ensemble, pas la même fonction d'annonce et de commentaire de manifestations en cours que Twitter.

Graphique 2.48. Dendrogramme représentant le résultat de la classification de Reinert sur le corpus d'articles de presse 'justice prédictive'

2010-20 (n=122)



L'autre constat majeur qui mérite d'être dressé porte sur la manière dont la 'justice prédictive' est définie ou plus exactement sur les débats qui se développent à ce sujet. Dans les corpus d'articles de presse et de *tweets*, une, voire deux classes sont réservées à l'expérimentation [PréviCompute] (classes 2 et 4 des *tweets* ; classe 3 des articles de presse). De plus sont apparus les positionnements divergents de [PréviCompute], [Anajur] et [Juri+] dans la classe 5 des *tweets* lorsqu'il s'agit de définir ce qu'est la 'justice prédictive' ou bien de qualifier ce que chacun de ces opérateurs propose. La polarisation entre [PréviCompute] et [Anajur] est confirmée et même accentuée dans les articles de presse, puisque si un univers de sens renvoie à l'expérimentation [PréviCompute], un autre renvoie à son concurrent direct [Anajur]. En effet, la classe 1 tournée vers un discours centré sur le calcul de l'issue d'un procès comprend aussi dans les mots les plus caractéristiques de la classe des références au patronyme d'un de ses fondateurs, le *Nantais* [Géraud] [Monta]. Tout ceci confirme qu'un des enjeux des discours sur la 'justice prédictive' est clairement la structuration d'un nouveau marché sur lequel certains opérateurs tentent de se positionner et sont en concurrence.

Les classifications respectivement obtenues pour les *tweets* et les articles de presse convergent également autour du double sens conféré à la 'justice prédictive'. Cette construction sémantique renvoie à la fois à des algorithmes de traitement des décisions de justice mobilisables en matière civile et à des dispositifs de justice actuarielle pour le pénal. Ils visent tous une meilleure anticipation : des jugements et arrêts pour les premiers, de la commission d'infractions et/ou de la récidive pour les seconds. La polarisation entre les deux, atténuée sur Twitter, est très vive dans la presse où deux classes sur six (2 et 4, avec 40,26 % des segments de texte classés) renvoient à la dimension assimilable à la justice actuarielle¹. La tension entre ces deux versants de la 'justice prédictive' réapparaît donc avec force et confirme que le périmètre de ce que cette expression recouvre varie aussi en fonction des supports de communication et des locuteurs qui s'expriment. Cette tension est au cœur de ce qui travaille les usages et appropriations de la 'justice prédictive' en tant que construction sémantique dans le contexte français.

Conclusion :

Il est maintenant clairement établi que la 'justice prédictive' est une expression récente, dont l'usage s'est considérablement développé, tant dans la presse que sur Twitter. Elle procède d'un tournant algorithmique plus général, comme le montre l'analyse de classification réalisée sur le corpus d'articles de presse global (n=1 622). La 'justice prédictive' y côtoie les sondages sur les intentions de vote aux élections, les tests de dépistage de certaines maladies par séquençage du génome, et le fait de créer des voitures dites autonomes dont la conduite est prise en charge par le véhicule lui-même. Deux points de contact principaux émergent dans la manière de penser ce tournant algorithmique : le caractère prédictif des approches par probabilités et modélisations qui permettent de

¹ Quatre classes (2, 3, 5 et 6), qui comptent pour 59,74 % de l'ensemble, sont consacrées aux dispositifs algorithmiques d'analyse de décisions de justice civile susceptibles d'être utilisés comme aide à la prise de décision juridique et judiciaire.

dessiner des projections et des anticipations d'une part, l'enjeu de déléguer des activités classiquement accomplies par les humains à des algorithmes, machines ou robots d'autre part.

Plus que dans d'autres secteurs, les questionnements philosophiques, anthropologiques et éthiques, relatifs aux effets de la « révolution algorithmique » sur la justice, mais aussi sur la société, son organisation et ses valeurs, sont spécialement prégnants. De même, les interrogations sur les transformations en cours et sur l'avenir sont un sujet de préoccupation majeur, suscitant parfois des inquiétudes. Ce trait distinctif a de quoi surprendre si l'on considère que les innovations médicales par exemple sont *a priori* porteuses d'enjeux aussi lourds. Cela est à mettre en lien avec le cadrage médiatique qui associe la justice à un monde a-technologique. Le fossé n'en est que plus vaste avec des perspectives futuristes présentées de manière hyperbolique. La place des universitaires dans le débat est un autre trait distinctif des discours sur la justice. Visiblement, les réflexions sur les enjeux de la 'justice prédictive' sont explorées par les juristes dans le cadre de manifestations scientifiques et de publications assez spécialisées, mais qui débordent parfois dans l'espace public.

Au fil de la décennie se dessine une controverse repérable à la fois dans la presse et sur Twitter. L'expression 'justice prédictive' est quasi absente au début des années 2010. De 2015 à 2016, elle monte progressivement en puissance. De janvier 2017 à février 2018 s'ouvre une période d'effervescence autour de ce sujet, qui confirme l'existence d'un moment 2017. À partir de mars 2018 et jusqu'à décembre 2020, l'intensité des articles et échanges décline ; elle se stabilise toutefois à un niveau plus élevé qu'avant 2017. Ceci suggère que le terme 'justice prédictive' s'est installé au fil de la controverse. En 2020, davantage qu'en 2015, la 'justice prédictive' désigne et labellise un ensemble de réflexions portant sur des dispositifs algorithmiques de traitement des décisions de justice en matière civile d'une part et d'évaluation des risques en matière pénale d'autre part. Cela ne préjuge toutefois pas de la suite de l'histoire. Il n'est pas exclu que la notion de 'justice prédictive' puisse décliner à plus ou moins long terme, même si nos résultats indiquent que des effets de labellisation existent et que des habitudes sémantiques sont en place.

La mise en relation de ce schéma chronologique en quatre séquences avec les classes de discours identifiées par l'analyse lexicale, complète l'approche diachronique. Twitter est d'abord la caisse de résonance d'une effervescence proprement médiatique sur la 'justice prédictive'. Une controverse se met ensuite en place durant l'année 2017. L'annonce de l'échec de l'expérimentation du logiciel [PréviCompute], en octobre de la même année, entraîne la fin de ce bouillonnement médiatique, mais les débats soulevés se poursuivent, notamment au sein du milieu universitaire et à travers l'organisation de manifestations scientifiques associant académiques et praticiens. Les débats de « fond » liés à cette innovation sociotechnique sont plus nombreux en aval de l'annonce de cet échec qu'en amont, ce qui indique le rôle important joué par cet épisode dans la structuration de la controverse et des débats.

Le cadrage des discours développés sur Twitter et dans la presse confère une part importante aux enjeux proprement économiques ouverts par le développement de la ‘justice prédictive’ et plus largement par les transformations numériques du droit, notamment sous l’angle de l’émergence de nouveaux marchés juridiques, du bouleversement de la profession et des cabinets d’avocats. *Startup* et acteurs de la *legaltech* sont très présents, à la fois mis en avant dans les débats et actifs dans l’effervescence autour de la ‘justice prédictive’. La logique de marché est au cœur des discours relatifs aux dispositifs algorithmiques de traitement et d’analyse des décisions de justice. L’espace académique et l’espace public sont aussi des lieux où se livre une guerre commerciale autour de la ‘justice prédictive’.

L’intervention d’acteurs privés aux enjeux économiques directs n’est pas une nouveauté ni dans la justice ni dans l’action publique de manière générale. En revanche, cela pointe la constitution d’un marché, à l’égard duquel les questions de labellisation et de légitimation des processus en cours sont essentielles, *a fortiori* dans un secteur régalién comme celui de la justice — et à la différence de l’industrie automobile par exemple. En réalité, les *startup* sont au premier plan de ces transformations. Elles surfent sur la vague de l’intelligence artificielle et s’inscrivent dans un contexte renouvelé d’action publique. La ‘justice prédictive’ est le révélateur, en même temps que le produit, d’une nouvelle « configuration de réforme administrative » (Bezès, 2009). Celle-ci a permis que s’élaborent de nouveaux arrangements institutionnels et économiques que le chapitre suivant mettra en évidence.

Chapitre 3 : La traduction de la politique d'*open data* dans le secteur de la justice

Si les années 2000 ont représenté pour les administrations et l'État français une période de conception et de mise en place d'une « administration électronique » (Dagiral, 2011), la décennie 2010 correspond à l'adoption et la mise en œuvre d'une politique européenne d'ouverture des données publiques, dont une des manifestations les plus emblématiques au plan national est la loi Lemaire de 2016 « pour une République numérique ». Le vote de ce texte représente un moment important symboliquement et confère au principe de publicité et de libre accès aux données publiques le statut de norme publique. La disponibilité et l'accessibilité des données est décisive pour l'alimentation et le fonctionnement d'algorithmes. La mise en place d'expérimentations de dispositifs algorithmiques de traitement et d'analyse des décisions de justice repose sur le fait que, en tant que données publiques, de grandes quantités de décisions de justice deviennent accessibles et utilisables (en texte intégral et format-type) et peuvent alimenter des algorithmes d'apprentissage machine. Les opérations de calculs dépendent directement des données à partir desquelles elles sont réalisées. Dans ce chapitre, nous allons donc faire un indispensable pas de côté qui nous emmènera vers la politique de l'*open data* et ce, afin de pouvoir comprendre ensuite (chapitre 4), comment la 'justice prédictive' doit être lue comme un « produit » de cette politique.

La politique d'*open data* s'enracine dans un double mouvement (Goëta, 2015, p. 27), d'accès du public aux informations produites par l'État et les administrations¹, mais aussi d'ouverture des données dans différents domaines du champ scientifique (botanique, astronomie, etc.) à partir des années 1990. Au milieu des années 2000, l'*open data* a émergé comme une revendication politique d'abord au Royaume-Uni, sous la forme de l'exigence que les données publiques, financées par les contribuables, soient restituées aux citoyens. Lors de la rencontre dite de Sébastopol qui a eu lieu en Californie en 2007, plusieurs acteurs clés du numérique, dont Lawrence Lessig (l'auteur de *The Code Is the Law*), Tim O'Reilly, éditeur américain, connu pour être à l'origine de l'expression « eb 2.0 », se sont réunis pour énoncer une série de principes fondateurs de l'*open data* dits « principes de Sébastopol ». Ils exigent la publication des données par les administrations qui les ont constituées, et ce, gratuitement, dès leur production et dans leur forme initiale c'est-à-dire telles qu'elles ont été collectées, sans modification ou agrégation ultérieures (Goëta, 2015 : 28). Des données tout de suite et des données « brutes », ainsi pourrait-on résumer le slogan des différents acteurs et groupes d'intérêt qui portent ces revendications.

L'exigence d'*open data* et les principes de Sébastopol ont connu une circulation rapide et large. D'abord, dans les États-Unis d'Obama puis en Europe avec la création de portails de

¹ S. Goëta rappelle que la Déclaration des droits de l'homme et du citoyen de 1789 prévoit que « la Société a le droit de demander compte à tout Agent public de son administration » (art. 15), « faisant de l'accès à l'information publique l'un des fondements de la démocratie naissante » (2015, p. 27).

données publiques largement accessibles et gratuits : data.gov.us (2007) ; data.gov.uk (2009) au Royaume-Uni et data.gouv.fr (2011) en France. En 2011, l'*Open Government Partnership* (OGP) ou Partenariat pour un gouvernement ouvert (PGO) a été fondé à l'initiative des États-Unis, par un groupe de huit pays que la France a rejoints en 2014. Cette initiative internationale multilatérale vise à promouvoir transparence des gouvernements et participation des citoyens, et ce, à travers le levier des technologies numériques. La France prend la présidence du PGO le 26 septembre 2016 de manière parfaitement concomitante avec le vote et la promulgation de la loi Lemaire.

La politique de l'*open data* est une politique constituante (Lowi, ; Baudot, Marrel et Nonjon, 2015 ; Bezès et Join-Lambert, 2010), qui porte sur les règles d'organisation de l'action publique et des rapports entre gouvernants et gouvernés. Posant les règles du jeu, elle le fait à un niveau transversal, c'est-à-dire pour l'ensemble des administrations et donc au-delà des secteurs d'action publique. Ouvrir largement les données administratives s'applique aussi bien aux transports, à la recherche, à la défense ou à la justice, mais ne suscite évidemment pas les mêmes questions d'un secteur à l'autre. La confidentialité ou la sensibilité de certaines données ne sont pas posées de façon uniforme.

Dans le cas qui nous intéresse ici, l'enjeu est donc bien celui de la mise en œuvre d'une politique de l'ouverture des données dans un secteur d'action publique largement organisé autour d'une « bureaucratie professionnelle » (Mintzberg, 1982), c'est-à-dire une bureaucratie conçue autour de la compétence des magistrats, lesquels assument à la fois la tâche centrale de l'institution (rendre des décisions de justice) et le contrôle de la qualité du travail accompli². À bien des égards, les jugements et arrêts sont au cœur de la justice : dans sa dimension institutionnelle et organisationnelle puisqu'ils sont à la fois centre de l'activité des magistrats, manifestation de l'autonomie de ceux-ci au sein de l'État, source de droit via la jurisprudence et pivot autour duquel toute la structure judiciaire est organisée. Ils jouent un grand rôle dans les principes, normes et règles qui structurent historiquement la justice. Conçue en dehors de la justice, comme d'autres politiques telles que l'introduction de la LOLF, du *lean management*, de la standardisation, etc. (Vigour, 2019), la politique d'*open data* doit être traduite dans le secteur de la justice, par ses acteurs et ses institutions, mais elle touche à ce qui est au fondement de la spécificité de la justice dans l'État.

Le questionnement central de ce chapitre porte sur la manière dont cette politique est traduite par les acteurs de la justice, pour passer de l'*open data* des données publiques, administratives à l'*open data* des décisions de justice. Comment réalisent-ils concrètement cette opération de médiation du global vers le sectoriel (Muller, 2015 ; Ravinet et Palier, 2019 ; Maillard et Kübler, 2015 : 191) ? Parviennent-ils à réaliser le couplage entre les deux et si oui, comment ? En effet, si certains acteurs individuels et collectifs sont porteurs de représentations nouvelles, en phase avec le paradigme dominant de l'ouverture des données, qu'ils s'efforcent de mettre en mots, en discours, en projets, ils doivent faire face aux croyances, représentations, normes — en un mot aux institutions — dominantes et ancrées

² Même si la montée d'une gestionnarisation de ce secteur d'action publique comme d'autres se traduit par une remise en cause de cette caractéristique de l'évaluation autonome (Vigour, 2019).

au sein de la justice et aux acteurs pour qui elles font sens. D'un autre point de vue, les acteurs centraux de la justice n'ont pas forcément un intérêt immédiat à s'adapter à ce nouveau système de croyances. Tentent-ils de le « subvertir » pour contenir les germes de changement et conserver leurs positions dominantes ? Restent-ils campés sur des représentations qui sont en décalage, de plus en plus grand, avec un environnement qui change ? Ou bien leur système de croyances évolue-t-il pour intégrer de nouvelles idées, de nouveaux outils ?

Ces questions se posent d'autant plus que la justice est encadrée dans le champ juridique qui fonctionne aussi comme un espace de lutte, structuré autour de valeurs cardinales, à partir desquelles se définissent des hiérarchies, positions et *habitus*. Le développement du numérique et l'*open data* des décisions de justice, en ouvrant la porte à de nouveaux savoirs et savoir-faire, fournissent de nouvelles ressources pour des acteurs entrant dans le champ. Il peut s'en suivre une recomposition à l'intérieur même du champ, voire un renversement de positions établies ; des conflits et controverses autour de la définition légitime du rôle du juge, de la nature même de l'activité de juger, des modèles d'excellence professionnelle chez les magistrats ou les avocats (Bourdieu, 1986 ; Vauchez, 2004, 2016). Qu'en est-il exactement de ces reconfigurations ?

Nous reviendrons d'abord sur le contenu de la politique de l'*open data* et la configuration de réforme qui la porte au sein de l'État (1). Nous analyserons ensuite comment cette politique est appropriée dans la justice, par un acteur qui tente de se positionner comme central et par des intermédiaires de politiques publiques (2), au fil d'une histoire marquée par trois temps forts (3). Enfin, nous nous attacherons à analyser les controverses et coalitions de cause qui émergent (4), pour brider ou au contraire maximiser l'*open data* des décisions de justice. À ce jour, le corpus documentaire constitué comporte essentiellement des informations publiques. En particulier pour la partie consacrée aux controverses, nous manquons encore de l'ensemble du matériau empirique nécessaire pour tracer finement les croyances et stratégies de coordination entre acteurs. Aussi est-ce prudemment que nous avancerons dans cette voie, avec des résultats encore provisoires et destinés à être vérifiés et réajustés selon ce qui sera collecté dans les entretiens et les archives privées.

I / La politique d'*open data*, résultat d'une nouvelle configuration de réforme

L'*open data* est une politique publique qui s'inscrit dans une économie de la connaissance fondée sur l'innovation et les services. Elle procède d'un nouveau répertoire de réforme administrative (1) soutenu par une nouvelle configuration (2) qui amorce des transformations importantes dans les relations et interdépendances entre l'État, le marché, la science et les citoyens (Baudot, Marrel et Nonjon, 2015). Elle est incarnée par la loi Lemaire de 2016 (3).

1.1. Le répertoire de réforme de l'*open data* : État-plateforme, méthode agile et *startup Nation*

Dans son ouvrage séminal, Philippe Bezès (2009) combine approche sociologique synchronique et diachronique des réformes administratives de l'État français qu'il étudie des

années 1960 aux années 2000. Il isole plusieurs « configurations de réforme » sur la période et s'intéresse à leurs trajectoires. L'enjeu n'est pas ici de les restituer, mais plutôt de s'approprier la notion, élaborée à partir des travaux de Norbert Elias. « Simplification du réel, la configuration de réforme peut être considérée comme un système d'interdépendances entre trois groupes d'acteurs collectifs majeurs qui participent de manière récurrente à la construction des réformes : les exécutif et législatif politiques ; les ministères et hauts fonctionnaires réformateurs ; les experts au sens large, offreurs de formulations du problème, de diagnostics et de solutions. » (Bezès, 2009 : 50) Un répertoire de réforme administrative — comme le « répertoire d'action » de Charles Tilly dont il est inspiré — est, quant à lui, « un script, un ensemble de contenus et de modalités d'intervention à travers lesquels se met en forme une intervention publique sur l'administration » (Bezès, 2009 : 48). Un répertoire parvient à s'imposer parce qu'il est porté par une configuration d'acteurs.

La politique de l'*open data* nous semble témoigner d'un nouveau « répertoire de réforme administrative » appuyé sur un cadre cognitif et normatif qui comprend des idées et des valeurs sur ce qui est et ce qu'il faudrait faire. En effet, il repose sur des principes et croyances, relatifs à la manière de concevoir l'action publique et en particulier de penser les rapports entre acteurs publics et privés. Le numérique apparaît comme le moyen de rendre les citoyens plus informés, plus libres et l'État plus réactif, plus proche d'eux et plus adapté à cette nouvelle économie. Parmi les acteurs qui plaident pour l'ouverture des données, certains conçoivent celles-ci comme un nouveau gisement de richesses à haut potentiel de rentabilité, pour les entreprises et les économies nationales.

Il ne s'agit pas ici d'uniformiser artificiellement ce contenu, loin d'être homogène. Au plan international et national, la revendication de l'ouverture des données publiques est portée par une grande diversité d'entrepreneurs de cause dont les philosophies, motivations et objectifs sont hétérogènes. Toutefois, trois dimensions fortes se dégagent : la recherche de transparence, la promotion des innovations et l'enjeu de modernisation administrative. Elles sont articulées à travers les logiques et réalisations pratiques de plateformes numériques.

1.1.1. *Transparence, innovation et modernisation de l'État*

Samuel Goëta et Jérôme Denis (2017) distinguent cinq facettes de la revendication d'ouverture des données qui ne se recouvrent pas forcément : la recherche de transparence envers les citoyens, au service notamment d'une plus grande redevabilité de l'action des gouvernants ; le principe de la libre circulation de l'information indispensable à une vie démocratique riche que défendent les penseurs de la cybernétique, mais aussi les mouvements favorables au logiciel libre ; l'exigence de données dites brutes qui répondraient à un format unique et seraient universellement appropriables, ce que certains voient comme une juste restitution aux contribuables ; la perspective d'une industrie de la donnée et du potentiel économique que représenterait l'exploitation de ce gisement informationnel pensé comme un nouvel or noir, source d'innovations pour les entreprises ; enfin, l'ambition de la modernisation des administrations qui seraient ainsi moins cloisonnées, plus orientées vers les citoyens et plus innovantes. Comme le souligne S. Goëta, l'*open data* a « pris la forme d'une injonction » reposant sur des dimensions plurielles, mais que les groupes d'intérêts concernés ont présentée comme des « biens en soi » (Dodier cité par Goëta, 2015 : 28), « des

objectifs dignes d'être poursuivis en tant que tels sans nécessiter de justification supplémentaire » (*ibid.*). Trois notions principales peuvent résumer ces « biens en soi » : « la transparence, l'innovation et la modernisation de l'État » (*ibid.*).

Dans le cas français, ce système de croyances s'exprime à travers l'enjeu de construire un écosystème de la donnée où interagissent un « État-plateforme » et une pluralité d'acteurs privés dont des *startup* (Flécher, 2019). Au-delà du côté slogan, il y a là un corps de doctrine qui, malgré une part de flou, définit priorités, principes, partenaires et manières pour l'État de fonctionner avec son environnement. Des méthodes (comme la méthode « agile »³) et des instruments d'action publique bien spécifiques (comme le portail data.gouv.fr) incorporent et véhiculent ces principes pour laisser « entrevoir la figure d'un État souple, agile, flexible, soucieux d'efficacité et féru d'innovation, servant de point d'appui et de relais aux initiatives citoyennes » (Chevallier, 2018 : 637). On notera toutefois que dans le cas français, la relation entre doctrine, méthode et outil n'est ni intellectuellement ni temporellement linéaire. Comme le souligne Jacques Chevallier, ce n'est qu'à partir de 2014 que le terme d'« État-plateforme » (2018 : 629) est mobilisé pour définir, englober et théoriser après coup une stratégie globale de modernisation administrative dont les modes de fonctionnement et instruments ont déjà été constitués peu à peu.

1.1.2. La notion d'État-plateforme

La paternité du terme est attribuée à Tim O'Reilly, éditeur américain, également connu pour être à l'origine de l'expression « Web 2.0 » et déjà présent lors de la rencontre de Sébastopol. En France, la notion est importée par Nicolas Colin et Henri Verdier (2012), deux entrepreneurs du numérique aux parcours hybrides entre secteurs public et privé, dans un ouvrage publié en 2012 et réédité en 2015. Le premier, inspecteur des finances, a une double formation d'ingénieur en informatique (Télécoms) et en administration (IEP, ENA). Le second est formé à l'ENS d'Ulm en biologie et en sociologie. Au moment où ils écrivent, ils sont tous les deux membres du pôle Compétitivité de Cap Digital. La même année, Henri Verdier devient directeur de la mission Etalab. Cette structure est une mission à l'intérieur de l'Agence pour le patrimoine immatériel de l'État (APIE), créée en 2007 et rattachée au ministère de l'Économie et des Finances. Elle est chargée, à partir de 2008, de créer un portail unique pour les données publiques, intitulé Etalab, avec le double objectif de donner accès à des données publiques réutilisables et de valoriser le patrimoine immatériel de l'État par le biais de redevances. Nous verrons plus loin qu'elle deviendra un acteur central de la politique d'ouverture des données publiques. Henri Verdier, lorsqu'il en devient directeur à partir de 2012, importe et acclimate la notion d'État-plateforme au sein de l'agence Etalab laquelle en devient ambassadrice et relais.

³ La ou plutôt les méthodes agiles relèvent d'une méthodologie de gestion de projet, élaborée au début des années 2000 par des développeurs de logiciels. Il s'agit de méthodes qui placent le client au centre du projet, et qui, contrairement à des approches plus planifiées, s'appuient davantage sur une logique d'adaptabilité tout au long du projet. Nées dans le secteur de l'informatique, elles s'appliquent au développement de produits, de services, de projets ou même d'organisations dans d'autres domaines (Reverdy, 2021 : 177-181).

L'ouvrage co-écrit par N. Colin et H. Verdier, intitulé *L'âge de la multitude : Entreprendre et gouverner après la révolution numérique* est entièrement tourné vers la révolution que représente Internet et les conséquences à en tirer tant pour l'économie que pour l'action publique. Il fait l'apologie de la « multitude », c'est-à-dire de la masse des individus instruits, équipés et connectés dont la créativité devient centrale dans les processus de création de la valeur (Benghozi, 2012 : 70). Dans un monde connecté et hyperfluide, les innovations se multiplient à un rythme rapide, les capacités d'adaptation des individus et des organisations sont intensément sollicitées et les risques de bouleversement des positions établies bien présents. L'action collective passe par de nouveaux supports, de nouvelles manières d'agir où la dimension participative, prégnante, redessine les frontières entre amateurs et professionnels, secteur marchand et non marchand. L'État-plateforme incarne cette vision de la gouvernance de l'action publique par la mise en réseau, sur des plateformes numériques, de tous types d'acteurs dont les capacités d'innovations seraient décuplées. L'enjeu pour les entreprises comme pour l'action publique est de développer des stratégies reposant sur la mise en place de telles plateformes qui permettent à la multitude des individus de s'en emparer pour concevoir et développer de nouvelles applications et de nouveaux services. Les politiques de soutien à l'innovation et à la créativité doivent accompagner ce processus, dont il est attendu des créations de valeur. Cette notion, qui répond au développement du capitalisme de plateforme, constitue une déclinaison d'un nouveau référentiel global indexé sur « le régime de capitalisme globalisé » (Muller, 2015 : 102), sur lequel on reviendra plus loin.

L'expression d'État-plateforme apparaît « dans les documents internes des services en charge de l'animation de la transformation gestionnaire de l'État à partir de 2014 » (Jeannot, 2020 : 166) ; elle est ensuite reprise dans le rapport annuel du Conseil d'État de 2017 et dans le rapport de la Cour des comptes de 2018. Pour Tim O'Reilly puis Henri Verdier, l'État-plateforme recouvre l'idée selon laquelle les bases de données publiques et les applications informatiques ouvertes (dites API pour *Application Programming Interface*) représentent un nouveau type d'infrastructures. Elles offrent des ressources pour que des acteurs, aussi bien des entreprises privées que des citoyens férus d'informatique, proposent des services innovants, vertueux sur le plan économique et pour la modernisation de l'État. Le fait que des tiers innovent à partir d'API est aussi ce qui garantit le succès des plateformes elles-mêmes et le fait qu'elles s'imposent sur un marché. Benoît Thieulin, conseiller politique devenu entrepreneur du numérique et qui sera à la tête du Conseil national du numérique de 2013 à 2016, donne l'exemple de l'API de Facebook qui a permis à ce réseau social de « répandre le bouton "Like" sur le Web et de dominer le marché de la recommandation » (2018 : 27).

Dans cette doctrine, une dimension participative et communautaire forte existe, étroitement liée à la culture de l'horizontalité développée par les acteurs du numérique. L'ambition affirmée de N. Colin et H. Verdier est de voir « s'inventer sous nos yeux des services publics sans administration, auto-organisés par des communautés de citoyens prenant leur part de leur opération par leurs contributions et leurs interactions » (cités par Thieulin, 2018 : 26). Outre les citoyens, cette doctrine confère également un grand rôle aux acteurs économiques privés à qui est aussi délégué le contenu des innovations et services proposés.

Dans ce modèle, les services publics deviendraient le résultat de l'auto-organisation des citoyens, mais aussi de la mobilisation d'acteurs privés, agissant en complète autonomie et non à travers des commandes et délégations formelles répondant à un cahier des charges dont la réalisation est contrôlée par l'État. Dans l'ouvrage collectif *L'État en mode start-up* dirigé par Yann Algan⁴ et Thomas Cazenave⁵ (2016), l'accent est mis sur l'urgence de la transformation de l'action publique, érigée en « dernière chance » de rétablir la confiance des citoyens. Le mode « agile » et la collaboration accrue avec le privé y sont présentés comme des opportunités à saisir et des solutions à privilégier, sous peine de voir la capacité de l'État à gouverner et sa centralité remises en cause. Si cette doctrine est portée par des acteurs académiques et non académiques, si elle est au carrefour entre l'économie, le management et les sciences administratives et n'est donc pas arrimée de manière forte à une discipline, elle produit de fait des savoirs managériaux mobilisables pour l'action publique qui évoquent les « sciences de gouvernement », bien étudiées dans les mécanismes de production et de légitimation de l'action publique (Ihl et Kaluszynski, 2002 ; Ihl, 2019 : 563). Les auteurs des ouvrages qui popularisent la notion d'État-plateforme utilisent en effet le registre de la connaissance théorique et pratique pour revendiquer une forme de scientificité et donc de légitimité à orienter les pratiques gouvernementales et les réformes de la bureaucratie.

Par ailleurs, il faut dire que dans la France du début des années 2010, ni les *startups* ni les citoyens ne s'emparent massivement de ces données pour concevoir de nouveaux produits et services — au contraire des géants du numérique. C'est ce qui explique que la stratégie de l'État-plateforme soit de fait progressivement infléchie vers le soutien à la création de *startups* d'État, ce qui, compte tenu de la centralisation classique dans le modèle d'administration français, peut être lu comme une réappropriation originale de cette doctrine. Des structures de petite taille, conçues pour une durée courte (moins de six mois) autour d'un objectif circonscrit, portées par des « intrapreneurs » — le terme désigne des agents issus de l'intérieur des organisations qui ont une idée originale et endossent les caractéristiques et le discours classiquement attribués aux entrepreneurs privés — auxquels sont associés des développeurs informatiques. À l'image d'Etalab, avec son soutien et celui de la mission beta.gouv.fr, des plateformes sont créées. Par exemple mes-aides.gouv.fr⁶, simulateur d'aides sociales destiné à lutter contre le non-recours (Alauzen, 2021), ou la Bonne boîte⁷, développée dans le cadre de Pôle Emploi par un de ses salariés et disponible sur le portail Emploi store (construit sur le modèle de l'Apple store). Cette dernière permet aux demandeurs d'emploi d'identifier les entreprises qui, dans un secteur d'activité et une région donnés, disposent d'un fort potentiel d'embauche. Celui-ci est déterminé par une application qui utilise des données

⁴ Il est à l'époque doyen de l'École d'affaires publiques de Sciences Po et professeur d'économie, spécialiste de l'économie numérique et collaborative.

⁵ Inspecteur des finances, il est alors directeur de cabinet adjoint du ministre de l'Économie, de l'Industrie et du Numérique, enseignant à Sciences Po et à l'Éna.

⁶ Devenue mesdroitssociaux.gouv.fr, <https://www.mesdroitssociaux.gouv.fr/accueil/>, consulté le 11 novembre 2021.

⁷ <https://www.pole-emploi.fr/candidat/vos-services-en-ligne/emploi-store/la-bonne-boite.html>, consulté le 31 mars 2021.

disponibles (relatives aux dynamiques de recrutement des entreprises les années précédentes, etc.). API et *startup* d'État se combinent pour proposer une transformation des organisations administratives « par les marges », via un décloisonnement effectué « dans une démarche gestionnaire de focalisation sur les attentes des usagers » (Jeannot, 2020 : 173).

1.1.3. Une doctrine à travers un instrument : data.gouv.fr

La capacité des instruments à encapsuler des idées, savoirs et axiomes et à les mettre en œuvre, dans et par leur technicité même, est bien connue (Lascoumes et Le Galès, 2004). Les plateformes numériques publiques, par la manière dont elles sont agencées, révèlent — en même temps qu'elles opérationnalisent — le corps de croyances que nous avons isolé : elles permettent à la fois l'empouvoirement⁸ des citoyens et l'investissement d'entrepreneurs privés, créateurs de *startups* ou de sociétés déjà installées.

Dans une perspective de sociologie des sciences, Clément Mabi (2015) s'est intéressé à la conception technique du portail data.gouv.fr dont il a montré le rôle dans la conciliation de ces objectifs disparates en direction de publics diversifiés. À partir d'une approche sensible au *design* des instruments et à la matérialité des dispositifs (organisation du contenu accessible en ligne, possibilités techniques offertes par le site, modalités d'intervention des concepteurs sur le site, manières de s'adresser aux publics de la plateforme), il a analysé cette interface pour faire ressurgir le projet politique sur lequel elle repose et la manière dont, en tant qu'instrument d'action publique, elle le traduit et le met en œuvre.

Il en ressort que c'est par « la valorisation d'un écosystème autour des données » que le compromis entre les différents objectifs de la politique française d'*open data* a été opéré par Etalab, la mission chargée de mettre en place data.gouv.fr. Sans définir ce terme d'écosystème, elle le mobilise beaucoup sur la plateforme. « Parler en terme "d'écosystème" permet de s'affranchir d'un projet de relation directe entre les citoyens et les données et de légitimer l'intégration des intermédiaires dans le processus pour créer de la richesse et favoriser les innovations. Si ses finalités citoyennes demeurent, l'*open data* doit permettre au passage, quasi mécaniquement, de produire de la richesse. » (Mabi, 2015 : 55) La plateforme est ainsi organisée autour des usages et réutilisations multiples des données via les applications informatiques ouvertes imaginées par des développeurs, offrant de nouveaux services, que la plateforme met à disposition des citoyens. Celle-ci valorise donc le rôle des intermédiaires privés que sont ces développeurs informatiques, devenus, à la grâce de l'*open data*, des « médiateurs légitimes au sein de l'espace public » (Ruppert, 2013 cité par Mabi, 2015 : 56). *A contrario*, le portail data.gouv.fr minimise le rôle de l'État. Très peu présent sur le site, sous l'angle de la seule garantie de validité et de traçabilité des données, Etalab, représentant de l'État, se positionne comme un acteur parmi d'autres dans la communauté fédérée par l'ouverture des données publiques. L'idée d'un « écosystème » de la donnée, qui sera développée ensuite dans la loi Lemaire, apparaît (Goëta, 2016 : 77), comme une résultante directe de l'accent mis sur la logique d'« APIisation ».

⁸ On sait que le terme *empowerment* est difficile à traduire tant ses significations sont riches en anglais. Le terme empouvoirement en est sa francisation. Il nous semble plus évocateur que ceux de « capacitation », « autonomisation », « responsabilisation », qui restent souvent assez abstraits en français.

1.2. Une configuration de réforme favorable à la « *startup Nation* »

Ce répertoire de réforme est porté par une configuration favorable, qui se met en place au début de la décennie 2010, et qui résulte de la transformation de trois groupes d'acteurs individuels et collectifs, au sein desquels les croyances dans le pouvoir libérateur des technologies (pour les citoyens et pour l'économie) sont globalement partagées.

1.2.1. *Des experts universitaires et venus de l'économie numérique*

Le premier comporte des spécialistes et experts, universitaires ou non, comme ceux que nous venons d'évoquer. La présence d'un économiste comme Yann Algan n'est pas surprenante quand on sait la place que la science économique est parvenue à occuper dans la structuration des savoirs sur l'État et sur sa réforme, et déjà bien présente dans les configurations de réforme étudiées par P. Bezès.

On rencontre aussi un autre profil et c'est une spécificité importante, celui d'acteurs de l'économie numérique, porteurs d'une connaissance et d'une expérience de ce secteur et convaincus du pouvoir libérateur des technologies, qu'ils transposent ensuite dans le secteur public. Pierre Pezziardi en est un bon exemple. Ingénieur formé à l'École centrale de Lyon, cofondateur d'un cabinet de conseil en architecture de systèmes d'information (Octo Technology) et créateur de plateformes de *crowdfunding*, devient en 2013 entrepreneur en résidence pour le Secrétariat général pour la modernisation de l'action publique et crée l'incubateur de *startups* d'État beta.gouv.fr évoqué plus haut. Il travaille en lien direct avec Henri Verdier alors à la tête de la mission Etalab et en 2017, publie avec lui *Des startups d'État à l'État plateforme*.

Chefs d'entreprises de la nouvelle économie, ils deviennent pour certains conseillers ministériels, selon des trajectoires qui hybrident le public et le privé. C'est le cas de Séverin Naudet qui après avoir créé des sites de musique en ligne est devenu conseiller du ministre de la Culture Renaud Donnedieu de Vabres pour les médias et Internet entre 2004 et 2006, puis conseiller du Premier ministre François Fillon sur l'innovation numérique et la politique de *l'open data*, pendant le quinquennat de Nicolas Sarkozy. Le cas d'Henri Verdier, déjà cité, est un peu différent au sens où après des allers-retours entre public et privé (Haut Comité de la santé publique, Groupe Lagardère, etc.) toujours dans le numérique, mais sans avoir créé sa propre entreprise, il devient directeur d'Etalab puis directeur d'administration centrale (Directeur interministériel du numérique) et ambassadeur du numérique pour le gouvernement.

Certains, aux profils de hauts fonctionnaires spécialistes du numérique, deviennent aussi des entrepreneurs : comme Nicolas Colin, rencontré précédemment. Inspecteur des finances, il est fondateur d'une *startup* (1x1connect) pour laquelle il s'est mis en disponibilité, il cofonde ensuite la société The Family qui a pour objet d'accompagner la création de *startups* européennes dans le champ de l'économie numérique. De même, Benoît Thieulin, après une formation à Sciences Po Paris, devient au début des années 2000 le responsable de la communication du gouvernement français sur Internet. Il s'engage ensuite comme directeur de la campagne présidentielle numérique de Ségolène Royal. Il partage d'ailleurs cet engagement pour la candidate socialiste avec Nicolas Colin qui avait lui-même monté un

groupe de soutien. Après l'échec à l'élection, il fonde une agence digitale de communication (La Netscouade). Quelques années plus tard, il sera nommé à la tête du Conseil national du numérique (2013-2016) par François Hollande.

Enfin, peut également être cité Gilles Babinet, qui combine au fil de sa trajectoire, des expériences dans le privé, un label de « *digital champion* » pour la France auprès de la Commission européenne, une fonction de professeur associé à Sciences Po Paris et une reconnaissance institutionnelle en tant que vice-président du Conseil numérique. Il est l'auteur d'un livre intitulé *Refondre les politiques publiques avec le numérique* dans lequel il reprend le corps de doctrine et les références évoquées plus haut (État-plateforme, méthode agile, référence à H. Verdier et Etalab, etc.), mettant l'accent sur l'impératif vital pour l'État de ne pas considérer le numérique simplement comme un moyen ou un secteur, mais de fonder transversalement et radicalement sur lui son gouvernement et sa stratégie (économique, de réforme de l'État, etc.) (Santiso, 2021).

Ce groupe d'acteurs — qui n'est pas exhaustif — a en commun des croyances dans le caractère révolutionnaire des technologies numériques et dans l'enjeu vital de prendre le tournant de la nouvelle économie (capitalisme de plateforme) et dans la validité des « solutions » évoquées plus haut (recours intense aux acteurs privés, dont les *startups*, État-plateforme).

1.2.2. Hauts fonctionnaires et structures souples tournés vers le potentiel des données

Le second groupe est celui d'acteurs étatiques transversaux « qui disposent d'assez de ressources et de légitimité pour revendiquer de fixer des règles transversales à tous les ministères et donc pour énoncer puis imposer un plan de réforme du système administratif » (Bezes, 2009). C'est le cas de hauts fonctionnaires qui occupent des positions au ministère de l'Économie et des Finances comme Thomas Cazenave, directeur de cabinet adjoint du ministre de l'Économie, de l'Industrie et du Numérique, lorsqu'il co-édite *L'État en mode start-up* (2016).

C'est le cas également de la structure Etalab laquelle, aux termes d'un processus dont il n'est pas utile de restituer ici tous les détails, est parvenue à se positionner au centre de la politique d'ouverture des données : notamment à travers la fonction d'administrateur général des données (AGD) créée en 2014, qui échoit à son directeur, Henri Verdier. Rattaché au secrétariat général pour la modernisation de l'action publique (SGMAP) comme Etalab, l'AGD a pour mission de « coordonner l'action des administrations en matière d'« inventaire, de gouvernance, de production, de circulation et d'exploitation des données » » (Goëta, 2016 : 80), ce qui le place véritablement au cœur de la politique de l'*open data*. En 2015, toujours au sein du SGMAP, l'AGD, Etalab et la mission du SGMAP qui deviendra l'incubateur de *startups* d'État (beta.gouv.fr) sont intégrés avec la direction interministérielle du système d'information et de communication de l'État (Disic) dans la nouvelle Direction interministérielle du numérique et du système d'information et de communication de l'État (Dinsic)⁹ à la tête de laquelle est placé Henri Verdier. Etalab participe alors directement à la

⁹ En 2019, la Dinsic devient la Dinum.

fabrique de la loi Lemaire, notamment à l'écriture de l'article sur le service public de la donnée où il peut faire juridiquement acter la notion d'écosystème de la donnée (Bellon, 2021 : 49-50). Cette position centrale, progressivement acquise, par la structure Etalab est très liée à l'activisme de son directeur. Dans le cas français, Etalab et la Dinsic, à travers les notions d'« État-plateforme » et d'« écosystème de la donnée » sont parvenus à devenir des acteurs incontournables et puissants de l'*open data* (Chevallier, 2018). Ils se sont appuyés sur les services de Matignon auquel ils sont rattachés directement (Alauzen, 2019), et en lien avec des ministères et secrétariats d'État qui portent une dimension transversale de réforme de l'appareil administratif, ainsi que du ministère de l'Économie et du Secrétariat d'État au numérique.

1.2.3. Disruption dans le pouvoir exécutif

Le troisième groupe d'acteurs importants dans la configuration concerne l'exécutif. Pendant le mandat de François Hollande, dans les gouvernements Ayrault, Valls et Cazeneuve, le numérique n'est plus une compétence parmi d'autres, mais la responsabilité bien identifiée soit d'une ministre déléguée soit d'une secrétaire d'État. Le projet de loi pour une République numérique sera d'ailleurs porté par Axelle Lemaire, secrétaire d'État au numérique, en lien direct avec Etalab (Bellon, 2021). Pour autant, les « réformateurs numériques » restent en position institutionnelle de marginalité par rapport à l'ensemble de la structure politico-administrative : parce que leurs parcours sont atypiques et qu'ils ont peu de ressources administratives, ils ne sont pas vraiment pris au sérieux ni ne parviennent à intéresser les autres ministères. Ils subissent aussi la concurrence de leur autorité de tutelle, le ministère de l'Économie alors dirigé par Emmanuel Macron. L'enjeu du numérique est bien compris par celui-ci qui prépare une loi de soutien aux entreprises du secteur de l'économie du numérique, dont l'effet est de dépouiller le projet législatif porté par Axelle Lemaire, lequel « se réduit comme peau de chagrin », passant « de 90 à 35 articles en quelques mois » (Bellon, 2021 : 35-36), alors qu'il était en gestation puis en préparation depuis plusieurs années au sein du secrétariat d'État. Emmanuel Macron recherche surtout la dimension de levier économique du numérique quand l'approche de la secrétaire d'État est plus large et mise aussi sur les perspectives civiques et politiques d'empouvoirement des citoyens par le numérique. Le recours à la méthode de la consultation citoyenne en ligne dans la préparation du projet de loi peut être analysé comme un marqueur de cet attachement aux enjeux civiques du numérique, mais aussi comme un « coup politique » (Bellon, 2021 : 38) pour parvenir à s'imposer dans un espace bureaucratique et politique concurrentiel.

Quoi qu'il en soit l'élection d'Emmanuel Macron à la présidence de la République, au printemps qui suit l'adoption de la loi Lemaire, place la stratégie de la disruption et de l'État-plateforme au premier plan. En tant que ministre de l'Économie, candidat à l'élection présidentielle puis Président de la République, Emmanuel Macron incarne lui-même la disruption chère aux mondes du numérique et de l'innovation. Mot d'ordre de sa campagne présidentielle, elle-même « disruptée et disruptive » (Martigny, Strudel, 2019 : 128), elle signale la rupture brutale qui menace les positions établies, rebat les cartes au profit de néophytes. L'effet miroir est évident. L'éloge de la disruption c'est l'éloge des entrants, jeunes entrepreneurs qui entendent profiter du tournant technologique pour s'imposer aux dépens

des entreprises établies et de ceux qui ont intérêt au *statu quo* ; tout comme Emmanuel Macron, les marcheurs entendent s'imposer en politique et opérer un renouvellement, qui prend de court les cadres et élus des partis politiques classiques (Rouban, 2018). L'éloge de la disruption c'est aussi une mythologie qui valorise les qualités associées à la jeunesse : audace, ambition, dynamisme et énergie, par celui qui sera lui-même le plus jeune président de la République française. Le mouvement En Marche ! peut, quant à lui, être qualifié de « *startup* politique » puisqu'il accumule les traits caractéristiques¹⁰ de cette forme entrepreneuriale (Gris, 2019 : 243). L'adhésion d'Emmanuel Macron et de son mouvement politique à l'idéal entrepreneurial et en particulier à celui des *startups*, s'explique à la fois par la trajectoire personnelle et les capitaux accumulés par E. Macron, mais aussi par la composition sociodémographique des soutiens économiques, électeurs et élus du mouvement (Offerlé, 2019 : 91). Ses ministres et entourages ministériels font la part belle à des « technocrates d'un nouveau genre, ayant cumulé les expériences au sein de la haute administration et dans le secteur privé », à tel point que la technocratie macronienne apparaît « assez éloignée de l'État bureaucratique des débuts de la V^e République » et « en phase avec un temps marqué par le New Public Management » (Behr et Michon, 2020 : 747 ; voir également Eymeri-Douzans, 2019), comme on a pu le pointer plus haut, à travers différents profils dont celui d'Henri Verdier, entrepreneur numérique devenu directeur d'administration centrale.

Dans tous les cas, il est clair que l'intérêt d'E. Macron et de ses équipes pour le numérique ne relève pas seulement d'un opportunisme politique ou d'un engouement éphémère. Comme ministre de l'Économie puis comme président de la République, il met l'accent sur la révolution des/par les technologies, à la fois sous l'angle de la dynamisation de l'économie nationale, de la valorisation du potentiel d'innovation français, de la réforme des administrations et des transformations culturelles et sociétales qui vont avec. Il épouse donc les valeurs et motivations plurielles qui ont amené à l'ouverture des données publiques, et qu'incarnent les *startups* d'État comme Etalab. Il a d'ailleurs préfacé l'ouvrage collectif *L'État en mode start-up* dont il a été question plus haut.

Tout juste élu, il participe au salon Vivatech¹¹ où il affirme : « Je veux que la France soit une "nation start-up", ce qui veut dire à la fois une nation qui travaille avec et pour les *startups* et une nation qui pense et bouge comme une start-up » (cité par Dolez, Frétel et Lefebvre, 2019, p.17). Le pays est comparé à petites structures, agiles, innovantes, lancées dans une compétition pour dénicher les gisements de profits de demain.

¹⁰ Mené par la présence d'un dirigeant « stratège visionnaire » qui s'appuie sur une équipe jeune possédant des compétences managériales, ce mouvement connaît une montée en puissance fulgurante. Il joue à plein de l'« effet de levier » pour accroître rapidement ses ressources électorales, militantes et financières ; et fonctionne selon « un mode de management participatif, faiblement hiérarchisé » (Gris, 2019 : 243).

¹¹ Salon consacré aux technologies, Paris, 15 juin 2017. Pour une analyse ethnographique de ce salon (Rosenthal, 2018).

D'ailleurs ses réalisations vont clairement dans ce sens. Des structures sont mises en place pour promouvoir la *startup nation* : la station F, incubateur géant¹² représente une pièce majeure dans l'élaboration de l'écosystème de l'innovation aux côtés « d'autres équipements, dispositifs et institutions : entre autres, le label [...] "French Tech" attribué à des "écosystèmes" innovants dans l'hexagone, les services de Business France à l'international, les mannes de la Bpifrance, les subventions des grands organismes de recherche publique comme le CNRS, ou encore les incitations fiscales [...] via le crédit d'impôt recherche ou le crédit d'impôt innovation » (Quijoux et Saint-Martin, 2020 : 16). La transformation des modes de financement et de structuration de la recherche est, de ce point de vue, menée depuis plusieurs années vers une plus grande interpénétration entre recherche publique et privée, à laquelle E. Macron et ses équipes adhèrent clairement (Guillaume et Macron, 2007).

Sur le plan de la réforme de l'État plus précisément, E. Macron lance un nouveau grand chantier intitulé *Action publique 2022*, visant la modernisation de l'action publique dans une optique d'efficacité et d'efficience, mais aussi de plus grande réactivité et proximité avec les citoyens. Un fonds de 700 millions d'euros sur 5 ans est dédié à la transformation numérique de l'administration (Santiso, 2021 : 249), au carrefour entre économie de moyens et plus grande horizontalité dans les rapports gouvernants-gouvernés.

La politique d'*open data*, politique constituante, se met donc en place autour d'un système de croyances marqué par : le caractère révolutionnaire des technologies numériques qui transforment non seulement l'économie, mais l'ensemble des rapports sociaux ; l'enjeu de l'ouverture des données perçues comme un nouvel or noir ; la place du secteur privé et de l'innovation via la forme de la *startup* comme solution pour prendre à bras-le-corps ce tournant ; la doctrine d'un État-plateforme qui favorise la réutilisation des données sans intention d'usages. Des méthodes et instruments d'action publique sont privilégiés, avec l'appui de structures étatiques qui deviennent propriétaires de ce problème public, tel Etalab. La loi en tant qu'instrument privilégié à la disposition des gouvernants énonce, affiche et met en scène ce système de croyances en organisant l'ouverture des données et en la promouvant au rang de nouvelle norme publique.

1.3. L'expression d'une nouvelle norme publique

Emmanuel Macron s'est imposé comme un président « disruptif ». Pour autant, lorsqu'il était ministre de l'Économie, il n'était pas parvenu à remporter un arbitrage présidentiel favorable au projet de loi sur les nouvelles opportunités économiques (NOÉ) que ses services avaient préparé courant 2015¹³. Le projet de loi porté par Axelle Lemaire et dont

¹² Il s'agit de la Halle Freyssinet à Paris, 34 000 m² rénovés. Station F est inaugurée en juin 2017 en présence de son fondateur, l'entrepreneur Xavier Niel et du Président de la République.

¹³ Le projet sera finalement abandonné, sans que les mesures liées à l'économie numérique, qui avaient été ôtées du projet de loi pour une République numérique, soient réintégrées dans celui-ci. Elles seront finalement injectées dans les projets de loi portés par Myriam El Khomri, ministre du Travail. Cf. Charles Foucault, « La France n'aura pas sa grande loi sur l'économie numérique et c'est moche », *L'usine digitale*, mis en ligne le 20 janvier 2016, consulté le 7 avril 2021. <https://www.usine-digitale.fr/editorial/macron-2-la-france-n-aura-pas-sa-grande-loi-sur-l-economie-numerique-et-c-est-moche.N374834>

la gestation avait été longue et tortueuse a finalement été inscrit à l’agenda parlementaire et voté en 2016.

La loi pour une République numérique énonce le principe de l’*open data* par défaut c’est-à-dire qu’elle proclame le caractère public, gratuit et ouvert de toutes les données publiques. « Nous avons fait de l’ouverture le principe, et du secret, l’exception : c’est un véritable changement culturel¹⁴ », explique Axelle Lemaire. Si cette législation marque la reconnaissance publique de l’impératif d’ouverture des données, elle approfondit un mouvement, entamé il y a une quarantaine d’années et marqué par l’empilage de différents textes constituant à partir des années 1970 un « droit des données et des informations publiques » (Yolka, 2016, p. 79). Elle s’inscrit aussi dans une forme de filiation avec le droit européen relatif à la réutilisation des données publiques (Teresi, 2017 : p. 88), qui envisage celles-ci sous l’angle du droit de la concurrence et à travers leur dimension marchande dans le cadre de l’économie numérique.

Cette loi comporte plusieurs chapitres traitant de différents aspects du numérique¹⁵ et modifie quatorze codes, ce qui marque bien sa dimension transversale. S’agissant du statut des données publiques, une certaine complexité demeure (Camus, 2018) puisqu’un « service public de la donnée » a certes été créé (Cluzel-Métayer, 2018), mais qu’à ce stade il est restreint aux seules « données de référence », lesquelles portent sur neuf jeux de données définis par décret¹⁶. Les décisions de justice n’en font pas partie.

En outre, le principe de l’ouverture par défaut comporte d’importantes limites qui sont liées à l’interdiction faite à l’administration de diffuser des données à caractère personnel et à des tiers de les réutiliser : « la protection de la vie privée prime en effet sur l’exigence d’information publique » (Cluzel-Métayer, 2017 : p. 104). L’effet croisé des lois CADA et CNIL retient trois exceptions seulement : le consentement de la personne, l’existence d’une obligation légale à la diffusion et l’anonymisation des informations diffusées (*ibid.*). Pour une diffusion de données massives telle que conforme à la logique de l’*open data* prévue par la loi Lemaire, et afin que l’administration échappe au recueil du consentement des personnes,

¹⁴ Interview, « *Open data* : “Le mouvement qui s’enclenche est inéluctable” – Axelle Lemaire », *La gazette des communes*, publié le 5 octobre 2018, consulté le 26 février 2021. <https://www.lagazettedescommunes.com/585196/open-data-le-mouvement-qui-senclenche-est-ineluctable-axelle-lemaire/>

¹⁵ En l’occurrence, la loi comprend trois chapitres. Le premier concerne la circulation des données et du savoir et traite notamment de la question du droit d’auteur. Le second porte sur le renforcement des droits des citoyens dans l’espace numérique en consacrant la dimension personnelle des droits. Le dernier prévoit des dispositions destinées à faciliter l’accès au numérique.

¹⁶ Aux termes du décret n° 2017-331 du 14 mars 2017 relatif au service public de mise à disposition des données de référence, il s’agit de la base Sirene des entreprises (Insee), du Répertoire national des associations (ministère de l’Intérieur), du Plan cadastral informatisé (DGFIP), du Registre parcellaire graphique (Agence des services et de paiement), du Référentiel à grande échelle (IGN), de la Base adresse nationale (IGN, La Poste, OpenStreetMap-France), du Référentiel de l’organisation administrative de l’État (DILA), du Répertoire opérationnel des métiers et des emplois (Pôle-emploi) et du code officiel géographique (Insee). Voir (Cluzel-Métayer, 2018 : p.494).

l'anonymisation des données représente le procédé privilégié, comme on le verra plus loin à propos de la justice.

La loi pour une République numérique pose donc un jalon important sur le fond, mais aussi sur la forme. En effet, l'avant-projet de loi a fait l'objet d'une concertation publique en ligne, qui, sans révolutionner complètement la fabrique législative, a représenté une innovation notable, signant « une victoire relative des réformateurs numériques au sein de l'État » (Bellon, 2021 : 50) et offrant un précédent pour de futures consultations numériques des citoyens (par exemple le Grand Débat national de 2019). Il ressort de la fabrique gouvernementale que, tout en adoptant des mesures destinées à ouvrir davantage l'accès aux données publiques, le texte finalement voté n'est pas allé jusqu'à instaurer la notion de « domaine commun informationnel », qui était présente au stade de l'avant-projet de loi,¹⁷ mais n'avait pas été reprise dans le projet de loi finalement déposé à l'Assemblée nationale¹⁸.

De ce point de vue, la loi Lemaire incarne et prolonge la tension entre ouverture et gratuité des données à destination des citoyens et pour des entreprises privées. Elle met en place le principe d'ouverture et de gratuité des données publiques pour créer les conditions favorables à l'émergence d'une économie numérique source de richesses pour des entreprises et *startups* proposant de nouveaux services payants. Ce paradoxe n'est pas surprenant dès lors que l'on sait combien la culture numérique est elle-même traversée par des tensions. L'enjeu de la liberté de circulation des connaissances et de leur gratuité est au fondement d'une culture numérique dans laquelle la technologie est vue comme émancipatrice, au service des individus et de la société (Cardon, 2019 : p. 37-54). Pourtant, dès les pionniers d'Internet, cette culture est « créative, mais marchande, communautaire, mais globalisée » (Cardon, 2019 : p. 77).

Ces tensions se retrouvent au sein de la politique de l'*open data* et parmi les entrepreneurs de cause de l'ouverture du droit, et ce, d'autant plus qu'en réalité, s'agissant des décisions de justice, la loi Lemaire énonce seulement un principe sans organiser du tout sa mise en œuvre. Avant de dessiner les grands moments de cette histoire aux multiples rebondissements, il convient de voir comment des acteurs de la justice se positionnent comme les traducteurs du nouveau répertoire de réforme administrative et apportent le discours du changement dans la justice, en conformité avec le référentiel global de l'*open data* centré sur l'État-plateforme.

II / Un répertoire de réforme promu par Open Law

Les premières traces de l'importation de ce nouveau répertoire de réforme dans la justice ont été pointées par F. Lebaron (2019 : 136-142). Il relève notamment le rôle d'acteurs entrepreneurs de changement — comme l'association Open Law. Historiquement, au moins

¹⁷ Alain Bensoussan, « Projet de loi numérique : vers la notion de communs informationnels ? », 30 décembre 2015, consulté le 3 mars 2021. <https://blog.lefigaro.fr/bensoussan/2015/12/projet-de-loi-numerique-vers-la-notion-de-communs-informationnels.html>

¹⁸ Projet de loi pour une République numérique n° 3318, déposé le 9 décembre 2015.

deux structures existent qui visent à ouvrir et mettre en réseau le droit à partir de l'information et l'informatique juridiques (1). Elles sont rejointes par Open Law, une association explicitement créée pour importer l'*open data* dans la justice (2). Elles ne participent pas forcément toutes à la « nébuleuse réformatrice » (Topalov, 1999) qui se forme autour de l'ouverture des décisions de justice, mais sont interconnectées et représentent des espaces d'échange d'informations, de production d'idées nouvelles et de mise en réseau. Elles jouent un rôle dans les différentes controverses ouvertes par la mise à disposition et la réutilisation de décisions de justice (3).

2.1. De Juriconnexion et l'Adij à Open Law

Juriconnexion et l'Adij réunissent surtout des professionnels de la documentation juridique (bibliothécaires, documentalistes, etc.), concernés par les développements de l'informatique puis du numérique, en particulier pour la constitution de bases de données permettant une recherche automatisée dans d'importants corpus de sources de droit (Bories, 2003 ; Bourcier, 1995 ; Catala, 1998). L'Adij, fondée en 1970 par un bâtonnier, devient un lieu où gravitent quelques personnages fameux de l'informatique juridique des années 1980 : le conseiller d'État Lucien Mehl, un moment président, et vraisemblablement aussi Pierre Catala. Après une période de mise en sommeil, elle semble aujourd'hui être revivifiée, et active, à travers l'organisation d'événements sur lesquels elle communique surtout sur Twitter (son site, lui, bien que de facture récente semble peu alimenté). Fondée en 1988¹⁹, Juriconnexion se présente, quant à elle, comme une association indépendante qui réunit des utilisateurs de produits d'information juridique²⁰, dont les objectifs visent à « contribuer à améliorer la qualité des produits et services offerts par les éditeurs juridiques grâce à une action de type consumériste ; assurer une veille sur les nouveaux produits édités et suivre l'évolution des bases existantes afin d'aider les utilisateurs à rester performants ; animer un réseau de professionnels de l'information juridique²¹ ».

Le projet « Open Law*, le Droit ouvert » a émergé en 2014 dans le cadre de la préparation de l'Open World Forum, événement international dédié aux acteurs de l'innovation ouverte, suite à la diffusion en *open data* des bases Légifrance par la Direction de l'information légale et administrative (Dila). Association de statut loi 1901, Open Law est formellement créée en 2015 à l'initiative de la Dila, d'Etalab et du Numa²², afin de « faire la démonstration de la pertinence des modèles ouverts, agiter le marché du droit et stimuler la création de nouveaux usages, produits et services²³ ». Concrètement, elle porte dans le domaine du droit la philosophie de l'écosystème de la donnée, développée par Etalab,

¹⁹ Par dix personnes sur lesquelles on ne dispose pas d'informations : Noëlle Choublier, Michèle Come, Francine de Doncker, Didier Frochot, Anny Maximin, Francine Nicot, Annie Plazanet, Patrice Sotero, Françoise Valot, et Clotilde Scemama.

²⁰ Art. 2 des statuts votés le 9 avril 2018, <https://www.juriconnexion.fr/statuts-de-lassociation-au-4-juin-2009/>, consulté le 2 avril 2021.

²¹ <https://www.juriconnexion.fr/a-propos/>, consulté le 2 avril 2021.

²² Lieu parisien de coworking, consacré à l'innovation et aux *startup*, fondé en 2013 à l'initiative de la Silicon Sentier, association d'accompagnement des jeunes pousses de l'innovation.

²³ <https://openlaw.fr/association/notre-gouvernance>, consulté le 31 mars 2021.

contenue dans le portail data.gouv.fr et instituée par la loi Lemaire. Regroupant juristes et experts des données juridiques, elle « pilote des programmes d'innovation collaborative avec des partenaires publics et privés et anime au quotidien la communauté française de la *legaltech* — la communauté des développeurs s'intéressant au marché du droit — et des juristes innovateurs » ([Pierre], [Lavaudan], Saint-Aubin, 2016 : 23). Comme l'indique cette citation, l'enjeu d'Open Law est d'explorer les nouveaux jeux de données devenus libres d'accès ; d'inciter à leur appropriation notamment en direction des développeurs informatiques et à l'articulation entre public et privé.

2.2. Des parcours hybrides et multipositionnés : les courtiers

Un regard cursif sur les instances de ces associations montre qu'au moins trois personnes sont actives dans plusieurs d'entre elles : [Georges Lavaudan], à la tête de la société de conseil [Conseil & droit], est administrateur de Juriconnexion, animateur de la liste électronique d'information de cette association, mais aussi membre et administrateur d'Open Law. [Nicolas Mariano], secrétaire général de l'ADIJ de 2017 à 2019, est aussi membre fondateur d'Open Law. De même, [Arthur Pierre] est successivement membre fondateur d'Open Law et administrateur de l'Adij (de 2017 à 2019). Bien entendu il faudrait aller plus loin, reconstituer les trajectoires, réaliser des entretiens, et travailler sur les archives de ces organisations, pour mieux en saisir les positionnements et relations réciproques, mais cela met sur la piste d'un milieu où se rencontrent et se socialisent des personnes aux parcours hybrides et partageant un système de croyances relatif aux bienfaits de l'ouverture du droit et à l'intérêt de l'informatique et des technologies numériques pour les professions qui travaillent avec lui.

Si ces associations, et en premier lieu Open Law, visent à la mise en relation d'acteurs publics et privés, elles sont animées par des personnes qui naviguent elles-mêmes entre structures privées et institutions publiques, par des allers-retours durant leurs carrières ou via des multiactivités et multipositionnalités synchrones. Caractéristiques de ce type de parcours, sont par exemple [Nicolas Mariano], [Morgan Forestier] (dont il sera question plus loin), [Arthur Pierre] ou bien encore [Georges Lavaudan].

Ils se positionnent et agissent doublement comme des intermédiaires : entre le monde de l'informatique et du numérique d'un côté et celui du droit de l'autre, mais aussi entre le monde de l'administration, du service public d'une part et celui de l'entreprise privée et marchande d'autre part. L'articulation de plusieurs mondes est typique du travail des « intermédiaires en politique » (Nay et Smith, 2002). Parmi eux, on peut distinguer analytiquement « l'activité du "généraliste" qui consiste à construire du "sens commun" entre des milieux institutionnels qui ne recourent pas aux mêmes savoirs et aux mêmes représentations [... de celle] du "courtier" qui consiste à rechercher des solutions acceptables entre des groupes éloignés qui peuvent trouver un avantage à coopérer même s'ils ne poursuivent pas les mêmes objectifs et n'ont pas les mêmes intérêts » (*ibid.* : 13).

Open Law est officiellement et explicitement positionnée comme une structure de médiation au sens où lui est dévolue la fonction de créer des ponts et des synergies, dans une

optique de médiation cognitive et de construction de compromis. Sa fonction relève à la fois de l'élaboration d'un sens commun autour de l'idée d'un droit ouvert par les technologies numériques ; de la mise en relation d'acteurs institutionnellement éloignés (informaticiens, *data scientists* ; juristes et acteurs administratifs) et de la construction de projets communs (où des hybridations et compromis sont concrètement réalisés). Forts de réseaux multiples qui leur offrent de nombreuses ressources, les animateurs d'Open Law circulent d'une arène à l'autre, et travaillent à l'émergence et à la réalisation d'idées passerelles.

[Nicolas Mariano] est un exemple particulièrement emblématique de ces profils. Formé en droit de la propriété intellectuelle et après avoir été consultant-formateur autour des questions de droit du numérique, de propriété intellectuelle, des données, des TIC, il devient chargé de recherche en valorisation de la donnée à Orange Labs. Il entre alors comme chef de bureau éditorial/numérique au ministère de la Justice (2007-2013), tout en assurant des fonctions de professeur en droit des données personnelles, *legaltech* et e-Justice à l'École de formation du barreau. Puis il devient responsable du pôle stratégie-innovation de la Dila (rattachée aux services du Premier ministre), pour le pilotage de la mise en *Open data* du droit français, notamment via le projet Open Law dont il est membre fondateur. Il restera à ce poste jusqu'en décembre 2016, date à laquelle il rejoint la MGEN (assurance privée qui remplit des missions de service public sur délégation du ministère de l'Éducation nationale), comme directeur du programme éducation numérique/innovation collaborative jusqu'en 2017. Il devient alors PDG de [toutnouveau].com, une solution *legaltech* spécialisée en propriété intellectuelle (janvier 2017-février 2020) puis il cofonde [Duprazleg] une autre *legaltech* à destination des juristes, développée avec des juristes et des développeurs informatiques. La société est fondée en décembre 2016 et il en est le PDG depuis mars 2017. Il devient également avocat-conseil sur les sujets liés au numérique, aux données, à l'IA, à l'innovation et à la propriété intellectuelle, depuis avril 2019²⁴.

Infrastructure de coopération orientée vers la création de communs numériques, Open Law se positionne en animateur central de la communauté : elle joue un rôle déterminant et très visible dans le montage d'une multitude de projets, comme on va le voir.

2.3. Compétition, collaboration et hackathon : le « catéchisme » de l'innovation numérique

Open Law peut être analysé comme un acteur qui contribue à une forme de gouvernement à distance, à travers la promotion de certaines manières de faire et de certains outils d'action publique.

Cette association est créée afin de fournir « un cadre de discussion et de rencontre » (*ibid.*) entre des acteurs épars pour définir des priorités communes sur « les Communs à faire émerger dans le domaine de l'information juridique ». Elle vise aussi à offrir une structure de support pour financer une telle entreprise, sur le modèle du « pot commun » alimentant ensuite un fonds de dotation ; et à assurer la gouvernance de cette nouvelle infrastructure.

²⁴ Sources : LinkedIn, consulté le 11 mai 2021.

Elle intervient aussi dans la « production de ressources communes (référentiels, ontologies, livres blancs, modèles de statuts, cahiers des charges, bases de données, développements Open Source, etc.), qui, à plus long terme, alimentent tous ceux qui cherchent à innover et développer de nouveaux services ou concepts²⁵ ». Open Law propose un intéressement à produire du collectif puisque le statut de « contributeur rémunéré aux communs », permet de rémunérer ceux qui élaborent des communs numériques, sans devoir créer des licences qui en limiteraient ensuite l'accès. Le patrimoine commun, développé par l'association et promu sur son site, comporte des *datasets* (base de données) d'apprentissage destinés à entraîner les algorithmes autoapprenants basés sur l'intelligence artificielle²⁶, une charte éthique pour un marché du droit en ligne et ses acteurs²⁷ ou encore des travaux de doctrine en libre accès²⁸.

Open Law se présente très clairement comme un lieu de rencontre et de valorisation des *legaltechs* et en particulier l'intelligence artificielle (IA) comme en témoigne le fait que celles-ci soient au cœur de la rubrique « Les travaux en cours²⁹ » de son site. Les objectifs comme les méthodes et le vocabulaire d'Open Law reposent sur le « catéchisme » de l'innovation, partagé avec Etalab. La compétition étant vue comme stimulant l'innovation et la performance, des concours sont organisés, couronnés de prix, valant labels et dotés de récompenses financières. Le *hackathon* en est une forme propre à l'univers du numérique importée dans l'action publique. « Concours d'innovation numérique se déroulant sur une courte durée » (Marquet, 2016 : 147), il consiste à créer des sessions de travail en commun (Lara et Lockwood, 2016), où, dans un temps réduit, plusieurs équipes relèvent un défi d'innovation et de programmation (Greselle-Zaïbet, Kléber et Dejoux, 2018), assorti ou non d'un prix. Le défi « Vulgarisation du droit » lancé avec le partenariat d'AXA protection juridique³⁰ ou le prix Open Law en sont de bons exemples. Le *hackathon* peut être analysé comme un outil de l'action publique (Lascoumes et Le Galès, 2004) au sens où ce dispositif incarne et traduit des principes et des normes, un corps de valeurs sur la manière de produire l'innovation dans le champ du droit. C'est aussi une forme de gouvernement par la mise en concurrence et l'attribution de prix et de labels mettant en scène et valorisant la dimension apparemment volontariste, proactive et non contrainte par laquelle l'action publique opère

²⁵ Site de l'association Open Law, <http://openlaw.fr/association/qui-sommes-nous>, consulté le 16 février 2021.

²⁶ <http://openlaw.fr/travaux/communs-numeriques/ia-droit-datasets-dapprentissage>, consulté le 11 mai 2021.

²⁷ <http://openlaw.fr/travaux/communs-numeriques/charte-ethique-pour-un-marche-du-droit-en-ligne-et-ses-acteurs>, consulté le 11 mai 2021.

²⁸ <http://openlaw.fr/travaux/communs-numeriques/open-doctrine>, consulté le 11 mai 2021.

²⁹ Qui sont : « Charte éthique pour un marché du droit en ligne et ses acteurs » ; « Programme 6 – former le juriste de demain » ; « Former le Juriste de demain » ; « I.A. Définitions » ; « IA & Droit » ; « Identité numérique : e-ID » ; « Interprofessionnalité » ; « Open Doctrine » ; « Open Law Badges » ; « Smart Contract Academy » ; « Think LeD: Legaldesign ».

³⁰ Les lauréats de produits innovants sont récompensés par deux prix (5 000 €/2 500 €). Information diffusée sur LinkedIn, mis en ligne le 10 juillet 2015, consulté le 9 avril 2021. <https://www.linkedin.com/pulse/open-law-top-d%C3%A9part-pour-le-d%C3%A9fi-vulgarisation-du-droit-pascal-mendak/>.

une régulation à distance et par l'incitation (Bergeron, Castel et Dubuisson-Quellier, 2014 ; Epstein, 2005).

Graphique 3.1. Les anglicismes dans le menu du site de l'association Open Law (bandeau en haut de page)



2.4. Une « communauté de la donnée » pour un service public des données juridiques

Comme on l'a vu, Open Law se positionne sur l'ouverture du droit de manière générale, mais il participe aussi directement à l'animation d'une « communauté de la donnée », dans la foulée des articles 20 et 21 de la loi Lemaire relatifs aux décisions de justice. Organisateur principal du programme court intitulé « Open "Case" Law », il rassemble acteurs sectoriels de la justice et généralistes de la politique de l'*open data* autour de l'enjeu de l'ouverture des décisions de justice. La mission Etalab, le ministère de la Justice, la Cour de cassation, la Cour des comptes, le Conseil d'État, la CNIL, l'INPI et l'Ordre des avocats de Paris sont conviés à travailler ensemble (Open Law, 2017 : 20) dans le cadre d'ateliers, d'un *fablab*³¹ et d'un *hackathon*, qui ont eu lieu en septembre et octobre 2016, c'est-à-dire dans la dernière ligne droite du processus de fabrication de la loi Lemaire.

Un livre blanc publié en janvier 2017 restitue la démarche : il envisage l'*open data* « comme une véritable opportunité de repenser la gouvernance de l'ensemble de la chaîne de production des données jurisprudentielles » (*ibid.* : 40). À court terme, les enjeux du programme portent sur la réalisation d'un diagnostic de ce qui existe (quel régime juridique des décisions de justice, quelles banques de données, etc.) et sur la création de conditions favorables au dialogue entre acteurs publics et privés... dans une perspective prospective visant à « anticiper de manière collective les enjeux techniques et juridiques relatifs à l'ouverture de la jurisprudence du fond », et ce, dans la perspective de dessiner « ce que pourrait être, demain, le service public de la donnée publique juridique. » (*ibid.* : 19) Un atelier du programme est d'ailleurs consacré explicitement à ce thème. En somme, l'accent est mis sur la construction d'une « communauté de la donnée », dont il s'agit d'éclairer le fonctionnement précis pour mieux le repenser ensuite. Les propositions avancées consistent notamment à faire qualifier les données jurisprudentielles de « données de référence » ou

³¹ C'est un tiers-lieu ouvert au public dans lequel sont mis à disposition des ordinateurs, outils et machines numériques (imprimantes 3D, etc.) pour permettre la conception et la production d'objets ou de réalisations.

leur équivalent (*ibid.* : 46), et ce, afin de donner lieu à la création d'un service public, comme c'est le cas pour d'autres jeux de données sur lesquels Etalab a déjà travaillé.

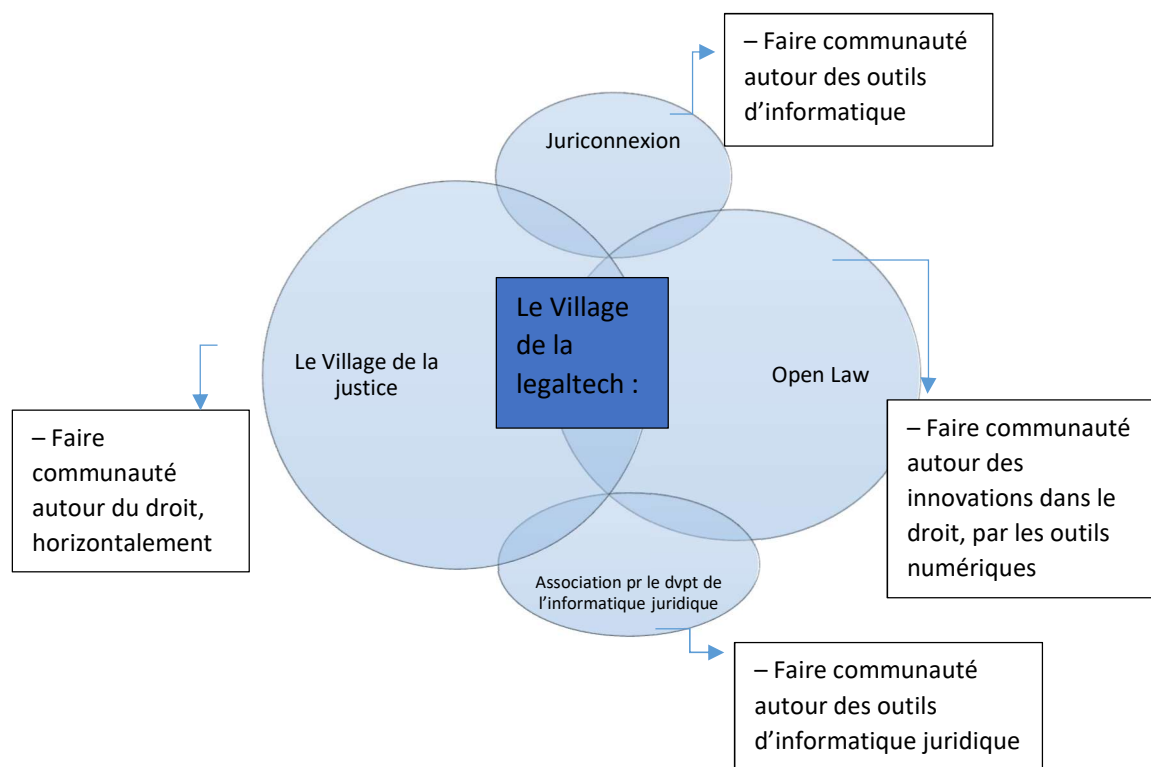
En réalité, dans ce livre blanc, un glissement est opéré d'un service public des données jurisprudentielles à un service public des données juridiques. Le discours d'Open Law maximise l'effet rupture de la loi Lemaire dans la justice sur le plan discursif en affirmant qu'elle « consacre un véritable changement de paradigme. » (*ibid.* : 17) et en imaginant un *open data* qui ne porterait pas seulement sur les décisions de justice, mais qui engloberait toutes les données juridiques : textes législatifs, etc. (*ibid.* : 62) Cet élargissement au droit emporte des effets importants comme celui de relativiser la spécificité des régimes existants de publicité des décisions de justice. L'ambition est, en outre, de se projeter dans un service public de la donnée juridique, que n'instaure pourtant pas la loi Lemaire — laquelle a certes créé ce service public applicable aux « données de référence », mais dont les décisions de justice ont été exclues. Open Law se positionne et intervient donc en faisant un travail de désépécification des décisions de justice et de médiation/mise en réseau d'acteurs généralistes et sectoriels d'une part ; privés et publics d'autre part.

2.5. Le Village de la legaltech : un événement-carrefour qui incarne le rapport public-privé et les valeurs d'ouverture du droit et d'innovation par les technologies

Open Law co-organise également de nombreux événements avec des partenaires de l'animation du monde juridique, comme le Village de la justice, avec lequel il partage des valeurs d'horizontalité³², d'où leur collaboration dans l'organisation du Village de la legaltech.

Graphique 3.2. : Les relations entre les différentes structures d'action collective autour de l'informatique juridique puis de l'open data du droit

³² Fondé en 1997 par une société privée, LegiTeam éditions, il est organisé autour de la circulation de l'actualité juridique *gratuite*. Sa spécificité à l'égard des outils de diffusion des grands éditeurs juridiques classiques — est double : la recherche d'une communication horizontale entre tous les types de praticiens du droit, professionnels ou non afin *de créer une communauté*, et ensuite celle d'une circulation de l'information juridique tout aussi *horizontale* à destination des citoyens. Cette absence de stratification et/ou de cloisonnement est revendiquée à travers ce qui sonne presque comme un slogan : « Nous ne pratiquons ni élitisme ni discrimination au titre ! [sic] », consulté le 2 avril 2021, <https://www.village-justice.com/articles/Mairie-Village,60>



Le Village de la legaltech est un événement qui existe depuis 2016 dont les organisateurs sont Open Law et le Village de la justice, appuyés par une série de partenaires et sponsors tous issus du secteur privé. Créé en 2016, année de vote de la loi Lemaire, c'est un rendez-vous annuel, destiné à créer une culture commune de l'innovation par les technologies dans le droit. L'édition de 2021 est rebaptisée *Les transformations du droit* avec le sous-titre « Le grand RDV des professions du droit qui innovent — Évolution 2021 du 5^e Village de la legaltech³³ », ce qui atteste d'un repositionnement vers un propos plus généraliste et donc d'une ambition de devenir plus central dans le monde du droit (les transformations du droit, pas seulement sous l'angle des technologies).

Des observations réalisées pendant les deux jours de la quatrième édition de ce salon, en 2019, il ressort que cette manifestation est conçue par ses organisateurs comme un rendez-vous important de construction de la communauté de la *legaltech*. Ces journées incarnent la dimension d'horizontalité et d'ouverture propre au Village de la justice : ciblant les professionnels, mais ouvertes à tous, y compris le public et complètement gratuites.

³³ La 5^e édition, qui aurait dû se tenir en novembre 2020, période de second confinement, n'a pas eu lieu. Il a été annoncé qu'en 2021, Le Village de la legaltech prendrait une nouvelle forme. Initialement programmé les 11-12 mars 2021, cet événement a lui-même été reporté aux 18-19 novembre 2021. Il est organisé en 5 villages concomitants définis en fonction soit des publics visés (le Village des innovateurs publics ; le Village de la regtech — destiné aux directions juridiques) soit de thématiques et sous-domaines (le Village de la legaltech ; le Village des trajectoires et carrières professionnelles ; le Village du legal design) <https://transformations-droit.com/>, consulté le 6 avril 2021.

Installé sur deux ensembles modulaires du sous-sol et du rez-de-chaussée de la Cité des sciences à Paris, le salon se présente comme une galerie des nouveautés commerciales : des stands se côtoient, qui proposent outils de gestion et analyse de contrats par IA, applications du protocole *blockchain*, logiciels de production automatique d'écrits authentifiés, plateformes de mise en relation de vendeurs et acheteurs de droits de propriété intellectuelle, etc. En matière d'algorithmes de traitement et d'analyse des décisions de justice, les sociétés [PréviCompute], [Prédroit], [Anajur], [Jurissimo], mais aussi [LegalEditions] (avec [[LegalDonnées+)], sont représentées. Un programme d'animations est également proposé, allant des conférences plénières, débats, ateliers de démonstration aux « sessions de pitch », lors desquelles des présentations de produits se succèdent toutes les cinq minutes. L'ambiance est celle d'une ruche bruyante (musique, annonces au micro) qui construit et diffuse son image en temps réel avec photographe, cameraman et espace studio pour les interviews.

La tonalité générale est aussi celle de la « cool attitude » et de la décontraction entretenue par les organisateurs, assez loin de l'image traditionnelle du monde du droit et de ses professions établies. La conférence d'ouverture lors de laquelle les deux responsables d'Open Law et du Village de la justice, jouent un sketch sur la scène du grand amphithéâtre, déguisés en personnages de jeux vidéo (Mario), en dit long sur la manière dont ce milieu cultive l'esprit potache et plus profondément sur ses valeurs (résistance aux traditions établies, rejet du statu quo). Rompre avec les codes du monde juridique et judiciaire vu comme protocolaire, guindé et poussiéreux, tel est aussi le message de cette communauté.

L'organisation pratique et symbolique du salon est au service de l'incitation au changement des professionnels du droit par l'intégration accrue de technologies numériques. Comme d'autres salons à forte dimension technologique, les « démos » sont une part importante de la mise en scène (Rosental, 2018) : elles sont conçues comme du « marketing technologique », à la manière des shows de Steve Jobs et de la culture de la Silicon Valley. Il faut pouvoir *pitcher* son produit et rendre le public témoin de ce qu'il permet d'accomplir. La « démo » fonctionne comme une épreuve à l'issue de laquelle le produit et la société qui le développe sortent grandis... ou pas. Dans un atelier de démonstration d'une solution commerciale, la technologie a posé des problèmes de connexion. Le public réagit alors diversement, entre interventions individuelles autour de l'enjeu de faire fonctionner la machine, et plaisanteries plus ou moins charitables sur les compétences technologiques de ceux qui, en exposant leur produit, s'exposent eux-mêmes. Cette situation montre que l'enjeu pour les *startups* qui participent activement au salon, c'est la capacité à tenir les promesses avancées dans leurs discours commerciaux. Elle manifeste également la tension permanente entre, d'une part le partage de valeurs sur lequel repose la solidarité de la communauté et, d'autre part les luttes et concurrences pour le prestige, les parts de marché et les clients.

Le salon recouvre en réalité trois dimensions qui ne sont pas naturellement articulées. La dimension communautaire d'abord, qui consiste à cultiver et exhiber des valeurs pour le groupe et à l'attention des autres, celles et ceux qu'il faut convertir au numérique et à l'innovation. La dimension commerciale ensuite, qui revient à faire « du business » : se faire connaître, attirer de nouveaux clients, conclure de nouveaux contrats. La dimension politique

enfin, qui passe par établir des contacts, faire du lobbying auprès d'acteurs politiques ou administratifs présents, en poste au ministère de la Justice, à la Cour de cassation ou encore au Conseil de l'Europe ; voire à aborder directement les enjeux de régulation de l'*open data*, comme ce fut le cas à propos du décret sur l'ouverture des décisions de justice, alors en préparation. Des représentants du ministère de la Justice ont ainsi participé à une table ronde, sur ce sujet.

À travers une variété d'initiatives, Open Law incarne, pratique et fait exister le répertoire de réforme de l'*open data* en se faisant le médiateur de son importation dans le monde du droit et le secteur de la justice. Malgré la coopération de certains acteurs, il se heurte toutefois à des résistances, qui transparaissent dans l'analyse des soubresauts de la mise en œuvre de l'*open data* dans la justice.

III / Trois temps forts dans l'*open data* des décisions de justice

Les articles 20 et 21 de la loi du 7 octobre 2016 pour une République numérique prévoient l'ouverture des décisions de justice. Ils proviennent d'un amendement qui n'a pas été présenté par le gouvernement ; vraisemblablement né du lobbying d'acteurs engagés dans l'*open data*, parmi lesquels Open Law. Comme ils n'ont fait l'objet d'aucune étude d'impact, ni donné lieu à un avis du Conseil d'État, ces deux articles livrent « une promesse de législation et non une législation immédiatement applicable³⁴ ». Cela explique en partie la lenteur et le caractère controversé de l'opérationnalisation de cette loi, en ce qui concerne les décisions de justice. Il faudra attendre 2020, c'est-à-dire quatre ans pour que le décret d'application soit publié. Le décalage est patent entre les acteurs pressés de tirer parti du principe de l'ouverture des décisions de justice — et forts de sa reconnaissance par la loi — et certains acteurs établis du champ qui tentent d'en limiter l'impact sur les pratiques réelles, dans la phase encore indéterminée de sa mise en œuvre.

Le droit est un instrument de gouvernement largement mobilisé dans les politiques publiques, à la fois langue historique de l'État, affirmation symbolique de l'autorité, expression d'une norme publique et moyen concret de mettre en forme des orientations et priorités politiques. Dans le cas de l'*open data*, trois types d'instruments juridiques sont utilisés en 2019-2020 qui dessinent une mise en œuvre contenue de l'*open data* des décisions de justice : un décret (1), une loi (2) et une circulaire (3).

3.1. Un décret qui conforte le pouvoir des hautes juridictions et des magistrats dans la mise en œuvre de l'*open data* dans la justice

Le décret d'application n° 2020-797 du 29 juin 2020 relatif à la mise à disposition du public des décisions de juridictions judiciaires et administratives en définit les conditions

³⁴ Emmanuel Poinas, « Le Garde des sceaux a-t-il diffusé une circulaire anti-[jurissimo] ? », mis en ligne le 10 avril 2019, mis à jour le 14 février 2020, consulté le 18 avril 2021. [https://www.village-justice.com/articles/garde-des-sceaux-diffuse-une-circulaire-anti-\[jurissimo\],31196.html](https://www.village-justice.com/articles/garde-des-sceaux-diffuse-une-circulaire-anti-[jurissimo],31196.html)

précises. Il crée un régime spécifique pour les décisions et les personnels de justice (1), avalisant les propositions contenues dans le rapport Cadiet (2).

3.1.1. Un régime spécifique pour les décisions et les personnels de justice...

Ce décret d'abord annoncé en mars 2018 a finalement été proposé à la concertation publique sur le site du ministère, à partir du 13 décembre 2019 (Mathis, 2019), pour être finalement promulgué à la fin du printemps 2020³⁵.

Il confie au Conseil d'État et à la Cour de cassation la responsabilité de la mise à disposition du public des décisions de justice sous format électronique. Il n'existe pas de restriction en matière administrative, mais pour le judiciaire, seules sont éligibles les décisions rendues publiquement et accessibles aux tiers, sans autorisation préalable. Les délais également différents³⁶, mais pour le reste, on peut dire que le même système est adopté pour les deux ordres de juridiction. Les plus hautes autorités qui les coiffent deviennent les responsables de la mise en œuvre de l'*open data*, ce qui conforte et élargit leur pouvoir établi, dans une logique où c'est le juridique qui l'emporte sur la logique administrative généraliste.

De même, si ce décret instaure l'occultation systématique des noms et prénoms des personnes physiques (les parties et les tiers comme les témoins), une procédure particulière est prévue lorsque l'occultation concerne le personnel de justice : magistrats, membres du Conseil d'État, agents de greffe. La décision est alors dans les mains des responsables de juridiction, à savoir pour l'ordre judiciaire, le président de la juridiction concernée, et pour l'ordre administratif, le président de la section du contentieux du Conseil d'État, de la cour administrative d'appel ou du tribunal administratif. C'est donc aux magistrats qu'il incombe d'évaluer, au cas par cas, l'opportunité de prévoir une occultation portant sur d'autres éléments d'identification que le nom et le prénom des personnes. En ce qui concerne cette fois l'identité des magistrats et greffiers, elle est soumise à l'appréciation des chefs de juridiction³⁷. Par ailleurs, le décret prévoit qu'un portail Internet placé sous la responsabilité du garde des sceaux mettra les décisions de justice à la disposition du public. Il viendra s'ajouter aux sites Internet du Conseil d'État et de la Cour de cassation.

Le cadre général qui est fixé est clairement favorable aux institutions de justice et à une conception juridique des décisions de justice au sens où il confère le pouvoir de mise en œuvre et d'arbitrage aux juridictions, magistrats et ministère de la Justice donc exclusivement à des acteurs sectoriels d'une part ; et où il prévoit des dispositions particulières pour les personnels de justice, dont l'anonymat est garanti d'autre part.

³⁵ Paul-Antoine Chevalier, « *Open data* : le décret est enfin publié », Actualités Dalloz en ligne, mis en ligne le 30 juin 2020, consulté le 16 avril 2021. <https://www.gazette-du-palais.fr/actualites-juridiques/jo-open-data-le-decret-est-enfin-publie/>

³⁶ La mise à disposition du public doit intervenir sous deux mois, à compter de la date de la décision pour l'ordre administratif ; et sous six mois, à compter de sa mise à disposition au greffe de la juridiction pour l'ordre judiciaire.

³⁷ Si un autre élément porte atteinte à leur sécurité, leur vie privée ou aux intérêts de la Nation, il pourra également être occulté, sur décision *ad hoc* de la juridiction de jugement l'ayant rendue (le président de la formation de jugement ou le juge en cas de juge unique).

3.1.2. ... recommandé par le rapport Cadiet

Le décret reprend en grande partie les conclusions du rapport de la « mission d'étude et de préfiguration sur l'*open data* des décisions de justice », commandité en mai 2017 par le ministre de la Justice de l'époque, Jean-Jacques Urvoas. La mission est confiée à Loïc Cadiet, professeur en droit privé et en sciences criminelles à Paris 1, spécialiste de procédure civile. Tous les membres de la mission sont des juristes et représentants des instances sectorielles des deux ordres de justice³⁸ : aucun membre d'Etalab ou de la DINSIC n'est présent. La lettre du ministre charge la commission de réfléchir au futur décret d'application des articles 20 et 21, c'est-à-dire de poser les prémisses d'un cadre juridique valable pour les ordres judiciaire et administratif, en entendant « les acteurs publics et privés de l'*open data* des décisions de justice » ainsi que « des représentants des organisations syndicales et professionnelles concernées » (Cadiet, 2018 : 8). L'accent est placé sur le double enjeu de la mise à disposition des décisions de justice et de l'approfondissement des garanties en matière de protection de la vie privée, ce qui contribue d'emblée à affirmer un cadrage davantage orienté vers les limitations possibles à l'*open data* des données publiques plutôt qu'aux aspects de leur pleine ouverture.

Dès l'introduction du rapport, la portée de la loi pour une République numérique est relativisée : elle introduit une rupture seulement dans le sens où elle lève toute restriction autre « que le respect de la vie privée des personnes concernées et, pour les décisions rendues par les juridictions de l'ordre judiciaire, des règles régissant la publicité et l'accès aux décisions ». Ainsi, est-il d'emblée estimé que la nouvelle loi ne remet pas en cause le « double régime » auquel sont soumises les données judiciaires en raison de leur caractère ambivalent. L'existence d'institutions et de textes historiquement antérieurs est mobilisée pour contenir l'impact de cette loi dans le cas des arrêts et jugements. Le cadre intellectuel et stratégique de la mission est défini par le triptyque : singularité des décisions de justice, potentiel économique et innovant qu'elles représentent, enjeux de protection des données qui leur sont liées. Le rapport dresse une liste de vingt propositions dont l'esprit repose toutefois en priorité sur l'affirmation de la spécificité des décisions de justice fondée sur trois caractéristiques propres au champ juridique : l'existence de juridictions suprêmes au sommet des ordres judiciaire et administratif ; la préexistence de règles précises en matière de publicité des décisions de justice avec lesquelles les dispositions de la loi Lemaire doivent être articulées ; enfin, la référence que représente le principe d'autonomie des juridictions lorsqu'il s'agit d'évaluer au cas par cas si la diffusion d'une décision de justice met en danger certains secrets protégés par la loi.

Le rapport est rendu à la nouvelle ministre de la Justice, Nicole Belloubet, le 29 novembre 2017 comme en témoigne la date de la lettre de conclusion de la mission. Son contenu n'a toutefois été rendu public que plus d'un mois après, début janvier 2018. Peut-être ce décalage est-il l'expression d'une volonté de donner un peu de temps à l'évolution des

³⁸ Elle est également composée d'un représentant de la CNIL, d'un avocat désigné par le CNB, d'un président de cour d'appel, d'un procureur de la République, d'un membre de tribunal administratif et d'un de la Cour administrative d'appel ainsi que d'un représentant de chacune des juridictions suprêmes (CA, CC).

positions respectives voire à la construction de compromis, face à plusieurs points restés sensibles. Si la publication du décret ne répond pas à toutes les questions posées par l'*open data* — certaines étant repoussées à des arrêtés et groupes de travail au sein de la Chancellerie — elle représente une clôture de cette séquence à l'issue de laquelle on peut dire que c'est la logique et les acteurs juridico-judiciaires qui priment. Ce constat est renforcé par le fait qu'un autre texte est adopté à la même période qui traduit lui aussi le « succès » des acteurs porteurs d'une logique juridique de l'*open data* sur une logique plus généraliste ou « administrative ».

3.2. Une loi qui bride les usages de décisions déjà accessibles

Deux ans avant la loi Lemaire, le contenu de Légifrance — plus de 350 000 décisions de justice — a été ouvert en libre accès avec l'appui d'Etalab³⁹. Une étude portant sur les décisions des cours administratives d'appel relatives au contentieux des OQTF a été réalisée dont les résultats sont parus dans un billet sur le Village de la justice ([Forestier], 2016). La comparaison des arrêts selon les ressorts des cours et en fonction des présidents de chambre, nommément identifiés, suggère qu'il existe un biais chez les juges, certains étant plus favorables à l'administration et plus sévères avec les particuliers requérants (cf. chapitre 1). L'impossible impartialité de juges humains est alors rendue visible par le recours aux méthodes algorithmiques et devient le support d'une remise en cause directe et individuelle des magistrats. Parce qu'elle incarne ce qui peut être fait concrètement à partir du traitement algorithmique des données contenues dans un grand nombre de décisions de justice, cette première expérience française sonne comme un coup de tonnerre dans un ciel d'été et déclenche une controverse autour du périmètre de l'*open data* des décisions de justice et de la définition des usages légitimes qui peuvent être faits des jugements et arrêts (Léonard et Moritz, 2020 : 41-42).

[Morgan Forestier] est l'auteur de cette étude, appuyée sur un dispositif algorithmique disponible gratuitement en ligne, [Legalplus], qu'il a cocréé. La trajectoire de cet acteur est intéressante. Il commence sa carrière comme expert-comptable, formé en droit des affaires et spécialiste de droit fiscal, puis il devient avocat en entreprise et ingénieur en *machine learning*. Après avoir cofondé [Legalplus] et suite à cette controverse, en avril 2017, il passe d'un poste d'avocat fiscaliste dans le cabinet Deloitte France (2013-2017) à un poste d'ingénieur en *machine learning* appliqué chez [Geoffroy-Maisons], groupe d'édition juridique de premier plan au niveau francophone et français (car il possède notamment des marques historiques comme [les Éditions du couchant], [les Éditions doctrinales] et [les Éditions Geoffroy]). La vive controverse autour de [Legalplus] lors de laquelle il a été très exposé ne lui a visiblement pas porté de tort, puisque c'est ensuite qu'il a été recruté chez [Geoffroy-Maisons]. On peut en déduire que cet épisode a aussi été perçu et traité dans le milieu de l'ouverture du droit comme l'expression d'une forme de virtuosité technique. D'ailleurs, le

³⁹ Ce sont toutes les bases de données juridiques et associatives gérées par la Dila qui ont été mises à disposition du public pour une réutilisation gratuite, sous licences elles-mêmes gratuites, par l'arrêté du 24 juin 2014 relatif à la gratuité de la réutilisation des bases de données juridiques de la direction de l'information légale et administrative.

site de la plateforme [Legalplus] est fermé, mais il est toujours répertorié sur le site data.gouv.fr⁴⁰ (et non passé sous silence), ce qui va dans le même sens.

En effet, de vives réactions sont suscitées par cette expérience, comme en témoignent les commentaires publiés à la suite de l'article initial ; ainsi que la série d'articles écrits par la *startup* qui en est à l'origine, afin de préciser les capacités et limites de l'algorithme. L'affaire est qualifiée de « scandale » par un magistrat à la Cour administrative d'appel de Lyon (Cadiet, 2017 : 151), dans un article transmis à la mission Cadiet⁴¹. Le rapport mentionne cette affaire et il est probable que la mission ait été un lieu de préparation d'une réponse, qui, en l'occurrence, prend la forme d'une riposte législative.

En 2019, à l'occasion du projet de loi de programmation 2018-2022 et de réforme pour la justice, une nouvelle disposition législative est introduite dans le code de l'organisation judiciaire (COJ) et dans le code de justice administrative (CJA) de manière à empêcher que « les données d'identité des magistrats et des membres du greffe » fasse « l'objet d'une réutilisation ayant pour objet ou pour effet d'évaluer, d'analyser, de comparer ou de prédire leurs pratiques professionnelles réelles ou supposées⁴² », et ce, sous peine de sanctions⁴³. Cette mesure revient à interdire explicitement certains usages des décisions de justice dans le but de protéger les magistrats de mises en cause individuelles et de fait, à un niveau agrégé, elle rend illégaux les dispositifs de classement des juges (outils dits de *ranking*), comme il en existe aux États-Unis. Les capacités des algorithmes et les velléités des innovateurs, ou de certains utilisateurs sont bridées juridiquement.

3.3. Une circulaire qui empêche la récupération de décisions de justice en masse

Concomitamment à la controverse [Legalplus], une autre situation conflictuelle apparaît, autour de [Jurissimo], plateforme en ligne créée en 2016 par trois jeunes entrepreneurs, qui fonctionne comme un moteur de recherche en matière juridique. L'image disruptive de [Jurissimo] est le résultat d'un positionnement et d'une stratégie de ses fondateurs, ainsi que des investisseurs qui la soutiennent (Léonard et Moritz, 2020 : 57-61). Cette stratégie est faite d'une double distanciation, à l'égard des pouvoirs publics d'une part et de l'entrepreneuriat classique d'autre part.

Dès l'automne 2016, la société [Magnum], éditrice du site, s'attèle à obtenir la mise à disposition des décisions de justice en grande quantité, directement auprès des greffes concernés, afin d'obtenir l'accès aux minutes et le droit d'utiliser les informations publiques qu'elles contiennent. La réponse étant négative, la même demande est formulée auprès du président du TGI de Paris, en juin 2017. Réponse également négative par ordonnance de juin

⁴⁰ [https://www.data.gouv.fr/fr/reuses/\[legalplus\]/](https://www.data.gouv.fr/fr/reuses/[legalplus]/), consulté le 1^{er} mai 2021.

⁴¹ Cet article, reproduit dans le rapport Cadiet a initialement été publié le 21 avril 2016 : Clément Marc, « Jurisprudence 2.0 », *Telos*, 21 avril 2016, <https://www.telos-eu.com/recherche/?q=jurisprudence+2.0> cité par Léonard et Moritz (2020 : 40-41).

⁴² Art. 33 al. 4, loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice, modifiant l'art. L. 111-13 du COJ.

⁴³ Prévues « aux articles 226-18, 226-24 et 226-31 du code pénal, sans préjudice des mesures et sanctions prévues par la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés »

2017. La Cour d'appel de Paris, saisie à son tour, en décide alors autrement, considérant qu'« il n'existe pas de raison juridique permettant de s'opposer à la communication des décisions rendues publiquement, qui devront être anonymisées » et « doivent être communiquées à tous requérants français ou étrangers, de sorte qu'il doit être enjoint au directeur des services de greffe judiciaires du tribunal de grande instance de Paris de communiquer au requérant, sous forme papier ou sous forme numérique ces décisions, à charge pour lui de les anonymiser, ou de le laisser accéder aux minutes dans les mêmes conditions que les autres opérateurs autorisés, à charge d'en faire un usage autorisé par la loi⁴⁴ ».

Le lendemain, le ministère de la Justice fait paraître une circulaire que certains ont qualifiée d'« anti-[Jurissimo]⁴⁵ » rappelant que, par principe, toute décision rendue publiquement peut être communiquée aux tiers en matière civile et qu'il en est de même pour les décisions pénales définitives. En revanche, en ce qui concerne « le traitement des demandes de masse », l'absence du décret d'application des articles 20 et 21 de la loi Lemaire « ne saurait être contournée par un procédé de diffusion massive qui n'offrirait pas les garanties utiles à la protection des données à caractère personnel. [...] La diffusion de décisions en masse répondant à des demandes dont il est manifeste qu'elles ne portent pas sur une ou plusieurs affaires en particulier, mais sur la jurisprudence de la juridiction dans une ou plusieurs matières sera en principe évitée. Le refus pourra trouver son fondement dans des considérations liées à l'objectif de valeur constitutionnelle de bonne administration de la justice et à la protection des données à caractère personnel⁴⁶ ». La circulaire ne suffira toutefois pas à éteindre la controverse comme on le verra plus loin.

Ces controverses et la manière dont elles sont réglées nous apparaissent décisives, car elles donnent à voir la constitution en cours de deux coalitions de cause engagées dans un affrontement aux multiples rebondissements.

IV / Coalitions de cause autour de l'alignement de la justice sur la politique d'*open data* ou l'incarnation d'un rapport sectoriel-global conflictuel

Dans le domaine de la justice, la loi pour une République numérique contribue à rebattre les cartes. Elle ébranle la situation des éditeurs juridiques puisque jusque-là, ces derniers se procuraient, généralement contre redevance, les décisions de justice qu'ils

⁴⁴ Extraits de la décision n° 17/22211 du 18 décembre 2018, cités par Fabien Drey, « Notation des avocats, algorithmes et *open data* des décisions de justice, les liaisons dangereuses », mis en ligne le 14 août 2019, consulté le 19 avril 2021. <https://www.village-justice.com/articles/notation-des-avocats-algorithmes-open-data-des-decisions-justice-les-liaisons,32216.html>

⁴⁵ Roseline Letteron, « L'accès aux décisions de justice, ou le dispositif "anti-[jurissimo]" », mis en ligne le 6 janvier 2019, <http://libertescherries.blogspot.com/2019/01/laces-aux-decisions-de-justice-ou-le.html> ; Emmanuel Poinas, « Le Garde des sceaux a-t-il diffusé une circulaire anti-[jurissimo] ? », mis en ligne le 10 avril 2019, mis à jour le 14 avril 2020, <https://www.village-justice.com/articles/garde-des-sceaux-diffuse-une-circulaire-anti-jurissimo,31196.html>, consultés le 18 avril 2021.

⁴⁶ « Note relative au traitement des demandes de copie de décisions judiciaires émanant de tiers à l'instance », circulaire du 19 décembre 2018, Bulletin officiel du ministère de la Justice. Citée par Fabien Drey, *ibid.*

numérisaient, nettoyaient, mettaient en forme, éventuellement enrichies de contenus supplémentaires (lien avec la doctrine, etc.) pour les rendre accessibles à leurs clients dans des bases de données numériques payantes via abonnements. Indirectement donc, les principes d'ouverture et de gratuité des décisions de justice mettent fin à un certain type d'économie organisée autour de ce qui ressemblait fort à un monopole des éditeurs sur l'accès et la diffusion de la législation et de la jurisprudence⁴⁷.

Inversement, cette nouvelle norme publique crée la possibilité pour de nouveaux acteurs, *startups* de la *legaltech*, de tirer parti de cet accès facilité aux données pour proposer de nouveaux services, créer de l'enrichissement économique et renverser la situation monopolistique. Comme on vient de le voir, cela conduit certains acteurs de l'innovation, qui se positionnent comme disruptifs, à manifester une certaine impatience et à « prendre les devants », dès que de premiers jeux de données sont disponibles ([Legalplus]) et sans attendre que les décrets d'application de la loi Lemaire ne soient publiés ([Jurissimo]).

À l'instar des éditeurs juridiques, les plus hautes juridictions des deux ordres de la justice française — Cour de cassation et Conseil d'État — font, elles, partie de ceux qui ont beaucoup à perdre : dans l'océan des décisions de justice désormais accessibles, le poids de leur jurisprudence pourrait être relativisé et leur autorité juridique s'en trouver affectée. Pour autant, le décret de 2020 réaffirme leur position dominante en leur confiant l'opérationnalisation de la mise à disposition des décisions de justice anonymisées. Enfin, les instances de représentation de la profession d'avocat craignent que l'expertise de la profession ne soit concurrencée par les dispositifs algorithmiques de traitement et d'analyse des décisions juridiques. Ils s'inquiètent plus largement que son rôle d'intermédiaire du droit ne soit mis sur la sellette par une nouvelle génération de plateformes algorithmiques associant traitement de l'information juridique et règlement en ligne des litiges. L'enjeu pour la profession est à la fois celui de la perte de parts de marché, en particulier sur les petits litiges, et celui de la montée en puissance de métiers et professionnels concurrents, porteurs de nouvelles normes et rapports à la justice, et susceptibles de menacer la profession établie sur le double plan des valeurs (déontologie) et de l'économie (monopole). À l'intérieur de la profession, pour autant, différents segments coexistent avec des visions et des intérêts différenciés sur ce plan (Léonard et Moritz, 2020). Les avocats du barreau d'affaires, par ailleurs très multi-positionnés, sont particulièrement engagés et proactifs dans la transformation des activités et marchés juridiques (Vauchez, 2016).

Nous avons établi assez nettement jusque-là que dans cette histoire qui couvre presque une décennie, coexistent deux façons de concevoir l'*open data* des décisions de justice qui dessinent deux directions assez clairement distinctes. L'une fondée sur l'autonomie de la justice, la spécificité des arrêts et jugements, ainsi que les règles juridiques qui leur sont applicables : la décision de justice comme jugement et jurisprudence incarnée par le rapport

⁴⁷ Déjà amorcé avec la création des sites Légifrance et Service-public.fr dont la responsabilité incombe à la Dila, direction d'administration centrale rattachée directement aux services du Premier ministre et placée sous l'autorité du Secrétaire général du gouvernement, née en 2010 de la fusion des directions de La Documentation française et des Journaux officiels.

Cadiet et le décret de 2020, et l'autre centrée sur la dimension administrative et publique des décisions de justice : la décision de justice comme production de données juridiques incarnée par Open Law et son livre blanc. La première agit de manière à limiter l'impact de la loi Lemaire au profit d'une internalisation de la mise en place de l'*open data* tandis que la seconde s'efforce d'en accroître la portée. La première est en position de domination, elle parvient d'ailleurs à mobiliser les outils juridiques (loi, décret, circulaire) pour contenir la seconde, laquelle est traversée par d'importantes fractures, notamment liées, comme on le verra, à la concurrence que se livrent les entrants du secteur entre eux.

À ce stade de la recherche et compte tenu des éléments empiriques à notre disposition, nous proposons d'envisager ces deux conceptions comme la manifestation de deux coalitions de cause (Sabatier et Jenkins-Smith, 1993, 1999) qui s'affrontent autour de la mise en œuvre de l'*open data* dans la justice. L'*advocacy coalition framework* repose sur le fait que « des acteurs sont regroupés en une ou plusieurs coalitions de cause, dont les membres partagent un ensemble de croyances normatives et de perceptions du monde, et qu'ils agissent de concert afin de traduire leurs croyances en une politique publique » (Sabatier, 2019 : 46). C'est donc la double condition du partage d'un système de croyances et de la coordination des comportements qui fonde l'existence d'une coalition de cause entre des acteurs qui peuvent être hétérogènes, mais cherchent ensemble à influencer les décisions et orientations de politique publique. À ce stade de l'analyse et toujours sous réserve du matériau recueilli, les premiers constats pointent une polarisation de la politique de l'*open data* des décisions de justice autour de deux positions. Ces premières constatations peuvent être raffinées et complexifiées en regardant plus en détail les controverses déjà évoquées et en considérant également d'autres événements non mentionnés jusque-là. « Chaque coalition développe un ensemble de stratégies, incluant l'utilisation de ses ressources politiques (argent, informations, autorité légale, etc.) dans le but de contrôler la politique publique. Les décisions politiques (lois, décrets, allocations budgétaires) produisent alors des résultats (*outputs*) et des impacts, qui entraînent des réactions (*feedback*) dans les stratégies et les systèmes de croyances de chaque coalition. » (48). Nous verrons se dessiner plus nettement les composantes de ces deux coalitions de cause (1), ce qui permettra de discuter le rapport sectoriel/global qui se joue à travers leur affrontement et en lien avec un régime d'action publique globalisé (2).

4.1. Deux coalitions de cause qui se constituent progressivement

Outre des positions différenciées sur la question de l'innovation, du numérique et de l'*open data*, ces coalitions reposent plus fondamentalement sur des manières contrastées de concevoir la place des décisions de justice dans l'ensemble des données publiques. Dans le cadre théorique de Sabatier et Jenkins-Smith (1999), sont distingués trois types de croyances et normes qui, ensemble, forment le système de croyances partagées au sein des élites d'une politique publique donnée. Nous allons rapidement les rappeler avant de revenir sur chacune des deux coalitions de cause.

4.1.1. Des systèmes de croyances en confrontation

Pour Sabatier et Jenkins-Smith, les changements majeurs de politique publique sont forcément liés au changement des acteurs dominants qui les portent. Autrement dit, à une politique publique donnée correspond à une coalition de cause hégémonique, c'est-à-dire parvenue à installer ses idées, croyances et normes. Ce qui cimente les acteurs de cette coalition, ce sont avant tout des conceptions et représentations convergentes sur la manière dont il faut organiser une politique publique.

Ces croyances se distinguent par leur caractère plus ou moins profondément ancré chez les acteurs. Au niveau le plus profond et le plus large se trouvent les croyances du noyau central (*deep core beliefs*) qui « comprennent des hypothèses normatives et ontologiques très générales » souvent acquises lors de la socialisation primaire (Sabatier, 2019 : 49). Il s'agit de visions du monde, de la société et de l'individu. Paul Sabatier donne l'exemple de l'orientation sur l'échelle gauche/droite, de la priorité donnée à la régulation par le marché ou par le gouvernement, ou bien encore du positionnement entre approche libérale (au sens de la liberté des individus) et égalitaire (au sens de l'égalité sociale).

Le second niveau est constitué des « croyances du noyau central appliquées à tout un sous-système de politique publique » (*policy core beliefs*) (Sabatier, 2019 : 49). Elles relèvent de la traduction des normes du noyau central dans une politique publique particulière, ce qui les rend tout aussi fondamentales, consistantes et difficiles à changer que les précédentes. Elles reposent sur le même type de hiérarchisation de valeurs et de priorités, appliqué à un secteur d'action publique. L'accent est mis sur certains problèmes considérés comme plus graves que d'autres ; l'interprétation privilégie certaines causes et solutions aux dépens d'autres. Par exemple la croyance dans la force supérieure du répressif ou de l'éducatif en matière de justice pénale ou du diptyque prévention/thérapie dans le secteur de la santé.

Enfin, un dernier niveau concerne « un ensemble "d'aspects secondaires" (*secondary aspects*) comportant une série d'idées sur les instruments, mesures et informations nécessaires pour la réalisation du *policy core* » (Maillard et Kübler, 2015 : 187). Parce qu'elles sont plus circonscrites, qu'elles touchent des dimensions plus opérationnelles, ces idées et croyances sont plus flexibles et peuvent plus facilement évoluer. Elles concernent par exemple règles, procédures, budgets, dans le cadre d'un programme de politique publique donné. Elles sont toutefois elles-mêmes plus secondaires dans l'*advocacy coalition framework* et plus difficiles à identifier, compte tenu des données dont nous disposons. Elles seront donc moins présentes dans notre propre analyse.

Les deux coalitions de cause que nous avons identifiées, et qui relèvent d'une première mise en intelligibilité du matériau réuni jusque-là, sont ainsi définies par des systèmes de croyances qui portent sur l'*open data* en lui-même, mais qui renvoient aussi plus profondément à la manière de concevoir les arrêts et jugements, et donc la politique de justice (*policy core beliefs*). Elles prennent assise sur des croyances plus générales et profondes encore sur l'individu, l'État, le marché et le public (*deep core beliefs*). La coalition dominante défend la singularité des décisions de justice considérées sous l'angle de la jurisprudence et appuyées sur la croyance dans le droit comme valeur supérieure et sur l'importance de la

régulation par les instances publiques. La coalition qui la challenge défend une politique d'*open data* généraliste, au sens d'appliquée à l'ensemble des données publiques (dont les décisions de justice), appuyée sur une conception du droit comme outil/ressource. Le marché et les acteurs privés sont vus comme des atouts. Ce sont donc deux approches de l'*open data*, particulariste versus généraliste qui s'affrontent.

4.1.2. Une approche particulariste de l'*open data*

L'approche particulariste est portée par les acteurs spécialisés et dominants de la justice. Elle est hégémonique comme le montrent les importantes ressources juridiques qu'elle détient et qu'elle a pu mettre en place pour riposter à l'adoption des articles 20 et 21 de la loi pour une République numérique. On en trouve la trace dans le rapport Cadiet et dans la reconnaissance des options qu'il propose, par le décret d'application de cette même loi.

4.1.2.1. ... fondée sur la place historique de la jurisprudence et des plus hautes juridictions

Le rapport Cadiet reconnaît le pouvoir établi de la Cour de cassation et du Conseil d'État dans la justice sur le double plan de la gestion des bases de données de jurisprudence et sur celui du contrôle de la jurisprudence elle-même. Il valorise le fait que ces instances sont explicitement en charge et ont historiquement géré les décisions de jurisprudence et les bases de données déjà en fonctionnement. « Cette expertise anciennement acquise par les deux juridictions les qualifie particulièrement pour assurer cette fonction [de pilotage de la mise en œuvre de l'*open data*]. La mission considère en outre que l'amélioration de la connaissance de la jurisprudence est une mission qui relève au premier chef des juridictions. » (Cadiet, 2018 : 64) C'est ainsi une « gestion centralisée de l'alimentation et de la collecte des décisions » (*ibid.*) qui est défendue, également appuyée sur un raisonnement économique et pragmatique : les architectures informatiques sont en place, gérées par des organisations efficaces et fortes de leurs expériences.

Or, ce raisonnement repose sur une confusion entre la seule jurisprudence (la sélection des décisions qui proposent une interprétation du droit devant faire référence) et les décisions de justice (l'ensemble des produits de l'activité juridictionnelle) (Serverin, 1993). Là où les principes dits de Sébastopol portés par Etalab et Open Law visent la « libération » de toutes les données et donc de toutes les décisions de justice de manière horizontale et indistincte, la référence à la jurisprudence permet de maintenir une logique hiérarchique à l'intérieur des décisions de justice et donc une vision plus restrictive et spécifique de l'*open data*. Cela explique que le point n° 16 du rapport suggère de « développer sur le site Internet de la Cour de cassation un canal de diffusion de la jurisprudence de l'ensemble des juridictions de l'ordre judiciaire assurant la mise en valeur de celle-ci, à l'instar de ce que pratique le Conseil d'État s'agissant des décisions de l'ordre administratif avec la base ArianeWeb ». C'est une manière de saisir l'occasion de l'*open data* pour réaffirmer la logique d'un système de droit écrit dans lequel les précédents obligatoires sont les « précédents d'interprétation de la Cour de cassation » et non les « précédents de cas », comme en *common law* (Serverin, 2018 : 23), et d'œuvrer à la consolidation de la place de la Cour de cassation dans le système juridique français, alors même qu'une des revendications de l'*open data* généraliste repose sur la

diffusion horizontale et « à égalité » de toutes les données. Cela signifie moins ici une rétention de certains jugements et arrêts par les juridictions supérieures que la possibilité pour celles-ci de maîtriser la fabrique de l'architecture des canaux de diffusion et leur orientation dans un sens qui exprime et réalise le caractère hiérarchique de certaines décisions sur d'autres. Définition des arborescences, présentation des décisions, fonctionnalités et fonctionnement des outils d'interrogation, c'est tout l'agencement de l'interface qui peut traduire et mettre en œuvre une conception horizontale ou verticale des décisions de justice et de leur force normative.

4.1.2.2. ... et sur l'*imperium* du juge et le pouvoir d'appréciation des juridictions

Lorsqu'il s'agit d'évaluer au cas par cas si la diffusion d'une décision de justice contrevient à la protection de certains secrets protégés par la loi, c'est à la juridiction responsable de la décision de s'en charger. Les recommandations du rapport Cadiet s'appuient ainsi sur l'existence d'un pouvoir d'appréciation des juridictions de jugement, juridiquement reconnu et au cœur de la pratique des magistrats. Cet *imperium* qui fonde leur légitimité à juger un cas, en toute indépendance et avec une forte autonomie (même si celle-ci tend à se réduire sous la pression de dispositifs gestionnaires), est étendu à l'appréciation du bien-fondé de sa mise à disposition auprès du public. Évaluer les risques de rendre publics non seulement le dispositif, mais aussi les motifs du jugement fait « naturellement » partie de l'office du juge.

Le principe même d'une diffusion systématique est donc remis en cause, ce qui, sous cet angle, pourrait conduire à une réduction significative du nombre d'arrêts et de jugements finalement accessibles. Il est toutefois difficile d'en évaluer le volume. La recommandation n° 12 prévoit la mise à disposition conditionnelle des décisions de justice, et ce, à l'appréciation des juridictions qui les ont émises, et la création d'une dérogation à l'obligation générale de mise à disposition si les conditions ne sont pas réunies : « Prévoir la possibilité, pour la juridiction prononçant la décision, de *conditionner sa délivrance* aux tiers à sa pseudonymisation ou à la suppression de tout ou partie de ses motifs lorsque cette délivrance est susceptible de porter atteinte à des droits ou secrets protégés, en modifiant notamment les articles 11-3 de la loi du 5 juillet 1972, R.156 du code de procédure pénale et R.751-7 du code de justice administrative. Dans l'impossibilité d'y parvenir, prévoir la possibilité pour la juridiction d'exclure, à titre exceptionnel, l'accès d'une décision aux tiers et sa mise à disposition du public en modifiant les mêmes dispositions. » Le raisonnement sous-jacent n'est pas justifié : ni par la référence au principe d'autonomie des juridictions, à l'*imperium* du juge, son office ou bien encore son pouvoir d'appréciation, ni par d'autres principes. L'absence de justification incite à penser qu'il y a une inférence juridiquement évidente qui traduit une croyance partagée propre à la justice (*policy core beliefs*) au sein des participants de la mission Cadiet.

La défense du caractère spécifique du secteur de la justice — entendue de manière globale c'est-à-dire s'appliquant aux institutions des deux ordres, administratif et judiciaire, dont les juridictions supérieures font alliance — est au cœur du rapport Cadiet et de la réflexion menée sur la manière d'adapter les dispositions de la loi Lemaire. Elle relève également des *policy core beliefs*. Des principes globalement identiques pour la justice

administrative et judiciaire sont préconisés ; des emprunts de l'une à l'autre sont recommandés ; des concertations et travaux en commun, s'agissant de préparer la collecte et la préparation des décisions de justice, sont préfigurés. Le rapport Cadiet (2018), parce qu'il donne l'avantage aux architectures en place via les bases de données déjà gérées par la Cour de cassation et le Conseil d'État, parce qu'il installe ces juridictions au cœur du dispositif de collecte, de pseudonymisation et de mise à disposition des décisions de justice peut être analysé comme une manière de garantir le *status quo* et le pouvoir actuel de ces instances dans le système juridique national.

Le premier président de la Cour de cassation et le vice-président du Conseil d'État lorsqu'ils interviennent publiquement sur la 'justice prédictive' dans le colloque organisé par l'Ordre des avocats aux conseils (Louvel, 2018 : 16 ; Sauvé, 2018 : 12) défendent la place de leurs institutions. Il est raisonnable de penser qu'ils ont adopté les mêmes positions lors des travaux de la mission, dont ils étaient membres, et qu'ils ont pu ainsi peser sur le schéma en cours d'élaboration, faisant alliance avec le rapporteur. Ils sont parties prenantes de cette coalition.

Le premier président de la Cour de cassation s'est, quant à lui, coordonné avec le Conseil national des barreaux pour effectuer une déclaration commune, solennelle et publique afin que l'encadrement de l'ouverture des décisions de justice, la pseudonymisation et la réutilisation des données qu'elles contiennent soient mises en œuvre dans l'esprit justement des recommandations du rapport Cadiet⁴⁸. Dans la déclaration du 25 mars 2019, c'est-à-dire à un moment où le décret n'est toujours pas publié, ils « rappellent l'urgente nécessité d'organiser l'*open data* des décisions de justice, en France, dans des conditions garantissant l'égal accès de tous à la donnée publique ». Ils considèrent « que les travaux à venir devront : – mettre la Cour de cassation en mesure de piloter et de contrôler de façon effective l'alimentation, le traitement, la pseudonymisation des décisions, ainsi que la mise à disposition du public d'une base de données unique des décisions de justice de l'ordre judiciaire ; – associer la Cour de cassation, les juridictions du fond et le Conseil national des barreaux à la mise en œuvre des dispositifs de régulation et de contrôle tant des algorithmes utilisés pour l'exploitation de la base de données des décisions de justice que de la réutilisation des informations qu'elle contient ; – aboutir à la constitution d'une instance publique chargée de cette régulation et de ce contrôle, dont devront, notamment, être membres la Cour de cassation et le Conseil national des barreaux ». La signature de cette déclaration commune se tient dans les locaux mêmes de la Cour de cassation, les symboles de l'autorité judiciaire y sont convoqués avec grandeur. La publicisation de cette action, en direction du ministère, est forte.

Le CNB nous semble donc être également à inclure dans cette coalition de cause qu'il rejoint en cours de route. Jusqu'en 2016-2017, le CNB et l'ordre des avocats de Paris étaient

⁴⁸ Déclaration relayée sur leurs sites respectifs :

https://www.courdecassation.fr/institution_1/intelligence_artificielle_7985/open_data_intelligence_artificielle_7821/archives_10021/cour_cassation_9200/ ; <https://www.cnb.avocat.fr/fr/actualites/retour-sur-lassemblee-generale-du-15-mars-2019#3>, consultés le 15 avril 2021.

présents et s'efforçaient de peser ensemble sur les enjeux du numérique pour la profession, mais à travers des tentatives assez timides, éparpillées (Bastard, 2016). À partir de 2018-2019, le CNB apparaît comme un acteur beaucoup plus offensif en matière d'*open data* des décisions de justice. Il exprime des prises de position publiques à travers motions et résolutions ; il commande une enquête auprès d'un cabinet de consultants spécialisés dans la transformation digitale⁴⁹ à l'issue de laquelle il publie quatorze préconisations et une charte sur la transparence et l'éthique de l'utilisation des données judiciaires. Il intervient solennellement pour recommander l'adoption des termes de « jurimétrie » ou « justice simulative » à l'exclusion de tout autre dont celui de 'justice prédictive'. Ces actions, qu'il entreprend seul et pour lesquels il mobilise des moyens, sont articulées à des interventions, en coordination avec la Cour de cassation, mais aussi, à la faveur de la controverse autour de Jurissimo, avec d'autres acteurs.

4.1.2.3. Juriconnexion et les éditeurs juridiques

L'analyse de la controverse autour de [Jurissimo] donne à voir l'alliance d'autres acteurs également membres de la coalition de cause, à savoir les grands éditeurs juridiques, mais aussi la mission Cadiet, le ministère de la Justice et Juriconnexion. Cette dernière association a lancé, courant 2018, une enquête comparative auprès des principaux éditeurs juridiques afin de recenser les fonds de décisions de justice dont ils disposent. Elle a mis en évidence de fortes disparités entre une grande partie de ces éditeurs qui ont répondu à l'enquête et [Jurissimo] qui revendique de disposer de plusieurs millions de décisions.

Dès 2017, une enquête d'une journaliste⁵⁰ du pôle économie du journal *Le Monde* avait révélé des pratiques possiblement douteuses de la *startup*⁵¹. Il y était question de « piratage massif » et donc de méthodes illégales et de concurrence déloyale. La journaliste relayait la parole des juridictions en place, et sur le mode de la révélation, dévoilait la manière dont [Jurissimo] s'attaquait au monopole des éditeurs juridiques. L'étude Juriconnexion a contribué à documenter la vraisemblance de pratiques peu éthiques, voire répréhensibles, dans la collecte d'arrêts et de jugements. Comment [Jurissimo] a-t-elle pu se procurer ce volume exceptionnellement élevé de décisions de justice alors même qu'elle s'est heurtée pendant plus d'un an et demi au refus des instances judiciaires de communiquer celles qui avaient été demandées et que le décret d'application de la loi Lemaire n'était toujours pas publié ?

Comme on l'a vu, le ministère de la Justice par une circulaire signée de ses trois grandes directions judiciaires (DSJ, DASC, DAGE) incitant à refuser les demandes massives de décisions adressées aux greffes, s'est aussi placé du côté d'un *open data* maîtrisé par les instances

⁴⁹ Évoqué dans l'introduction et le premier chapitre.

⁵⁰ Ses biographies indiquent qu'elle a suivi une école de commerce et un master de journalisme. Passée par *Les Échos*, *le Figaro*, elle est entrée au *Monde* en 2013, où elle est chargée des rubriques défense, réforme de l'État, PME à la section généraliste et industrie du service économie et entreprise. <https://www.linkedin.com/in/isabelle-chaperon-60b2205/?originalSubdomain=fr> et https://www.leguidedupouvoir.fr/biographie/11047/isabelle_chaperon/index.htm, consultés le 16 septembre 2021.

⁵¹ Isabelle Chaperon, « Piratage massif de données au tribunal », *Le Monde*, 28 juin 2018, consulté le 16 septembre 2021.

juridictionnelles. La mission Cadiet, quant à elle, s'est clairement prononcée en faveur de « mesures visant à permettre aux juridictions de rejeter les demandes de copies de décisions lorsque ces demandes sont abusives ou lorsqu'elles ont pour objet ou pour effet la délivrance d'un nombre important de décisions » (recommandation n° 9 ; 2018 : 57), faisant implicitement référence au cas [Jurissimo].

La controverse est ensuite devenue contentieux judiciaire puisque sur la base de l'enquête menée par Juriconnexion, [Jurissimo] a été assignée en justice devant le tribunal commercial par ses concurrents, cinq des plus gros éditeurs juridiques⁵², mais aussi devant une juridiction pénale par les organes de représentation de la profession d'avocat, au niveau local, le barreau de Paris — vu la localisation de [Jurissimo] — et national, avec le Conseil national des barreaux⁵³. Il lui est notamment reproché de s'être procuré des décisions par *typosquatting*, c'est-à-dire en faisant des demandes de communication aux greffes par mail à partir d'adresses pouvant laisser croire qu'il s'agissait de demandes formulées par des universités, d'autres juridictions, etc. Des décisions de justice contrastées ont été rendues ; il est difficile de savoir si toutes les voies de recours sont purgées⁵⁴ ; les futurs développements seront intéressants à suivre. Mais il est certain que cette controverse autour des formes légitimes de récupération de ces décisions a donné lieu à un nouveau domaine contentieux, celui des données juridictionnelles, et a suscité des articles d'analyse de la jurisprudence⁵⁵.

Autour de ce cas, la mission Cadiet, les directions d'administration centrale du ministère de la Justice, les éditeurs juridiques, l'association Juriconnexion, les organes de représentation de la profession d'avocat, en s'appuyant sur le journal *Le Monde*, se sont ainsi mobilisés pour mettre une pression sur les entrants du secteur et en particulier sur celui qui se présente probablement comme le plus généraliste et le plus disruptif. On ne peut toutefois dire dans quelle mesure précise ces différentes actions et prises de position ont été concertées entre les protagonistes⁵⁶. À ce stade il est raisonnable de dire *a minima* que cette coalition de cause repose sur des valeurs et intérêts communs que l'ouverture des décisions de justice au public et l'arrivée d'acteurs, jeunes, généralement non issus du milieu juridique viennent bousculer.

⁵² Il s'agit de [Les Éditions du Couchant], [Legdata], [LegalEditions], [Legato] et [Brandt], pour « concurrence déloyale », « publicité trompeuse » et « paritarisme » (usurpation de notoriété), cf. [https://www.usine-digitale.fr/article/la-legaltech-\[jurissimo\]-attaquee-en-justice-par-cinq-editeurs-juridiques.N939476](https://www.usine-digitale.fr/article/la-legaltech-[jurissimo]-attaquee-en-justice-par-cinq-editeurs-juridiques.N939476), mis en ligne le 12 mars 2020, consulté le 20 avril 2021.

⁵³ Delphine Iweins et Déborah Loye, « La guerre s'intensifie entre [Jurissimo] et les avocats », *Lesechos.fr*, mis en ligne le 27 juin 2019, [https://business.lesechos.fr/entrepreneurs/actu/0601476339638-la-guerre-s-intensifie-entre-\[jurissimo\]-et-lesavocats-330194.php](https://business.lesechos.fr/entrepreneurs/actu/0601476339638-la-guerre-s-intensifie-entre-[jurissimo]-et-lesavocats-330194.php) ; Actualités CNB, [https://www.cnb.avocat.fr/fr/actualites/le-cnb-et-le-barreau-de-paris-ont-depose-une-plainte-penale-contre-\[jurissimo\]](https://www.cnb.avocat.fr/fr/actualites/le-cnb-et-le-barreau-de-paris-ont-depose-une-plainte-penale-contre-[jurissimo]) mis en ligne le 10 septembre 2019, consultés le 6 avril 2021.

⁵⁴ Source : « L'information juridique en procès », *Le Monde*, 6 août 2019, p. 10.

⁵⁵ Il est intéressant de noter que ces décisions et leurs commentaires sont en libre accès sur le site de [Jurissimo] à l'exception de quelques commentaires doctrinaux, [https://www.\[Jurissimo\]/d/CA/Paris/2018/CBD628309EEBD087B10BA](https://www.[Jurissimo]/d/CA/Paris/2018/CBD628309EEBD087B10BA), consulté le 20 avril 2021.

⁵⁶ Des entretiens seront réalisés dans le cadre de l'ANR Just-IA.

À cet égard, un autre indice de cette coalition de cause peut être relevé à travers l'interview donnée par Loïc Cadiet dans le Livre blanc de [LegalEditions]⁵⁷. Dans cette brochure intitulée *Open data & jurisprudence, Les données de jurisprudence au service de l'intelligence juridique* (2019), invité à répondre à la question, assez orientée, « Quel est le rôle des acteurs de la *legaltech* et, notamment, des éditeurs ? », il explique : « L'augmentation de la masse de données en *open data* devrait en effet stimuler l'activité de nouvelles entreprises de la *legaltech*, ces start-ups du domaine juridique qui développent leur activité à partir de l'exploitation des données juridiques. Avec la mise à disposition du public des décisions de justice, les opérateurs historiques de la diffusion du droit, que sont les éditeurs juridiques, devront certainement trouver de nouvelles sources de création de valeur économique et diversifier les services qu'ils proposent en s'appuyant sur l'extraordinaire expérience qu'ils ont acquise. [LegalEditions] œuvre déjà en ce sens avec [LegalDonnées+]. Ils ont pour eux la légitimité que confère leur très longue expérience experte et leur régulation au sein du Syndicat national de l'édition. Leur souhait d'être reconnus comme des tiers de confiance dans l'organisation en cours de l'*open data* des décisions de justice n'est donc pas injustifié. » (*ibid.* : 10) Loïc Cadiet rappelle ainsi l'expérience et la légitimité historique des éditeurs juridiques en matière de *legaltech* et leur apporte clairement son plein soutien.

Si les *startups* sont mentionnées, c'est pour les placer du côté d'une forme d'opportunisme économique (tirer parti de l'*open data*) alors qu'est tout de suite rappelée la fiabilité des éditeurs juridiques, désignés comme les acteurs historiques et expérimentés du champ de la diffusion et de l'analyse des décisions de justice. L'argumentation repose sur la préservation d'un monde à part, celui de la justice ; comme c'était déjà le cas pour justifier de confier aux plus hautes juridictions la mise en œuvre de l'*open data* dans la justice. C'est une référence indirecte au poids de la tradition dans la définition de la justice, qui représente une croyance fondamentale dans la définition de la politique de justice et qui repose sur des croyances du noyau central (*deep core beliefs*) relatives au désintéressement, à la force de la tradition et de la sacralité des institutions de manière plus générale.

4.1.2.4. La Cepej ou l'appui du supranational

Dans le secteur de la justice, on connaît l'importance de réseaux composés de professionnels et d'experts du monde judiciaire (Dallara et Piana, 2015) dans l'activité consistant à fixer des normes et des standards au niveau européen. Comité du Conseil de l'Europe, la Commission européenne pour l'efficacité de la justice (Cepej) en est un maillon essentiel, notamment en ce qui concerne la production d'instruments d'évaluation de la qualité de la justice (Cappellina, 2017). Compte tenu de sa composition, à la fois plurinationale et pluridisciplinaire, et de la variété de ses auditoires, cette arène adopte des principes, standards et instruments, qui sont le résultat d'une production incrémentale et empirique, propre à chaque sujet ou instrument en cours d'élaboration.

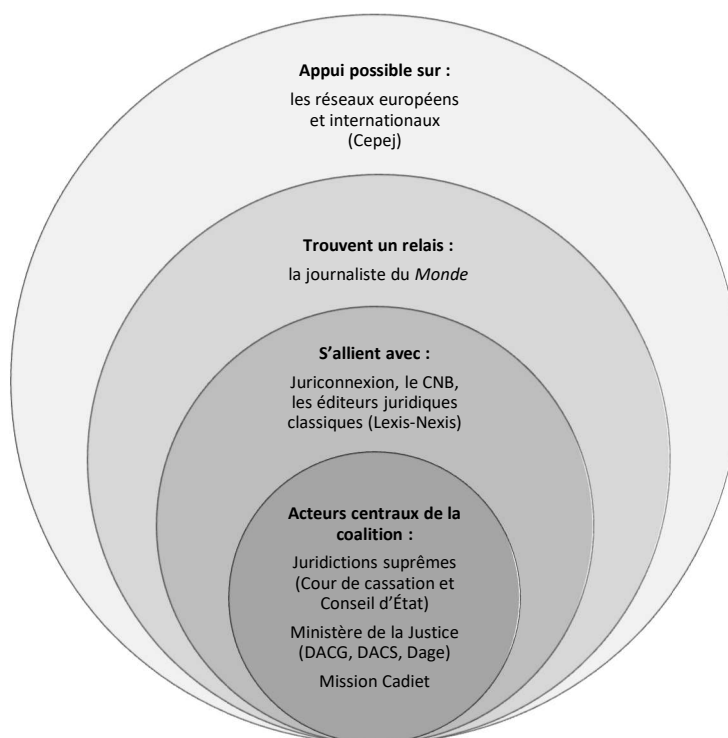
⁵⁷ Groupe d'édition juridique international installé dans le paysage français via le rachat d'éditeurs historiques nationaux, il est au plan mondial un acteur puissant de l'édition juridique et a racheté plusieurs *startup* du genre de [Jurissimo].

La Cepej n'est pas intervenue sur l'*open data*, mais sur l'usage des données par des dispositifs d'intelligence artificielle dans la justice. La charte⁵⁸ qu'elle a établie en 2018 c'est-à-dire au plus fort de la controverse sur la 'justice prédictive' en France, rappelle que les dispositifs d'intelligence artificielle doivent être soumis au respect des droits fondamentaux, et en particulier au principe de non-discrimination. Les recommandations affirment également que c'est aux systèmes judiciaires et aux professionnels de justice qu'il appartient de contrôler les données de justice et les usages qui en sont faits. La priorité est ainsi clairement donnée aux acteurs et institutions de justice régaliennes par rapport à des acteurs généralistes de l'action publique ou aux acteurs privés. Les normes qu'elle adopte sont clairement assises sur une logique juridique de protection des droits, dont les systèmes judiciaires nationaux sont les garants. Il semble que les porteurs d'une logique généraliste de l'*open data* qui existent pourtant au sein de la Cepej (qui comprend aussi des spécialistes de l'administration) ne soient guère parvenus à se faire entendre sur ce dossier. Il faut dire que parmi les acteurs de premier plan du groupe de travail *ad hoc* figurait le premier président de la Cour d'appel de Rennes, auteur du communiqué sur l'échec de l'expérimentation [PréviCompute] auprès de magistrats.

Cela reste évidemment à préciser, mais à ce stade, nous proposons de reprendre l'image de cercles concentriques pour décrire les parties prenantes de ces coalitions de cause. La Cepej partage le même cadre cognitif et normatif que la coalition particulariste, même si cela demande encore à être vérifié du point de vue de la coordination concrète à l'intérieur de la coalition.

⁵⁸ Cepej (2018a), European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their Environment, adoptée lors de la 31^e réunion plénière de la Cepej. Strasbourg, 3-4 décembre 2018, consulté le 19 septembre 2021. <https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c>

Graphique 3.3. Les acteurs de la coalition particulariste de l'open data



Pour récapituler, on pourrait dire que cette coalition de cause fait cohésion autour de croyances plus ou moins profondes et centrales synthétisées dans le tableau ci-dessous.

Tableau 3.1. Les croyances des acteurs de la coalition particulariste de l'open data

	Coalition particulariste
Position	Hégémonique
<i>Policy core beliefs</i>	<p>Les décisions de justice = sources de jurisprudence,</p> <p>Les données de justice = problème public appartenant aux systèmes judiciaires et aux professionnels de justice</p> <p>Importance de l'autonomie et de l'interprétation du juge comme fondatrice de la notion même de justice</p> <p>Le système juridique de droit écrit comme patrimoine à entretenir</p> <p>Droit comme valeur en soi</p> <p>Exceptionnalité de la justice (judiciaire et administrative)</p> <p>Tempérance — Désintéressement</p> <p>Valorisation du temps long</p>
<i>Deep core beliefs</i>	<p>Vertus de la régulation par le droit et l'État</p> <p>Importance des statuts, de la tradition et de la sacralité des institutions</p>

Cette première coalition contraste fortement avec la seconde, où se rassemblent notamment les entrantes dans la justice que sont les *startups*, dont les fondateurs sont eux-mêmes des outsiders du droit.

4.1.3. Une approche généraliste de l'*open data* dans la justice

Cette seconde coalition se fédère autour des valeurs fortes déjà identifiées dans le répertoire de réforme administrative de l'*open data* et parmi les acteurs qui participent à la configuration qui le porte. Ces croyances touchent à la vertu du progrès technologique et au pouvoir libérateur du numérique ; mais aussi à la logique de l'individu, créateur de richesse, qui s'exprime à travers la figure du *startupeur*. Open law, créé et conçu par Etalab comme un ambassadeur, un relais et un organisateur de la mise en œuvre d'une vision extensive de l'*open data* dans la justice, est clairement acteur de cette coalition. Cela est très clair dans la réaction d'un de ses membres à la lecture du rapport Cadiet.

4.1.3.1. Confier la mise en œuvre de l'*open data* aux acteurs généralistes

Consultant indépendant et membre-fondateur d'Open Law, Bruno Mathis (2018) s'exprime toutefois en son nom personnel pour regretter, dès le titre de l'article « une vision frileuse de l'*open data* des décisions de justice ». Il estime que « le choix de faire apparaître en tête des recommandations celle qui consiste à réguler la *Legal Tech* témoigne [...] d'une attitude défensive » (*ibid.* : § 1). La déception est provoquée par un apport insuffisamment opérationnel du rapport et par le choix de « propositions, inspirées par un principe de précaution plutôt que par une approche fondée sur les risques recour[a]nt de façon excessive aux instruments juridiques comme mode de protection des données personnelles. » (*ibid.* : § 27).

L'orientation principalement juridique de certaines analyses de la mission est d'ailleurs une critique récurrente. Outre la composition de la mission « constituée exclusivement de juristes » (Mathis, 2018, §11), il est relevé, concernant l'articulation des règles de publicité, d'accès et de mise à disposition des décisions de justice, que la mission « appliqu[e] un *raisonnement strictement juridique* », alors que « s'il applique, en revanche, un *raisonnement économique*, l'État ne sera pas indifférent aux coûts que cette juxtaposition des modes d'accès lui occasionnerait » (*ibid.*). Le dispositif retenu qui consiste à s'assurer que les demandes de diffusion en ligne, éventuellement refusées, ne seront pas contournées par des demandes de communication des décisions directement auprès des greffes « revient à sanctuariser des règles de délivrance de copie conçues à l'ère du papier. Il est donc nécessaire de réexaminer la pertinence d'un maintien des modes existants. » (*ibid.* : § 12). Même dans les cas où les décisions ne peuvent être diffusées aux tiers, elles devraient être numérisées pour communication aux parties et administrations concernées et leurs métadonnées mises à disposition, avec « un champ reprenant le motif de non-publicité » (*ibid.* : § 17). Selon cette vision, l'enjeu est de systématiser la diffusion numérique des données et décisions, en refusant que la non-publicité du contenu devienne un motif pour ne pas en rendre accessibles les données communicables. C'est un changement radical et profond de logique qui est ici défendu, sur le mode de la révolution numérique : ne plus réfléchir à partir des règles,

procédures et façons de faire conçues « à l'ère du papier » ; repenser entièrement l'ouverture des données selon le double principe de l'exhaustivité et de la systémativité propre à la politique de l'*open data*.

Plusieurs paragraphes de cet article suggèrent que la Dila eut été un meilleur choix que celui des plus hautes juridictions pour administrer et gérer, y compris techniquement, la mise à disposition des décisions de justice. Il est regretté qu'« aucun schéma alternatif, par exemple qui conférerait à la Dila un rôle central, ou impliquerait le Service interministériel des archives de France (Siaf), conformément au modèle suisse d'*open data* » ne soit examiné (*ibid.* : § 19). « La connaissance de la jurisprudence entre, certes, dans la vocation des juridictions, et la mission en tire argument pour placer les juridictions suprêmes dans un rôle d'opérateur technique, comparable à celui de la Dila en matière de contenus législatifs. » (*ibid.* : § 20) Or, l'on comprend que si la Dila était aux manettes, les décisions de justice seraient davantage considérées comme des données juridiques comportant certaines particularités, ce qui aurait pour effet de minimiser leur spécificité par rapport aux autres données publiques. D'autant plus que la Dila est un acteur fort de la politique d'*open data*, qui s'est positionné tôt dans la mise en ligne des données de Légifrance, qui comprennent à la fois lois, décrets, codes, et décisions de justice. Les décisions de justice seraient ainsi fondues dans les termes généraux de la politique de l'*open data* gouvernemental, ce que souhaite aussi [Nicolas Mariano] (2017). Ce qui est reproché au rapport Cadiet, c'est donc de ne pas souscrire à l'idée que la libération des données et le recours aux technologies numériques sont la solution à privilégier ; de préférer des arguments et considérations juridiques aux dépens de perspectives économiques ; d'empêcher la constitution d'un acteur fort, généraliste et au pouvoir centralisateur à l'intérieur de l'État dans le domaine des données juridiques (la Dila).

4.1.3.2. Ouvrir le droit, radicalement

L'enjeu c'est donc bien l'ampleur de l'ouverture du droit et la place conférée à une logique de révolution numérique. D'ailleurs, les *startups* se reconnaissent effectivement dans le projet de « libérer » les données de justice pour les réutiliser de manière économiquement rentable. Le projet de l'ouverture du droit est clairement un mantra dans le discours de [Jurissimo]. Ses fondateurs expliquent ainsi : « Notre mission c'est l'ouverture et la transparence du droit ! [...] [Ils] ne comprennent pas pourquoi il est aussi difficile d'accéder aux sources du droit sur Internet alors que l'adage dit : "nul n'est censé ignorer la loi". La réponse est simple : le système n'a pas évolué depuis 200 ans. Les informations sont compilées dans des livres alors que le droit se met à jour toutes les secondes... Ils sont convaincus que la tech et l'intelligence artificielle peuvent changer les choses.⁵⁹ » Cette vidéo réalisée pour le compte de [Jurissimo] met en scène un discours, une stratégie et des arguments marketing, qui attestent d'une volonté radicale d'« ouverture du droit », à la suite de la loi Lemaire. Les fondateurs de [Jurissimo] « attaqu[ent] un marché monopolistique », détenu par « les éditeurs juridiques traditionnels ». Ils insistent sur le fait qu'ils ont mis en ligne 7 millions de décisions de justice, là où seulement 900 000 étaient disponibles, puis bâti

⁵⁹ « [Jurissimo], la legaltech qui bouscule un marché vieux de 200 ans : l'accès aux sources du droit », 23 octobre 2018, consulté le 15 septembre 2021. <https://www.youtube.com/watch?v=62LYOrivRK0>

un outil technologique permettant de numériser et d’anonymiser automatiquement ces décisions, et qu’ainsi ils ont « atteint [leur] premier objectif d’ouverture du droit⁶⁰ ». L’accent est mis sur la simplicité et l’accessibilité de la plateforme, y compris en termes de prix. C’est aussi par « l’offre commerciale » qu’ils ont créé la rupture : « on a divisé les prix par deux [...] et on rajoute une offre résiliable à tout moment [là où les éditeurs proposent un engagement à un ou trois ans].⁶¹ », comparant leur offre à l’abonnement à la plateforme Netflix. La *startup* a réuni des capitaux considérables, une première levée de fonds de 2 millions d’euros puis une seconde de 10 millions, lui ayant permis de recruter une centaine de salariés en trois ans ; et revendique plus de 5 000 clients professionnels du droit, « de l’avocat individuel aux grands cabinets et directions juridiques de toutes tailles⁶² ». La logique, par son caractère ouvertement marchand et axé sur la rentabilité et le calcul économique, est donc aussi en décalage avec les valeurs traditionnelles de la justice et tranche avec le désintéressement comme valeur supérieure du champ juridique (Bourdieu, 1986). En revanche, on perçoit clairement les affinités avec les avocats de droit des affaires (Vauchez, 2016) beaucoup plus proches de l’économie⁶³. Le discours du renouveau, de la révolution par les technologies numériques est partagé par les autres *startups* comme on aura l’occasion d’y revenir dans le chapitre suivant.

L’étude de F. Lebaron (2019) relève que les fondateurs d’entreprises comme [PréviCompute], [Jurissimo], [Prédidroit] partagent plusieurs caractéristiques. La plupart d’entre eux sont jeunes (*ibid.* 115) et animés par la volonté de dépoussiérer le monde du droit, de le moderniser et de le rendre plus ouvert. Ils se retrouvent dans une vision critique voire négative de l’état et du fonctionnement de la Justice en France. Leur « *discours permettrait aisément de les classer parmi les réformateurs, voire réformateurs radicaux* ». Ils font le constat d’une justice archaïque, qui a besoin de se moderniser, ce que leurs algorithmes peuvent faciliter (*ibid.* 123). « L’utilité des algorithmes promus par les acteurs du monde des start-up est indissociable de ces constats sur l’état actuel du monde judiciaire. » (*ibid.* 124).

Cet extrait d’entretien tiré de la même enquête atteste de la conception radicale de ce jeune chef d’entreprise, orientée vers la possibilité d’automatiser les décisions de justice quitte, pour ce faire, à sacrifier la part d’interprétation et donc d’autonomie des juges et acteurs du droit. « – *Je trouve que c’est absurde qu’on ait autant de procédures qui ne soient pas automatisées. Je trouve que c’est absurde que le code fiscal soit écrit à la main, et que ce ne soit pas un vrai code, avec des maths dedans, des algorithmes. Je trouve que c’est absurde que l’on ait des notions qui ne soient pas univoques, que l’on ait des trucs qui doivent encore être interprétés par des gens. Pourquoi est-ce que tout n’est pas écrit en maths en fait ? Que tout soit clair, univoque, cohérent, qu’un même mot veuille dire le même mot dans tous les pans du droit... Ce genre de choses en fait. Que les personnes qui écrivent le code aient fait un*

⁶⁰ Interview BFM-Business, 28 septembre 2017, consulté le 15 septembre 2021.

<https://www.youtube.com/watch?v=BvNxMeAldus>

⁶¹ *Ibid.*

⁶² <https://www.jurissimo/>, consulté le 19 avril 2021.

⁶³ Voir sur ce point la thèse en cours réalisée par Héloïse Eloi-Hammer, « Justice et algorithmes : vers une transformation du métier d’avocat ? », codirigée par Dominique Cardon et Laurence Dumoulin.

petit peu de maths avant quoi, pour assurer la cohérence de tout ça. Non mais sérieusement, c'est un bordel innommable, et c'est ça qui me gêne dans la fabrique du droit. – Il faut des maths pour réorganiser le tout ? – Pour moi, il faut que les gens qui vont faire de la codification de certains pans du droit travaillent avec des logiciens et des mathématiciens, même des informaticiens, pour arriver à construire un droit de manière cohérente. Ça le rendra beaucoup plus facilement traductible en algorithmes, ce qui permettra d'automatiser un certain nombre de procédures. » (Lebaron, 2019 : 123). L'argumentation repose sur la réduction du droit codifié à un code au sens informatique du terme.

4.1.3.3. Une tribune en réponse à l'affaire [Jurissimo]

L'affrontement entre le CNB, le barreau de Paris et les grands éditeurs juridiques avec [Jurissimo] a suscité des prises de position d'autres acteurs qui sont également à placer du côté d'une volonté d'ouverture du droit. Le lendemain même du dépôt de plainte, une tribune est publiée⁶⁴ conjointement par Open Law, l'Association française des juristes d'entreprise (AFJE), l'Union nationale des huissiers de justice (UHNJ), le réseau interprofessionnel Eurojuris, l'association Juriconnexion, le Cercle Montesquieu et l'Association des avocats numériques. L'article reconnaît que les *startups* doivent avoir « une éthique exemplaire », mais rappelle l'inertie et le retard pris par le décret d'application de la loi Lemaire, relevant ainsi la responsabilité des pouvoirs publics dans la situation. Le positionnement de Juriconnexion, qui joue un rôle central dans la mise en évidence d'irrégularités qui auraient été commises par [Jurissimo], est ambivalent parce qu'il est proche des éditeurs juridiques traditionnels, tout en faisant pression pour que l'accès aux décisions de justice soit effectif pour tous, y compris les *startups*. On note par ailleurs que l'attelage ici formé fait la part belle aux juristes d'entreprises, aux huissiers de justice et aux avocats numériques, c'est-à-dire à des professions (huissiers) ou segments de la profession d'avocat tournés vers l'activité économique et vers une échelle transnationale concernant Eurojuris.

Eurojuris est une association française membre d'un collectif européen de cabinets d'avocats (incluant en France les professions de notaire et d'huissier) qui propose un réseau de professionnels de qualité certifiée, afin de répondre au contexte de la globalisation économique : fournir aux clients et pouvoirs publics des interlocuteurs fiables et facilement identifiables en matière de conseil juridique. Eurojuris était représenté au Village de la legaltech dont il est partenaire, il fait partie des espaces de promotion du recours aux technologies numériques, et a notamment encouragé l'utilisation du dispositif algorithmique de traitement et d'analyse des décisions de justice [PréviCompute]⁶⁵. Il n'est pas étonnant de ce fait qu'il fasse alliance autour de la défense d'une ouverture large des données de justice.

⁶⁴ Publiée sur *lesechos.fr*, le 17 juillet 2018, consultée le 26 avril 2021. <https://business.lesechos.fr/directions-juridiques/droit-des-affaires/contentieux/0301983616768-legaltech-construire-le-monde-de-demain-dans-le-dialogue-et-la-collaboration-322309.php>

⁶⁵ Eurojuris a collaboré directement avec [PréviCompute] : il a contribué à l'élaboration du *Guide de la justice prédictive* et au second semestre 2017, il a équipé une centaine de cabinets d'avocats de l'outil [PréviCompute], publié le 14 juin 2018, consulté le 19 septembre 2021. <https://www.eurojuris.fr/categories/actualites-eurojuris-2200/articles/guide-justice-predictive-37649.htm>

D'autres acteurs supranationaux font probablement partie de cette coalition de cause, même si, à ce stade, nous manquons d'éléments empiriques.

On sait en effet que data.gouv.fr et la mission Etalab sont nés à la suite d'opérations de *benchmarking*, qui ont valorisé le modèle américain de portail unique de mise en ligne gratuite des données, aux dépens du modèle de redevances abandonné en France vers 2010-2011 (Goëta, 2016 : 60-61). Le Partenariat pour un gouvernement ouvert, espace international de collaboration, est certainement un lieu où ce modèle et sa dimension généraliste (intersecteur) sont promus et où s'échangent idées et solutions d'action publique autour de l'ouverture des données. Il est probable que d'autres acteurs s'y rattachent.

Graphique 3.4. Les acteurs de la coalition généraliste de l'open data

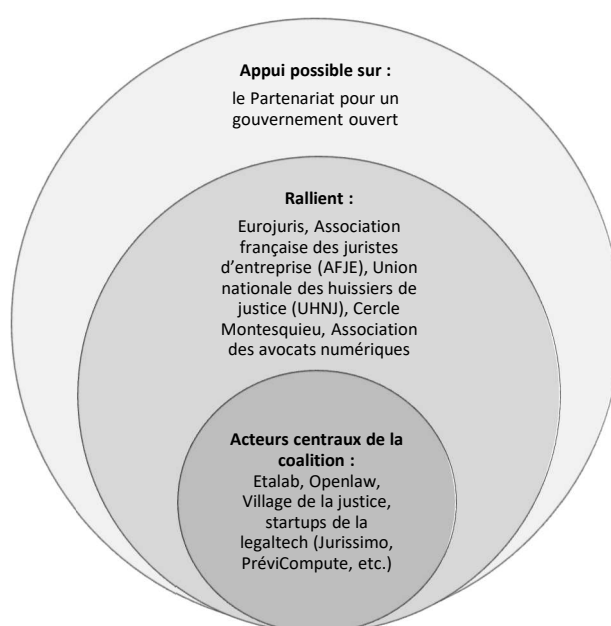


Tableau 3.2. Les croyances des acteurs de la coalition généraliste de l'open data

	Coalition généraliste
Position	Challenger
<i>Policy core beliefs</i>	<ul style="list-style-type: none"> - Les décisions de justice = des données publiques - L'intérêt des partenariats public-privé - Le potentiel libérateur des technologies - La justice, un secteur comme un autre - Le droit comme moyen
<i>Deep core beliefs</i>	<ul style="list-style-type: none"> - Internationalisation et globalisation - Vertus du marché et de l'économie - Croyance dans l'horizontalité des rapports sociaux, la décontraction - Croyance dans la liberté et l'initiative individuelles - Valorisation de la rapidité

Entre les deux coalitions des causes, des acteurs individuels ou collectifs jouent le rôle d'intermédiaires.

4.2. Des *policy brokers* qui font le lien entre les coalitions de cause

Nous avons remarqué que Juriconnexion est en position de charnière entre les deux coalitions de cause. Cette association, proche des éditeurs juridiques classiques, a entrepris l'enquête qui a permis d'objectiver le décalage entre [Jurissimo] et les autres acteurs positionnés sur le marché du traitement, de l'analyse et de la diffusion des décisions de justice. Mais elle a aussi signé la tribune publiée par Open Law et d'autres pour que l'accès aux décisions de justice soit effectif pour tous, y compris les *startups* et dans un délai raisonnable. Sa position est donc clairement à la jonction des deux coalitions et mériterait évidemment d'être mieux cernée.

De la même manière, plusieurs acteurs individuels ont des profils mixtes et effectuent une forme d'intermédiation entre les deux coalitions. C'est le cas de [Valère Desjardins] déjà rencontré dans le chapitre 1. Magistrat, il a fait sa carrière à la cour d'appel de Paris, au ministère de la Justice et à la Cour de cassation, et s'est spécialisé sur des postes comportant une dimension informatique et numérique. À la Chancellerie, il a été adjoint à la sous-direction de l'organisation judiciaire et de l'innovation de la Direction des services judiciaires (DSJ), et à ce titre, a eu à connaître du fonctionnement et de la modernisation de l'informatique des métiers de la magistrature et du greffe (applications civiles et pénales, infrastructures). Il est ensuite devenu responsable du service des études et de la documentation de la Cour de cassation où il a porté des projets liés à l'*open data* des décisions de justice. Il a apporté son concours à la Mission Cadiet de même que Louis [Bourg], autre magistrat avec lequel il a travaillé⁶⁶ et co-écrit un article sur l'*open data* des décisions de justice ([Desjardins] et [Bourg], 2017).

Le discours développé dans cet article est très proche de celui du rapport Cadiet (qui sera rendu quelques mois plus tard) : c'est au ministère de la Justice qu'il appartient de définir le périmètre exact de l'*open data* des décisions de justice, compte tenu du fait qu'il existe déjà des règles juridiques en matière de communication des décisions à des tiers. L'ouverture large et inconditionnelle des décisions de justice est écartée d'emblée pour des raisons juridiques. Par ailleurs, l'expérience de la Cour de cassation est valorisée puisqu'elle possède déjà un savoir-faire en matière de pseudonymisation de ses arrêts, de constitution et d'administration de sa propre base de données (Jurinet), de celle des cours d'appel (Jurica) diffusées sur le site Légifrance. À l'inverse, l'apport réel des outils de traitement et d'analyse proposés par les *startups* est relativisé. Il se limite à l'analyse statistique de contentieux indemnitaires sans entrer dans l'analyse lexicale du contenu des décisions. Enfin, l'attachement à l'interprétation et à une forme d'irréductibilité de la justice à la prévision est revendiqué : « La décision judiciaire implique, en effet, une part de casuistique, d'interprétation et d'adaptation au cas d'espèce impossible à anticiper. [...] Envisagée strictement, l'idée d'une sécurité juridique

⁶⁶ Chef du pôle « Juridictions de droit commun » du bureau du droit de l'organisation judiciaire à la sous-direction de l'organisation judiciaire et de l'innovation, [L.Bourg] a été le rapporteur de la mission Cadiet, remercié par L. Cadiet dans son courrier de fin de mission.

absolue aboutirait à la négation de l'idée même de justice » ([Desjardins] et [Bourg], 2017 : 1487). [V. Desjardins] peut ainsi être rattaché à des éléments forts de la coalition particulariste de l'*open data*.

Toutefois, il adhère aux perspectives d'évolutions ouvertes par les outils et modèles de prévision : ce sont des « perspectives d'évolutions particulièrement attendues », que ce soit pour accroître le règlement amiable des conflits, prendre en compte « des tropismes injustifiés qu'il serait, dès lors, possible de corriger (ex. : tendance à être plus compréhensif envers telle catégorie de justiciables sans raison objective) » ou pour améliorer la qualité de la justice par une plus grande transparence. ([Desjardins], 2019 : 275) De plus, ce magistrat fait très certainement partie de ceux qui ont mis en place l'expérimentation d'une cellule d'anonymisation des décisions de justice par le recours à l'apprentissage machine au sein de la Cour de cassation. Il partage donc avec la coalition généraliste une confiance dans le potentiel libérateur des technologies. Il est probable qu'il se trouve à la charnière entre les deux espaces cognitifs isolés. Il s'avère toutefois difficile d'aller plus loin sans disposer d'entretiens semi-directifs avec les personnes concernées ni d'éléments plus fins sur leurs parcours et entreprises concrètes.

Les deux coalitions de cause identifiées s'affrontent pour tenter de maximiser ou de minimiser l'impact pour la justice de l'adoption d'une nouvelle norme de l'*open data* pour l'ensemble des politiques publiques. C'est d'une certaine manière un rapport global-sectoriel conflictuel qui est à l'œuvre.

4.3. À travers deux coalitions, un rapport global-sectoriel conflictuel

Le rapport global-sectoriel est envisagé par Pierre Muller comme « un opérateur de causalité permettant d'expliquer le changement — non pas *tout* le changement, mais la dialectique entre le changement global tel qu'il est perçu et "parlé" et le changement "nécessaire" des secteurs » (2015 : 73). Deux dimensions du global peuvent être distinguées (Palier et Ravinet, 2015 : 27-28 ; repris par Muller, 2015 : 181-182). Le premier sens est celui du « référentiel global », un cadre d'interprétation du monde qui, dans la pensée de P. Muller, fait couple avec la sectorisation de l'action publique, à travers le rapport global-sectoriel (RGS). Ensuite le global désigne également « l'état du monde qui constitue le cadre au sein duquel les sociétés se pensent et agissent sur elles-mêmes », et qui renvoie donc à « un jeu d'échelles multiples dont l'autre extrémité est le "local" » (Muller, 2015 : 182).

La première coalition repose sur un système de croyances relatif aux spécificités de la justice dans l'État et les politiques publiques, appuyé sur les règles juridiques. Le droit est vu davantage comme une référence que comme une ressource (Commaille, 2015) ; le système de droit écrit comme un patrimoine à entretenir, les décisions de justice comme des sources de jurisprudence dont l'importance relative est définie par la pyramide des juridictions. Plus largement, sont recherchés, voire célébrés, les statuts, la tradition et la sacralité des institutions et les vertus de la régulation par le droit et l'État. Pour autant, cette coalition n'est pas en soi hostile au changement ou aux nouvelles technologies, comme le montre le fait que la Cour de cassation soit elle-même assez moteur dans le recours à des dispositifs numériques et algorithmiques.

Dans cette coalition, figurent en bonne place les grandes directions centrales du ministère de la Justice, les hautes juridictions et les hauts magistrats, porteurs dans l'ensemble de la culture d'un droit légiféré et contrôlé par les hautes instances juridiques selon la perspective pyramidale déjà évoquée. Cette coalition réunit des acteurs qui sont en position dominante dans le secteur de la justice. Dans le contexte de la mondialisation, de la mise en concurrence des différents systèmes juridiques (via les processus de forum shopping, la création de juridictions supranationales, etc.), ses acteurs partagent la défense d'un système juridique de droit écrit, de tradition romano-germanique ainsi que la recherche d'une régulation publique pour limiter l'impact des transformations en cours. Les professions judiciaires et leurs organes de représentation participent à cette coalition, mais certains segments peuvent adopter des positions différentes, c'est le cas des avocats d'affaires qui sont très impliqués et favorables au développement de dispositifs algorithmiques de traitement et d'analyse des décisions de justice comme T. Léonard et M. Moritz l'ont bien montré.

La seconde coalition identifiée correspond aux acteurs qui visent à acclimater dans le secteur de la justice le nouveau référentiel global de l'ouverture des données publiques, en lien avec un monde qui fonctionne de manière globalisée et où, de fait, les données circulent à toutes les échelles. Ces acteurs assurent la médiation entre le global et le sectoriel en s'investissant dans « mettre le changement en mots et en dispositifs d'action publique » (Muller, 2015 : 179). Ils croient au potentiel libérateur des nouvelles technologies, à leur capacité de levier pour bousculer le monde du droit. En outre, pour certains d'entre eux à l'intérieur de l'État comme Etalab et la Dinsic, la politique de l'*open data* représente une occasion de gagner en centralité, et donc en pouvoir, dans le système politico-administratif national. C'est aussi pour l'État l'occasion de tenter de reprendre la main dans un jeu devenu de plus en plus complexe et interdépendant, d'abord du fait de la montée en puissance des règles et instances supranationales comme l'UE, puis de l'intrication de plus en plus grande du niveau globalisé des échanges dans la production d'un ordre mondial, et enfin de la concurrence d'acteurs infra-étatiques. S'ils sont porteurs d'un cadre de référence global, ces acteurs le traduisent et le reconstruisent aussi à partir de spécificités nationales. La prééminence du concept de *startup* d'État rappelle la force des structures centralisatrices et la place de la référence à l'État dans le système français, y compris dans l'importation d'un référentiel et via une configuration de réforme qui agissent pour aménager une place importante aux acteurs privés.

Conclusion :

Déclarer la mise à disposition gratuite des décisions de justice comme le font les articles 20 et 21 de la loi Lemaire pose un principe dont l'implémentation ouvre une nouvelle séquence d'action publique dont nous avons vu les acteurs et temps forts. De ce chapitre, il ressort que sa mise en œuvre dans la justice est marquée par une tension entre deux manières de concevoir l'*open data* qui sont aussi deux manières de voir les décisions de justice, soit comme des données publiques comme les autres en vertu d'une approche généraliste, soit

comme des productions de l'activité juridictionnelle de magistrats indépendants, dont certaines font jurisprudence, et dont la communication est déjà soumise à un régime juridique spécifique, en vertu d'une approche plus particulariste. Dans un cas, l'enjeu est d'étendre le régime d'*open data* aux décisions de justice le plus largement possible ; dans l'autre, de contenir, réguler et encadrer l'*open data* pour le subordonner à des logiques spécifiquement juridiques et juridictionnelles.

Cette tension, isolée empiriquement ici, fait écho à d'autres politiques et instruments d'action publique développés dans la justice. Juristes d'une part et experts de l'administration d'autre part, parce qu'ils sont porteurs d'ethos différents, divergent (Vigour, 2015) sur la manière de concevoir et d'évaluer ce qu'est la qualité de la justice. Dans le cadre de la Cepej en particulier, des hybridations empiriques et incrémentales sont effectuées par des acteurs qui ont institutionnellement une vocation de traduction (Cappellina, 2017).

Nous avons isolé des profils d'acteurs individuels ou collectifs mixtes qui, par leurs positions et trajectoires, mais aussi probablement par leurs cadres cognitifs et normatifs et leurs discours, jouent un rôle d'intermédiaires. Ils contribuent à la circulation d'idées et d'instruments nouveaux qui sont, de fait, vecteurs d'apprentissages et de transformations incrémentales à l'intérieur du secteur de la justice. Dans ce chapitre, nous en avons repéré surtout du côté du droit en général (Open Law). En ce qui concerne la justice à proprement parler, nous allons en identifier davantage dans le chapitre suivant, à travers les acteurs qui participent à l'émergence et aux expérimentations de dispositifs de traitement et d'analyse algorithmique des décisions de justice.

Finalement, qu'en est-il ici de la trajectoire de cette nouvelle politique publique et du changement produit dans la justice ? Jusqu'à présent le développement de l'*open data* des décisions de justice semble être opéré par des acteurs et institutions déjà hégémoniques qui parviennent à le maîtriser à leur profit et dans une logique restrictive. La coalition de cause particulariste, du fait de sa position dominante, capitalise des ressources considérables et est parvenue jusque-là à défendre efficacement sa vision. Le principe de l'ouverture par défaut de toutes les données publiques, acté par la loi de 2016, n'a pas pu être effacé, mais le pouvoir de ce texte concernant les décisions de justice a été singulièrement endigué. Le décret de 2020 qui confère un rôle central aux juridictions suprêmes dans la gestion de l'ouverture des décisions de justice, les dispositions de la loi de 2019 qui limitent les usages possibles des informations relatives aux magistrats et greffiers dans les décisions rendues accessibles, et enfin la circulaire de 2020 qui empêche la récupération de décisions de justice en masse ont considérablement bridé la portée des articles 20 et 21 de la loi Lemaire. Est-ce à dire que c'est le *statu quo* qui l'emporte et que cette coalition triomphera à plus long terme ? Il serait bien imprudent de le penser.

Philippe Bezès souligne en effet qu'une des caractéristiques de l'administration française est que « sa transformation relève de chaînes causales lentes, incrémentales et graduelles ou bien d'effets de cumulativité de petits changements ou d'effets cliquets » (Bezès, 2009 : 57). La justice n'y échappe pas, bien au contraire. À travers l'implémentation de l'*open data* dans ce secteur nous avons vu se constituer une autre coalition de cause qui fait

la promotion d'un monde globalisé, où l'on s'exprime en anglais, où la capacité à être connecté et performant est une valeur en soi, où le droit est plus un moyen qu'une fin. Les avocats d'affaires, déjà les plus internationalisés et les plus connectés, en sont logiquement des soutiens, et se font les médiateurs d'un référentiel global, lui-même marqué par la mondialisation des échanges et la difficulté accrue pour l'État de construire un sens commun. S'il s'avère que le système de croyances défendu par cette coalition de cause est bien le plus en phase avec le référentiel global, il est probable que les multiples changements, petits ou grands, intervenus en une décennie vont faire boule de neige. Création d'Open Law, du Village de la legaltech, structuration de nouveaux acteurs économiques autour de la réutilisation des données et constitution d'un nouveau marché, développement du discours sur le tournant que doit prendre le droit, autant d'éléments à partir desquels ces acteurs pourraient imposer leur système de croyances dans le secteur de la justice, d'autant qu'ils y disposent de certains relais...

Chapitre 4 : Un nouveau répertoire de réforme en action : la fabrique d'une expérimentation de 'justice prédictive' auprès de magistrats

Le précédent chapitre a montré comment une nouvelle configuration de réforme administrative s'est mise en place depuis les années 2010 pour favoriser l'*open data* des décisions de justice et encourager l'innovation dans le droit. Le présent chapitre va maintenant étudier la manière dont est née une expérience concrète qualifiée de 'justice prédictive' qui manifeste que le nouveau répertoire de réforme de l'*open data*, porté par la nouvelle configuration de réforme administrative identifiée, est en action dans le secteur de la justice. Dès le programme du candidat Macron¹, sans qu'il soit explicitement question d'intelligence artificielle (Villani, 2018), les propositions pour la justice font la part belle aux technologies du numérique, concernant des innovations parfois déjà en cours d'expérimentation (plateformes de résolution des conflits en ligne, numérisation des pièces des dossiers, etc.). Les dispositifs algorithmiques de traitement et d'analyse des décisions de justice font clairement partie des propositions du candidat à l'élection présidentielle même si le terme de 'justice prédictive' n'est pas employé².

Nous sommes alors en 2016-2017, période pendant laquelle la 'justice prédictive' émerge dans le débat public, comme que nous l'avons établi dans le second chapitre ; où elle acquiert une consistance et surtout une dimension concrète à travers l'expérimentation [PréviCompute] dans les cours d'appel de Douai et Rennes. Cette expérimentation, qui a contribué à faire exister le « moment 2017 » pour la 'justice prédictive', a consisté à faire tester un dispositif par des magistrats, des juges civils, dans le cadre de leurs activités d'élaboration des jugements. Revenir sur la genèse de ce test et son montage institutionnel permet de suivre le répertoire de l'*open data* en action et de retrouver les acteurs du précédent chapitre.

Rappelons que nous y avons identifié deux coalitions de cause, l'une qualifiée de particulariste et l'autre de généraliste. La première rassemble des acteurs et institutions en position hégémonique dans le monde juridique et le secteur judiciaire (cour de cassation, certains segments du ministère de la Justice, éditeurs juridiques, CNB, etc.), qui partagent un système de croyances centré autour de l'exceptionnalité de la justice (fondée sur l'importance de l'autonomie et de l'interprétation des magistrats) et sur un système de droit écrit organisé autour de la jurisprudence comme collection de précédents d'interprétation de la Cour de cassation. Ces acteurs valorisent la sacralité, la tradition et le temps long. Cela les amène à se coordonner autour d'actions visant à présenter et défendre l'*open data* dans la justice comme un problème public relevant principalement des systèmes judiciaires et professionnels de justice.

¹ <https://en-marche.fr/emmanuel-macron/le-programme/justice>, consulté le 7 avril 2021.

² « Les juges bénéficieront d'outils numériques d'aide à la décision. », *ibid.*

La seconde coalition de cause est composée d'acteurs génériques de l'action publique et de professionnels du droit (Etalab, Openlaw, Village de la justice, *startups* de la *legaltech* : [Jurissimo], [PréviCompute], etc.) qui tentent d'introduire une logique d'ouverture maximale des données dans le secteur de la justice. Ils partagent un système de croyances qui touche à la vertu du progrès technologique et au pouvoir libérateur du numérique ; mais aussi à la logique de l'individu, créateur de richesses, qui s'exprime à travers la figure du *startupeur*. Ils sont favorables à ce que l'*open data* dans la justice s'effectue selon une logique généraliste, c'est-à-dire déployée et organisée de manière comparable aux autres secteurs de l'action publique.

Dans le présent chapitre, ces acteurs sont appréhendés dans des séquences temporellement plus resserrées et à un niveau d'analyse plus microsociologique que dans le précédent. Cela nous conduit à les apercevoir sous un jour en partie nouveau, ou plus exactement à nuancer l'analyse en matière de coalition de cause. Dans l'*Advocacy Coalition Framework*, la coalition de cause repose sur deux dimensions principales : le partage d'un système de croyances par certains acteurs et la coordination entre eux pour agir, dans le sens de la défense de leurs croyances et à travers des mobilisations visant à influencer les décisions publiques. Dans les pages qui suivent, nous allons observer que la coalition généraliste est en réalité traversée par des fractures et des lignes de faille qui se dessinent sur le double plan des croyances et des coordinations, et qui sont en partie liées à la situation concurrentielle dans laquelle certains acteurs de cette coalition sont plongés.

Cela implique que notre approche, pour être pertinente, ne relève plus seulement des outils de la sociologie de l'action publique, mais aussi de ceux de la sociologie économique. Les concepteurs de dispositifs algorithmiques, même s'ils partagent un certain système de croyances, sont aussi des acteurs économiques en concurrence sur un marché. Ils ne sont pas naturellement alignés, mais pourraient travailler à cultiver leurs points communs et leurs intérêts partagés ; et pour certains aspects, élaborer un front commun. Or, les actions de coordination qui pourraient être entreprises ne le sont guère et nous verrons qu'au contraire, autour de cette expérimentation, ces acteurs affichent et exacerbent leurs différences : sur le plan des valeurs qu'ils promeuvent, des dispositifs algorithmiques qu'ils proposent et du discours qu'ils développent.

Cela étant, l'expérimentation résulte des valeurs et méthodes promues par la coalition des challengers — même si elle ne « profite » directement et dans un premier temps qu'à une des *startups*, celle dont le dispositif algorithmique est testé. La mise en place de l'expérimentation résulte des efforts de certains pour tenter de séduire des acteurs installés du système judiciaire, appartenant plutôt à la coalition particulariste et de les intéresser à leur dispositif : des magistrats et des cours d'appel en lien avec le ministère de la Justice. L'enjeu d'innovation suppose en effet de construire des dispositifs d'intéressement (Callon, 1986) avec les acteurs en place, pour les associer à des projets précis. Ces montages sont opérés par-delà les coalitions de cause. Toutefois, le ministère de la Justice et la cour d'appel de Rennes vont retourner cette expérimentation à leur avantage, pour renforcer la place d'acteurs et dispositifs existants (les bases des données et outils d'interrogation déjà proposés par la Cour

de cassation), confortant la logique et les intérêts de certains acteurs de la coalition de cause particulariste.

La genèse de cette expérimentation illustre la manière dont une pluralité de nouveaux outils d'action publique (défis, *hackatons*, etc.) est mobilisée autour d'un projet conçu comme ciblé, limité et temporaire. La jeune pousse [PréviCompute] remporte le prix valant contrat d'expérimentation dans deux juridictions. Si la controverse se focalise sur cette expérimentation et ensuite sur l'annonce de son « échec », [PréviCompute] diffuse largement le terme de 'justice prédictive', en s'en faisant le porte-parole et même l'incarnation (1). Dans le contexte de l'émergence d'un nouvel arrangement marchand sur les données de justice, sur lequel ils sont en concurrence, les autres opérateurs de dispositifs algorithmiques ne s'alignent pas : au contraire, ils s'opposent frontalement à la fois à [PréviCompute] et au terme de 'justice prédictive', la *startup* et la notion apparaissant désormais comme inséparables. Le moment 2017, dont nous avons vu qu'il marque un tournant, peut alors aussi être interprété comme une épreuve, pour [PréviCompute], pour la coalition généraliste et pour la 'justice prédictive' (2). Est-ce à dire que la 'justice prédictive' pourrait être résumée à une tentative de faible ampleur et qualifiée d'échec par certains de ses protagonistes ? Probablement pas puisque cette expérimentation, et la controverse publicisée qui l'a accompagnée, ont été porteuses d'apprentissages cognitifs et de rapprochement d'acteurs qui peuvent produire des effets indirects et profonds sur le positionnement de l'offre de dispositifs algorithmiques à destination des magistrats (3). Les ratages ou « replis » comme Marie Alauzen (2019) les qualifie sont au cœur des tentatives et démarches de modernisation de l'action publique (Weller, 1999). De plus, ils ne sont pas isolés : ils font écho et système avec d'autres tentatives et réflexions portant sur la rationalisation de l'activité juridictionnelle des magistrats.

I / L'expérimentation [PréviCompute], fruit d'un concours et incarnation de la 'justice prédictive'

Pour qui s'intéresse à la 'justice prédictive', le nom de [PréviCompute] est étroitement attaché à la notion, ce qui s'explique à la fois par l'expérimentation de l'outil [PréviCompute] auprès de magistrats (1), par la publicisation de cette expérimentation et de son bilan mitigé (2) et par la stratégie de la *startup* pour incarner la 'justice prédictive' (3).

1.1. Un module de calcul et de data-visualisation né d'un hackaton

L'outil de calcul et de data-visualisation de [PréviCompute] qui sera ensuite testé dans les deux cours d'appel est né dans le cadre favorable d'un programme conçu par Etalab et Open Law. Le montage de cette expérimentation est en lui-même typique de la façon de procéder et des effets attendus, en matière de réforme administrative, qu'il s'agisse du prix Open Law (1) ou bien du jury réuni pour le décerner (2).

1.1.1. [PréviCompute], heureux gagnant d'un prix du programme « Open Case Law »...

Un programme « Open Case Law » est amorcé à l'automne 2016. Co-organisée par Open Law et la Dila, cette opération rassemble de nombreux acteurs publics impliqués dans

la politique de l'*open data* des décisions de justice. Ils appartiennent d'une part à des institutions gouvernementales généralistes comme le Secrétariat général du gouvernement, et d'autre part à des instances sectorielles que ce soit le ministère de la Justice (via la DSJ) ou des juridictions suprêmes (Cour de cassation, Conseil d'État, Cour des comptes, Conseil constitutionnel). Celles-ci tentent en effet de rester centrales dans le processus de mise à disposition au public des décisions de justice, comme nous l'avons vu précédemment. La CNIL et l'INPI y sont également associés, de même que l'ordre des avocats de Paris et une structure publique des politiques sociales, Pôle emploi. Plusieurs opérateurs de dispositifs algorithmiques y participent dont [Anajur], fraîchement créée.

Pendant quatre mois, se succèdent ateliers, rencontres et *hackatons*. Deux défis sont proposés dont un relatif à la jurisprudence³ dont l'objectif est ainsi défini par les organisateurs : « proposer des services innovants susceptibles d'être utiles au ministère de la Justice dans le cadre de sa mission d'administration des tribunaux. Il s'agira notamment de développer des solutions de visualisation des données autour des informations publiques disponibles sur les juridictions civiles et pénales (données statistiques de la justice, décisions du fond, data-visualisation, etc.). »⁴

Il n'y aura qu'un seul gagnant de ce prix : [PréviCompute], une jeune *startup* créée quelques mois auparavant (janvier 2016) par quatre jeunes gens, qui se présentent alors comme « deux élèves avocats technophiles [...] et deux data-scientists⁵ », issus de grandes écoles parisiennes. L'un d'eux, formé d'abord à l'université Paris 2 Panthéon-Assas, puis à Sciences po, a 24 ans et est en école d'avocat lorsqu'il cofonde cette *startup*⁶. C'est lui qui assume la partie relations publiques et qui, de ce fait, incarne [PréviCompute]. L'entreprise s'adresse d'abord aux avocats et directions juridiques vers lesquels l'outil est orienté. À son démarrage, elle est « incubée » par une école d'ingénieurs. Elle est ensuite sélectionnée par NUMA pour un programme d'accélération ; remporte plusieurs récompenses (prix Open Case Law, 24 h de l'innovation juridique). À l'occasion de ce défi, les fondateurs de [PréviCompute] développent un algorithme et un outil de data-visualisation, pour calculer et interroger le « montant des indemnités allouées par une juridiction » dans le cadre de prestations compensatoires et d'indemnités de licenciement. Ils se fondent sur des jeux de données fournies par le groupe [Geoffroy-Maisons]. La cible est d'emblée plus large que celle des

³ L'autre est un défi ECLI. Il s'agit d'une norme européenne qui définit pour les données de jurisprudence des identifiants (URI) et des métadonnées. Il existe l'équivalent pour les données tirées de la législation et l'enjeu est de construire une interopérabilité entre les différents systèmes.

⁴ <https://www.dila.premier-ministre.gouv.fr/actualites/toutes-les-actualites/open-case-law-2016-remise-des-prix-le-droit-ouvert-jurisprudence>, mis à jour le 18 novembre 2016, consulté le 28 avril 2021.

⁵ http://www.lexweb.fr/la-revolution-de-la-justice-predictive-interview-du-fondateur-du-site-PréviCompute/?utm_source=Sociallymap&utm_medium=Sociallymap&utm_campaign=Sociallymap, consulté le 29 juin 2021. Apparemment, un des fondateurs a rapidement quitté l'entreprise, puisque le blog PréviCompute évoque seulement trois d'entre eux.

⁶ Marion Robert, « [Victor Chevalet] ([PréviCompute]) : "Il ne s'agit en aucun cas d'automatiser le rendu de la justice" », *Décideurs magazine*, mis en ligne le 22 novembre 2016, consulté le 26 avril 2021. [https://www.magazine-decideurs.com/news/\[victor-chevalet\]-\[PréviCompute\]-il-ne-s-agit-en-aucun-cas-d-automatiser-le-rendu-de-la-justice](https://www.magazine-decideurs.com/news/[victor-chevalet]-[PréviCompute]-il-ne-s-agit-en-aucun-cas-d-automatiser-le-rendu-de-la-justice)

avocats : c'est un outil « pour la performance des professionnels du droit » et « la prévisibilité et [...] la transparence de la justice ». On comprend que les magistrats sont étroitement liés au dispositif, puisque ce sont leurs décisions qui sont analysées et qu'ils peuvent être utilisateurs de l'outil proposé s'il est pensé comme une aide à la décision.

[PréviCompute] remporte donc le « Prix Jurisprudence ouverte et services associés » avec cette application, alors qualifiée de « solution de data-visualisation ». Les implications de ce prix sont doubles puisque la dotation financière qui lui est associée (8 000 €) fournit le financement nécessaire au développement de l'application⁷, mais aussi la perspective d'une expérimentation par des juridictions, menée avec le concours du ministère de la Justice et dans le cadre d'un contrat *ad hoc* (Licoppe et Dumoulin, 2019 : 542). Il s'agit d'une sorte d'appel d'offres, particulièrement original du point de vue de la fabrique administrative de l'innovation. La signature de la convention d'expérimentation avec le ministère de la Justice intervient en août 2017⁸ et concomitamment [PréviCompute] s'associe avec [Brandt], groupe néerlandais d'édition juridique, qui lui donne accès aux ressources documentaires juridiques dont il dispose⁹. La *startup* se dote ainsi d'un accès propre et durable à la documentation juridique, pour remplacer le fonds [Geoffroy-Maisons] qui avait été momentanément ouvert à tous les participants du concours, par le biais du programme d'Open Law.

1.1.2.... décerné par un jury typique de la nouvelle configuration de réforme en place

La composition du jury qui attribue ce prix, très proche d'ailleurs du groupe d'acteurs associés au programme, n'est pas ordinaire (cf. tableau). Certes, parmi les quinze membres, on observe une dominante de personnalités issues d'institutions publiques. Se mêlent toutefois des acteurs, institutions et profils caractéristiques de cette nouvelle configuration de réforme administrative tournée vers l'*open data*. Ainsi, croise-t-on des acteurs et institutions généralistes porteurs de l'*open data* (Dila, SGG) et des questions de propriété intellectuelle qu'il pose (INPI) : un acteur du secteur des politiques sociales qui a de l'expérience dans les innovations digitales (le directeur Expérience utilisateur et digital de Pôle Emploi), les responsables des services de documentation et d'information juridique des juridictions suprêmes, en matière judiciaire, administrative, financière et constitutionnelle, un magistrat de la Cour des comptes très engagé dans l'ouverture des données publiques sur le modèle de la gratuité et lui-même programmeur et membre de la communauté des logiciels libres, des acteurs privés dont l'un est chef d'une entreprise de conseil en informatique ayant lui-même participé à des programmes Etalab, une autre, la directrice d'un établissement d'enseignement supérieur privé, spécialisé dans la rencontre entre le droit et le milieu des

⁷ Pauline Bousch, « PréviCompute gagne l'Open Case Law 2016 », mis en 22 novembre 2016, consulté le 28 avril 2021. <https://blog.PréviCompute.com/PréviCompute-gagne-lopen-case-law-2016>

⁸ [Victor Chevalet], « Deux cours d'appel testent une solution de justice prédictive », mis en ligne le 10 août 2017, consulté le 28 avril 2021. [https://blog.\[PréviCompute\].com/deux-juridictions-fran%C3%A7aises-testent-une-solution-de-justice-pr%C3%A9dictive](https://blog.[PréviCompute].com/deux-juridictions-fran%C3%A7aises-testent-une-solution-de-justice-pr%C3%A9dictive)

⁹ [Victor Chevalet], « Justice prédictive en France : [Brandt] et [PréviCompute] s'allient », mis en ligne le 16 août 2017, consulté le 28 avril 2021. [https://blog.\[PréviCompute\].com/\[brandt\]-et-\[PréviCompute\]-sallient-pour-d%C3%A9ployer-la-justice-pr%C3%A9dictive-en-france](https://blog.[PréviCompute].com/[brandt]-et-[PréviCompute]-sallient-pour-d%C3%A9ployer-la-justice-pr%C3%A9dictive-en-france)

affaires¹⁰, un troisième, le lauréat d'une édition antérieure du concours et lui-même directeur d'une entreprise spécialisée dans la gestion de données. Dans ce jury, les fonctions de responsabilité sont assumées par des acteurs propriétaires du problème public de l'*open data* juridique : la directrice adjointe de la Dila en est la présidente et les deux coprésidents sont, l'un, le président d'Open Law et l'autre, représentant du ministère de la Justice ; dans les deux cas des personnalités au carrefour entre la haute administration, la réforme de l'État et la justice¹¹. Étrangement, l'ordre des avocats de Paris n'est pas représenté par son bâtonnier ou par un de ses membres élus, mais par un juriste, consultant spécialiste d'informatique juridique, multipositionné dans Open Law, Juriconnexion — et qui n'est pas lui-même avocat.

Tableau 4.1. Composition du jury Open Case Law 2016¹²

	Statut dans le jury	Fonction	Rattachement institutionnel	Type de structure	Domaine de spécialité
Véronique LEHIDEUX	Présidente	Directrice adjointe Dila	Secrétariat général du gouvernement	Public/gouvernemental	Information juridique
Thomas LESUEUR	Coprésident	Adjoint à la sous-direction DSJ	Ministère de la Justice	Public/administration centrale Justice	Gestion et pilotage de réformes
[Arthur Pierre]	Coprésident	Président d'Open Law	Association créée par un pool d'institutions	Secteur associatif	Ouverture du droit et innovations par les technologies
Eudes CHIGÉ	Membre	Chef du service documentation	Secrétariat général du gouvernement	Public/gouvernemental	Information et documentation juridique

¹⁰ L'École des hautes études de droit appliqué (Head) se définit ainsi sur sa page d'accueil : « Issue de la rencontre entre des avocats et des professeurs souhaitant innover dans l'enseignement du droit des affaires en France, l'école HEAD offre un pont entre le cursus universitaire et le monde des affaires. », consulté le 28 avril 2021. <https://www.ecolehead.fr>

¹¹ Devenu magistrat à la Cour des comptes depuis 2005, après une première partie de carrière d'officier de l'armée de l'air, Thomas Lesueur a été en poste à la Direction interministérielle des systèmes d'information et de communication de l'État (Disic) et au Secrétariat général pour la modernisation de l'action publique, comme chef du service Pilotage et maîtrise des risques de 2011 à 2013, avant de passer à Direction des services judiciaires du ministère de la Justice comme chef de service et adjoint au directeur, de 2013 à 2017. Il a ensuite été directeur adjoint du cabinet du ministre de la Transition écologique et solidaire, François de Rugy, puis a été nommé en avril 2019 commissaire général et délégué interministériel au développement durable. <https://www.institut-economiepositive.com/thomas-lesueur/> ; <http://www.afite.org/nominations-details.html?id=372>, 28 septembre 2018, consultées le 29 avril 2021.

¹² <https://www.dila.premier-ministre.gouv.fr/actualites/toutes-les-actualites/open-case-law-2016-remise-des-prix-le-droit-ouvert-jurisprudence>, mis à jour le 18 novembre 2016, consulté le 28 avril 2021.

Ronan GUERLOT	Membre	Conseiller référendaire, adjoint au directeur, Service de documentation, des études et du rapport	Cour de cassation	Public/juridiction suprême	Information et documentation juridique
Pierre-Yves MARTINIE	Membre	Responsable du centre de diffusion de la jurisprudence	Conseil d'État	Public/juridiction suprême	Information et documentation juridique
Lionel BRAU	Membre	Chef du service documentation	Conseil constitutionnel	Public/juridiction suprême	Information et documentation juridique
Mohammed Adnène TROJETTE	Membre en tant que « personnalité qualifiée »	Conseiller référendaire	Cour des comptes	Public/juridiction	Haute administration et technologies numériques ¹³
Julie LADANT	Membre	Directrice de la documentation	Cour des comptes	Public/juridiction	Information et documentation juridique
Françoise QUINAULT	Membre	Responsable des publications	Observatoire de la propriété intellectuelle, INPI	Public	Propriété intellectuelle
Reynald CHAPUIS	Membre	Chargé de mission ¹⁴	Pôle emploi	Public	Digital
Florent ANDRÉ	Membre	Directeur	Société Ooffee de conseil en informatique	Privé	Informatique et innovation numérique avec Etalab ¹⁵

¹³ Magistrat à la Cour des comptes, il est décrit comme « ingénieur en informatique et télécoms, ancien de Sciences po Paris, de l'ENA et développeur » dans le domaine des logiciels libres sur le blog d'Etalab auquel il participe activement (comme mentor de certains programmes), il est l'auteur du rapport sur l'ouverture des données publiques, « Les exceptions au principe de gratuité sont-elles toutes légitimes ? » Rapport au Premier ministre, 2013, <https://entrepreneur-interet-general.etalab.gouv.fr/communaute/2019/adnene-trojette.html>, consulté le 28 avril 2021.

¹⁴ Il est qualifié de manière vague « chargé de mission », mais le poste exact qu'il occupe à l'époque est celui de directeur Expérience utilisateur et digital. Nomination le 1^{er} juillet 2016 pour une prise de poste le 10 juillet 2016. Information tirée de <https://business.lesechos.fr/directions-numeriques/metier-et-carriere/nominations-digital/172466>, consulté le 28 avril 2021.

¹⁵ Il a développé le projet « Mon dashboard d'entreprise », co-développé dans le cadre du programme Open Law avec la Dila, Pôle emploi et l'Institut Mines-Telecom. Il semble qu'il ait également contribué au projet « Répertoire intelligent des procédures silence vaut accord » (Ripsa), développé dans le cadre du programme Open Law, puisqu'il l'a présenté lors d'une journée organisée par Juriconnexion. Information tirée de Miren Lartigue, « Big data juridique : enjeux et opportunités », 29 décembre 2015, consulté le 28 avril 2021.

Estelle SEGONDS-DOMART	Membre	Directrice	École HEAD		Formation droit/affaires
[Georges LAVAUDAN]	Membre	Pas de statut mentionné	Ordre des avocats de Paris	Public	Activité des avocats
Marc SALLIERES	Membre	Chief Data Officer et directeur général	Synaltic Group, société spécialisée dans le data management	Privé	lauréat du prix Dila Le droit ouvert 2014 pour Data asso

Cet espace est très nettement celui d'une communauté qui se constitue autour de la réforme de l'État et de la justice par l'ouverture des données publiques et l'encouragement à ce que des initiatives venues du privé proposent de nouveaux services à destination du public, de clients et/ou de l'administration elle-même. Ici, c'est le ministère de la Justice qui se positionne comme recherchant des innovations pouvant naître de « l'écosystème de la donnée », innovations que ce type d'événement fait vivre et incarne concrètement. Mais il est notable que la place de ce ministère dans le jury soit à la fois privilégiée (avec la coprésidence) et très limitée proportionnellement (1 membre sur 15 seulement). Ce jury est original dans la mesure où il dépasse les logiques administratives classiques, au profit d'un accent mis sur la dimension d'innovation numérique en matière d'information et de documentation juridiques. Les institutions présentes et les personnalités choisies pour les représenter témoignent du fait que c'est une logique de « problème » et de compétence thématique — et pas statutaire — qui prévaut. À travers ce prisme de l'innovation numérique, la justice et le droit sont connectés avec d'autres secteurs de l'action publique ; le public et le privé sont associés. Ce jury incarne la fabrique d'une communauté singulière au croisement de plusieurs institutions, autour de la promotion de l'*open data* par le recours à des innovations numériques. On peut également remarquer que l'aspect concours du hackaton est réinvesti à travers l'inclusion d'un ancien récipiendaire du prix dans le jury, manière de cultiver une dimension transversale et généraliste à travers la force de l'exemple et du parcours original.

À la suite de ce concours, l'expérimentation de ce dispositif est donc lancée dès l'année suivante auprès de deux cours d'appel, Douai et Rennes. Licoppe et Dumoulin (2019) ont analysé le déroulement de ce test dans une des deux cours d'appel. Il y est soutenu par le premier président, et une magistrate, désignée par ce dernier, s'engage activement au service de l'expérimentation pour assurer une médiation entre la *startup* et la cour d'appel. Il est fait appel aux magistrats sur le mode du volontariat, pour tester l'application de calcul des indemnités en matière civile. Seuls quelques-uns s'y intéressent. Une réunion est organisée au printemps 2017 pour présenter le dispositif aux magistrats concernés et lancer l'expérimentation proprement dite. Quelques semaines auparavant, le barreau local avait

<https://business.lesechos.fr/directions-juridiques/droit-des-affaires/contrats-et-clauses/big-data-juridique-enjeux-et-opportunités-205736.php>

annoncé un test du même logiciel par une quinzaine de cabinets d’avocats pilotes dans des matières autres que pénales¹⁶.

1.2. [PréviCompute], l’expérimentation et son « échec » au cœur de la controverse sur la ‘justice prédictive’

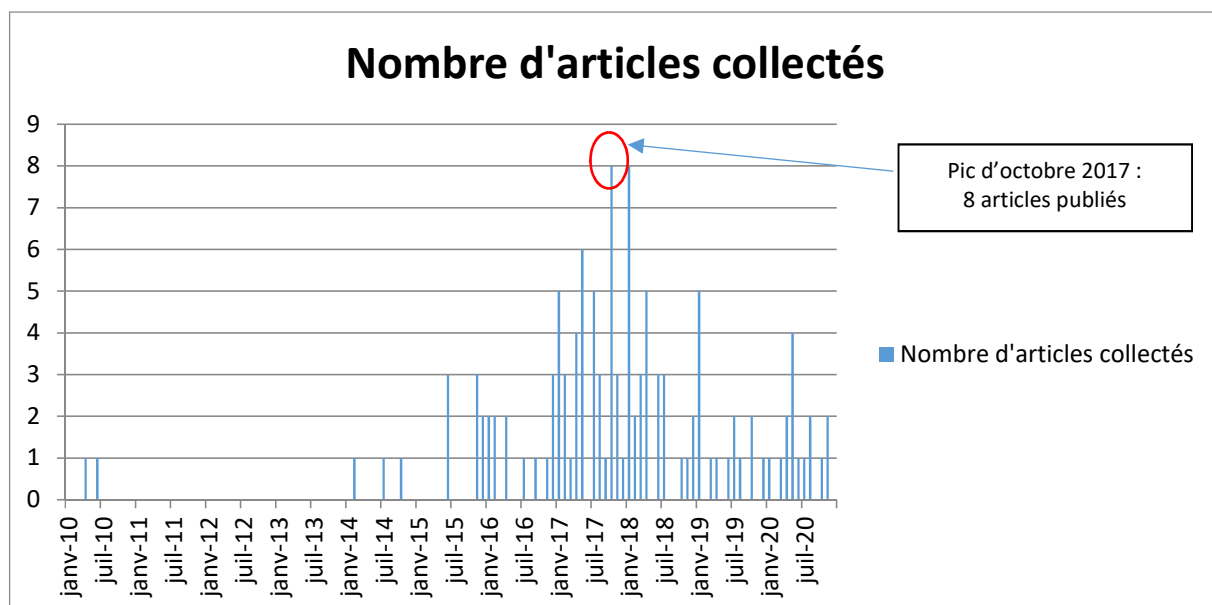
Le bilan de l’expérimentation par les acteurs de la justice — ministère, président de la cour d’appel de Rennes et magistrats — est plus que mitigé (Licoppe, Dumoulin, 2019). À l’automne 2017, le premier président de la cour d’appel de Rennes, en lien avec le ministère de la Justice, dresse le bilan de l’expérimentation en des termes très durs. Le communiqué connaît un important retentissement dans les médias et sur les réseaux sociaux. Nous avons vu dans le chapitre 2 que la date de publication de ce communiqué constitue un tournant, repérable dans les *tweets* et articles de presse. Je vais montrer ici comment cette annonce se répercute sur ces différents supports (1) et revenir sur la place que l’expérimentation occupe dans la structuration des discours sur la ‘justice prédictive’ (2) pour établir enfin combien [PréviCompute] est au cœur de la dynamique propre à Twitter (3).

1.2.1. *Une annonce qui marque un tournant quantitatif et qualitatif*

À l’issue de la période de test, les magistrats volontaires indiquent être plutôt déçus par l’outil ; les premières manipulations effectuées ne les ont pas vraiment convaincus de son intérêt, comparativement aux autres dispositifs qui équipent habituellement leur activité décisionnelle (barèmes, référentiels, bases de données de la Cour de cassation, etc.). Ils l’ont, de ce fait, et malgré leur intérêt initial, progressivement délaissé. Le caractère contre-productif du logiciel est également mis en avant : il peut, dans certains cas, conduire à des résultats absurdes (liés à la non significativité sur un plan statistique des résultats obtenus sur un trop petit nombre de cas par exemple). L’annonce de l’échec de l’expérimentation par communiqué de presse intervient le 9 octobre 2017 marque un tournant quantitatif : c’est un point culminant de l’intensité du débat en ce qui concerne la presse comme le montre le graphique ci-dessous, avec un pic à 8 articles publiés en octobre 2017.

¹⁶ Stéphanie Smatt Pinelli « Le barreau de Lille se lance dans la justice prédictive », *Revue pratique de la prospective et de l’innovation*, n° 1, mars 2017, p. 7.

Graphique 4.1. Distribution mensuelle de la publication des articles de presse du sous-corpus 'justice prédictive', 2010-2020



Le tournant est également qualitatif puisque, à partir du 9 octobre, les articles de presse qui provoquent des pics d'intensité traitent de la 'justice prédictive' d'un point de vue bien plus critique que ce ne fut le cas auparavant. Deux articles en sont emblématiques et sont repris sur Twitter, générant en écho des pics d'intensité sur la plateforme. Le premier, publié dans *Les Echos* traite des limites de la 'justice prédictive' en revenant notamment sur l'échec de l'expérimentation du logiciel [PréviCompute]¹⁷, alors que ce journal vantait jusque-là le potentiel économique des *startups* de 'justice prédictive'. Le second, paru le même jour, évoque les biais de ce type de dispositif et les premiers tests manqués pour le logiciel [PréviCompute]¹⁸.

Le même type de constat s'impose quand on se penche sur le corpus de *tweets*. Le graphique suivant montre que les principaux pics d'intensité repérés à l'automne 2017¹⁹ sont relatifs non seulement à l'expérimentation, mais plus encore à son échec. Sur les sept pics d'intensité repérables, trois d'entre eux²⁰ sont directement liés au communiqué officiel du

¹⁷ Philippe Rozec et Louise Thiebaut, « Intelligence artificielle : les limites de la justice prédictive », *Les Échos*, 2 novembre 2017, consulté le 18 mai 2020 via Europresse.

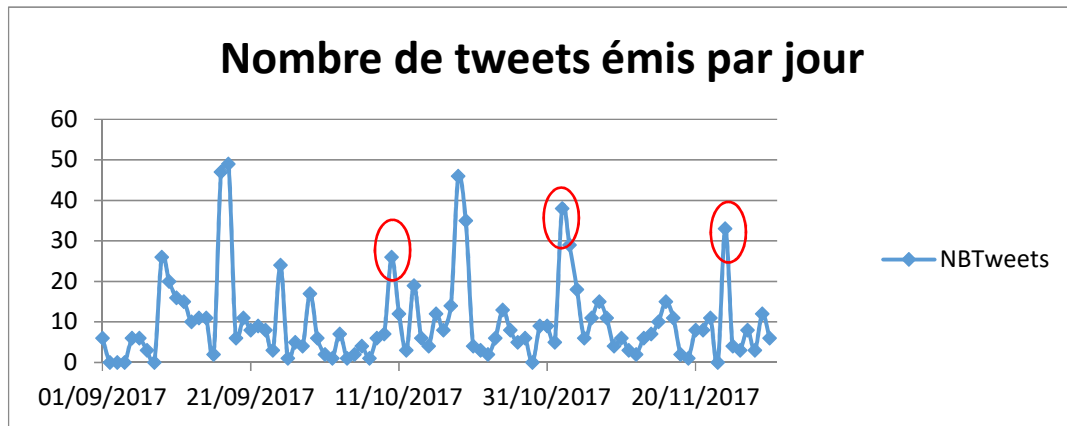
¹⁸ Perrine Signoret, « Justice prédictive : les algorithmes ne sont pas prêts de remplacer les pros », *L'Express*, 2 novembre 2017, consulté le 18 mai 2020. https://lexpansion.lexpress.fr/high-tech/justice-predictive-les-algorithmes-ne-sont-pas-pret-de-remplacer-les-pros_1956703.html

¹⁹ Dans un moment de reprise de l'activité, avec pour les mois de septembre, octobre et novembre 2017 respectivement 330, 281 et 296 *tweets* contre 147 et 64 *tweets* pour les mois de juillet et août 2017.

²⁰ Nous considérons des pics d'intensité dès lors qu'ils dépassent 20 *tweets* par jour. Il s'agit des pics du 10 octobre (26 *tweets*), des 2-3 novembre (38 et 29 *tweets*) et du 24 novembre (33 *tweets*).

ministère de la Justice et de la cour d'appel de Rennes rendant public l'échec de l'expérimentation du logiciel [PréviCompute].

Graphique 4.2. Nombre de tweets émis par jour au cours des mois de septembre, octobre et novembre 2017.



Le pic du 10 octobre 2017 (le premier cerclé sur le graphique) intervient au lendemain de la publication de cette annonce. Il est provoqué par de nombreux *tweets* qui relaient celle-ci et la commentent. Le pic des 2-3 novembre (le deuxième cerclé) est, quant à lui, lié à la parution simultanée de deux articles de presse, critiques sur la 'justice prédictive' et l'expérimentation – déjà évoqués plus haut²¹. Enfin, le dernier pic d'intensité identifié sur cette période intervient le 24 novembre (le troisième cerclé). Il correspond au partage d'un article dédié à la 'justice prédictive' paru sur le site *Frenchweb*. À chaque fois, face aux critiques, [Victor Chevalet] réajuste sa position en reconnaissant que son produit est plus à même de satisfaire les avocats et les directions juridiques²² que les magistrats.

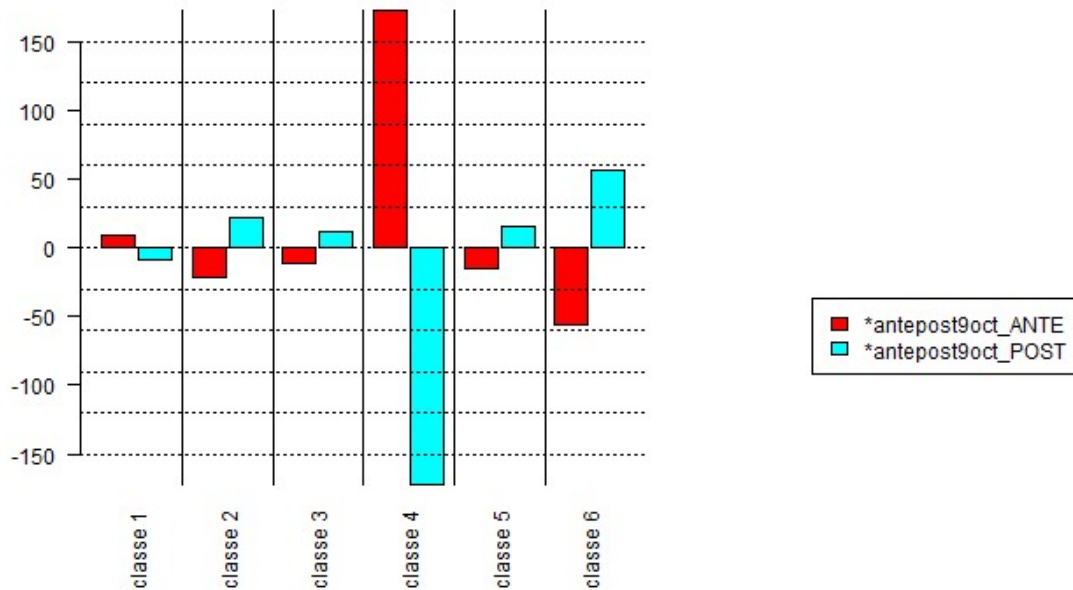
Si l'on s'intéresse désormais à l'analyse de classification des articles de presse réalisée avec le logiciel Iramuteq, il apparaît que toutes les classes traitant de la 'justice prédictive' comme d'un dispositif algorithmique de traitement des décisions de justice (c'est-à-dire les classes 2, 3, 5 et 6)²³ présentent une surreprésentation significative après l'annonce de l'échec de l'expérimentation de [PréviCompute].

²¹ Philippe Rozec et Louise Thiebaut, « Intelligence artificielle : les limites de la justice prédictive », *Les Echos*, 2 novembre 2017 ; Perrine Signoret, « Justice prédictive : les algorithmes ne sont pas prêts de remplacer les pros », *L'Express*, 2 novembre 2017.

²² Anonyme, « PréviCompute, l'analyse prédictive dédiée à la justice », *frenchweb.fr*, 24 novembre 2017, consulté le 18 mai 2020. <https://www.frenchweb.fr/fw-radar-PréviCompute-lanalyse-predictive-dediee-a-la-justice/308360>

²³ Que nous distinguons de celles qui renvoient davantage à la justice actuarielle (classes 1 et 4). Cette interprétation découle la fois de la structure de la classification (1/4 vs 2/3/5/6) et de la projection de celle-ci sur un plan factoriel (cf. chapitre 2). La classe 4 étant appuyée sur une conception de la 'justice prédictive' essentiellement actuarielle et pénale, il est logique qu'elle apparaisse massivement sous-représentée par rapport à la variable construite autour de la date de l'annonce de l'échec de l'expérimentation de [PréviCompute].

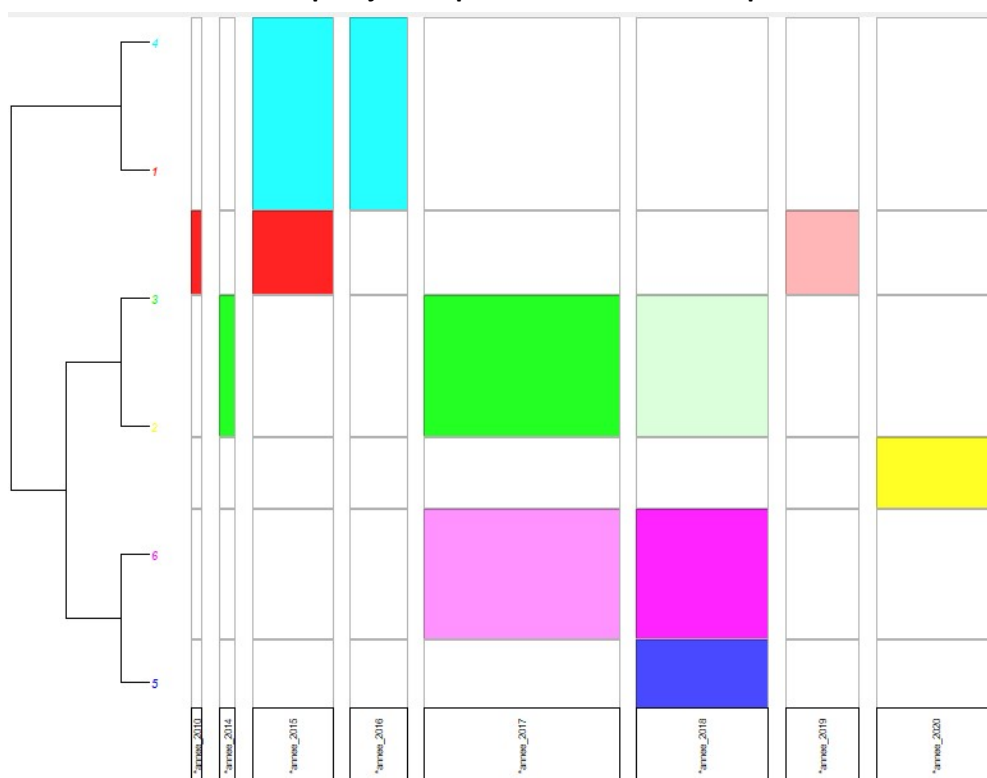
Graphique 4.3. Distribution de la variable antepost9oct au sein de chacune des classes, articles de presse, sous corpus ‘justice prédictive’ 2010-2020



La surreprésentation signifiante la plus marquée concerne la classe 6, laquelle porte justement sur la ‘justice prédictive’ comme aide à la décision des magistrats et l’expérimentation à Douai et Rennes. Même les discours relatifs aux potentialités de la ‘justice prédictive’ (classe 2) sont plus présents après l’annonce qu’avant. Cela tend à signifier à la fois que cette expérimentation nourrit le débat *a posteriori* et ne constitue pas la fin du parcours de cette innovation mais aussi que cette expérimentation, soutenue notamment par la Chancellerie, a été lancée à un moment où la ‘justice prédictive’ ne faisait pas vraiment l’objet de débat public. D’ailleurs, comme le montre le graphique suivant²⁴, au sein de la classe 6 du corpus d’article de presse, si l’année 2017, correspondant à l’expérimentation, est légèrement surreprésentée, c’est surtout 2018 qui est significativement surreprésentée. Les discours de la classe 2 relative aux potentialités de la justice prédictive semblent paradoxalement être aussi nombreux après qu’avant l’échec de l’expérimentation de [PréviCompute].

²⁴ Ce graphique met en évidence de manière synthétique et visuellement parlante l’évolution chronologique des classes de discours surreprésentées, en donnant à voir la surreprésentation de chaque année de publication au sein de chacune des classes. Les différentes échelles permettent de rendre compte à la fois de la part d’articles publiés chaque année, du poids de chacune des classes étudiées et bien évidemment du χ^2 d’appartenance à la classe de chaque année. Les surreprésentations sont identifiées par un remplissage en couleur.

Graphique 4.4. Evolution chronologique des surreprésentations des classes de discours dans le sous corpus 'justice prédictive' d'articles de presse



Tous ces éléments convergent donc pour établir que l'annonce de l'« échec » est marquante dans la controverse autour de la 'justice prédictive'. À l'intérieur du moment 2017, elle fait événement et représente un point de bifurcation dans l'intensité comme dans la nature de la controverse²⁵. L'expérimentation [PréviCompute] apporte donc une contribution particulière au « moment 2017 » précédemment identifié.

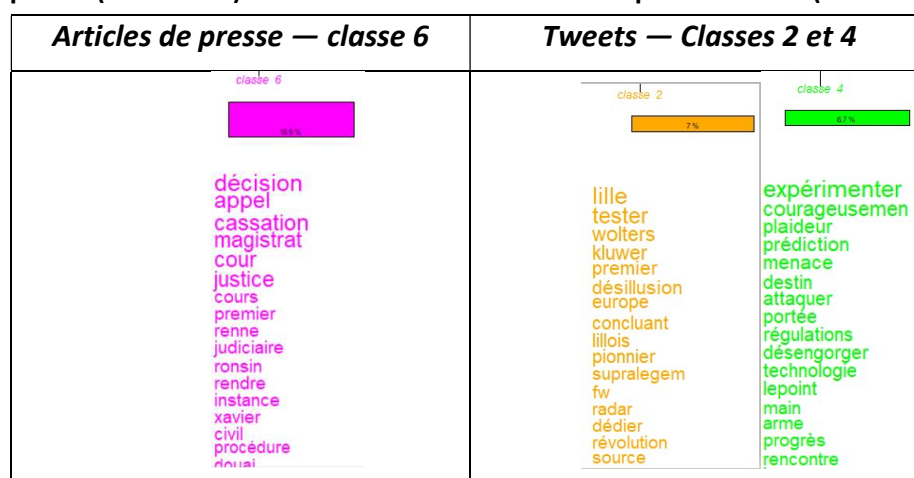
1.2.2. L'expérimentation au cœur de la structuration des discours sur la 'justice prédictive'

Si l'annonce de l'échec représente un temps fort de l'activité médiatique, l'expérimentation est elle-même structurante dans les discours sur la 'justice prédictive', et ce, tous supports confondus, puisque dans les deux corpus 'justice prédictive' (articles et tweets), la classification par la méthode Reinert fait apparaître à chaque fois, une classe de discours qui correspond à l'utilisation de « la justice prédictive » par les magistrats et à l'expérimentation [PréviCompute] en juridiction. La classe 6 contient 18,88 % des segments de texte classés pour les articles de presse c'est-à-dire presque un cinquième de ce corpus et donc une surreprésentation de cette classe par rapport à la moyenne. Pour ce qui est des tweets, c'est la classe 4 qui incarne les discours sur l'expérimentation par les magistrats ; elle est nettement plus modeste (6,7 % des tweets classés), mais la classe 2 (7 % des tweets

²⁵ Une étude de contenu systématique des tweets et des articles de presse, réalisée via un codage et des analyses sous N-vivo, est programmée dans le cadre de l'ANR Just-IA. Elle permettra d'être plus précis sur la nature du changement de discours qui s'opère à partir de l'annonce du bilan mitigé de l'expérimentation dans les deux cours d'appel.

classés) concerne également l'expérimentation concomitante du même outil par des avocats du barreau lillois (*tester, Lille*), comme vu dans le chapitre 2.

Graphique 4.5. Les mots les plus significatifs et le poids de la classe 6 du sous-corpus d'articles de presse (2010-2020) et des classes 2 et 4 du sous-corpus de tweets (2010-2020)



Il est probable (et ce serait à tester) que si l'on constituait un corpus sur « justice quantitative », « jurimétrics » ou d'autres termes présentés comme des alternatives à celui de 'justice prédictive', cet épisode de l'expérimentation serait bien moins déterminant dans les discours. Dans tous les cas, ce lien qui semble exister entre l'expérimentation et la controverse sur le terme 'justice prédictive', passe par la *startup* [PréviCompute].

1.3. [PréviCompute], point de focalisation et porte-parole de la 'justice prédictive'

L'analyse de la littérature juridique sur la 'justice prédictive' fait apparaître une place spécifique pour quelques noms d'outils et de sociétés souvent mentionnés, comme [PréviCompute] et [Anajur] en France. On relève le même phénomène dans le corpus de *tweets* avec [PréviCompute], étroitement associé à la 'justice prédictive' (1). C'est au moins en partie le résultat de la stratégie de l'entreprise pour promouvoir et incarner la notion de 'justice prédictive' (2).

1.3.1. [PréviCompute] au cœur de la dynamique Twitter

Ce nom propre figure à 340 reprises dans le corpus (contre seulement 2 fois dans le corpus de presse de la même période), que ce soit pour désigner le logiciel où est la société qui l'a créé, ce qui le place dans le top 20 des formes les plus présentes. Les discours et débats autour de la 'justice prédictive' sur Twitter se polarisent donc bien sur cet actant/acteur, ce qui n'est pas le cas des débats dans la presse²⁶.

On peut penser que cela est lié au fait que les réseaux sociaux offrent un accès direct à l'expression publique, via la médiation technologique de la plateforme certes, mais sans les intermédiaires que sont les journalistes. Les *startups* peuvent s'y exprimer à volonté, en les investissant comme un espace de débat et de communication promotionnelle. De manière

²⁶ On remarque le même phénomène pour le vocabulaire lié aux *legaltech*, davantage présent dans les *tweets* que dans les articles de presse, sur la même période.

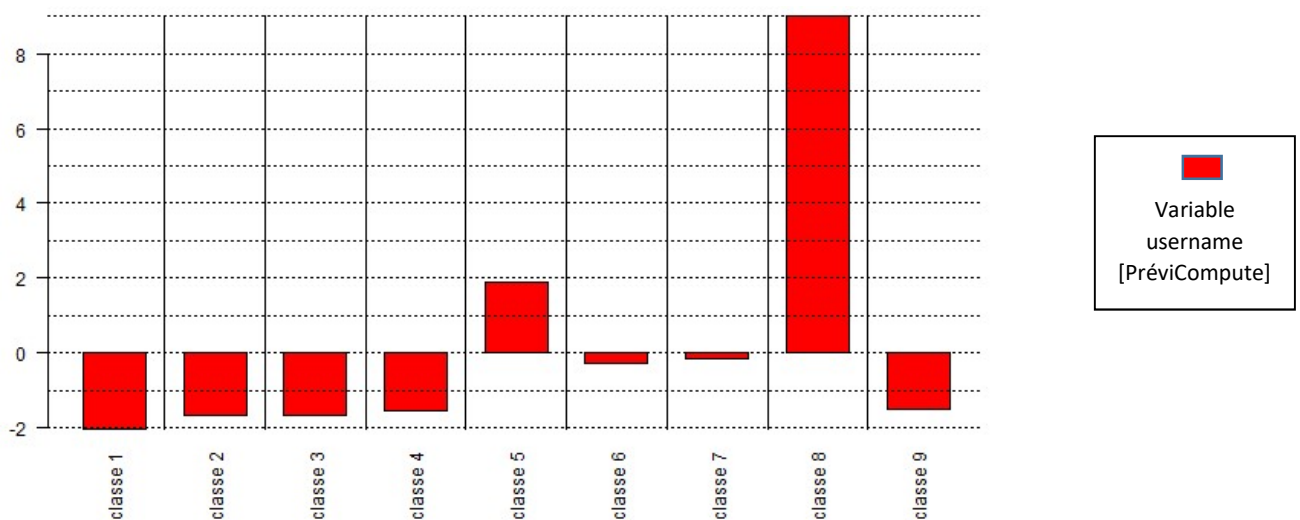
générale, les réseaux sociaux (LinkedIn, etc.) sont très pratiqués par les entreprises de l'innovation technologique et juridique : leur dimension horizontale, l'aspect construction et entretien de réseaux, l'immédiateté sur laquelle ils reposent cadrent bien avec la culture et l'image que veulent donner les créateurs d'entreprises innovantes.

[PréviCompute], comme ses concurrentes, est présente sur les réseaux sociaux. Toutefois, s'agissant précisément du corpus de tweets comportant les mots 'justice prédictive', une forte disparité apparaît. Si l'on compare les opérateurs de dispositifs algorithmiques de traitement et d'analyse des décisions de justice, on s'aperçoit qu'[Anajur] est quatre fois moins mentionné que [PréviCompute], avec seulement 79 occurrences (toutes formes cumulées). [Jurissimo] l'est huit fois moins avec 41 occurrences quand d'autres, [Prédidroit] et [Juri+], ne sont présents que de manière résiduelle, avec quelques occurrences.

Cette saillance particulière de [PréviCompute] est-elle le reflet de la forte médiatisation de l'expérimentation et/ou le résultat d'une stratégie de positionnement de cette *startup* pour diffuser cette expression ? Cela est difficile à dire, mais à titre indicatif, 54 tweets issus de notre corpus proviennent du compte officiel @[PréviCompute] quand c'est le cas de seulement 28 pour [Anajur]. Les tweets postés par @[PréviCompute] et @[Anajur] sont surreprésentés au sein de cette classe.

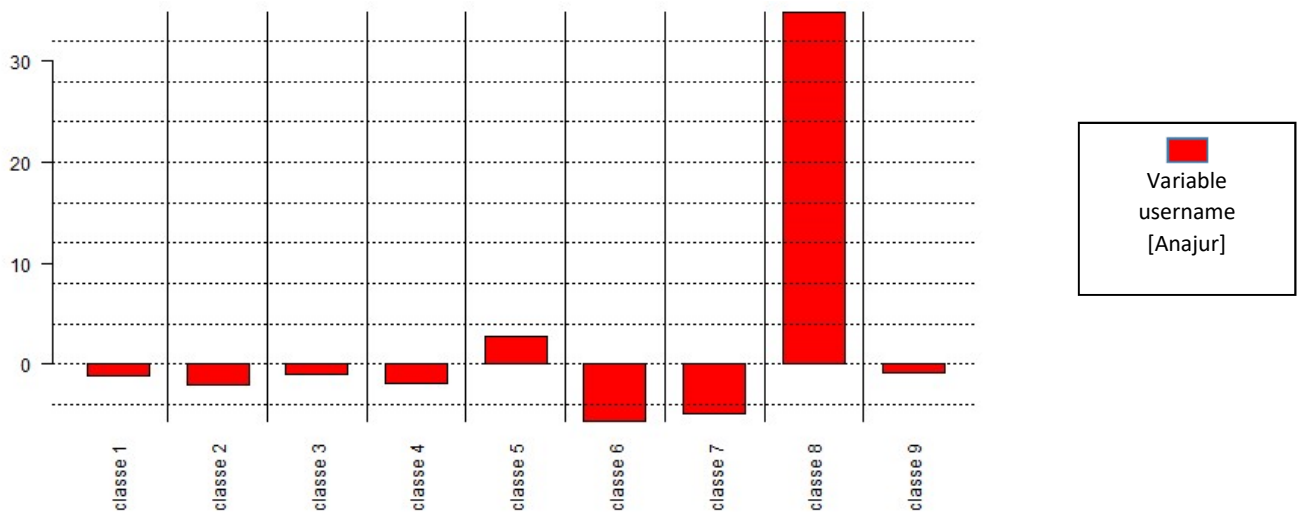
D'autant que l'analyse des tweets publics émis par le compte officiel de [PréviCompute], montre une surreprésentation de ceux qui relaient une activité universitaire (colloque, etc.) liée à la 'justice prédictive'. Ceci suggère que cet acteur promeut les événements de ce type et alimente le *buzz* autour d'eux²⁷.

Graphique 4.6. La distribution des tweets émis par [PréviCompute] au sein des classes.



²⁷ Les autres résultats ne sont pas statistiquement significatifs.

Graphique 4.7. La distribution des tweets émis par [Anajur] au sein des classes.



Il conviendrait d'aller plus loin, notamment à partir d'une analyse de contenu des *tweets*. D'ores et déjà, toutefois, cela est à mettre en relation avec d'autres éléments indiquant que [PréviCompute] s'est efforcé d'incarner la 'justice prédictive'.

1.3.2. La consolidation de l'avantage compétitif conféré par le prix Open Law

[PréviCompute] s'est d'emblée présenté comme faisant de la 'justice prédictive'. Si cette jeune entreprise communique peu dans sa première année d'existence (janvier à décembre 2016), l'expression 'justice prédictive' est déjà présente dans les interviews et *tweets* qu'elle publie. L'ambition poursuivie apparaît même être celle de fonder une nouvelle discipline : « la justice prédictive » comme on le lit dans l'encadré suivant.

Encadré 4.1. L'analyse prédictive par [PréviCompute]

« Processus et fonctionnalités

Pour faire simple, l'analyse prédictive se décompose en trois étapes, qui reviennent à répondre à trois questions :

1 — “Que s'est-il passé ?” : grâce à l'enrichissement des données, c'est-à-dire à l'ajout de métadonnées décrivant les caractéristiques du litige, cette solution est d'abord un formidable moteur de recherche, permettant de proposer aux avocats toutes les informations (textes, jurisprudences, doctrines, informations diverses) dont il aura besoin.

2 — “Que va-t-il se passer ?” : puisqu'il n'existe jamais deux litiges parfaitement identiques, *l'enjeu est alors de comprendre* l'impact d'un facteur ou d'une combinaison de facteurs (analyse multivariée) sur la résolution. C'est ici qu'entrent en jeu les algorithmes de Machine Learning, capables de croiser les observations pour créer des modèles prédictifs complexes. En appliquant ces modèles aux caractéristiques de son litige, l'avocat est ainsi capable d'évaluer les probabilités de résolution.

3 — “Comment l'optimiser ?” : la technologie est capable d'évaluer et de comparer plusieurs stratégies contentieuses, habilitant l'avocat à choisir l'option qui a le plus de chance de succès statistique, en fonction des caractéristiques variables de l'affaire.

Cette nouvelle discipline, qui consiste à appliquer les technologies et méthodologies prédictives à la justice, *est appelée “justice prédictive”, ou encore “[PréviCompute]”*. »²⁸

Surtout, [PréviCompute] a d'emblée breveté la notion afin d'en protéger l'usage au titre de la propriété intellectuelle. Dans un entretien donné en mai 2017, à la question « Comment peut-on définir la justice prédictive ? », la réponse de [Victor Chevalet] est en effet la suivante : « La justice prédictive est avant tout une marque déposée, propriété de [PréviCompute] SAS. Le terme désigne pour nous une pluralité de fonctionnalités qui ont nécessité neuf mois de R&D et de co-innovation avec des cabinets d'avocats partenaires. »²⁹ Cette information est de nature à expliquer, au moins en partie, l'insistance de [PréviCompute] dans la promotion de la ‘justice prédictive’. L'identité de cette *startup* est chevillée à la ‘justice prédictive’ et sa stratégie consiste à faire la promotion de ce terme pour le faire accepter. À partir du moment où [PréviCompute] remporte le prix Open Case Law, avec la perspective d'une expérimentation par des magistrats en juridiction, elle amplifie sa politique de communication visant à être assimilée à la ‘justice prédictive’³⁰. La veille de l'annonce du test du logiciel dans deux juridictions, [PréviCompute] annonce la création du comité éthique et scientifique qu'elle vient d'installer et qui s'intitule « Comité éthique et

²⁸ Tiré de [Victor Chevalet], « La justice prédictive : nouvel horizon juridique », *Le Petit Juriste*, mis en ligne le 11 juillet 2016, consulté le 16 juillet 2021. <https://www.lepetitjuriste.fr/justice-predictive-nouvel-horizon-juridique/>

²⁹ Interview du 3 mai 2017, http://www.lexweb.fr/la-revolution-de-la-justice-predictive-interview-du-fondateur-du-sitePréviCompute/?utm_source=Sociallymap&utm_medium=Sociallymap&utm_campaign=Sociallymap, consulté le 29 juin 2021. On note qu'une série de *tweets* a relayé l'information de ce brevet, mais que l'information est restée cantonnée à un cercle très restreint.

³⁰ Il ne s'agit pas pour autant d'un tournant discursif.

scientifique de la justice prédictive »³¹. L'utilisation de l'article défini « la » suggère que ce comité a vocation à se réunir pour l'ensemble de la 'justice prédictive', voire qu'il fait autorité en la matière.

Ce positionnement stratégique, et la rhétorique qui va avec, se poursuivent ensuite, même lorsque la controverse sur le terme 'justice prédictive' bat son plein. Le titre du *Guide de la justice prédictive* (sous-titré : *À destination des juristes et des avocats du XXI^e siècle*) qui paraît en juin 2018, coréalisateur avec des étudiants de Sciences po Paris le montre bien ; de même que celui du livre blanc intitulé *Les Enjeux éthiques de la justice prédictive* (Huss, Legrand, Sentis, 2018), également rédigé par des étudiants de Sciences po Paris. Ces deux publications relèvent d'une sorte de publiccommunication qui allie la synthèse sur le sujet — y compris en mentionnant par exemple les désaccords sur ce terme³² — et la présentation d'orientations et de dispositions qui sont en réalité spécifiques à [PréviCompute]. Elles manifestent une stratégie bien étudiée par Moritz et Léonard (2020 : 98-105) consistant à s'allier à des acteurs académiques prestigieux qui font partie du réseau personnel des fondateurs de la *startup* ou ont des prédispositions à cette alliance parce qu'ils fréquentent les mêmes réseaux. Marcel Moritz et Thomas Léonard (2020 : 75) relèvent le pouvoir légitimant de ce type de pratique : « Parler de "justice prédictive" plutôt que de [PréviCompute] [...] vient implicitement mettre en scène une unité fictive, le produit de cette présentation déniait alors les antagonismes qui existent [...] entre les différentes conceptions de la "justice prédictive". » [PréviCompute] anime également un blog dans lequel écrivent des étudiants de Sciences po Paris, et se dote d'une « académie ». L'enjeu de validation par la communauté scientifique, juridique essentiellement, est au cœur de sa stratégie de légitimation.

À ce stade, nos analyses, bien que parcellaires, permettent déjà de montrer par ce faisceau d'indices que l'expression de 'justice prédictive' est directement liée, entre autres, à l'activisme de [PréviCompute] qui joue un rôle déterminant dans cette construction sémantique. Le soutien d'avocats du barreau d'affaires ainsi que d'universitaires et académiques a été montré par ailleurs (Moritz et Léonard, 2020 : 30-34, 72-82). Pour autant, malgré ces stratégies plurielles, l'expérimentation du dispositif [PréviCompute] s'avère être une épreuve pour la 'justice prédictive' et pour l'entreprise qui la porte.

³¹ Élise Maillot, « Le comité éthique et scientifique de la justice prédictive », 9 août 2017, consulté le 30 avril 2021. [https://blog.\[PréviCompute\].com/le-comit%C3%A9-%C3%A9thique-et-scientifique-de-la-justice-pr%C3%A9dictive](https://blog.[PréviCompute].com/le-comit%C3%A9-%C3%A9thique-et-scientifique-de-la-justice-pr%C3%A9dictive)

³² Par exemple en ouverture du chapitre 2 du *Guide de la justice prédictive*, consacré aux « Précisions technologiques et sémantiques », où l'on peut lire le paragraphe suivant : « La querelle sur le terme de justice prédictive n'est pas éteinte. "Justice quantitative", "justice analytique" ou "justice prédictive" ? À la limite, peu importe ce débat sémantique, car ces différentes appellations couvrent une réalité unique : celle de l'utilisation d'outils de compréhension du langage naturel pour scanner la jurisprudence et d'outils statistiques pour en ressortir une information claire et compréhensible de tous. ». À la suite de quoi quatre termes techniques faisant consensus (*Machine Learning*, *NLP*, *Dependency Parser* et *Legal Analytics*) sont définis puis trois étapes de « l'analyse prédictive » sont décrites, sans plus revenir sur la controverse (p. 14-15).

II / L'expérimentation [PréviCompute], une épreuve pour la 'justice prédictive'

La notion d'épreuve est entendue ici, et dans la tradition pragmatiste, comme le dispositif social de sanction d'une controverse. L'expérimentation [PréviCompute] est une épreuve, au sens où elle représente un moment critique au sein de la controverse sur la 'justice prédictive'. La controverse est reconfigurée par le fait qu'à l'occasion de ce moment critique, les positions des acteurs sont déplacées, les alliances reconfigurées et les termes du débat requalifiés³³. Le passage de cette épreuve se solde par des questionnements accrus sur la pertinence de l'expression (1). Il concourt également à un désalignement des opérateurs de dispositifs algorithmiques concurrents entre eux et en désaccord sur la 'justice prédictive' et à de nouvelles alliances, entre le Conseil national des barreaux (CNB) et [Anajur] par exemple (2). La coalition généraliste qui promeut l'ouverture large et inconditionnelle des décisions de justice est bien travaillée par des lignes de fractures auxquelles s'ajoutent des concurrences économiques.

2.1. Des startups pas tellement alignées

Les startups partagent globalement le même système de croyances autour des vertus du numérique. Cela étant, coexistent des positionnements variés s'agissant des motivations et manières de faire, qui traduisent aussi des formes de concurrence parfois exacerbées. Tous ne se retrouvent pas dans l'image et la volonté du coup de pied dans la fourmilière, ne serait-ce que parce que leurs profils et parcours (1), leurs motivations et intérêts économiques sont différents (2). Probablement faut-il ajouter que l'épreuve de l'expérimentation [PréviCompute] les a amenés à faire évoluer leurs positionnements et leurs prises de parole.

2.1.1. Des orientations différentes : « entrepreneurs de l'innovation » vs « concepteurs scientifiques »

Pour Marcel Moritz et Thomas Léonard (2020 : 28), la disparité parmi les acteurs opérateurs d'algorithmes peut être appréhendée sous l'angle d'une opposition « entre les "entrepreneurs de l'innovation", fortement mobilisés dans la démarche de diffusion et de légitimation des "démarches innovantes", d'un côté, et, à l'opposé, les "concepteurs scientifiques", davantage tournés vers la recherche de scientificité et de résolution des problèmes techniques. » Selon eux, [PréviCompute] et [Juri+] incarnent la démarche des « entrepreneurs de l'innovation » tandis que [Anajur] et [Legalplus] reflètent davantage l'idéal type de la démarche des « concepteurs scientifiques », qui privilégie « une approche réflexive et critique vis-à-vis des outils de justice prédictive, laquelle s'observe notamment par le discours que leurs concepteurs portent à l'égard des potentialités et des limites de leurs outils respectifs. » (*ibid.*, 36).

Ils associent [PréviCompute] au profil de l'entrepreneur de l'innovation notamment à partir de l'analyse des interviews données par l'un de ses fondateurs, [Victor Chevalet]. Ils relèvent des fluctuations dans la manière dont il se positionne par rapport à la 'justice prédictive' (Moritz et Léonard, 2020 : 44). Il adapte son discours en fonction des publics et des

³³ Danilo Martucelli, « Les deux voies de la notion d'épreuve en sociologie. », *Sociologie*, vol. 6, n° 1, 2015, p. 43-60.

circonstances, et manifeste un *ethos* plus « communicationnel » que « scientifique » (*ibid.*, 46-47) ; davantage orienté vers des cibles marketing et le souci de préserver une image valorisante de l'entreprise et du produit qu'elle commercialise que d'aborder d'un point de vue scientifique et critique la notion de 'justice prédictive' mise en cause. De fait, si [Victor Chevalet] ne renie pas la 'justice prédictive' au moment de l'annonce de l'échec de l'expérimentation, il opère toutefois une inflexion de discours, reconnaissant que le module de calcul est davantage conçu et performant pour les avocats et les directions juridiques que pour les magistrats, comme on l'a vu plus haut. Cela place les fondateurs de la *startup* et singulièrement [Victor Chevalet], davantage du côté de « l'innovateur entrepreneur » que du « concepteur scientifique » (*ibid.*, 47).

2.1.2. Des profils et motivations hétérogènes

Ces disparités se retrouvent aussi dans des profils variés du point de vue de leur statut d'innovateur ; de leur discipline et spécialité d'origine (droit ou mathématique, informatique) ; de leur propre projet d'innovation plus ou moins conçu comme une activité marchande.

En sociologie économique, trois figures ou profils d'innovateurs entrepreneurs sont distingués (Cloutier, 2014 : 389), en fonction du contexte à l'origine de l'activité innovatrice — notamment l'existence au pas d'un lien de subordination — et des modes de valorisation du projet d'innovation. Les académiques sont issus de structures de l'enseignement supérieur et de la recherche et peuvent déposer un brevet, développer une activité de valorisation marchande dans leur établissement puis en dehors à travers des mécanismes et fonds d'aide à l'incubation d'entreprises. Les intrapreneurs, quant à eux, appartiennent à une entreprise à partir de laquelle ils développent leur innovation et peuvent, à un certain moment, quitter celle-ci pour développer leur propre activité commerciale. Les autonomes, enfin, « portent un projet d'invention sans lien de subordination explicite » (Cloutier, 2014 : 392), ils peuvent être professions libérales, retraités ou étudiants et doivent mener un processus de recherche de ressources hétérogènes que facilitent depuis quelques années les dispositifs publics d'aide aux *startups*.

Les deux cofondateurs d'[Anajur] illustrent grossièrement les deux premiers profils. L'un est chercheur mathématicien dans une structure publique. Son parcours est le reflet d'un processus d'encouragement public à ce que les académiques transfèrent et valorisent leurs résultats à travers la création d'entreprises (Lamy, 2014 : 104). Le second est un magistrat, en disponibilité pour se consacrer au projet d'[Anajur]. Certes, l'institution judiciaire n'est pas une entreprise et les magistrats n'ont pas pour fonction explicite de développer des innovations comme dans le service R&D d'une grande entreprise. Toutefois, la justice est aussi une organisation et, en tant que magistrat, même en disponibilité, [Jérôme Monta] dispose de ressources spécifiques liées à son statut de membre de cette structure (légitimité, réseaux, connaissance des besoins). Les fondateurs de [PréviCompute] ou de [Prédidroit], étudiants ou doctorants lorsqu'ils se lancent dans leur projet d'innovation, correspondent au troisième profil, le plus répandu parmi les concepteurs de dispositifs algorithmiques de traitement et d'analyse des décisions de justice.

En ce qui concerne les appartenances disciplinaires, plusieurs acteurs de l'innovation algorithmique dans le droit sont eux-mêmes juristes, plutôt venus du droit des affaires, c'est le cas de [Morgan Forestier] de [Legalplus], d'[Axel Johnson] de [Prédroit] ou de [Victor Chevalet] et de [Guillaume Chapuis], cofondateurs de [PréviCompute], ou bien encore de [Jérôme Monta], magistrat spécialisé en droit des affaires et cofondateur d'[Anajur]. Ils s'allient avec des informaticiens, ingénieurs, mathématiciens pour travailler à la conception de dispositifs algorithmiques de traitement et d'analyse des décisions de justice et tirer parti de l'ouverture de ce type de données. Cette proximité ne doit toutefois pas masquer d'importantes différences dans la manière de construire leurs projets et de les inscrire ou pas dans un projet radical d'ouverture et de libération des données. Leurs rapports à l'open source divergent, lorsqu'il s'agit non plus de leur propre accès aux décisions de justice, mais bien de l'accès aux outils qu'ils ont créés.

Certains profils s'inscrivent davantage dans la recherche d'un positionnement sur ce marché. Plusieurs fondateurs de *startups* viennent d'écoles de commerce et d'ingénieurs et entendent préserver la valeur marchande et capitalistique de leurs innovations, c'est le cas par exemple des fondateurs de [Jurissimo].

[Victor Chevalet], avant d'être élève avocat et de cofonder [PréviCompute] en janvier 2016, avait contribué à créer, l'année précédente, une autre *startup* numérique dans le secteur de l'art³⁴ — *a priori* assez éloigné du droit et suggérant un profil plus innovateur et commercial généraliste que juriste innovateur. Il décrit son parcours dans les termes suivants : « *Après un cursus universitaire classique, de la licence de droit jusqu'à l'obtention d'un master 2 en droit public des affaires à l'université Paris I Panthéon-Sorbonne, j'ai intégré l'EFB. Je n'ai pourtant jamais prêté serment. Je ne me suis pas engagé dans l'avocature car j'ai choisi la voie de l'entrepreneuriat, en créant avec deux associés et amis une première entreprise, The Wall, qui proposait une galerie de street-art en ligne.* »³⁵ À l'évidence, la démarche au fondement de [PréviCompute] est entrepreneuriale : réalisation d'une étude de marché, création d'une structure de commercialisation du produit et investissement à temps plein dans l'activité³⁶ ; dépôt d'un brevet protégeant l'innovation ; efforts pour faire de [PréviCompute] la marque qui monopolise la 'justice prédictive'. [Victor Chevalet] explique

³⁴ En l'occurrence la société The Wall, 25 juin 2015, consulté le 1^{er} mai 2021, <https://blog.horyou.com/the-wall-par-victor-chevalet/>

³⁵ *Verbatim* tiré de l'autoportrait qu'il réalise lors d'une interview pour l'association IDPA, Gazette n° 37/avril 2019. Propos recueillis par Gaspard Terray et Martin Charron, mis en ligne le 23 avril 2019, consulté le 1^{er} mai 2021. <https://association-idpa.com/blog-invites/2019/4/victor-chevalet>

³⁶ Il explique en effet : « *Avec un ami juriste, nous avons commencé à faire de l'analyse statistique, mais de manière très artisanale. [...] Pour améliorer notre approche, nous nous sommes rapprochés de deux copains ingénieurs, l'un spécialiste des solutions de maintenance prédictive et l'autre développeur full stack. Ensemble, nous avons réalisé une étude de marché, créé une société en janvier 2016 et démissionné de nos travaux respectifs pour se concentrer à plein temps sur ce projet.* », *verbatim* tiré de l'interview par Marion Robert, « [Victor Chevalet] ([PréviCompute]) : "Il ne s'agit en aucun cas d'automatiser le rendu de la justice" », *Décideurs magazine*, mis en ligne le 22 novembre 2016, consulté le 13 février 2022. <https://www.magazine-decideurs.com/news/victor-chevalet-PréviCompute-il-ne-s-agit-en-aucun-cas-d-automatiser-le-rendu-de-la-justice>

encore, dans un article sur l'éthique, qu'il n'est pas possible de rendre les codes transparents pour des raisons de compétition économique : « dans un monde où la donnée est libre — c'est le cas des décisions de justice —, les logiciels ne peuvent l'être complètement, au risque de tuer les avantages compétitifs acquis en construisant la technologie pour les exploiter. Contraindre les entreprises à mettre leurs algorithmes à la disposition de tous reviendrait à bloquer les cycles d'innovation. » ([Chevalet], 2018 : 88)

Une trajectoire comme celle de [Géraud Monta] est assez nettement différente. Docteur en droit avec une thèse dans le domaine de l'espionnage industriel, il devient juriste d'entreprise en droit des contrats et de l'informatique. Il entre ensuite à l'ENM, et alors magistrat, il prend des responsabilités en administration centrale dans le domaine du droit économique. Pendant une mise en disponibilité qu'il a demandée pour des raisons familiales, il investit un projet de *startup* avec un mathématicien issu d'une structure de recherche publique : « [Philippe Sinclair], que je connaissais de longue date, m'a ensuite sollicité pour travailler sur le projet [Anajur], ce que j'ai fait avec beaucoup de bonheur au sein de [structure publique]. Aujourd'hui, j'interviens comme avocat en droit du numérique avec, toujours, un goût prononcé pour les domaines du droit où il n'y a rien ou presque et où tout est à construire. »³⁷. Il s'est retiré apparemment d'[Anajur] vers 2018, est effectivement devenu un temps avocat en droit du numérique puis a réintégré la magistrature.

Enfin, [Morgan Forestier], juriste fiscaliste devenu *data scientist*, déjà rencontré dans le précédent chapitre, valorise l'open source, et se positionne en faveur de la transparence des codes sources, s'agissant des algorithmes traitant des décisions de justice. Il est à l'origine de l'algorithme de [Legalplus] dont l'accès est libre et gratuit. Sa perspective est plutôt celle de l'ouverture du droit, du défi scientifique, voire de l'engagement civique, apparemment relativement déconnecté des aspects financiers et marchands — ce qui ne l'empêche pas de faire carrière dans le domaine et d'avoir aussi accès à d'autres rétributions plus indirectes.

Tous ces profils sont clairement ceux de passeurs entre les technologies et le droit. Le droit des affaires apparaît comme le point nodal de cette articulation. Un autre point commun est le caractère très masculin de ces profils, ce qui est une caractéristique qui emprunte davantage au monde des *startups* et des technologies qu'aux métiers de la justice, en majorité exercés par des femmes. En revanche, une certaine diversité des métiers et statuts économiques apparaît : avocature et magistrature, activité libérale ou de service public. L'expérience et la socialisation au milieu judiciaire sont elles-mêmes inégales : la bascule vers une *startup* intervient avant même d'avoir terminé sa formation dans un cas ou bien à l'issue de plusieurs années d'expérience juridique ou judiciaire dans d'autres. Cela dessine des profils et des rapports variés non pas tant aux innovations, au numérique qu'à la dimension économique et marchande de l'activité créée et qu'au monde judiciaire lui-même. Nous retrouvons ici une hétérogénéité des motivations et revendications déjà identifiée dans le précédent chapitre à partir de l'*open data*.

³⁷ Verbatim tiré de l'autoportrait qu'il réalise lors d'une interview pour Inria Alumni, à l'occasion de son entrée dans le comité d'administration de cette association, 29 janvier 2018, consulté le 1^{er} mai 2021. [https://www.inria-alumni.fr/\[geraud-monta\]-avocat-comite-administration-inria-alumni/](https://www.inria-alumni.fr/[geraud-monta]-avocat-comite-administration-inria-alumni/)

2.1.3. Différentes lignes stratégiques

Se positionner sur un marché économique, pour ceux qui privilégient cette dimension, suppose d'entrer dans un monde concurrentiel où les spécificités des différents produits commercialisés sont stratégiques. Les dispositifs algorithmiques de traitement et d'analyse des décisions de justice ont en commun de fonctionner à partir de bases de données de jugements et d'arrêts. Mais comme on l'a vu plus tôt avec l'épisode [Jurissimo], le nombre et la provenance de ces décisions de justice, variables suivant les opérateurs, sont au cœur d'affrontements : 2 millions sur les Editionsducouchant.fr, 2,9 millions sur [LegalEditions]³⁸, entre 7 et 8 millions sur [Jurissimo]³⁹. Ce dernier ne manque pas de valoriser l'avantage concurrentiel obtenu avec le slogan « Toute l'information juridique » ou bien encore « Toute l'information juridique disponible est sur [Jurissimo]. Quelle que soit votre spécialité. »⁴⁰

Depuis, [PréviCompute] a investi ce registre de la compétition en instaurant un programme collaboratif intitulé « 100 000 décisions de justice » qui vise à convaincre ses clients de contribuer à alimenter sa base en fournissant les décisions dont ils disposent. L'objectif annoncé est de « créer la plus grande base de jurisprudence inédite française » sans attendre l'horizon de 2025 annoncé par l'arrêté du 28 avril 2021. « Contribuer à cette collecte [...] permet d'accéder à toutes les décisions fournies par les autres utilisateurs et de stocker et de valoriser [les] décisions sur [PréviCompute]. » En juillet 2021, le compteur affiche un volume de 11 007 décisions jusque-là inédites⁴¹. D'autres acteurs du marché n'entrent pas dans la course au plus grand nombre de décisions. Ainsi [Anajur] insiste au contraire sur le passage par des équations mathématiques, lesquelles rendent les résultats moins tributaires du volume de décisions collectées ([Monta] et [Sinclair], 2017 : 501).

Cet exemple montre bien la concurrence qui est au cœur des rapports entre opérateurs. Chaque acteur économique est en quête d'une position sur le marché des algorithmes de traitement et d'analyse des décisions de justice, ce qui implique de se démarquer de ses concurrents. Mais il faut préciser que cette situation objective de concurrence pourrait faire l'objet de stratégies différenciées que les approches de management distinguent selon que la compétition est plus ou moins agressive et la coopération plus ou moins développée (Chiambaretto, Fernandez, Le Roy, 2019). On pourrait imaginer que des conventions soient mises en place sous la forme de chartes, de création d'associations professionnelles, etc. autour des outils d'aide à la décision juridique et judiciaire par exemple, et pour défendre l'idée même de développer ces outils dans une perspective éthique. Au-delà de la participation concomitante à des événements organisés par Open Law par exemple, il n'y a pas de formes de coopérations ou de constructions de conventions communes : la compétition est ici directe et frontale et la coopération apparemment inexistante. Quand des formes de régulation éthique interviennent, elles sont soit propres à

³⁸ <https://www.cnb.avocat.fr/fr/actualites/le-cnb-et-le-barreau-de-paris-ont-depose-une-plainte-penale-contre-jurissimo>, consulté le 17 juillet 2021.

³⁹ Selon les sources, cf. « Piratage massif de données au tribunal », *Le Monde*, suppl. Éco & Entreprise, 29 juin 2018.

⁴⁰ <https://www.jurissimo.fr/fonctionnalites#section-search>

⁴¹ <https://blog.previcompute.com/previcompute-decisions-de-justice>, consulté le 16 juillet 2021.

un acteur et n'associent pas d'autres opérateurs (le guide et la charte [PréviCompute] par exemple), soit issues d'un acteur de la coalition particulariste (par exemple la charte du Conseil national des barreaux), et visent alors à canaliser les concepteurs d'algorithmes pour leur faire reconnaître et intégrer les normes spécifiques au monde judiciaire. L'enjeu est de les « domestiquer », après l'amère expérience de l'affaire [Jurissimo].

En matière de positionnement par rapport à des propositions de politiques publiques, on perçoit également une distinction entre, d'un côté, [PréviCompute] qui développe des formes d'autorégulation et, de ce fait, s'inscrit davantage dans une vision dans laquelle la régulation est confiée aux acteurs économiques du domaine chargés de la production (comme c'est le cas pour le respect de nombreuses normes par exemple d'hygiène et de sécurité confiées aux industriels et opérateurs privés) et, de l'autre, [Anajur] qui semble davantage prête à défendre la nécessité de normes et régulations publiques, par exemple européennes. Tout cela reste bien sûr à investiguer plus en profondeur.

2.2. L'opposition autour de l'expression 'justice prédictive' : [PréviCompute] vs [CLA]

Une ligne de fracture supplémentaire concerne explicitement la notion même de 'justice prédictive', ce qui est d'une certaine manière un angle mort du travail proposé par Moritz et Léonard. [PréviCompute] promeut cette expression et s'efforce d'en tirer le monopole. Ce n'est toutefois pas le cas de ses concurrents qui, pour la plupart, rejettent ce terme pour lui en préférer d'autres, et qui travaillent ainsi, eux aussi, à définir leur identité. L'appartenance à une même communauté de l'innovation technologique dans le droit n'exclut pas la concurrence qu'incarne le débat développé par les concepteurs eux-mêmes sur la manière de désigner les dispositifs algorithmiques qu'ils mettent sur le marché. La 'justice prédictive' est une notion que [PréviCompute] promeut avec une grande vigueur ; ce qui alimente la controverse et fait naître une ligne de fracture avec d'autres innovateurs du droit, en particulier avec [Anajur].

2.2.1. Deux projets d'emblée en concurrence...

Les prises de parole du côté de [PréviCompute] et d'[Anajur] peuvent être lues en miroir. Dans un article publié dès juillet 2016, c'est-à-dire lorsque la notion même de 'justice prédictive' est alors à peine présente dans le débat public, [Victor Chevalet], explique le mode de fonctionnement de l'algorithme dont il est le cofondateur. La spécificité et la « nouveauté » de ce dispositif reposent sur le fait que les algorithmes peuvent travailler en langage naturel, grâce au *machine learning*. À la question « Où en est le marché de la justice prédictive aujourd'hui en France ? », il répond : « *Il y a un an, la justice prédictive était perçue comme farfelue, voire comme une arnaque. Aujourd'hui, elle génère 100 % de satisfaction auprès de nos clients. Quelques personnes travaillent sur la thématique de la quantification du risque ou de la prévisibilité, mais le font d'une manière très différente de la nôtre. C'est le cas d'une société créée par un ancien magistrat et un mathématicien. Ces derniers réalisent des régressions linéaires sur des modèles schématisés à l'avance. Ils réalisent un travail en profondeur sur une question de droit, comme la pension alimentaire, et essaient de trouver tous les facteurs qui peuvent jouer. Leur outil est plus complexe que le nôtre, ne peut traiter que quelques problématiques juridiques et présente le désavantage de devoir être remodelé*

à chaque apparition d'un nouveau critère ou revirement de jurisprudence. *Chez nous*, tout est automatisé. Et l'intégralité du contentieux est traitée.⁴² » Cette longue réponse, très argumentée, montre que l'enjeu de positionnement par rapport à la concurrence et singulièrement par rapport à [Anajur] est présent dès le démarrage de la *startup*.

L'inverse est également vrai. [Anajur] insiste sur le fait que son produit ne repose pas sur de simples statistiques dont les limites sont connues (via les *spurious correlations*). La démarche est détaillée dans ces termes : « une fois que des corrélations ont été repérées entre des variables, on recherche des équations qui permettent, étant donné certaines de ces variables, de calculer celle qui nous intéresse. [...] Ces équations obtenues, on évalue leurs paramètres sur des données du passé, et on teste le tout sur des données indépendantes de celles utilisées pour construire le modèle. En cas de succès, *c'est seulement ensuite que l'on se risque à faire des prédictions*. Il est facile de comprendre que *cette approche, plus exigeante, fournit des résultats beaucoup plus robustes que le simple calcul de corrélations statistiques*. Elle règle aussi le premier problème de l'approche statistique : une fois le modèle mis au point, on peut l'appliquer dans des cas où le nombre de données dans la base correspondant aux critères entrés par l'utilisateur est faible, voire nul, car on ne se réfère plus à ces données, mais seulement aux équations pour obtenir les résultats » ([Monta] et [Sinclair], 2017 : 501). Ce faisant, [G. Monta] et [P. Sinclair] opèrent à demi-mot la critique de dispositifs concurrents fondés sur le seul calcul statistique qu'ils estiment fragile et donc peu fiable lorsqu'il porte sur de petites quantités de décisions et parce qu'il opère une confusion entre corrélation et causalité. Ils cherchent à se démarquer de leurs concurrents à la fois sur le plan de la méthode et sur celui de l'ambition, puisqu'ils rejettent l'expression de 'justice prédictive' et son caractère excessivement déterministe.

2.2.2.... qui divergent sur la notion de 'justice prédictive'

Le décalage entre les deux approches apparaît à travers des manières différentes de se positionner : [PréviCompute] revendique de faire de la 'justice prédictive' et de transformer en profondeur le système juridique et judiciaire là où le discours d'[Anajur], qui met en avant la notion de « justice quantitative », insiste sur un enjeu moins révolutionnaire et plus en phase avec la coalition de cause particulariste. Les outils algorithmiques peuvent permettre de tirer pleinement parti des décisions de justice. Dans un article de doctrine pour la revue Dalloz, les cofondateurs d'[Anajur] expliquent les mécanismes par lesquels l'intelligence artificielle peut être mise « au service de la valorisation du patrimoine jurisprudentiel ». La notion de « patrimoine jurisprudentiel », sans être définie en tant que telle, fait référence dans leurs propos, aux « millions de décisions rendues par toutes les juridictions chaque année » qu'il s'agit « de donner à voir [...] sur un sujet défini, ce qui est informatif sans être prescriptif » [(2017 : 500). La notion permet de penser l'*open data* des décisions de justice et

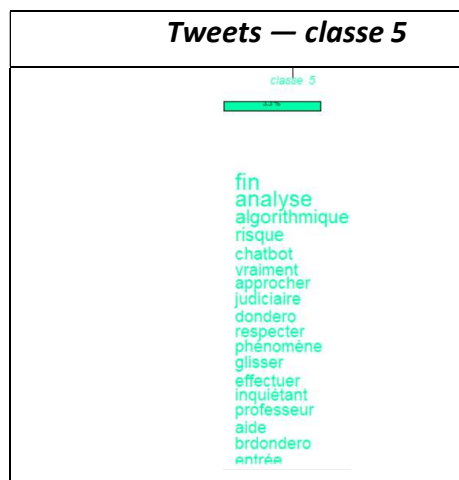
⁴² Verbatim tiré de l'interview par Marion Robert, « [Victor Chevalet] ([PréviCompute]) : "Il ne s'agit en aucun cas d'automatiser le rendu de la justice" », Décideurs magazine, mis en ligne le 22 novembre 2016, consulté le 13 février 2022. [https://www.magazine-decideurs.com/news/\[victor-chevalet\]-\[PréviCompute\]-il-ne-s-agit-en-aucun-cas-d-automatiser-le-rendu-de-la-justice](https://www.magazine-decideurs.com/news/[victor-chevalet]-[PréviCompute]-il-ne-s-agit-en-aucun-cas-d-automatiser-le-rendu-de-la-justice)

les dispositifs qui traitent ces données comme étant un prolongement du travail jurisprudentiel des juridictions, qui ne menace pas la liberté d'appréciation des magistrats.

A contrario, dans un article publié par les *Archives de philosophie du droit*, [V. Chevalet] présente la rencontre entre le droit et l'intelligence artificielle comme un « big bang », une « explosion créatrice », provoquée par ce qu'il qualifie de « technologie de rupture » (2018 : 287). Il y défend la thèse selon laquelle « l'avènement de la science de la donnée judiciaire [...] est en train de faire voler en éclat la hiérarchie des juridictions », mais aussi de rapprocher les systèmes de *common law* et de droit écrit vers un nouveau paradigme qu'il qualifie de « droit isométrique », c'est-à-dire « issu de la synthèse de l'étude exhaustive des décisions de justice, ce système conduirait à l'émergence d'une autre source de droit : celle résultant de la synthèse de toutes les décisions collectives, sorte de pensée collective de tous les magistrats !⁴³ » ([Chevalet] 2018 : enjeux du numérique).

Cette différence de positionnement à l'égard de la notion de 'justice prédictive' se retrouve dans le corpus de *tweets* que nous avons étudié. Dans un premier travail sur les *tweets* de 2010 à 2019 (et non 2020), une classe de discours était construite autour du vocable *risque* dans le double sens de l'analyse du « risque juridique » ce qui est l'objet de la 'justice prédictive', et des *risques* que le développement de celle-ci ferait courir au système de droit français.

Graphique 4.8. Corpus de tweets 'justice prédictive' 2010-2019



Dans cette classe, comme on peut le voir ci-dessus, étaient particulièrement saillantes deux expressions alternatives de celle de 'justice prédictive' : « l'analyse algorithmique » et le calcul de « l'aléa juridique ». La première fait écho à celle utilisée par [Philippe Sinclair], Lémy Godefroy et Frédéric Lebaron (2019) dans le rapport rendu dans le cadre de la mission de

⁴³ [https://blog.\[previcompute\].com/quest-ce-que-la-justice-pr%C3%A9dictive](https://blog.[previcompute].com/quest-ce-que-la-justice-pr%C3%A9dictive), consulté le 24 juillet 2021. Dans un autre texte, une définition sensiblement différente est donnée. Le droit isométrique est plutôt assimilé à « un nouveau système juridique [...] dans lequel l'importance normative des décisions des hautes juridictions ne disparaîtra pas, mais sera progressivement dissoute dans la mesure où toutes les décisions, de fond comme de hautes juridictions, peuvent se valoir dans une statistique donnant une idée extensive de l'interprétation du droit par les juges ». ([Chevalet], 2018, enjeux du numérique : 91)

recherche droit et justice proposant la notion de « modes algorithmiques d'analyse des décisions (MAAD) » (Godefroy, 2019 : 38). La seconde est utilisée dans plusieurs interviews et articles des fondateurs de [Anajur] [(2017 : 500) : « La justice prédictive, que nous préférons pour notre part nommer *justice quantitative* » ; « Ne parlons plus de justice prédictive, mais de *justice quantitative*, car *l'aléa* est inhérent au jugement »]. [Anajur] a beaucoup communiqué autour de l'idée d'une « justice quantitative », se démarquant de la 'justice prédictive' et récusant même l'expression et de ce qu'elle dit des représentations et démarches à la fois scientifiques et juridiques qu'elle véhicule.

Dans cette même classe, l'opposition entre [Anajur] et [PréviCompute] est observable à travers les mentions de deux articles de presse qui les concernent et qui sont ensuite largement relayés via des *tweets* et des *retweets*, mais aussi au détour de conversations, typiques du mode d'interaction propre aux réseaux sociaux. Dans l'échange de *tweets* ci-dessous, la différence d'approche entre [PréviCompute] et [Anajur] est mentionnée par un premier locuteur auquel [Anajur] répond en se démarquant plus nettement de [PréviCompute] et de la 'justice prédictive' (« [Anajur] *ne fait pas de "justice prédictive"* »), et en s'y opposant par la convocation de l'aléa judiciaire (« *Au contraire, notre approche est fondée sur le fait que l'aléa judiciaire est à la fois inévitable et souhaitable.* »).

Encadré 4.2. Échange de tweets quant à l'enjeu de définition de la 'justice prédictive'.

@MarrauddesGrot : « @BrDondero : 3 types de #justicePredictive — @[PréviCompute] qui fait de la #statistique — @[Anajur] : méthode du #juge virtuel : "leur approche a l'air plus prudente" — Analyse #sémantique : expérience des arrêts de #CEDH Ces outils ne sont pour l'instant q des supports »

@[Anajur] : « Merci de nous mentionner. Toutefois [Anajur] ne fait pas de "justice prédictive". Au contraire, notre approche est fondée sur le fait que l'aléa judiciaire est à la fois inévitable et souhaitable. »

Lorsque nous avons étendu notre corpus à l'année 2020, nous n'avons pas retrouvé la même saillance d'une classe de discours qui incarne le discours d'[Anajur] en réaction à celui de [PréviCompute]. Nous reviendrons plus loin sur l'interprétation de ce changement.

2.2.3.... qui enrôlent différents acteurs et partenaires

L'enjeu de compétition entre [PréviCompute] et [Anajur] transparait également à travers l'enrôlement de cercles d'acteurs autour de chacune de ces *startups*. Nous avons vu précédemment comment [PréviCompute] s'est entourée d'institutions académiques et universitaires prestigieuses, notamment avec l'École de droit de Sciences po Paris. [Anajur], quant à elle, se prévaut des résultats très favorables de l'enquête commanditée auprès d'un cabinet d'études par le CNB — déjà mentionnée (cf. *supra* et chapitre 1). Sur son blog, elle publie un long billet sous le titre « Seule [Anajur] présente les caractéristiques d'une technologie de rupture » dans lequel elle cite longuement les *verbatim* de l'étude qui « a permis de démontrer que le terme de justice prédictive ne repose sur aucun fondement et ne correspond pas à la réalité du paysage technologique actuel dans le champ des Legaltechs. Pour autant, ce champ est marqué par une importante diversité d'approche, et les modules de simulation créés [sic] par [Anajur] sont les seuls éléments identifiés pouvant se rapprocher du terme. Pour décrire cette ambition, l'expression de justice simulative semble plus adéquate.⁴⁴ » La *startup* ne manque pas d'expliquer longuement la position et les recommandations du CNB, autour de l'adoption des termes « jurimétrie » ou « justice simulative » à l'exclusion de tout autre dont celui de « justice prédictive⁴⁵ », indiquant ainsi combien elle marque des points sur le registre central du débat avec [PréviCompute].

Cette alliance entre le Conseil national des barreaux et [Anajur] fait sens si l'on considère le positionnement moins révolutionnaire de cette *startup* (valorisation de la diversité du patrimoine jurisprudentiel en laissant la place pour l'interprétation des magistrats ; discours plus « prudent » sur les résultats comme sur les effets du dispositif) et par l'accent mis sur le service public (un magistrat, un chercheur d'une structure publique de recherche) dans sa communication. La dynamique de la controverse autour de la 'justice

⁴⁴ Les passages en gras sont le fait de l'auteur du billet. Laurine Tavitian, « Seule [Anajur] présente les caractéristiques d'une technologie de rupture », 30 octobre 2020, consulté le 30 avril 2021. <https://www.anajur.com/rapport-cnb/>

⁴⁵<https://www.cnb.avocat.fr/fr/actualites/preconisations-dactions-pour-les-legaltechs-du-domaine-de-la-jurimetrie>, consulté le 30 avril 2021.

prédictive' et de [Jurissimo] a montré la sensibilité de ces sujets et les risques pour l'institution judiciaire que des sociétés perçues comme complètement *outsiders* (au sens d'étrangères au domaine) deviennent incontournables dans le bouleversement du contexte lié à l'*open data*.

Derrière les dispositifs algorithmiques de traitement et d'analyse des décisions de justice que certains associent à l'expression de 'justice prédictive', se dessine donc une communauté d'innovateurs qui partagent un certain regard sur la justice, mais que traversent aussi lignes de fracture, polarisations et concurrences. Dans la coalition de cause généraliste, non seulement les *startups* ne se coordonnent pas entre elles, mais au contraire elles entretiennent des rapports de concurrence économique dans un marché en construction. La controverse autour de la 'justice prédictive' creuse l'écart d'alignement sur les croyances, notamment avec une différence d'appréciation de l'aléa propre au jugement que traduit le fait d'accepter ou pas la validité intellectuelle de la notion de 'justice prédictive' (sur le plan de ce que c'est que rendre justice). La prolifération de termes visant à remplacer celui de 'justice prédictive', pointée dans le premier chapitre, n'est donc pas seulement le fruit d'une inventivité particulière, mais aussi le résultat d'un travail de riposte d'acteurs individuels et collectifs, pour certains coordonnés entre eux, afin de lutter contre ce terme qui nie une spécificité importante de l'activité de justice que défendent les acteurs de la coalition particulariste de la justice. Des liens et alliances se développent donc qui transcendent les coalitions, ainsi entre le Conseil national des barreaux et [Anajur] autour de la promotion du terme de 'jurimétrie'.

III / Quels effets de la controverse ?

Finalement, quelles conclusions tirer de cette expérimentation et de la controverse qui s'est nouée autour de la 'justice prédictive' ? Celle-ci se résume-t-elle au test infructueux par quelques magistrats d'un outil inabouti de traitement et d'analyse des décisions de justice ? Il importe de restituer l'épaisseur historique de cet épisode pour aller au-delà de cette première impression. L'expérimentation contribue à l'histoire de la traduction de la politique de l'*open data* dans le secteur judiciaire et à ce titre, les prises de position et débats intervenus, produisent des effets. On repère en premier lieu des effets de consolidation du mot-valise et de la catégorie de 'justice prédictive' (1). Ils nécessitent en retour un travail de détricotage de la part de ceux qui ne se reconnaissent pas dans ce que cette expression dit du rôle des outils algorithmiques de traitement et d'analyse des décisions de justice dans le processus judiciaire (2). Par ailleurs, cet épisode s'accompagne d'autres défis et innovations mis en place qui ont, eux aussi, un succès modéré, mais manifestent l'ambition du ministère de la Justice de se doter de ses propres outils (3).

3.1. La 'justice prédictive' existe

Ce que la 'justice prédictive' recouvre exactement reste très flou et la validité même de l'expression a été fortement discutée, comme montré au début de ce mémoire. Pour autant, l'expérimentation [PréviCompute] et les débats qui l'ont suivie ont contribué à installer le terme dans les esprits (1), à susciter des repositionnements de la part des différents

acteurs (2) et participent enfin d'un bouquet d'opérations qui tentent une mise en comparabilité accrue des décisions de justice (3).

3.1.1. Un mot-valise qui s'impose

Le terme 'justice prédictive' est devenu plus banal et courant, ce que traduit le fait qu'il existe en tant que *hashtag*, largement repris — et pas seulement par [PréviCompute] — y compris pour étiqueter des *tweets* qui portent en réalité sur [Anajur] ou émanent d'elle, comme c'est le cas des cinq *tweets* les plus fréquents de la classe 5 du corpus 2010-2020, consacrée à la révolution numérique.

Encadré 4.3. Les 5 tweets les plus caractéristiques de la classe 5, corpus de tweets 'justice prédictive', 2010-2020

« **[Anajur]** au **service** de la **quantification** du **risque juridique** #IA #legaltech #justicepredictive »

« **RT** kbratings : **@[anajur]** qui **développe** et propose aux **professionnels** du **droit** des **outils d'anticipation** des **aléas juridiques** lève 2 millions d'**euros** pour **financer** son **développement** à l'**international** #justicepredictive #legaltech #regtech #fren »

« **@[anajur]** propose aux **professionnels** du **droit** des **solutions** de **quantification** du **risque** qui se fondent sur une collaboration entre juristes et **mathématiciens** #1000startups #justiceprédicative #legaltech #regtech #IA #INRIA »

« **@contact[anajur]** Les **mathématiques** au **service** de la prise de décision **juridique**. À lire aussi pour mieux comprendre : Justice prédictive, vers une **analyse** très **fine** du **risque juridique** #legaltech »

« **@[anajur]** qui **développe** et propose aux **professionnels** du **droit** des **outils d'anticipation** des **aléas juridiques** lève 2 millions d'**euros** pour **financer** son **développement** à l'**international** #justicepredictive #legaltech #regtech #frenchtech »

Le terme a ainsi acquis un certain pouvoir d'étiquetage — lequel n'est toutefois pas définitif. De plus, on peut interpréter l'étude Sopra Steria commanditée par le CNB comme un effort pour contrer et miner l'influence acquise par ce terme. L'organe de représentation de la profession d'avocat recommande en effet explicitement de bannir cette expression pour la remplacer par celui de jurimétrie voire de justice simulative. Une telle recommandation n'a de sens précisément que dans un contexte où ces acteurs perçoivent la tendance à l'institutionnalisation de ce terme, amplement repris comme montré en chapitre 1.

3.1.2. Des usages difficiles à objectiver, mais qui tendent à se développer

Pour les magistrats, acteurs publics qui entretiennent un rapport spécifique à l'État, le terme s'institutionnalise davantage que les pratiques puisqu'à l'issue de l'expérimentation [PréviCompute], il ne se passe rien, sinon ce constat d'échec et une apparente absence de développement de l'outil dans les juridictions.

Parfois gratuitement mis à la disposition des magistrats dans les tribunaux et cours — ce que fait par exemple [Anajur] — ces dispositifs algorithmiques ne sont pas pour autant appropriés. Il est en réalité difficile d'avoir accès à ces informations d'usages. D'autres expérimentations sont tentées, portées par exemple par [LegalEditions] à Strasbourg. Elles ne sont guère médiatisées, et donc là encore, peu d'informations sont disponibles.

Toutefois, les dispositifs algorithmiques de traitement et d'analyse des décisions de justice continuent à s'implanter auprès d'autres professions juridiques : avocats, directions juridiques de grandes entreprises, compagnies d'assurances. [Anajur] ne communique pas sur ses clients, hormis la mise en ligne de 26 interviews, principalement des avocats. [PréviCompute] communique davantage et affirme avoir pour clients 2 000 avocats exerçant en solo⁴⁶ et « 56 des 100 plus grands cabinets d'avocats⁴⁷ ». De grands noms de l'assurance sont également cités sur le site de [PréviCompute]⁴⁸ : Allianz, Maaf, Civis, Macsf, Covéa, Matmut, Groupama Océan Indien, Crédit agricole Pacifica. On peut donc penser — sous réserve d'informations plus précises — que l'implantation dans différentes professions juridiques progresse.

3.3. Détricotages, ripostes, apprentissages et repositionnements

Les effets de l'expérimentation et de la controverse qu'elle a nourrie se font ressentir de différentes manières : à travers le retournement de cet échec en remise en selle des instances traditionnelles de la coalition particulariste et le repositionnement des concepteurs d'algorithmes dans une logique plus traditionnelle, faite d'héritages.

3.2.1. L'échec de l'expérimentation retourné en faveur de la coalition particulariste

L'annonce de « l'échec » de l'expérimentation intervient dans un communiqué conjoint du ministère de la Justice et du premier président de la cour d'appel de Rennes⁴⁹. Y est affirmé le caractère non pleinement satisfaisant de l'algorithme utilisé, en décalage avec les promesses avancées : « Il est apparu que le logiciel, participant d'une approche dont la modernité était reconnue, méritait d'être sensiblement amélioré.⁵⁰ » La comparaison avec les outils existants n'est pas à son avantage puisqu'il « ne présentait pas en l'état de plus-value pour les magistrats qui disposent déjà d'outils de grande qualité d'analyse de la jurisprudence de la Cour de cassation et des cours d'appel » (*ibid.*).

L'expérimentation est doublement saisie comme l'occasion de décrédibiliser une *startup* généraliste et de valoriser le travail déjà entrepris par les acteurs spécifiques du secteur de la justice comme la Cour de cassation. La coalition particulariste retourne ainsi les choses en sa faveur, ce qui apparaît très clairement dans l'interview du premier président de la cour d'appel de Rennes dans laquelle il enfonce le clou en pointant le caractère parfois aberrant des résultats obtenus : « Le logiciel n'apporte aucune plus-value par rapport à d'autres moteurs de recherche qui permettent déjà une analyse très satisfaisante de la

⁴⁶ « [PréviCompute] est utilisé par 2 000 avocats exerçant à titre individuel, partout en France », [https://blog.\[PréviCompute\].com/avocat-individuel](https://blog.[PréviCompute].com/avocat-individuel), consulté le 6 mars 2022. Selon le CNB, la profession compte environ 70 000 avocats en France en 2020 : <https://www.cnb.avocat.fr/fr/les-chiffres-cles-de-la-profession-davocat>, consulté le 6 mars 2022.

⁴⁷ [https://blog.\[PréviCompute\].com/cabinet-davocat](https://blog.[PréviCompute].com/cabinet-davocat), consulté le 6 mars 2022.

⁴⁸ consulté le 6 mars 2022.

⁴⁹ Mis en ligne sur le site du ministère de la Justice le 9 octobre 2017, il a depuis été enlevé.

⁵⁰ Perrine Signoret, « Justice prédictive : les algorithmes ne sont pas prêts de remplacer les pros », mis en ligne le 2 novembre 2017, https://lexpansion.lexpress.fr/high-tech/justice-predictive-les-algorithmes-ne-sont-pas-pret-de-remplacer-les-pros_1956703.html, consulté le 6 mars 2022.

jurisprudence des cours d'appel et de la Cour de cassation. On se situe plus dans un projet d'approche statistique et plus quantitative que qualitative. Parfois, les résultats peuvent même être aberrants. Prenons l'exemple d'un calcul d'indemnités de licenciement sans cause réelle et sérieuse. Les montants sont calculés par le juge en mois de salaire. Or le salaire d'un cadre n'est pas le même qu'un ouvrier. Il suffit qu'une décision de justice concerne un cadre pour fausser complètement l'analyse, selon l'affaire ou le bassin d'emplois, et pour que le résultat proposé ne dise rien du cas à résoudre. Il serait plus pertinent de privilégier une approche selon le nombre de mois de salaire alloué, plutôt que savoir si le résultat probable sera de 4 000 ou 8 000 €. ⁵¹ » Et de conclure plus loin sur l'avenir radieux des avocats et magistrats : on est encore loin du fantasme de la « justice bouton ».

L'année suivante, dans le colloque de 2018 largement évoqué dans le premier chapitre (Ordre des avocats au Conseil d'État et à la Cour de cassation, 2018), le premier président (Louvel, 2018) et le procureur général (Marin, 2018) de la Cour de cassation prennent tous deux de la distance avec la notion de la justice prédictive. Ils réaffirment en outre l'importance que « l'autorité judiciaire organise elle-même l'alimentation des bases de données issues de son activité, et donc la diffusion des décisions de justice » (Louvel, 2018 : 16). Il est probable que l'expérimentation [PréviCompute] les a renforcés dans l'idée qu'il convient de mettre en place des dispositifs qui soient parfaitement ajustés aux valeurs et modes de fonctionnement de la justice et qu'ils sont en réalité les mieux placés pour le faire.

3.2.2. Les concepteurs d'algorithmes se repositionnent dans une filiation historique

Suite à cette expérimentation, certains opérateurs de dispositifs algorithmiques révisent leur position. C'est le cas de [PréviCompute]. Dans un premier temps, cette *startup* a développé un discours de l'hypermodernité en présentant son produit comme issu de l'intelligence artificielle et radicalement disruptif, mais sans guère s'appuyer sur des références scientifiques fondatrices contrairement aux expérimentations pluridisciplinaires internationales (Aletras *et al.* ; Katz *et al.* présentées dans le chapitre 1). Ces dernières, à la même période, s'inscrivent explicitement dans le champ scientifique des *judicial behavior studies* et se réclament d'une filiation avec le courant académique *Law & Informatics*. Guillaume Zambrano (2015) a resitué les outils d'analyse prédictive dans la continuité de ceux de Lee Loevinger (1963), père de la jurimétrie (*jurimetrics*). On peut penser que la controverse autour de l'expérimentation [PréviCompute] a amené les dirigeants de la *startup* à mieux appréhender l'importance de la légitimité historique pour le secteur de la justice qui cultive, entre autres, la référence au passé, aux grandes figures tutélaires et aux traditions (Bancaud, 1993).

En effet, en 2018, [PréviCompute] publie son *Guide de la justice prédictive* (2018 : 11) et dans une partie historique, reprend les analyses de Guillaume Zambrano ; cite Lee Lovinger, Fred Kort et Reed C. Lawlor, se plaçant dans cette filiation et la légitimité scientifique et historique qu'elle procure — au profit toutefois de « la justice prédictive » (et non de la

⁵¹ Interview du premier président de la cour d'appel de Rennes, « L'utilisation de l'outil [PréviCompute] déçoit la cour d'appel de Rennes », Dalloz actualités, [https://www.dalloz-actualite.fr/interview/l-utilisation-de-l-outil-\[PréviCompute\]-decoit-cour-d-appel-de-rennes](https://www.dalloz-actualite.fr/interview/l-utilisation-de-l-outil-[PréviCompute]-decoit-cour-d-appel-de-rennes), mis en ligne le 16 octobre 2017, consulté le 6 mars 2022.

jurimétrie). Ces auteurs et références sont assez peu connus en France où le mouvement *Law & Informatics* s'est beaucoup moins développé qu'aux États-Unis notamment. En revanche, l'analyse informatisée de la documentation juridique est une autre tradition intellectuelle et scientifique plus connue en France et dans laquelle puise également [PréviCompute]. Son guide mentionne l'émergence de bases de données devenues indispensables dans la pratique des juristes de la fin du xx^e siècle, à l'image de « celle ayant pris le nom de Désiré Dalloz, créateur du Répertoire méthodique et alphabétique de législation, de doctrine et de jurisprudence et fondateur de la maison d'édition juridique éponyme » (*ibid.* : 12), prenant soin de s'appuyer sur une figure très symbolique et consensuelle, et sur une maison d'édition qui ne propose pas alors de produit concurrent, en matière d'analyse algorithmique des décisions de justice. Des figures historiques nationales du mouvement Droit et Informatique comme Pierre Catala sont absentes, de même que leurs réalisations (Catala, 1998 ; Bories, 2013). Les exemples donnés sont pris à l'étranger : États-Unis (Lex Machina), Espagne (Jurimetria).

Il est probable, et cela reste à vérifier, que la montée en puissance de la rhétorique de la 'justice prédictive' a été un aiguillon pour certains éditeurs juridiques qui se sont aperçus que ce qu'ils avaient contribué à construire était en train de disparaître, invisibilisé par d'autres récits.

Le contraste du Guide [PréviCompute] avec la présentation de [LegalEditions], groupe d'édition juridique international installé dans le paysage français depuis longtemps via le rachat d'éditeurs historiques, est frappant. En effet, [LegalEditions] a publié, l'année suivante, une brochure intitulée *Open data & jurisprudence, Les données de jurisprudence au service de l'intelligence juridique* (2019). Elle est conçue comme une présentation didactique de la mise en œuvre de l'*open data* pour les décisions de juridictions françaises. Une partie est consacrée à l'histoire de l'informatisation juridique et Pierre Catala, professeur de droit civil, y figure en pionnier. « La question de l'informatisation de la jurisprudence se pose depuis les années 1960. Pour faire face au volume croissant de jurisprudence, les professionnels du droit ont besoin de nouveaux outils pour organiser et retrouver aisément les informations qui les intéressent. Dans ce contexte, le professeur Pierre Catala voit dans les progrès de l'informatique documentaire une solution prometteuse. » Ses liens historiques avec [LegalEditions] sont mis en exergue à travers une frise chronologique et des dates clés.

Le fait que Pierre Catala ait été cofondateur dès 1972 de LegalDonnées « pour développer un outil sur mesure de rédaction d'abstracts et résumés des décisions par des juristes » (*ibid.* : 15) confère une légitimité traditionnelle et historique à cet acteur, [LegalEditions], lequel consacre les pages suivantes de la brochure à présenter ses bases de données, son expertise en matière d'analyse et d'enrichissement des décisions de justice ainsi que ses offres commerciales, dont un module [LegalDonnées+] qui propose le même type de prestations que celles de PreviCompute. L'accent mis sur les ateliers régionaux de jurisprudence, sur la collaboration d'un réseau d'universitaires juristes entretenu depuis plusieurs décennies, contribue aussi à inscrire les dispositifs de traitement et d'analyse des décisions de justice dans une histoire de la documentation et de l'édition juridiques que les termes mêmes de 'justice prédictive' ou de « *startups* de la *legaltech* » ont contribué à passer

sous silence et à invisibiliser⁵² : « Le développement en 2016 de [LegalDonnées+], solution d'analyse de la jurisprudence en contentieux indemnitaire, a été facilité par les données extraites et catégorisées depuis 1989 par les ateliers régionaux de jurisprudence. » (*ibid.* : 16) est-il rappelé⁵³. Ces acteurs affinent donc leurs discours et surtout, concernant [PréviCompute], qui appartient à la coalition généraliste, lui ajoutent une dimension historique qui vise à les ancrer dans une perspective historique et à les rattacher à un passé pas seulement appréhendé en termes de rupture, de disruption⁵⁴.

3.4. Un faisceau de réalisations visant à utiliser et rendre comparables les décisions de justice

L'expérimentation [PréviCompute] vise à faire avancer de manière incrémentale et concrète l'utilisation des décisions de justice en anticipant sur le décret d'application de la loi Lemaire — lequel n'intervient qu'en 2020, c'est-à-dire trois ans plus tard. Ce test n'est pas la seule tentative visant à exploiter les décisions de justice. Plusieurs opérations concomitantes sont déployées, qui ont en commun de procéder de collaborations entre différentes institutions juridiques (la Cour de cassation, le ministère de la Justice à travers une de ses directions) et de faire avancer l'idée selon laquelle les décisions de justice devraient être rendues comparables. Etalab a mis en place des défis consistant à cibler un objectif précis, qu'il s'agit d'atteindre sur un temps court, en mettant à disposition de l'administration concernée une équipe de *data scientists*. Les deux premières éditions du « défi Openjustice » ne comportaient pas d'actions relatives à l'ouverture des données, ce qui a changé lors de la troisième campagne (2019) avec un défi sur la pseudonymisation des décisions de justice (1) et un autre sur l'utilisation de méthodes d'apprentissage automatique pour construire un référentiel d'indemnisations du dommage corporel (2).

3.3.1. Des « défis » et un accord de partenariat pour faire avancer la pseudonymisation

Le « défi » portant sur la pseudonymisation des données de jurisprudence est porté par la Cour de cassation. Il a été organisé d'abord en 2019 (Promotion EIG3-2019, durée 10 mois) puis réédité en 2020 sous le nom de Label (Promotion EIG4-2020, durée 10 mois) visant à « accélérer l'ouverture des décisions de justice⁵⁵ ».

La première édition a été consacrée à la mise en place d'un moteur de pseudonymisation reposant sur de l'apprentissage automatique. Comme il s'avère que « les

⁵² Une interview de Loïc Cadiet, déjà mentionnée au chapitre 3 comme indice de l'alliance avec les acteurs historiques de l'édition juridique, est également mise en avant.

⁵³ Pour mémoire, il faut préciser que l'héritage de Pierre Catala est aussi revendiqué par une autre approche qui, bien que non marchande, se positionne en opposition à la 'justice prédictive' (Négron, Bories, Da Silva, 2017), cf. *infra*.

⁵⁴ Pour être parfaitement symétrique, il faut préciser qu'[Anajur] ne s'inscrit dans aucune de ces filiations. Les fondateurs de cette *startup* ne revendiquent pas les héritages français ou internationaux mentionnés par leurs concurrents, à l'exception du dialogue entre mathématiques et droit alimenté notamment par Condorcet et sa célèbre analyse des probabilités pour réduire les erreurs dans les jurys d'assises ([Monta] et [Sinclair], 2017). Ils citent aussi l'étude de Danziger *et al.* sans toutefois pointer un lien particulier.

⁵⁵ Page Etalab sans date de mise en ligne, consultée le 17 avril 2021. <https://entrepreneur-interet-general.etalab.gouv.fr/defis/2020/label.html>

performances de ce moteur dépendent de la possibilité d’entraîner l’algorithme sur des décisions de justice dont les données à pseudonymiser ont été labellisées » (*ibid.*), la seconde édition a été dédiée à la conception d’un algorithme d’annotation des décisions de justice. C’est une interface adaptée au moteur de pseudonymisation⁵⁶, permettant d’automatiser le processus – pour une analyse de la boîte noire, voir (Girard-Chanudet, 2021). Entre les deux, un accord de partenariat a été conclu entre acteurs privés et publics pour construire une collaboration et capitaliser au plus vite leurs avancées respectives. Ainsi plus d’un an avant la promulgation du décret, le groupe [Geoffroy-Maisons] (qui possède notamment les [Éditions du couchant], les [Éditions doctrinales] et les [éditions Geoffroy]) et la direction interministérielle du numérique et du système d’information et de communication de l’État (Dinsic) annoncent la signature d’un accord de partenariat d’un an visant à permettre à Etalab, de travailler à partir des ressources de [Geoffroy-Maisons] pour commencer à mettre en œuvre l’anonymisation nécessaire à l’ouverture en *open data* des décisions de justice⁵⁷. Lorsque le décret paraît, le groupe [Geoffroy-Maisons] ne manque pas l’occasion de « réaffirmer son engagement et son action concrète en faveur de l’ouverture des données publiques⁵⁸ », rappelant l’accord passé dès 2019 avec la Cour de cassation via la convention avec Etalab. Ce défi marque bien la volonté des acteurs de la coalition généraliste non pas de s’opposer à l’ouverture des décisions de justice, mais de se positionner comme les points centraux par lesquelles passe ce processus et qui définissent la manière dont il sera mené.

3.3.2. Data-Just : un « défi » sur l’indemnisation du préjudice corporel

L’opération Data-Just, quant à elle, est assez comparable au test réalisé par [PréviCompute] au sens où elle est également positionnée comme une expérimentation. En revanche, elle incarne la tentative du ministère de la Justice de se doter de son propre outil face à l’augmentation d’une offre déjà proposée par des acteurs privés comme [PréviCompute], [Anajur] et d’autres, dans le domaine de l’indemnisation du préjudice corporel. [T.Vial], magistrat, en poste au Conseil de l’Europe où il a mené les travaux d’encadrement de l’intelligence artificielle pour la justice (cf. chapitre 1), estime que c’est dans le contexte de « la création d’une nouvelle génération de référentiels inférés par le traitement massif d’une quantité considérable de décisions de justice » que « la Direction des affaires civiles et du Sceau a porté, avec l’appui d’enthousiastes “entrepreneurs d’intérêt général”, l’expérimentation d’un système à destination des victimes, des magistrats, des avocats, des assureurs et des fonds d’indemnisation. » L’enjeu est bien celui de la création

⁵⁶ Il est précisé que, jusque-là, « la labellisation des données se fait via une interface d’annotation développée par un prestataire extérieur et inadaptée aux interactions avec le moteur de pseudonymisation, rendant la vérification humaine systématique » (*ibid.*).

⁵⁷ Paul-Antoine Chevalier, « La mission Etalab et le groupe [Geoffroy-Maisons] collaborent pour accélérer l’ouverture des décisions de justice » billet de blog Etalab mis en ligne le 26 avril 2019, consulté le 16 avril 2021. [https://www.etalab.gouv.fr/la-mission-etalab-et-le-groupe-\[geoffroy-maisons\]-collaborent-accelerer-louverture-des-decisions-de-justice](https://www.etalab.gouv.fr/la-mission-etalab-et-le-groupe-[geoffroy-maisons]-collaborent-accelerer-louverture-des-decisions-de-justice)

⁵⁸ Communiqué de presse du groupe [Geoffroy-Maisons], 3 juillet 2020, [https://www.\[geoffroy-maisons\].eu/2020/07/03/communiquede-presse-parution-du-decret-open-data-de-la-jurisprudence-\[geoffroy-maisons\]-reaffirme-son-role-en-faveur-de-louverture-des-decisions-de-justice/](https://www.[geoffroy-maisons].eu/2020/07/03/communiquede-presse-parution-du-decret-open-data-de-la-jurisprudence-[geoffroy-maisons]-reaffirme-son-role-en-faveur-de-louverture-des-decisions-de-justice/), consulté le 17 avril 2021.

d'un outil public propre au ministère, en réaction à l'éclosion d'une offre issue d'entreprises privées. On peut penser que la controverse autour de l'expérimentation [PréviCompute] (qui couvrait la réparation du dommage corporel) a contribué à mettre en lumière les enjeux de ces dispositifs et l'intérêt pour le ministère à garder une certaine maîtrise. Cela reste toutefois à documenter plus précisément.

Ce projet, porté donc par la Direction des affaires civiles et du Sceau du ministère de la Justice, en lien avec Etalab, vise à construire un référentiel d'indemnisation des préjudices corporels (Promotion EIG3-2019, durée 10 mois). Pour atteindre cet objectif, trois étapes sont définies. D'abord, constituer un fond comprenant l'intégralité des décisions de justice d'appel prononcées en la matière, de 2017 à 2019⁵⁹. Ensuite, « recenser, grâce à un algorithme, les montants alloués aux victimes à titre d'indemnisation des différents chefs de préjudice corporels listés dans la nomenclature dite "Dintilhac"⁶⁰ ». Enfin, toujours par le recours à des techniques d'intelligence artificielle supervisée, établir un « référentiel indicatif d'indemnisation » destiné à être diffusé au public et aux professionnels, autrement dit un barème construit à partir de la modélisation des précédentes décisions.

Un décret *ad hoc*⁶¹ a décrit ce processus et autorisé le développement de cet algorithme « chargé d'extraire de manière automatique et d'exploiter les données contenues dans les décisions de justice portant sur l'indemnisation des préjudices corporels⁶² ». Publié en pleine pandémie, au plus fort du premier confinement ; conçu sans consultation des organisations professionnelles du secteur (ni des instances de représentation des avocats ni des syndicats de magistrats), ce texte réglementaire a suscité un tollé parmi les professionnels de justice⁶³.

⁵⁹ Ces décisions sont extraites des bases de données Jurica pour la Cour de cassation et Ariane pour le Conseil d'État.

⁶⁰ Page Etalab sans date de mise en ligne, <https://entrepreneur-interet-general.etalab.gouv.fr/defis/2019/datajust.html>, consultée le 17 avril 2021. La nomenclature Dintilhac est une classification qui répertorie, de manière non limitative, mais précise, les types de préjudices qui peuvent donner lieu à indemnisation, en distinguant selon leur nature (patrimoniaux ou extrapatrimoniaux) et leur durée. Figurent par exemple les dépenses de santé, les souffrances endurées, le déficit fonctionnel temporaire ou permanent, le préjudice d'agrément, etc.

⁶¹ Il s'agit du décret n° 2020-356 du 27 mars 2020, pris après avis de la CNIL.

⁶² Page ministère de la Justice mise à jour le 9 septembre 2020, consultée le 17 avril 2021. <https://www.justice.fr/donnees-personnelles/datajust>

⁶³ La raison de ce tollé est aussi à chercher dans le fait que ce décret dispense le ministère de la Justice de se mettre en conformité avec les exigences du RGPD, considérant que, « *compte tenu des efforts disproportionnés que représenterait la fourniture des informations mentionnées aux paragraphes 1 à 4 de l'article 14 du règlement (UE) 2016/679 du 27 avril 2016 (...), le droit d'information prévu à ce même article ne s'applique pas au présent traitement* ». De plus, il exclut toute mise en œuvre du droit d'opposition pour « *garantir l'objectif d'intérêt public général d'accessibilité du droit* » (*ibid.*). La ministre a répondu par voie de presse qu'il s'agissait d'une « erreur temporelle », sans aborder la question de fond, relative à l'exception que s'autorise le ministère. Marie Beau, décret « DataJust » : stupeur et mécontentements », *Le monde du droit*, mis en ligne le 6 avril 2020, <https://www.lemondedudroit.fr/decryptages/69428-decret-datajust-stupeur-mecontentements.html>, consulté le 17 avril 2021. Un collectif d'avocats et de victimes a formé un recours devant le Conseil d'État, portant sur

L'expérimentation a porté sur la conception d'une plateforme et d'un outil d'annotation destinés à extraire les données pertinentes (41 critères relatifs aux montants alloués, caractéristiques de la victime et du dommage) des arrêts concernés. Il s'agit d'algorithmes qui effectuent une préannotation automatique des arrêts, ensuite validée, rectifiée ou complétée par des annotateurs humains. Un outil de data visualisation a également été créé, qui cartographie les différentes informations statistiques, faisant clairement apparaître les disparités territoriales, qu'elles concernent le poids de ce type de contentieux par ressort judiciaire ou administratif, ou bien la répartition des montants d'indemnisation alloués en matière de réparation du dommage corporel. Les perspectives sont ensuite de « modéliser l'issue du jugement pour identifier et corriger d'éventuels biais injustifiés dans l'attribution des montants (genre, origine géographique, etc.), et d'améliorer la préannotation en entraînant un algorithme sur les données annotées⁶⁴ ». Si les termes « prédictif.ve » et « prédiction » sont absents du décret et des communications institutionnelles d'Etalab ou du ministère de la Justice, il est en revanche explicite que l'objectif est de compléter cet outil d'analyse et de modélisation par des fonctionnalités complémentaires de simulation pour des cas en cours : « À terme, si cette phase de développement était satisfaisante, le ministère souhaiterait mettre en place un outil pérenne de mise à disposition des résultats au public (un référentiel indicatif d'indemnisation devant notamment permettre une meilleure information des victimes de dommages corporels et fournir *aux juges une aide à la décision*).⁶⁵ »

Cette tentative a fait l'objet de nombreuses critiques vu l'ampleur de la tâche au regard de la faiblesse des moyens consacrés par le ministère de la Justice. Data-Just a depuis été abandonné, confirmant les réserves qui avaient d'emblée été formulées⁶⁶. L'expérimentation a été menée quasiment à son terme (deux ans), mais au moment du bilan, le ministère de la Justice y a finalement mis un terme en raison de « la complexité du chantier⁶⁷ », liée au caractère non formaté des décisions de justice — très disparates dans leur

l'argument de la protection des données personnelles. Il a été rejeté, cf. L'Argus de l'assurance, « Data-Just, un outil contesté dans l'évaluation du préjudice corporel », mis en ligne le 28 avril 2021, consulté le 6 mars 2022. <https://www.argusdelassurance.com/juriscope/analyses-d-experts/datajust-un-outil-conteste-dans-l-evaluation-du-prejudice-corporel.181704>

⁶⁴ Page Etalab sans date de mise en ligne, consultée le 17 avril 2021. <https://entrepreneur-interet-general.etalab.gouv.fr/defis/2019/datajust.html>

⁶⁵ Page ministère de la Justice mise à jour le 9 septembre 2020, consultée le 17 avril 2021. <https://www.justice.fr/donnees-personnelles/datajust>

⁶⁶ Alice Vitard, « C'est la fin pour Data-Just, l'algorithme d'évaluation des préjudices corporels », mis en ligne le 14 janvier 2022, consulté le 6 mars 2022. <https://www.usine-digitale.fr/article/c-est-la-fin-pour-datajust-l-algorithme-d-evaluation-des-prejudices-corporels.N1773872>

⁶⁷ Décision d'abandon annoncée dans la presse le 13 janvier 2022 et confirmée ensuite par le ministère. Olivia Dufour et Aurélie Corviaux, DataJust : « Plutôt que de faire de la justice prédictive, il faut engager une démarche d'indexation et de tri des décisions », mis en ligne le 27 janvier 2022, consulté le 27 février 2022. <https://www.actu-juridique.fr/civil/responsabilite-civile/datajust-plutot-que-de-faire-de-la-justice-predictive-il-faut-engager-une-demarche-dindexation-et-de-tri-des-decisions/>

contenu comme dans leur forme — et à la multiplicité des règles applicables en matière de préjudice corporel.

En dépit de ratages (Alauzen, 2019), ces différentes expérimentations témoignent de la progression de l'idée qu'il convient, notamment en matière de dommages corporels, d'avancer dans le sens d'une comparabilité et d'une harmonisation accrues des décisions de justice. À travers la controverse sur la 'justice prédictive', c'est notamment la question du jugement et des manières plurielles de l'appréhender qui sont posées.

Conclusion :

Ce chapitre s'est concentré sur une expérimentation d'ampleur limitée : un dispositif algorithmique testé par une poignée de magistrats de deux cours d'appel pendant quelques mois. L'expérience n'en est pas moins intéressante si on la considère à travers sa genèse, son déroulement et ses effets.

Sur le plan de la genèse, elle donne à voir le nouveau répertoire de l'*open data* des décisions de justice en action. Le format du concours technologique, lui-même propice à l'irruption d'une certaine improvisation, est renforcé par la mise en place d'un jury dont la composition manifeste la réunion d'acteurs disparates, publics et privés d'une part, généralistes de l'action publique et acteurs sectoriels de la justice d'autre part. Le fait que ce qui est attendu des innovateurs et des innovations en compétition soit à peine formalisé confère un rôle majeur à ce jury de concours.

Le produit du concours est un dispositif algorithmique, officiellement expérimenté, sous l'égide du ministère de la Justice, et ce, bien que celui-ci y ait pris part sans avoir défini au préalable des besoins, une commande ou un cahier des charges ni supervisé ou suivi la réalisation de celui-ci. La première conclusion que l'on peut en tirer, en ce qui concerne la manière de concevoir l'action publique en matière de justice, c'est que ce qui prévaut ici n'est pas tant un contenu d'innovation qu'une manière d'innover (un certain type de processus dans lequel participent des acteurs privés, dans une logique d'État plateforme et une diversité de savoirs et compétences : droit, informatique, science des données) et un objectif (l'innovation numérique au service de l'*open data* des décisions de justice). La seconde, c'est qu'autour de l'*open data*, le ministère de la Justice ne se place pas tellement en chef de file : il accompagne les initiatives d'acteurs venus de l'extérieur et dans le cadre de programmes génériques de transformation de l'administration publique comme Etalab, directement rattaché aux services du Premier ministre. Il collabore avec des acteurs privés, comme les *startups* auxquelles il laisse le soin de proposer un service d'aide à la décision pour les magistrats, sans avoir clarifié ses attentes. Dans tous les cas, autour de l'expérimentation officielle de [PréviCompute] par des magistrats, se dessine une innovation limitée dont l'impulsion née au niveau central est concrétisée par des acteurs aux statuts diversifiés (public/privé), aux appartenances institutionnelles plurielles (Etalab / DSJ).

Sur le plan de sa mise en œuvre, l'expérimentation est elle-même une tentative de faire converger suffisamment d'acteurs différents qui ne sont pas porteurs des mêmes

croyances, cadres cognitifs et intérêts, pour faire tenir l'innovation et démontrer la viabilité de dispositifs algorithmiques d'aide à la décision pour les magistrats. L'innovation est ainsi construite de manière à dépasser les coalitions de cause existantes puisqu'elle associe par exemple des représentants de la Cour de cassation, de la DSJ du ministère, etc. d'une part et Open Law, Etalab, etc. d'autre part. Mais elle est réappropriée par des acteurs judiciaires qui l'utilisent de manière stratégique pour reconfigurer le rapport de force entre coalitions de cause et défendre une gestion sectorielle de l'*open data* des décisions de justice (et des outils algorithmiques qui s'appuient dessus). En outre, les opérateurs de dispositifs algorithmiques en exacerbant leurs identités et différences, alimentent la controverse au point que l'expérimentation apparaît aussi comme une mise à l'épreuve de la 'justice prédictive' à la fois en tant que dispositif destiné à équiper l'activité concrète de juger et en tant que construction sémantique.

De ce fait, si [PréviCompute] est une expérience d'ampleur limitée dont la trajectoire tendrait au premier abord à montrer un délitement à travers ce que les acteurs ont désigné comme un échec de l'expérimentation, le chapitre a permis de dégager les effets plus profonds qu'elle a produit ou contribué à produire. Elle a joué un rôle déterminant dans la diffusion de la notion de 'justice prédictive', mais aussi dans sa contestation, sa remise en cause et l'émergence de vocables alternatifs ainsi que dans l'ouverture d'un questionnement sur ce que c'est que rendre justice, en premier lieu pour les professionnels comme les magistrats et magistrates.

Point(s) de suspension

Du jugement « produit fini » au jugement comme action équipée

Ô temps suspend ton vol...

Avant de clore ce voyage et d'en arriver à la conclusion, il importe de prendre le temps d'une réflexion conçue non pas tant comme un chapitre de plus, un pas de côté ou un *post-scriptum*, mais davantage comme un objet en suspension, au double sens d'objet maintenu en (relative) apesanteur et d'objet en devenir. Les pages qui suivent sont d'une nature légèrement différente de celles des précédents chapitres : un peu moins empiriques et plus prospectives peut-être, elles dessinent des perspectives qui demandent à être prolongées, consolidées, ce à quoi invitent les... points de suspension !

Je me propose donc d'ouvrir un espace de réflexion sur la manière dont le jugement est thématiqué dans les débats sur la 'justice prédictive'. Cette question, à elle seule, aurait pu être le cœur de ce mémoire original et nourrira mes futurs travaux. L'idée est née du constat qu'il existe des parallèles entre les dispositifs algorithmiques de traitement et d'analyse des décisions de justice ici étudiés et les rapports d'expertise judiciaire, explorés dans mes travaux antérieurs, vieux de vingt ans. Alors que je ne les avais pas forcément anticipés, un effet d'écho et d'étranges résonances se sont fait entendre. De là ont resurgi des notions déjà mobilisées pour mettre en lumière le rôle complexe des rapports d'expertise dans la production des jugements et que je mets ici à l'épreuve.

La controverse sur la 'justice prédictive' est traversée par des constantes sur la manière dont les concepteurs d'algorithmes et les médias présentent les jugements et en débattent. Hormis quelques nuances qui seront précisées, les jugements sont principalement abordés à partir d'une conception décisionniste, qui place le juge en majesté et se focalise sur la décision comme produit fini. Le premier mouvement de cette réflexion sera consacré à identifier les manifestations empiriques de cette conception (1). Il conviendra ensuite de saisir ce que cette représentation socialement partagée révèle et d'établir le biais de perspective qu'elle introduit par rapport au fonctionnement réel du processus décisionnel et des juridictions (2). Dans un troisième mouvement, je dessinerai les contours d'une approche qui, par sa conceptualisation, son questionnement et son protocole de recherche, met à distance cette conception décisionniste du jugement (3).

I / La relation entre algorithmes et jugements : éléments empiriques

Cette partie s'intéresse à la relation entre jugements et algorithmes. L'objectif est d'approcher, à partir de données empiriques, la façon dont sont thématiquées les décisions de justice dans les débats sur les dispositifs algorithmiques qui en assurent le traitement et

l'analyse. Quelle conception des arrêts et jugements est proposée, implicitement ou explicitement? Je m'appuie ici sur les données issues des discours des concepteurs d'algorithmes, des articles de presse et des *tweets*. Je montrerai que parmi les différents acteurs qui débattent de la 'justice prédictive' prévaut une conception implicite de la décision judiciaire que je qualifie de décisionniste.

1.1. Une focalisation sur l'adjudication exprimée sous forme quantifiée par le juge...

Les activités de justice peuvent être entendues comme « un ensemble de méthodes de traitement des différends et des conflits, parmi lesquels figure le jugement *stricto sensu* » (Assier-Andrieu, 1996 : 143). Par le terme de jugement, on désigne en général ce qui en droit est défini comme un acte juridictionnel, accompli par un organe juridictionnel (i.e. une juridiction). « On attend de l'organe qui tranchera le différend qu'il se prononce en droit, c'est-à-dire en recherchant si la situation, objet du litige, respecte ou transgresse les règles juridiques qui lui sont applicables, et en tirant les conséquences juridiques de leur violation éventuelle. [...] Ces organes chargés de trancher les différends en droit présentent un caractère juridictionnel : ils sont organisés de manière à être indépendants des parties, ce qui contribue à assurer leur impartialité ; ils sont astreints à respecter des règles de forme et de procédure. Des recours sont d'ailleurs organisés, en règle générale, contre leurs sentences, si elles ne respectent pas les textes ou les conventions qui s'imposent à eux. » (Guinchard, Varinard et Debard, 2021 : 117).

Lorsqu'ils jugent en droit, via des décisions exécutoires, les cours et tribunaux exercent une activité juridictionnelle que l'on peut aussi qualifier – par anglicisme – d'adjudication. Ces mêmes institutions peuvent aussi résoudre les différends en amenant les parties à des solutions mutuellement acceptées. Cela correspond cette fois à des procédures comme la conciliation, la médiation, et plus globalement ce que l'on appelle aux Etats-Unis les *Alternative Dispute Resolution* (ADR) et en France les modes alternatifs de règlement des conflits (MARC) / différends (MARD) / litiges (MARL). Il existe un certain flou conceptuel quant à ce que recouvrent ces techniques dites alternatives qui ont connu un nouvel essor à partir des années 1970-80¹. Elles peuvent en effet être pratiquées dans le cadre d'une procédure judiciaire, sous l'égide du juge, ou de manière privée dans un cadre extra-judiciaire. Les modèles de médiation et les raisons de les promouvoir sont variées (De Munck, 1995). Aujourd'hui, le règlement amiable des litiges est incité voire contraint par les nouveaux textes législatifs en vigueur : le juge peut imposer aux parties le recours à une mesure de conciliation,

¹ Ce sont des formes de justice multiséculaires qui ont été dominantes dans certaines périodes historiques où la norme était l'autogestion communautaire des conflits (Garnot 1996 ; Sbriccoli 2007). Les conciliations étaient le quotidien des tribunaux de la famille (Commaille, 1982) et de la justice de paix, tous deux instaurés après la Révolution française ; ainsi que des Conseils des Prud'hommes créés en 1806 (Cottureau, 1987). Ces formes non adjudicationnelles sont souvent présentées comme extra-judiciaires au sens de pratiquées à l'extérieur de l'institution judiciaire ; et/ou informelles au sens de non réglées par des procédures juridiques ; et/ou non juridictionnelles au sens de ne faisant pas l'objet d'une décision qui s'impose aux parties (Bonafé-Schmidt 1993). La réalité est complexifiée par le fait que ces techniques sont absorbées par le système judiciaire étatique qui peut les rendre obligatoires, soit comme une première étape conditionnelle (cf. la conciliation aux Prud'hommes), soit comme l'unique moyen de régler certains litiges (plus de moyen de saisir le juge).

médiation ou procédure participative pour des conflits de voisinage et des petits litiges² (voir Guinchard, Varinard et Debard, 2021 : 73-78). La médiation interentreprises est également promue (Deffains et Doriat-Duban, 2019) et bien sûr de longue date, la médiation pénale via les solutions dites de troisième voie. L'évolution est européenne et pas seulement française (Vigour, 2018).

La plupart des discours et dispositifs de 'justice prédictive' sont marqués par une approche juridictionnelle du jugement au sens où ils assimilent l'intervention du juge au fait de régler un litige en tranchant par le droit entre des prétentions contradictoires qui s'expriment sous forme quantifiée. Ce constat peut être dressé à partir de l'analyse des sites Internet de plusieurs sociétés commercialisant des dispositifs algorithmiques et de celle des articles de presse.

1.1.1. *Les discours des concepteurs d'algorithmes*

Sur le site de PreviCompute, un article destiné aux utilisateurs potentiels donne le mode d'emploi de l'algorithme (mots-clefs et modes d'interrogation de la base de données). Après avoir présenté les différents filtres disponibles (date, juridiction et chambre, textes juridiques concernés), celui relatif au dispositif est mis en avant : « PreviCompute vous propose de *filtrer les décisions en fonction de la solution retenue par le juge*, sur une demande précise. [...] Vous pouvez non seulement *filtrer les décisions en fonction des demandes qui y sont formulées, mais également selon si elles ont été acceptées* (le cas échéant selon le montant d'indemnités associé) *ou si elles ont été rejetées par le juge*.³ » L'explication est illustrée de captures d'écran. L'exemple donné, concernant le filtre par dispositif, a trait à des chefs de demande relatifs au préjudice distinct et au préjudice moral, accompagné du commentaire suivant : « Ainsi dans l'illustration ci-dessus sont sélectionnées les décisions dans lesquelles les demandes suivantes ont été formulées et acceptées par le juge. ». Ce cas illustre que le dispositif algorithmique est focalisé sur la décision d'adjudication (Dans quel sens l'affaire a-t-elle été tranchée ? Dans quelles conditions a-t-elle été favorable à telle partie ? Etc.), ce qui permet ensuite des extrapolations pour fonder la décision juridique (d'un avocat, d'un assureur, etc.) et d'aller ou pas au contentieux.

De même, [Anajur] cible les professionnels du droit et du chiffre auxquels sont proposées des analyses de risques relatives aux décisions des juridictions : « [Anajur] a vu le jour en septembre 2017 avec pour objectif de fournir aux professionnels du droit (avocats, directions juridiques d'entreprises ou des collectivités territoriales, notaires, huissiers de justice...), mais aussi aux assureurs, aux fonctions supports de l'entreprise (DRH...), et enfin aux professionnels du chiffre (experts-comptables, commissaires aux comptes...), des solutions de quantification du risque juridique et judiciaire.⁴ » Cette société se positionne comme « l'alliance de l'intelligence juridique et de l'intelligence artificielle au service des professionnels du droit ». Comme sa concurrente [PréviCompute], *elle s'adresse ainsi*

² Article 4 de la loi du 18 novembre 2016 de modernisation de la justice au XXI^e siècle.

³ [https://formation.\[previcompute\].com/fr/articles/1029030-recherchez-une-decision-de-justice](https://formation.[previcompute].com/fr/articles/1029030-recherchez-une-decision-de-justice), consulté le 26 juillet 2021.

⁴ [https://www.\[anajur\].com/a-propos/](https://www.[anajur].com/a-propos/), consulté le 22 juillet 2021.

d'emblée à des utilisateurs qui sont dépendants des décisions juridictionnelles et de leur contenu adjudicatoire. En page d'accueil du site web, sa proposition est formulée comme suit : « Analysez votre dossier en quelques minutes », ce qui est décliné en trois items : « Simulez les issues possibles d'une procédure contentieuse : évaluez la probabilité d'entrée en voie de condamnation et calculez les montants possibles de la condamnation » ; « Adoptez la meilleure stratégie pour votre client : mesurez l'impact des critères de votre dossier sur les décisions de justice et optez le cas échéant pour un MARD [mode alternatif de règlement des différends] » ; enfin, « Trouvez la jurisprudence la plus proche de votre dossier⁵ ».

Ce ciblage de professionnels qui sont directement intéressés à l'issue de la décision de justice est opéré à partir d'une focalisation sur l'expression quantifiée des jugements : prestation compensatoire, pension alimentaire, indemnité pour licenciement injustifié sont les domaines de prédilection des dispositifs algorithmiques de traitement et d'analyse des décisions de justice. Tous les trois renvoient à des contentieux de masse et qui font intervenir, pour reprendre la typologie de Marc Galanter judiciaire (2013 [1974]), soit de grands nombres de justiciables « occasionnels » (les divorçants, les salariés licenciés), soit des « joueurs répétés » (entreprises, syndicats, professionnels du chiffre), soit encore des professionnels du droit (avocats, etc.) auxquels il est proposé de rationaliser aussi le rôle d'intermédiaire en systématisant l'accès à l'information juridique pertinente et en offrant la possibilité de facturer auprès de leurs clients les rapports obtenus à l'issue de la recherche effectuée sur les plateformes algorithmiques.

Dans son livre blanc, [LegalEditions] présente le dispositif algorithmique qu'il propose, conçu comme une brique supplémentaire dans la gamme des outils de recherche et d'analyse documentaire juridique : « Lancé en 2016 et intégré dans la solution de recherches en ligne [100%Legal], [LegalDonnées+] est un outil d'aide à la décision qui permet d'évaluer rapidement les montants susceptibles d'être accordés aux clients en s'appuyant sur des arrêts comparables. Ce service d'analyse quantitative de la jurisprudence repose sur le savoir-faire des travaux de l'équipe d'analystes [LegalDonnées+] combinés au traitement semi-automatique des données chiffrées des décisions de cours d'appel en matière d'Aliments, Bail, Divorce, Dommage corporel, Licenciement, Troubles de voisinage. Une exploitation rapide et facile des résultats est proposée grâce aux 2 modes de recherches depuis une sélection de critères objectifs (entre 10 à 15) de la décision. Les résultats de la recherche mettent en perspective les datas présentées sous forme de tableaux ou graphiques interactifs (datavisualisation). » ([LegalEditions], 2018 : 11).

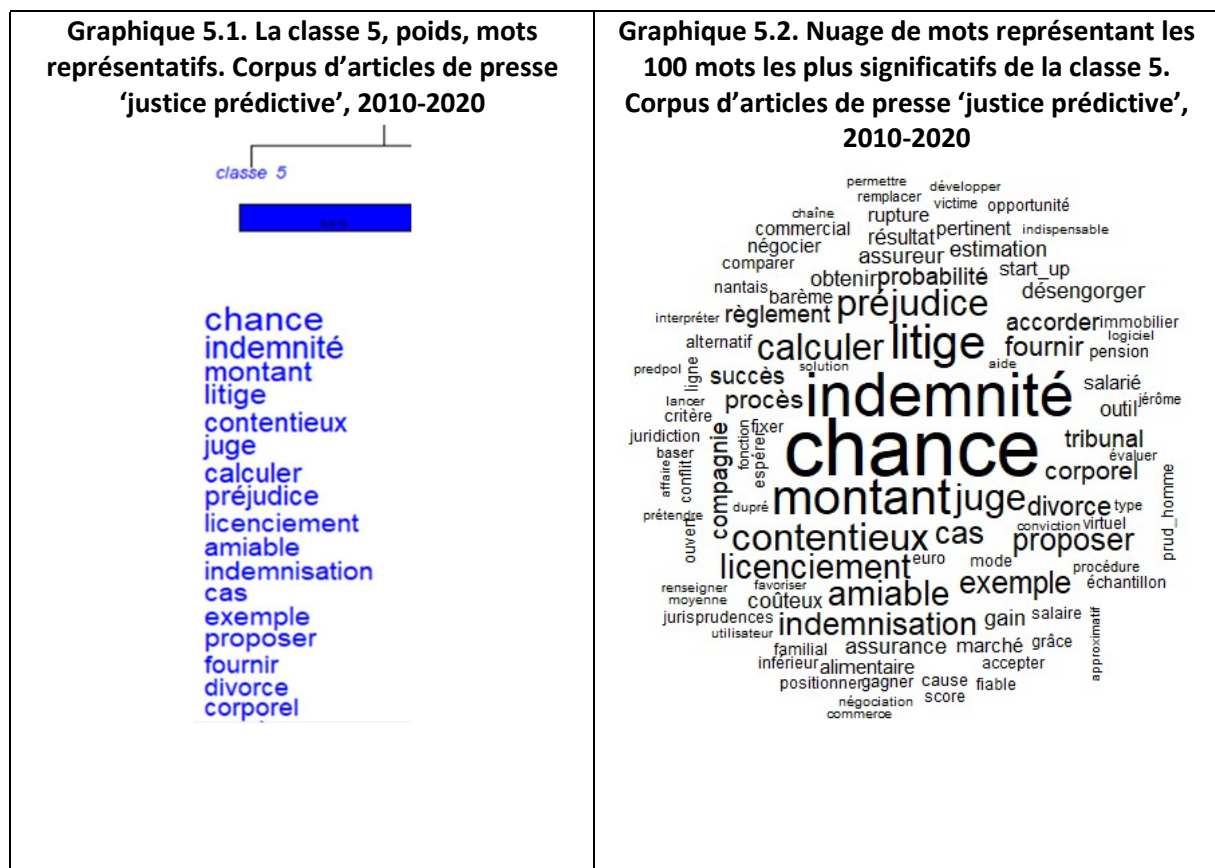
Ces dispositifs utilisent donc des décisions dont l'expression est quantifiée par des montants (d'indemnisation, etc.) pour proposer des mesures, facilement accessibles et compréhensibles via des outils de datavisualisation. La diversité des métriques apparaît dans la présentation de [Legdata] de même que la focalisation sur l'issue de la décision : « Liste des décisions, durées moyennes des procédures, taux de recours, montant du préjudice. Obtenez des statistiques fiables et inédites sur le contentieux d'une entreprise ou d'un domaine juridique. Bénéficiez d'un outil statistique d'aide à la prise de décisions et à la stratégie

⁵ [https://www.\[anajur\].com/](https://www.[anajur].com/), consulté le 26 juillet 2021.

contentieuse. [...] Éditez en quelques secondes une vue à 360° du contentieux qui vous intéresse et analysez : *la durée des procédures, les indemnités obtenues par poste de préjudice, les solutions de confirmation (rejet) ou d'infirmité (cassation) des solutions qui vous intéressent, et les taux de recours.*⁶ » Le point de focalisation est donc bien l'analyse par données agrégées des probabilités associées à une situation donnée et en particulier l'anticipation des décisions du juge, surtout lorsqu'elles sont exprimées par des montants.

1.1.2. Les discours dans le corpus d'articles de presse

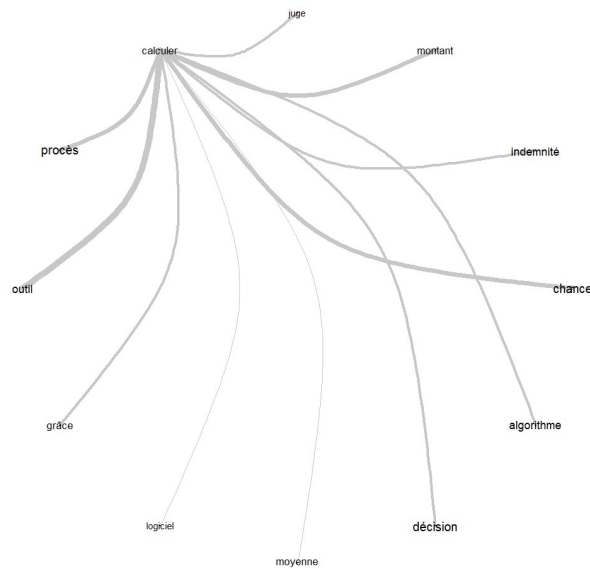
La sémantique médiatique investit également le jugement comme décision du juge, exprimée quantitativement et dont la probabilité est elle-même quantifiable. Dans le corpus d'articles de presse, une classe de discours (classe 5) se déploie ainsi à la confluence entre le champ sémantique du conflit judiciaire et de son résultat quantifié (*litige, contentieux, juge, licenciement, amiable, cas, exemple, divorce, corporel, préjudice, indemnité, montant, indemnisation*) et celui du calcul de probabilité (*chance, calculer, probabilité, estimation*).



L'analyse du mot *calculer* et de ses liens avec d'autres mots (cf. graphique ci-dessous) illustre le caractère quantifié des *décisions* de justice que les dispositifs algorithmiques (*outil, logiciel, algorithme*) permettent (*grâce*) d'approcher ou de simuler. L'objectif porte sur *calculer* les *chances* de gagner un *procès* et le *montant* probable de l'*indemnisation* possiblement déterminée par la *décision* du *juge*. On voit bien ici que le registre de discours

⁶ [https://www.\[legdata\].fr/login](https://www.[legdata].fr/login), consulté le 26 juillet 2021.

se concentre sur l'anticipation de la décision du juge, dans une optique d'évitement du contentieux.



Graphique 5.3. Graphe du mot calculer au sein de la classe 5. Corpus d'articles de presse 'justice prédictive', 2010-2020

Les segments d'articles de presse les plus caractéristiques permettent de vérifier le sens général de cette interprétation, en ajoutant que ce discours est en partie le résultat des prises de parole des concepteurs et de leurs *verbatim* (« notre outil permet de traiter » ; « nous fournissons les probabilités de succès d'une procédure »).

Encadré 5.1. Les 5 segments de texte les plus caractéristiques de la classe 5. Corpus d'articles de presse 'justice prédictive', 2010-2020

**** *annee_2017 *AnneeMois_2017-02 *Date_2017-02-24 *Journal_Liberaton_PAPIER *sourcesconfondues_Liberation *typepresse_PQN *support_PAPIER
*couleurpol_GAUCHE *antepost9oct_ANTE *phases_3 *20 621
score : 1135.75

plusieurs **start up** se **lancent** sur le **marché** du **big data** **proposant** aux **professionnels** du droit d **évaluer** leurs **chances** de **succès** dans un **litige** de dégoter les **arguments** les plus **pertinents** ou de prévoir le **montant** d **éventuelles indemnités**

**** *annee_2019 *AnneeMois_2019-01 *Date_2019-01-30 *Journal_Le_Point_WEB *sourcesconfondues_Le_Point *typepresse_HEBDO *support_WEB
*couleurpol_DROITE *antepost9oct_POST *phases_4 *74 1854
score : 1075.04

reste que la justice dite **prédictive** ces **algorithmes** qui **calculent** les **chances** de **succès** d une **procédure** ou les **montants** d **indemnités** en **fonction** des **décisions** passées voire de la composition du **tribunal** et de sa situation géographique **lance** un vrai défi aux **juges**

**** *annee_2016 *AnneeMois_2016-12 *Date_2016-12-27 *Journal_Le_Point_WEB *sourcesconfondues_Le_Point *typepresse_HEBDO *support_WEB
*couleurpol_DROITE *antepost9oct_ANTE *phases_2 *12 460
score : 999.24

grâce aux **algorithmes** couplés à des **outils mathématiques** et au **big data** c'est_à_dire l ensemble des données **issues** des **décisions** de justice on peut **désormais** **calculer** ses **chances** de **gagner** son **procès** et **connaître** le **montant approximatif** des **indemnités** à en **attendre**

**** *annee_2016 *AnneeMois_2016-12 *Date_2016-12-27 *Journal_Le_Point_WEB *sourcesconfondues_Le_Point *typepresse_HEBDO *support_WEB
*couleurpol_DROITE *antepost9oct_ANTE *phases_2 *11 455
score : 953.29

nous travaillons aussi sur le **préjudice corporel** et la **rupture** brutale des **relations commerciales** notre **outil permet** de **traiter** tous les domaines dès lors que l on peut s appuyer sur des **critères** objectifs et quantifiables comme le **type** de **licenciement** le **montant** d une **indemnité** le **tribunal saisi etc**

**** *annee_2016 *AnneeMois_2016-12 *Date_2016-12-27 *Journal_Le_Point_WEB *sourcesconfondues_Le_Point *typepresse_HEBDO *support_WEB
*couleurpol_DROITE *antepost9oct_ANTE *phases_2 *11 449
score : 944.94

nous **fournissons** les **probabilités** de **succès** d une **procédure** et plus précisément les **chances** d **obtenir** tel ou tel **montant** d **indemnités** devant telle ou telle **juridiction**

Comme nous allons le voir, l'accent sur la dimension adjudicatoire du jugement et sa meilleure prévisibilité va de pair avec une perspective de déjudiciarisation de certains conflits, pour développer le règlement amiable (hors juge) et le transférer à des acteurs privés.

1.2. ... pour mieux déjudiciariser et développer le règlement amiable

L'essor des modes alternatifs de règlement des conflits, des processus de déjudiciarisation (Jeammaud, 2004) et du traitement à l'amiable dans divers domaines font l'objet de politiques soutenues en France et dans d'autres pays. Les réformes de la procédure civile entreprises depuis plusieurs années s'inscrivent dans ce mouvement, marqué par une profusion de textes juridiques depuis la loi pour la modernisation de la justice⁷ (2016) jusqu'à celle pour la confiance dans l'institution judiciaire⁸ (2021), en passant par la loi dite Belloubet⁹ (2019). Simplification des procédures, déjudiciarisation des petits litiges, fortes incitations à l'amiable en amont du contentieux, autant de mesures visant à désengorger les juridictions et recentrer les magistrats sur l'activité qui consiste à trancher des litiges en droit (Guinchard, Varinard et Debard, 2021). Dans ce contexte, les plateformes de règlement en ligne des différends se multiplient (Amrani-Mekki, 2021) et les concepteurs de dispositifs algorithmiques de traitement et d'analyse des décisions de justice, s'ils ne proposent pas de la médiation ou de l'arbitrage en ligne, assument l'objectif de délestage des juridictions.

D'emblée, la notion de « justice quantitative » est associée à l'enjeu de développer la gestion amiable des conflits, comme les fondateurs d'[Anajur] l'expriment dans cette longue

⁷ Loi du 18 novembre 2016 dite « loi J21 ».

⁸ Loi n° 2021-1729 du 22 décembre 2021 pour la confiance dans l'institution judiciaire.

⁹ Loi n° 2019-222 du 23 mars 2019 de programmation 2018-2022 et de réforme pour la justice (loi de programmation pour la justice)

citation : « Cette technologie permet de valoriser mathématiquement le patrimoine jurisprudentiel, en permettant que les millions de décisions rendues par toutes les juridictions chaque année servent à orienter vers des solutions amiables. [...] L'intérêt de pouvoir refléter par anticipation les différents résultats possibles, et la probabilité de leur survenance, est de pouvoir guider les parties vers une négociation en évitant à la fois l'écueil d'un contentieux long, coûteux et incertain et la prescription d'une solution rigide. Les avocats peuvent y recourir notamment en vue de négocier des transactions sur des bases solides. Un justiciable, sachant qu'il a 90% de chances d'obtenir 10 000 €, peut ainsi être incité à transiger à 9 000 €... Les entreprises peuvent provisionner leur risque de manière beaucoup plus solide. Les assureurs peuvent mieux gérer leur risque, singulièrement dans le domaine de la protection juridique. Cette approche est particulièrement pertinente pour les contentieux de masse compte tenu des données accessibles et de l'intérêt à désengorger les tribunaux : moins de procédures, plus de négociation, pour un résultat plus rapide... Quoiqu'il en soit, il faut naturellement préserver la faculté de remettre en cause la solution proposée en saisissant une juridiction si on l'estime plus adaptée. Alors que la déjudiciarisation consiste à sortir tel ou tel contentieux des missions du juge, la justice quantitative ne fait qu'inciter les parties à la voie amiable, tout en préservant l'accès au juge si nécessaire. » ([Monta] et [Sinclair], 2017).

De même, dans le *Guide de la justice prédictive*, [PreviCompute] arrime les dispositifs algorithmiques aux modes amiables de règlement des conflits : « La médiation, une externalité positive des algorithmes de justice prédictive ? Avoir une idée claire et quantifiée du résultat probable d'une action contentieuse est un levier de réussite des modes alternatifs de règlement des litiges. [...] Le déploiement rapide des technologies de justice prédictive entraînera progressivement un mouvement de désengorgement des juridictions par la déjudiciarisation des litiges les moins sujets à variation. Par un effet de vases communicants, la médiation, la conciliation, la transaction et les autres types de règlements amiables de conflits vont se multiplier, développant un nouveau système économique de résolution des litiges rapide, peu onéreux, et parfois même dématérialisé. » (*Guide de la justice prédictive*, 2018 : 48-49)

Ces discours mettent en avant le jugement en tant que décision qui tranche et les concepteurs positionnent leurs produits comme des solutions pour développer le recours aux transactions dans le cadre d'une conception médiatrice de la justice. *Les dispositifs algorithmiques de traitement et d'analyse des décisions de justice suscitent donc des discours qui associent deux conceptions de la justice : la justice adjudicatrice et la justice médiatrice, mais en incitant à une dissociation des deux. Au juge placé en surplomb échoit la première ; et à des acteurs privés la seconde.* Le magistrat est renvoyé à la seule dimension adjudicationnelle de son activité décisionnelle tandis que la fonction de tiers médiateur amiable est dirigée principalement vers d'autres acteurs.

Dans une célèbre typologie, François Ost (2007) a distingué trois modèles de juge : le juge Jupiter qui incarne la transcendance du droit et à ce titre décide et tranche ; le juge Hercule qui se place au niveau des rapports sociaux et cherche à ajuster le droit ; et enfin le juge Hermès qui met l'accent sur la communication et se positionne comme un acteur au sein d'un réseau juridique. Les discours sur les dispositifs algorithmiques confortent les magistrats

dans leurs fonctions de juges Jupiter, agissant en surplomb, tandis que les fonctions de médiation sociale sont davantage associées à d'autres types d'intermédiaires. L'autonomie décisionnelle des magistrats reste toutefois un point sensible.

1.3. Une vision instrumentale des dispositifs algorithmiques qui préserve la liberté du juge

On a vu dans le précédent chapitre que la controverse sur la 'justice prédictive' ouvre des débats sur la place à conférer à l'aléa judiciaire et donc sur la nature même de l'acte de juger, y compris parmi les concepteurs d'algorithmes — en particulier [PréviCompute] et [Anajur]. Ils se situent différemment sur le point de savoir s'il est possible d'une part, et souhaitable d'autre part, que le jugement soit complètement prévisible et automatisé ou complètement indéterminé et aléatoire (on se trouverait alors en dehors du cadre juridique qui agit comme un réducteur d'incertitude).

Graphique 5.4. Les jugements entre deux pôles



Même si des variations sont observables, tous les concepteurs présentent leurs dispositifs sous un jour instrumental. Ce sont des « outils » qui facilitent mais ne contraignent pas la décision, et laissent place à l'autonomie des utilisateurs. C'est une manière de soutenir que lorsque ceux-ci sont des juges, ils restent souverains.

[Laurent Moline] cofondateur de [Prédroit] estime ainsi que « la justice prédictive aura seulement la place qu'on lui accorde. Ce n'est pas parce que les ordinateurs sont capables de produire des statistiques qu'il faut se soumettre à ces statistiques. » (2017 : 527) et plus loin : « En bref, l'utilisateur reste maître tant de la question posée que de l'interprétation du résultat. » (2017 : 527) Sur son site, [Prédroit] affirme ne pas discriminer ou pondérer les décisions les unes par rapport aux autres, pour laisser *une entière liberté à l'utilisateur* : « les algorithmes de [Prédroit] sont : [...] *objectifs* : pas de modification des résultats en fonction du nombre de clics ou des consultations des décisions ». Pendant mon observation de la 4^e édition du Village de la legaltech, sur le stand de [Prédroit], cet argument m'a été donné par un commercial. Il a insisté sur l'engagement des concepteurs à ne pas « biaiser » les résultats tandis que des concurrents survaloriseraient les décisions de la Cour de cassation pour reconstituer une architecture juridique pyramidale.

Pour [Anajur], « Avec les technologies de justice quantitative, un juge peut connaître au sein de sa cour d'appel ou au niveau national la position de ses pairs, et prendre sa décision en se plaçant sur la distribution proposée, ou en s'en écartant s'il le souhaite. La justice quantitative permet ainsi aux juges de se situer en leur tendant un miroir. » ([Monta] et [Sinclair], 2017). Elle livrerait le reflet fidèle de la pratique des magistrats, sans orienter les décisions futures.

Pour [LegalEditions], il est important de « faciliter la prise de décision » (*Open data & jurisprudence*, 2018 : 20) et « d'éviter de créer des boîtes noires, des algorithmes opaques qui prendraient des décisions à leur place [celle des utilisateurs] ou fourniraient un service de

justice prédictive. À l'image du rôle joué par les outils d'aide au diagnostic dans le secteur médical, les outils technologiques dans le domaine juridique doivent servir l'utilisateur dans sa réflexion et certainement pas l'en déposséder » (*ibid.*). L'autonomie décisionnelle du magistrat est présentée comme intacte, pas plus impactée que le pouvoir de décision du médecin qui utilise scanners et autres formes d'imagerie médicale. La comparaison de ces deux types de professionnels se retrouve aussi chez [Prédroit]¹⁰ et [PréviCompute] (*Guide de la justice prédictive*, 2018 : 23)¹¹, pour qui « la justice prédictive ne vise pas à automatiser la justice, mais seulement à fournir un outil d'aide à la décision à destination des professionnels du droit, et à contribuer à la performance et la transparence du service public de la justice. » (*Guide de la justice prédictive*, 2018 : 16).

Dès lors, les opérateurs partagent très largement une vision instrumentale des dispositifs de traitement et d'analyse des décisions de justice qu'ils fournissent — ou au moins développent un discours de ce type. Ils relativisent l'impact sur les jugements et arrêts futurs.

1.4. Twitter, une arène où le jugement apparaît comme la chose du juge

Twitter est une arène dans laquelle convergent et débattent à la fois des scientifiques, des spécialistes de la communication juridique, des acteurs de la communauté des politiques publiques déjà repérée et une part importante d'acteurs économiques (en particulier les opérateurs de dispositifs algorithmiques comme [Anajur], [PréviCompute]). Twitter peut être assimilé à une arène du débat public (Hilgartner et Bosk, 1988 ; Céfaï, 2007 ; Douay et Reys, 2016). C'est un espace de discussion qui suppose une audience (le ou les publics), repose sur ses propres règles, sa propre technicité et culture du débat public (Badouard, Mabi et Monnoyer-Smith, 2016 : 10). C'est un lieu d'élargissement de la visibilité des questions et des problèmes publics comme le sont les arènes médiatique et judiciaire. Cette plateforme fonctionne aussi telle une caisse de résonance des médias traditionnels ainsi que pointé dans le chapitre 2. Toutefois, elle a ceci de spécifique, qu'elle permet à différentes catégories d'acteurs de participer au débat public sans devoir passer par l'intermédiation classiquement effectuée par les médias et journalistes. Dans cet espace, les échanges et débats font la part belle à une conception décisionniste de l'activité judiciaire.

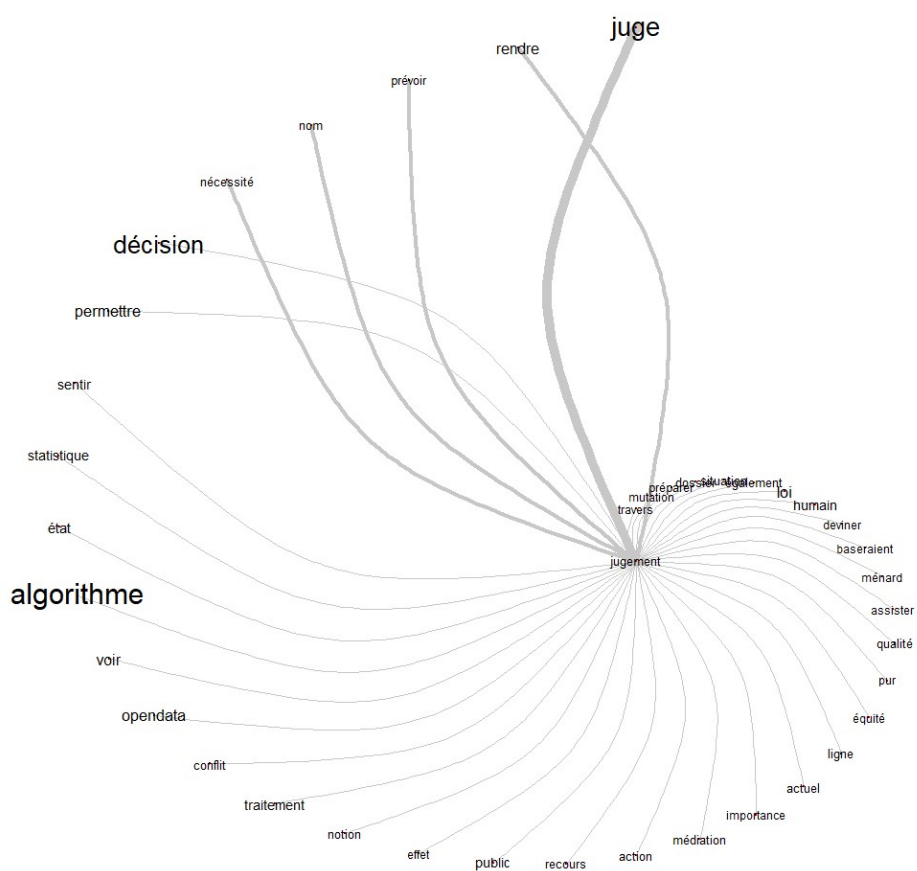
Nous analysons ici les liens lexicaux entre chacun des mots « jugement », « décision » et « procès » et d'autres termes utilisés dans les *tweets* : les graphes indiquent quels mots présentent la plus forte co-occurrence avec le mot sélectionné. Cette première approche

¹⁰ « Un avocat augmenté, dirait-on, qui comme son homologue médical commencerait à s'acculturer au fait de s'équiper d'outils capables de "booster" sa productivité et son chiffre d'affaires. » PrediDroit : <https://www.Predidroit.com/fr/blog/id-13-justice-predictive>,

¹¹ La comparaison avec le secteur médical vise à établir que si les deux professions de médecin et d'avocat reposaient sur des pratiques artisanales il y a un siècle, « le corps médical a très largement profité du progrès technologique [... alors que du côté de l'avocat] le fax et le traitement de texte ont été adoptés. Le mail et le téléphone, dont la profession fait un usage souvent effréné, aussi. Mais sinon, rien. Cette comparaison met en lumière quelque chose de simple : le monde du droit n'a pas suffisamment tiré profit de l'innovation ». Tiré du chapitre 3 du guide *Un outil de performance pour les cabinets d'avocats* rédigé par Sophie Clanchet, présentée comme « avocat associé du cabinet d'avocats Clanchet & Rivière, présidente d'Eurojuris France, ancienne présidente de l'UJA [Union des jeunes avocats] ».

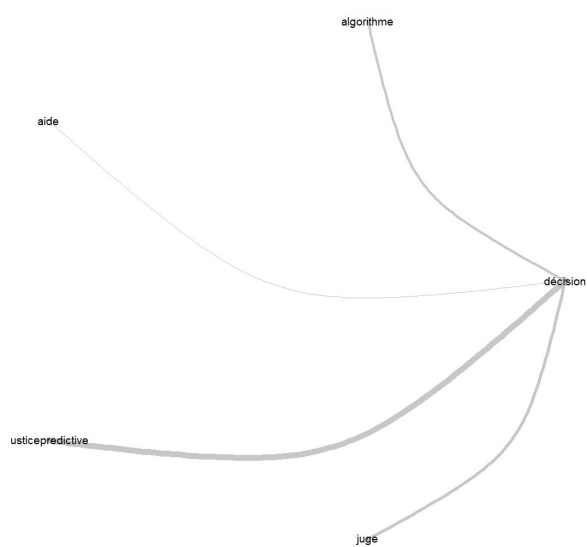
demande à être précisée et affinée dans nos travaux ultérieurs — notamment parce qu'à ce stade, ces données lexicales n'ont pas été croisées avec les informations sociodémographiques dont nous disposons pour leurs locuteurs. Le graphe ci-dessous, montre que, dans les *tweets*, le jugement apparaît directement et est principalement associé aux mots *juge*, *algorithmes* et à l'enjeu de *décision*.

Graphique 5.6. Analyse de graphe du mot jugement. Tweets 'justice prédictive', 2010-2020



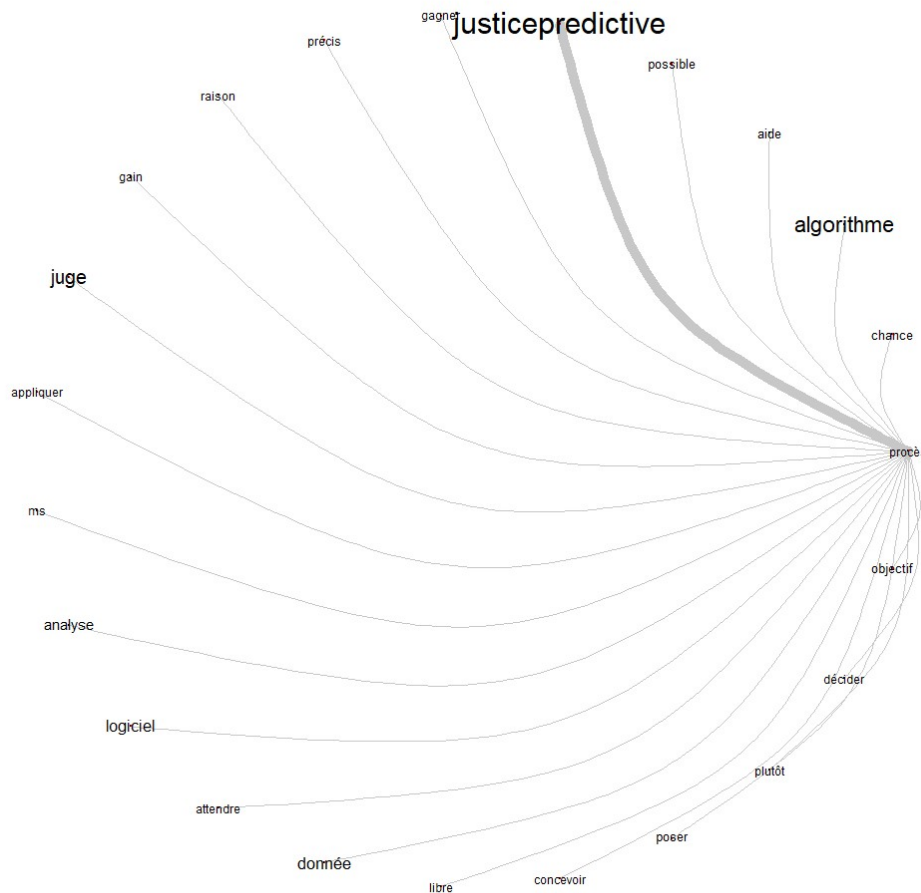
En outre, le jugement est surtout associé à la décision que le juge doit rendre ce que confirme le graphe du mot « décision ».

Graphique 5.7. Analyse de graphe du mot décision. Tweets 'justice prédictive', 2010-2020



Dans ce dernier graphe, la 'justice prédictive' apparaît comme un *algorithme* qui apporte une *aide* au *juge* dans la prise de *décision*. L'analyse du terme « procès » montre que celui-ci est réduit au jugement rendu, au sens de décision finale du juge, en occultant les autres dimensions de l'événement et la dimension processuelle de sa fabrique. L'intérêt des dispositifs algorithmiques se concentre dans la potentialité à anticiper le résultat d'un procès, connaître ses *chances* de *gagner*, d'obtenir un *gain*.

Graphique 5.8. Analyse de graphe du mot procès. Tweets 'justice prédictive', 2010-2020



La conclusion que je tire de ces quelques éléments lexicaux est que les discours sur Twitter associent le jugement au juge et au pan décisionnel de son activité. Cela indique une représentation centrée sur l'acte décisionnel (trancher, décider) d'une part et sur le juge (à l'exclusion d'autres acteurs) d'autre part. Cela donne un peu plus de poids à mon hypothèse de circulation d'une représentation décisionniste du jugement, dans le cadre des débats sur les dispositifs algorithmiques de traitement et d'analyse des décisions de justice.

II / Une conception décisionniste en décalage avec la réalité du fonctionnement judiciaire

Je voudrais à présent montrer que cette vision décisionniste subsiste en dépit de son incapacité à exprimer avec justesse la fabrique des jugements. C'est un mythe qui survit alimenté par un certain nombre d'acteurs.

Dans différentes sciences sociales, l'analyse de la décision a été longtemps dominée par la figure du décideur rationnel, issue de la conception wébérienne de la décision. Celle-ci peut être résumée comme reposant sur « le principe du choix de la solution optimale par un décideur clairement identifiable qui agit selon une rationalité instrumentale » (Hassenteufel, 2011 : 66). Les théories de l'acteur rationnel ont nourri ce modèle (Balme, 2019). Celui-ci repose sur quatre postulats principaux : il existe un décideur identifiable ; ce dernier s'appuie sur un système de préférences stables, il est cohérent dans ses objectifs et les moyens pour les atteindre ; il décide en parfaite connaissance de cause, c'est-à-dire en disposant de toutes les informations sur les options disponibles, leurs conséquences, etc. ; enfin, le décideur recherche la solution optimale, à savoir celle qui maximise ses profits et minimise ses coûts (1). Le droit et la science juridique ont également été profondément marqués par la figure du décideur rationnel, appliquée au juge. La conception du jugement qui prévaut dans les discours sur les dispositifs algorithmiques de traitement et d'analyse des décisions de justice peut être qualifiée de décisionniste (2). Cette représentation, quoique socialement partagée, n'en est pas moins fantasmatique et erronée au regard de ce que nous savons du fonctionnement concret des juridictions et de la réalité de l'activité de juger des magistrats (3).

2.1. Décision, rationalité et décisionnisme

La question de qui décide et comment est au cœur de la science politique. La théorie et la philosophie politiques, les relations internationales, la sociologie politique, l'analyse des politiques publiques ont partagé les pièces de ce puzzle académique. Au sein de l'analyse des politiques publiques, l'existence même d'un décideur et d'une décision identifiables a été interrogée, et ce, alors qu'elle repose classiquement sur le postulat selon lequel « l'action de l'État se matérialise dans l'étude de la décision publique » (Halpern, 2019 : 200). Ce n'est pas le lieu de reprendre ici les bibliothèques entières qui ont contribué à ce travail de déconstruction de la décision. Il suffit de dire qu'il est maintenant établi par de très nombreux travaux, inscrits dans des écoles et des modèles théoriques variés, que la décision n'est ni nécessairement un choix entre plusieurs options ; ni forcément rationnelle ; ni toujours attribuable à un acteur identifié (pour une synthèse, voir Eymeri-Douzans, 2021). Dans le flou et l'absence de lisibilité caractéristiques des anarchies organisées, les décisions se prennent insensiblement (March et Olsen, 1976). La maîtrise du contenu des décisions est toute relative, au vu des jeux d'actions qui se déploient non seulement au moment de l'adoption d'une loi par exemple, mais aussi de sa mise en œuvre. Cette dernière apparaît comme un processus non pas d'application mécanique d'une décision, mais bien de redéfinition des objectifs de la politique elle-même à travers des cadres cognitifs, des systèmes et jeux d'acteurs déjà en place. Cela conduit à relativiser l'importance des décisions de politiques

publiques et à les replacer dans un ensemble d'actions et de processus qui ne sont pas maîtrisés par un décideur — fût-il repérable.

Dans le contexte francophone, *Critique de la décision*, publié par Lucien Sfez au début des années 1990 a joué un rôle séminal, même s'il est aujourd'hui un peu oublié. Cet ouvrage passe au crible « le paradigme de la décision autonome, rationnelle expression de la liberté individuelle [qui] est au cœur de notre conception occidentale de la décision » (Sfez, 1992). Il isole et décrit cette conception qu'il qualifie de « décisionniste », qui combine les qualités supposées de linéarité, de rationalité et de liberté du décideur. Il montre que le résultat en est la survalorisation des décisions comme produit de l'action d'un décideur, omniscient et omnipotent, agissant de façon autonome et rationnelle. Les actions sont réduites à leur caractère décisoire condensé dans un temps T, isolable et repérable sous la forme d'un événement. Bref, c'est la conception que l'histoire événementielle des grandes décisions et des grandes figures (généralement masculines) a longtemps mise en avant, et que les acteurs politiques et les médias continuent d'alimenter, car elle fournit des clefs de lectures simples et efficaces de la réalité sociopolitique. À partir de l'étude du cas du RER francilien, il montre qu'il est impossible d'établir quand la décision de construire celui-ci a été prise exactement : le RER est le produit d'un ensemble illisible de décisions apparemment anodines intervenues dans les années 1960.

Le constat de Lucien Sfez est largement partagé par l'analyse des politiques publiques : une part importante des études dans ce domaine s'est précisément attachée à remettre en cause ce mythe du décideur et de la décision repérables. Aux États-Unis, des travaux ont également considéré la décision à travers son envers, la non-décision et montré que celle-ci consiste à la fois à décider de ne rien faire, mais aussi à ne pas décider de faire certaines choses et décider de ne pas en faire d'autres (Bachrach et Baratz, 1963). La décision prend ainsi « la forme d'un flux continu de décisions et d'arrangements ponctuels, pris à différents niveaux du système d'action, qu'il faut analyser comme un ensemble de processus décisionnels » (Muller et Surel, 1998, cités par Hassenteufel, 2011 : 82-83). De fait, ces approches mettent en évidence combien les politiques publiques résultent de décisions et d'accommodements largement achevés, processus dans lesquels intervient une diversité d'acteurs disposant de ressources et moyens d'action inégaux.

Le terme même de « décisionnisme » est polysémique et s'inscrit dans des débats très larges en théorie politique et en théorie du droit, notamment en lien avec la discussion sur le positivisme et l'approche de Carl Schmidt (Chaniel, 2001). Ce qui est discuté ici n'est pas sans rapport avec ce débat, mais en concerne un aspect plus limité. Dans *La Technique et la Science comme idéologie*, Jürgen Habermas (1973) distingue en effet trois formes idéales-typiques des rapports entre science et pouvoir : les modèles décisionniste, technocratique et pragmatique. Dans le premier, le pouvoir est dans les mains du décideur qui détient l'autorité ; le politique décide seul, en fonction d'une rationalité par valeur, même s'il peut s'appuyer sur un savoir scientifique qu'il a sollicité et qui intervient de manière informative et subordonnée. Dans le modèle technocratique, la réalité du pouvoir est inversée : les spécialistes et les techniciens dictent les décisions politiques au nom d'une supériorité du savoir scientifique et technique, d'une prévisibilité des effets de la décision, etc. Dans le modèle pragmatique enfin, les

scientifiques et les connaissances qu'ils produisent viendraient en appui d'une autorité politique en difficulté.

En somme, le modèle décisionniste tel que dessiné par Sfez ou par Habermas est la concaténation de tout un ensemble de présupposés : la décision serait identifiable ; elle serait détenue par un acteur ; lequel agirait rationnellement (du point de vue de l'adéquation objectifs et moyens pour les atteindre) ; et serait légitime à décider en fonction d'une rationalité par valeur (choix de valeurs), laquelle inféoderait d'autres justifications et légitimités à décider, comme la logique scientifique. La science politique a questionné cet agglomérat, à différents degrés et via plusieurs angles. La conséquence en est que cela a relativisé l'intérêt porté à la décision en tant qu'objet analytique : en science politique, peu de travaux portent désormais directement sur elle. Les analyses de la décision ont finalement contribué à dissoudre leur objet, mais elles ont posé un certain nombre de jalons qui, à mon sens, sont utiles pour penser le jugement en contexte de justice.

2.2. Décisionnisme, droit et justice

La science politique a certes contribué à une sociologie des décisions publiques, mais sans intégrer le cas des jugements en contexte de justice, restés en dehors de son questionnement (Boucher, 2019, Saint-Martin, 2014). Or, l'approche décisionniste est particulièrement ancrée dans la doctrine et la science du droit (1), comme le révèlent de manière comparable le cas de l'expertise judiciaire et celui des dispositifs algorithmiques de traitement et d'analyse des décisions de justice (2).

2.2.1. La vision mythifiée du juge

La vision mythifiée du jugement et du juge est fondatrice et vivace. Elle comprend les deux faces d'une même médaille : le juge « bouche de la loi » et le juge créateur de droit. Les deux conduisent à une focalisation comparable sur la décision comme acte central, linéaire et univoque, et sur le juge comme point nodal de cette action – à l'exclusion des autres acteurs impliqués : avocats, experts, parties et justiciables, médias, etc.

Dans la théorie de Montesquieu (1758), où la séparation des pouvoirs est envisagée comme un rempart nécessaire contre le despotisme, le juge est réduit au rôle de rouage du système judiciaire. C'est le modèle de la « bouche de la loi », dans lequel Montesquieu lui-même décrit le jugement comme « un texte précis de la loi » (cité par Timsit, 1993 : 159). Les philosophes des Lumières idéalisent la législation : les compétences des tribunaux sont limitées, l'interprétation de la loi par le juge, considérée comme dangereuse (Beccaria, 1764 : 52-53). Dans cette veine, les approches positivistes voient le juge comme étroitement lié par « la volonté du législateur » et nécessairement placé sous sa subordination. Le principe du syllogisme judiciaire¹² incarne cet effacement du juge puisque cette opération déjà séquencée et rationnelle, lui impose quasiment le verdict. Issu de la logique et importé dans la justice, ce mode de raisonnement assimile le travail judiciaire à un procédé *mécanique* d'application de

¹² Le syllogisme est une forme de raisonnement qui comprend trois propositions : deux prémisses — la majeure et la mineure — et la conclusion. Ainsi cet exemple bien connu : Tous les hommes sont mortels ; or Pierre est un homme ; donc Pierre est mortel.

normes générales, plus ou moins abstraites, à des faits concrets. Pour la doctrine formaliste, le jugement serait le résultat de la mise en regard d'une majeure (le texte juridique) et d'une mineure (les faits). Le ou les juges, en cas de collégialité, sont certes les personnages centraux de ce raisonnement mais ils apparaissent quasiment dépourvus de latitude d'appréciation.

La critique du modèle légicentrique et de la méthodologie classique centrée sur l'exégèse de la loi et le syllogisme, est passée par différents chemins, en droit et en théorie du droit. Les écoles réalistes, en particulier, ont marqué un tournant (Michaut, 2000 ; Millard, 2014 ; Serverin, 2000). Dès lors, certains juristes souvent juges eux-mêmes¹³, surtout dans le contexte de pays de *common law* se sont intéressés aux comportements effectifs des juges dans une perspective sociologique et psychologique. Ils insistent sur les facteurs extra-juridiques du jugement et sur le rôle déterminant de la personnalité du juge. Dans cette lignée, bien des modèles ont ensuite été construits pour rendre compte des facteurs multiples qui interagissent dans la fabrication des jugements (pour une synthèse : Delpeuch, Dumoulin et de Galember, 2014, chap. 2).

Une autre voie de critique du modèle positiviste et formaliste, s'appuie sur l'idée que l'exercice du jugement ne relève pas du syllogisme, mais d'une intuition qui dépend nécessairement d'un sens du juste qui est découvert en même temps qu'il s'impose au juge. Juger comporterait une part de mystère propre à tout processus de découverte. La valorisation de ce sens du juste dont les juges seraient porteurs, du fait de leur formation, leur expérience, mais aussi leurs qualités morales (la tempérance, la sagesse, etc.), est un lieu commun chez les juges eux-mêmes. Alain Bancaud (1993), dans une étude des hauts magistrats français, a mis en évidence cet *ethos*, qui fonde leur identité professionnelle. La survalorisation du décideur est ainsi bien présente en contexte judiciaire pour décrire le rôle du juge et discutée par différents mouvements de la doctrine et de la théorie du droit. Elle n'est toutefois guère connectée aux travaux qui, dans les sciences sociales et notamment en science politique, ont entrepris la critique du décisionnisme.

2.2.2. De l'expertise judiciaire... aux algorithmes : même combat ?

Le décisionnisme m'apparaît aujourd'hui encore bien vivace dans la manière de penser les rapports des juges aux jugements. Dans un précédent travail, j'ai analysé le cadre juridique mis en place aux XIX^e et XX^e siècles autour de l'expertise judiciaire (Dumoulin, 2007), à travers codes, lois, décrets et ordonnances, commentaires de doctrine et décisions de jurisprudence. En restituant ce processus d'encadrement, j'ai montré comment un modèle décisionniste de l'expertise judiciaire, centré sur l'idée d'une compétence ou d'une spécialité strictement technique venant éclairer la prise de décision d'un juge maître de l'issue finale, s'est progressivement constitué. La longue tradition de recours aux experts dans le champ judiciaire, la force du corps de doctrine élaboré pour normer les pratiques d'expertise ont très largement façonné une vision idéale de l'expertise, comme procédure accessoire, strictement

¹³ Pour Holmes (1897), le droit n'est rien d'autre que « la prophétie de ce que les juges feront » (The prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law) et pour Llewellyn (1951), ce que les acteurs de la chaîne civile et pénale « font à propos des conflits » (What these officials [judges, sheriffs, clerks, jailers, lawyers, etc.] do about disputes is, to my mind, the law itself. »).

technique et dénuée d'effets sur la décision. L'expertise judiciaire incarne, de manière emblématique, la figure de l'expertise mandatée par un décideur en position de surplomb. Selon les textes juridiques, en France, les experts judiciaires ne sont ni professionnels du droit, ni professionnels de justice, ni même professionnels de l'expertise puisqu'ils n'ont pas le droit d'en faire leur activité principale. Leurs conclusions ne peuvent engager les magistrats qui restent libres de les apprécier et de les utiliser ou non dans le jugement. L'expertise est construite comme une procédure secondaire qui n'engage pas le juge, commanditaire du rapport d'expertise, et dont la position est sublimée au nom d'une légitimité fondée sur le droit et le lien avec le système étatique et politique.

Au fil de mes investigations actuelles, est apparu un parallèle inattendu entre le rôle des systèmes algorithmiques et celui des rapports d'expertise dans la prise de décision du juge. Dans les deux cas je vois se dessiner un rôle d'aide à la décision dont il est estimé qu'il ne doit pas empiéter sur l'*imperium* du juge. Un auteur comme Emmanuel Jeuland (2017 : 9 ; 2019 : 194) fait explicitement la comparaison. « Il faudrait aussi bien mesurer que la mise à jour des faits par algorithme ne devra jamais remplacer la décision du juge, pas plus qu'une expertise ne doit se substituer au juge. » Et dans un autre texte : « Il existe déjà des barèmes locaux en matière pénale ou de dommages et intérêts, mais pas plus la décision calculée que l'avis de l'expert ne s'impose au juge. [...] Une expertise ne dit pas qui est responsable d'un dommage, mais ce qui en est la cause la plus probable. Elle se rapproche de la qualification de faute. » (Jeuland, 2017 : 9-10, 11)

Je retrouve ici cette même représentation sociale que pointait Lucien Sfez dans ses travaux, et qui consiste à réaffirmer la figure du décideur, la réalité de sa capacité à décider et à agir, et son autonomie à le faire, loin des menaces technocratiques d'une prise de pouvoir d'acteurs ou d'actants extérieurs.

Le juge est donc hypertrophié au sens où il apparaît comme le personnage central (unique ?) de la décision. Ce n'est pas sans faire écho à la figure du « législateur », dont la doctrine fait grand cas et scrute les intentions et argumentations, avec force anthropomorphisme. Lorsque j'ai analysé plus haut les discours sur les mots « jugement », « décision », « procès », à partir du corpus de *tweets*, ce biais de perspective est apparu : le jugement semble se résumer à une décision qui serait la chose d'un juge seul, dans sa tour d'ivoire. Cette double focalisation sur le juge et la décision comporte une série de raccourcis et de simplifications, qui relèvent d'une vulgate qui mêle la figure du juge, acteur souverain de la décision et créateur de droit, largement diffusée par les théories réalistes du droit, et celle du juge, bouche de la loi, classique dans les approches formalistes et positivistes. Dans les deux cas, il y a une erreur de perspective : le juge apparaît à travers son seul pouvoir de trancher et comme celui vers qui convergent toutes les attentions. La valorisation de la dimension instrumentale des dispositifs algorithmiques fait en outre écho à ce qui est connu et documenté pour toute une série de technologies. L'accent mis sur la capacité à dominer « l'outil » technologique est une manière de rassurer quant à ses effets possibles sur l'activité qu'il équipe. Cela peut aussi être relié à une façon de dichotomiser et de prioriser le droit sur les autres savoirs et techniques classiquement pensés dans la doctrine juridique comme accessoires ou subordonnés.

Or, à propos de l'expertise judiciaire, j'ai aussi montré que les rapports d'expertise représentent à la fois une ressource et une contrainte (Dumoulin, 2000), ce qui donne lieu à des rapports à géométrie variable entre expertise et décision. J'ai expliqué la variabilité de ces relations à partir d'une pluralité de facteurs, parmi lesquels la nature des opérations scientifiques réalisées, la présence ou pas d'autres éléments de preuve dans le dossier, etc. Dans tous les cas, j'ai mis en évidence des jeux complexes autour de la fabrique de la décision judiciaire qui laissent la place à une vision coproduite de celle-ci. Elle a été prolongée par d'autres auteurs qui confirment le caractère mythe décisionniste est bien... un mythe.

C'est le cas des travaux de Romain Juston-Morival (2016, 2018, 2021a, 2021b) qui sont allés plus loin, en adoptant un cadre théorique centré sur la coproduction entre science et droit (Jasanoff et Leclerc, 2013). En matière pénale, il observe des « configurations locales de la science et du droit » (2021b) entre médecins légistes, enquêteurs et magistrats du parquet. Dans un contexte de procédures de traitement en temps réel (TTR) et d'accélération du temps judiciaire, les parquetiers voient leur autonomie professionnelle réduite par le rôle déterminant pris par les rapports médicolégaux dans le façonnage des mesures qu'ils ordonnent. R. Juston-Morival souligne combien les configurations locales pèsent concrètement et en situation sur le jeu de qui décide. Elles interviennent dans l'évaluation de l'opportunité de réaliser un acte — une autopsie ou des examens complémentaires — mais aussi dans la mobilisation des rapports d'expertise — par exemple concernant le caractère plus ou moins automatique de la qualification juridique à partir du nombre de jours d'incapacité temporaire de travail (ITT) établi par le médecin légiste. De nombreux éléments influencent ces configurations : la manière dont est organisé le travail parquetier, la culture scientifique régionale en matière de médecine légale. Les « attentes formulées à l'endroit du dispositif médicolégal par ses différents protagonistes » (342) d'une part, « les attributs des magistrats, les éléments d'enquêtes policières, les normes inscrites dans les codes juridiques et les protocoles locaux » (336) d'autre part interagissent. Ces différents composants produisent des assemblages caractérisés par leur modularité, qui, par conséquent, livrent une image kaléidoscopique des façons dont le jugement des parquetiers est orienté. Dans tous les cas, à partir de l'expertise judiciaire, il est démontré la non-pertinence sociologique d'une représentation décisionniste du jugement. Plus largement, une diversité de travaux de sociologie du droit et de la justice, démentent eux aussi la validité de cette représentation du juge et du jugement.

2.3. Le décisionnisme malgré la sociologie du droit et de la justice

Les discours des opérateurs de systèmes algorithmiques et les débats relayés par la presse et sur Twitter, font référence au juge, généralement au singulier, comme figure désincarnée et considérée en dehors de la réalité concrète. Le jugement est pensé comme la décision *d'un* juge « hors sol ». Cette vision est en décalage profond avec ce que les sciences sociales nous ont appris de l'activité de juger, que ce soit à propos des décisions de contrôle de constitutionnalité dans les *judicial behavior studies*, des décisions pénales avec les *sentencing studies*, ou bien dans les études qui portent sur jugements en matière civile.

2.3.1. Des juges incarnés et porteurs d'idéologies dans les judicial behavior studies

Aux États-Unis, de nombreux travaux ont analysé la manière dont les juges prennent leurs décisions¹⁴. Les théories développées dans cette littérature¹⁵ examinent trois facteurs principaux de la prise de décision des juges : l'état du droit (les doctrines et décisions de jurisprudence disponibles, etc.), les préférences politiques individuelles des juges et celles d'autres acteurs — autres juges, acteurs de l'exécutif, élites ou encore grand public (Maveety, 2003 ; Baum, 2017 : 114). Le modèle attitudinal met l'accent sur les préférences politiques des juges. Le modèle dit stratégique insiste sur le rôle combiné des préférences politiques des juges avec celles des autres acteurs de l'environnement social et politique. L'institutionnalisme historique souligne la combinatoire entre ces éléments et un certain état du droit. Dans tous les cas, le juge n'est pas seul ni en position de surplomb ; il n'est pas non plus désincarné, mais considéré comme un entrepreneur moral et un acteur politique, à travers sa trajectoire, ses orientations idéologiques et son environnement sociopolitique. Plus largement encore, le caractère éminemment social du processus décisionnel des juges est établi, que ce soit en référence aux différents publics auxquels s'adresse la décision et dont les juges doivent anticiper les réactions (Baum, 2006) ou bien en lien avec la construction des positions idéologiques (conservatrice et libérale) elles-mêmes. L'enjeu d'aller au-delà de l'analyse du processus de décision judiciaire (*judicial decision making*) pour prendre en compte la « production sociale des décisions judiciaires » (*the social production of judicial decisions*) est à l'agenda scientifique (Owens, 2016 : 503).

2.3.2. Les juges, maillons des chaînes civile et pénale

Dans cette bureaucratie professionnelle qu'est la justice (Heydebrand, 1977 ; Mintzberg, 1982), les magistrats disposent d'une grande autonomie d'action, renforcée par l'organisation en petites cellules de travail relativement dissociées les unes des autres, et l'exercice dans un cadre géographique donné. On sait pourtant que les magistrats ne décident pas de manière isolée, mais interviennent en tant que maillons d'une chaîne décisionnelle qui en comporte beaucoup d'autres, tant en amont (plaignants, requérants, organes administratifs, policiers, avocats, procureurs...) qu'en aval (cours d'appel, de cassation, constitutionnelles, internationales, etc.)¹⁶. Cette chaîne, agit comme conditionnement objectif des décisions possibles et comme support d'anticipations stratégiques. Bien que possédant un pouvoir d'appréciation, dont la nature varie selon la position qu'ils occupent dans la chaîne, les juges sont obligés de tenir compte des décisions qui ont précédé celle qu'ils doivent prendre. Cela relativise l'image d'un juge solitaire et omnipotent (Bordonaba et Chouchan, 2016 : 111). Ce qui, dans le jugement, est présenté comme une structure logique

¹⁴ Compte tenu du fait que, dans le système américain contrairement au droit constitutionnel continental, le contrôle des lois (*judicial review*) est exercé par l'ensemble des tribunaux. « N'importe quel juge de n'importe quel tribunal, quel que soit le cas, à n'importe quel moment, à la requête de n'importe quel justiciable, a le pouvoir de déclarer une loi anticonstitutionnelle » (Stone Sweet, 2000, p. 32).

¹⁵ (*inter al.* Howard et Randazzo, 2017 ; Epstein et Lindquist, 2017 ; Segal, 2011).

¹⁶ (*inter al.* Danet, 2013 ; Faget, 2008 ; Hawkins 2003 ; Jobard et Nevanen, 2007 ; Kaminski, 2015 ; Léonard, 2010 ; Philippe, 2022 ; Van Hamme et Beyens, 2007).

et des attendus, peut être analysé comme le résultat d'une reconstruction *a posteriori* qui informe surtout sur le système judiciaire et son organisation (Robert, Faugeron et Kellens, 1972 ; Lenoir, 1996). De plus, l'amont engage l'aval. De nombreuses enquêtes le confirment, par exemple avec l'effet d'entonnoir constaté à partir d'une étude empirique des sanctions pénales prononcées par les juridictions correctionnelles françaises (Saas, Lorvellec et Gautron, 2013).

Cela est également vrai en matière civile. Les effets du choix de la procédure sur la suite du processus de séparation conjugale sont déterminants (Collectif Onze, 2013 : chap. 2 ; Biland, Gollac *et al.*, 2021 : 126)¹⁷. Si les jugements judiciaires emportent avec eux un effet de clôture des disputes, en raison du principe de l'autorité de la chose jugée, ils résultent aussi d'une chaîne d'actions et d'autres décisions juridictionnelles, intervenues en amont et qui encadrent de fait l'espace d'autonomie décisionnelle des magistrats. Pour reprendre l'analogie avec les décisions de politiques publiques, elles créent des effets d'inertie et rendent plus difficile un changement radical de cap, conformément au modèle de la dépendance au chemin emprunté (*path dependency*) (Pierson, 2000).

Les juges prennent aussi en considération les répercussions probables de leur jugement sur les étapes ultérieures de l'affaire, sa réception sociale et politique, mais aussi les logiques et processus institutionnels qui en affecteront la mise en œuvre. Des mécanismes complexes d'anticipation ou de compensation *ex post* sont relevés (Beyens, Snacken, et Van Zyl Smit, 2013). Par exemple, dans le contexte de la réforme du droit pénal belge où de nouveaux textes prévoient que les peines de moins de trois ans ne feront pas l'objet d'incarcération (mais d'autres formes d'exécution de la peine), les juridictions sont moins réticentes à placer les accusés en détention provisoire, considérant que cela compense le fait qu'en cas de courte peine, celle-ci ne sera pas exécutée ou qu'elle sera aménagée, notamment par le port d'un bracelet électronique. Certains juges prononcent des peines plus longues que celles qu'ils estiment justifiées au regard des faits, pour s'assurer qu'à l'issue des réductions de peine et aménagements, un temps minimum d'incarcération soit effectué.

2.3.3. Les juges, porteurs de cultures professionnelles et parties prenantes de systèmes d'action locaux

Si en France l'ensemble du système judiciaire national est soumis aux mêmes règles de droit, aux mêmes exigences éthiques et déontologiques, les pratiques peuvent différer singulièrement d'une juridiction à l'autre. Chaque tribunal produit sa propre culture organisationnelle, en fonction de son histoire et de son environnement (Ackermann et Bastard, 1993). Des politiques locales déclinent les politiques nationales (politiques pénales, accès aux droits, modes alternatifs de règlement des litiges, etc.) ; des modalités particulières de coopération avec les partenaires (services sociaux, services enquêteurs, médecine légale, etc.) dessinent des configurations locales spécifiques. Ce contexte façonne la manière dont les

¹⁷ Déjà, à partir de l'analyse de dossiers d'affaires familiales, Irène Théry (2001[1993] : 307) avait établi que les premières décisions concernant le droit de garde et d'hébergement des enfants ont une portée particulière, même prises à titre provisoire. Plus le temps passe, moins elles sont remises en cause par les juges amenés à statuer au fond.

juges travaillent et produit des effets sur les orientations prises dans le règlement des dossiers (Juston-Morival, 2021a). De plus au fil de relations professionnelles régulières, des groupes de travail se stabilisent qui développent valeurs, normes et routines partagées¹⁸.

L'activité décisionnelle des magistrats repose également sur un ensemble d'interactions personnelles des juges, en interne et en externe de la juridiction. Souvent membres de juridictions collégiales¹⁹, ils et elles s'appuient sur leurs collègues locaux, notamment à leurs débuts dans la fonction (Bastard et Mouhanna, 2010 : 80-81) ainsi que sur des réseaux personnels de pairs (anciens collègues de promotion ENM ou de précédentes juridictions, etc.). Ils s'adosent également à des réseaux plus ou moins institutionnalisés : associations et syndicats de magistrats, listes de discussion généralistes ou spécialisées, etc. (Licoppe et Dumoulin, 2019). Ils et elles utilisent également des trames de jugement, des abaques, des barèmes plus ou moins précis²⁰. À bien des égards, le jugement apparaît comme le résultat d'une activité sociale et collective (à propos du Brésil, voir Prates, 2011), inscrite dans un environnement local, une culture organisationnelle, mais aussi dans des échanges et réseaux propres à la profession de magistrat, ou de segments et sous-groupes à l'intérieur de celle-ci.

Il convient donc de substituer au portrait du juge la peinture d'un vaste paysage où réapparaissent collègues, collaborateurs, greffiers, partenaires. Le juge est certes un personnage central de l'organisation judiciaire, mais il n'est pas seul : il interagit avec de nombreux autres intervenants, dans le tribunal ou à l'extérieur. Même si elle tend à reculer, l'audience est une scène d'interactions institutionnellement réglées sur laquelle se croisent, voire se confrontent différents acteurs du drame judiciaire²¹. C'est une épreuve et un cadre (Codaccioni, Puccio-Den et Roussel, 2015) où les juges procèdent à un certain nombre d'opérations cognitives et discursives, pendant les échanges oraux : interprétation des éléments portés à leur connaissance, évaluation et classification des personnes auxquelles ils ont affaire et aux discours qu'elles avancent. Ces opérations reposent sur des modes de raisonnement fondés sur une rationalité et un formalisme spécifiques, sur des routines de pensée, préjugés, intuitions et croyances, et enfin sur des discussions, confrontations linguistiques (Komter, 2012).

Le jugement, comme d'autres décisions, est donc le résultat de collaborations, de négociations, d'ajustements et d'arrangements entre acteurs ou groupes d'acteurs. Il peut arriver que comme d'autres décisions publiques, il se dissolve au fil de temporalités et d'interactions multiples. « Le moment du jugement devient indiscernable. [...] La décision s'est prise, peu à peu, dans un continuum auquel les parties ont été invitées à participer. » conclut

¹⁸ (Eisenstein et Jacob, 1977, à partir du cas de la négociation du *plea bargaining* dans les procédures correctionnelles de trois ressorts du nord-est des États-Unis).

¹⁹ Malgré une nette tendance à la progression du juge unique au pénal et au civil, « l'irrésistible ascension de l'unicité » pour reprendre le mot qui n'a pas vieilli de G. Giudicelli-Delage (2003 : 140).

²⁰ (Gerry-Vernières, 2019 ; Sayn, 2019 ; en matière de droit du travail : Bourreau-Dubois, 2019 ; dans le cadre du traitement en temps réel en matière pénale : Bastard et Mouhanna, 2007 ; Juston-Morival, 2021b).

²¹ (En matière correctionnelle : voir Herpin, 1977 ; Christin, 2008 ; en matière familiale, voir Le Collectif Onze, 2013).

Irène Théry (2001), toujours en matière d'affaires familiales. Vingt ans plus tard, une enquête estime également que l'accord des parents s'impose dans les décisions : « il est extrêmement rare que la décision du juge n'entérine pas ce consensus » (Biland, Gollac *et al.*, 2021 : 126). Ainsi, ce que l'on a étudié plus haut à propos des experts judiciaires spécifiquement, s'applique aussi dans une certaine mesure, et suivant les domaines, à d'autres catégories d'acteurs : professionnels du droit (greffiers, avocats) et de la justice ou destinataires des décisions. En matière pénale et civile, bien d'autres partenaires interviennent aussi dont on ne détaillera pas le rôle ici, mais qui suggèrent que le jugement n'est pas le fait du seul juge, mais bien le résultat d'une coproduction et qu'il ne relève pas 'un choix tranché mais comporte aussi, en fonction des configurations, des négociations, compromis, arrangements, tout comme les autres décisions publiques (Eymeri-Douzans, 2021).

La controverse sur la 'justice prédictive' a donné à voir une conception du jugement que j'ai qualifiée de décisionniste. Mes analyses, fondées sur les acquis de la sociologie du droit et de la justice, font apparaître un décalage entre les représentations sociales de l'activité de jugement et la réalité du processus de fabrique de celui-ci. Le risque existe, dès lors, de développer une image erronée des effets et appropriations possibles des dispositifs algorithmiques dans le contexte juridictionnel. À ce stade, l'enjeu est de proposer une lecture des dispositifs algorithmiques de traitement et d'analyse des décisions de justice, qui repose sur une autre représentation du jugement, plus conforme à la réalité des pratiques, et qui éclaire d'un nouveau jour les effets de leur développement.

III / Comment élaborer une analyse différente des dispositifs algorithmiques de traitement et d'analyse des décisions de justice qui ne repose pas sur cette conception décisionniste du jugement ?

À la lumière de ce qui vient d'être présenté, comment élaborer une analyse des dispositifs algorithmiques de traitement et d'étude des décisions de justice qui organise la mise à distance de ce déjà-là repéré comme une évidence, cette conception socialement partagée qu'est la vision décisionniste du jugement ? Clarifier et élargir la notion de jugement, rapprocher les jugements juridictionnels des jugements ordinaires, prendre en considération la fabrique du jugement dans sa dimension processuelle et matérielle sont autant de manières que j'avance pour problématiser l'objet— qui dessinent une voie différente, traductible dans un protocole méthodologique.

3.1. Clarifier la notion de jugement : l'épaisseur du jugement contre l'irruption de la décision

Un premier geste consiste à clarifier ce que l'on entend par jugement, à conceptualiser la notion, de manière à échapper à sa représentation réductrice. Cela suppose d'éclairer ce que le jugement juridictionnel véhicule avec lui.

3.1.1. La double légitimité du jugement en justice

Au Moyen Âge, la justice en tant qu'institution est née de la récupération par les hommes d'une « fonction transcendante » longtemps réservée aux dieux, détenteurs du

pouvoir de procéder au jugement dernier (Jacob, 1995 : 12). Symboliquement, le serment a pour fonction de transférer au juge la capacité de dire le vrai et le juste, qui ne sont plus dans les mains de Dieu. Cela est à rapprocher du mouvement, plus tardif, d'autonomisation des sociétés à l'égard de toute transcendance. Les politiques publiques sont nées de la volonté d'affirmation d'une société devenue, au XIX^e siècle, autoréférentielle et autoréflexive (Muller, 2015). Les décisions peuvent reposer sur différentes rationalités (instrumentale, par valeur) ; mais le modèle rationnel les conceptualise comme choix entre plusieurs options, pris au terme d'une délibération. En ce sens, elles représentent la rationalité scientifique et humaine portée à son paroxysme, dans un contexte où les sociétés sont responsables de leur administration et de leur régulation, en un mot de leur destin, c'est-à-dire qu'elles refusent une transcendance (la fatalité, Dieu, etc.). Le régime de la démocratie trouve son assise dans la société elle-même — via la souveraineté populaire — et non à l'extérieur de celle-ci. Les politiques publiques, théoriquement fondées sur des décisions rationnelles, préparées et délibérées, incarnent cette prise de pouvoir de la société sur elle-même.

Dans le mouvement d'affranchissement de la justice (à l'égard d'une puissance supérieure), l'établissement du vrai est toutefois fragilisé par le fait même qu'il n'est plus révélé par cette transcendance, mais qu'il est délégué à des humains. Dans ce contexte, le droit, la loi et les procédures confèrent une « objectivité nouvelle » (Jacob, 1995 : 12) à la fonction de juger. De ce fait, l'existence du cadre procédural et son respect sont une part importante de ce qui fonde l'autorité et la légitimité du jugement. Si ce dernier est accepté, c'est parce que son prononcé intervient à l'issue d'une procédure dont les formes et conditions sont reconnues comme reposant sur des principes acceptés et offrant des garanties juridiques. « Le juge a donc deux façons de “dire le droit” » (Bernabé, 2016 : 65). D'une part, « celle qui nous vient spontanément à l'esprit est celle qui est contenue dans le jugement, l'arrêt, la déclaration de vérité judiciaire » (*ibid.*). L'énoncé public du verdict, en présence des parties, propre au pénal, matérialise le pouvoir décisionnel détenu par l'instance de jugement. Cet acte de langage, accompli dans un contexte institutionnel donné, celui du tribunal, incarne la capacité à faire rupture et à transformer le réel, que la conception décisionniste associe généralement au jugement. Dans tous les cas, en matière civile comme pénale, il y a un avant et un après le rendu de la décision, laquelle rompt le cours ordinaire des choses et signe la force du jugement et de celles et ceux qui en sont les acteurs. L'écrit du jugement ou de l'arrêt comporte la même force, appuyée sur son caractère exécutoire. C'est un « acte d'État », une décision qui, sitôt prise, est « dotée d'une autorité qui nous oblige » (Weller, 2018 : 4^e de couverture).

Mais d'autre part, « le jugement (*judicatum*) ne saurait exister sans son procès de fabrication (*judicium*) » (Bernabé, 2016 : 65). Dire le droit, c'est exprimer un jugement (*ius dicare*) qui n'a de sens que parce qu'il repose en amont sur une phase où le travail de préparation est accompli, et se coule dans le respect de formes procédurales ayant valeur juridique (*ius dicere*). Le jugement est certes un énoncé performatif, selon la théorie des actes de langages de John Austin. Mais, pour être pleinement performatif, le prononcé de la décision doit être précédé d'une phase préparatoire, qui lui confère une part de sa valeur. La puissance de juger et la force du jugement sont certes liées à la légitimité du juge lui-même,

c'est-à-dire à la reconnaissance de sa compétence, aux conditions de sa nomination et au lien qu'il entretient avec l'autorité politique qui l'a fait juge. Elles le sont également au processus de jugement incarné par la notion de *procès*. Le terme ne désigne pas seulement l'audience publique, mais plus largement le dispositif à travers lequel le dossier est patiemment constitué et l'affaire instruite selon le respect de « bonnes formes »²².

La distinction entre ces deux dimensions du dire droit fonde l'idée selon laquelle l'acte de juger ne consiste pas seulement à prendre une décision. Certes, le prononcé du verdict incarne la puissance décisionnelle, faisant événement, créant une rupture et entraînant des conséquences plus ou moins immédiates. Mais en arrière-plan de ce moment marquant et hautement symbolique, bien des éléments ont été concaténés à travers le fil procédural suivi lors du « lent et patient labeur » (Weller, 2018 : 46) de sa préparation. Le jugement gagne en épaisseur : il est replacé dans une sacralité, restée très présente via les rituels de justice, et réintégré dans une temporalité plus longue, celle de sa fabrication pas à pas.

3.1.2. Deux manières d'étudier les jugements

Cette double dimension du jugement fait écho aux deux conceptions, distinguées par Cyrus Tata (1997) dans un article fondateur portant sur les *sentencing studies*. La première est qualifiée d'approche « juridico-analytique » (*legal-analytical*, *ibid.* : 395). Elle nourrit une tradition de recherche positiviste qui s'intéresse aux produits de la décision des juges et juridictions. Le jugement y est envisagé comme un produit fini, considéré *ex post* comme un ensemble d'opérations rationnelles, logiques. La seconde met l'accent sur le jugement comme processus social en construction, toujours relativement ouvert, et dans lequel les indéterminations et les bifurcations apparaissent au grand jour.

À l'occasion d'un autre état de l'art, Françoise Van Hamme et Kristel Beyens (2007) ont prolongé cette analyse pour dissocier le jugement en tant que « résultat de différents facteurs d'influence dont l'impact est mesuré de façon quantitative » ; de celui qui relève d'« une pratique contextualisée à étudier comme un tout » (*ibid.* : 200 ; voir également Faget, 2008). La première approche se focalise sur la décision juridictionnelle en tant que « produit fini » à partir duquel sont menées des investigations qui recherchent les caractéristiques statistiques de facteurs identifiés au préalable. À partir de l'agrégation d'un grand nombre d'arrêts ou de jugements sont repérés les critères qui influencent la prise de décision. Ils peuvent être dégagés de façon déductive ou hypothético-déductive ; et concerner la personnalité d'un juge, la gravité des faits, le profil d'un accusé ou bien encore les conditions dans lesquelles l'audience intervient. L'enjeu est de construire des corrélations et des modèles statistiques prédictifs afin de nourrir des anticipations sur les décisions futures.

À l'évidence, les dispositifs algorithmiques de traitement et d'analyse des décisions de justice reposent sur cette première conception (1997 : 395). En effet, leurs bases de données regroupent de grandes quantités de jugements et arrêts, dont les données sont triées, classées, croisées, pour réaliser différents types de traitements statistiques et algorithmiques.

²² Sur les bonnes formes du procès face aux mobilisations politiques, voir (Codaccioni, Puccio-Den et Roussel, 2015).

Ce sont bien des décisions, en tant que produits finis écrits, qui sont « moulinées » pour faire apparaître variables, corrélations, tendances et modèles explicatifs. Il en va de même pour les recherches sociologiques ou économétriques qui mettent en évidence les barèmes dits constatés, c'est-à-dire ceux qui sont repérables, après coup, dans les pratiques des juridictions et magistrats (cf. chap. 1), au pénal comme au civil.

La seconde approche s'intéresse à la décision en tant que processus social complexe et que pratique sociale qui intervient dans un contexte social et politique spécifique. Centrée sur le jugement en train de se faire (Dupret, 2006), elle suppose des méthodes qualitatives : entretiens, observations, interprétation de données. L'enjeu est de saisir la manière dont se forment les jugements, dans une approche compréhensive ouverte aux mécanismes et logiques sous-jacentes, aux incertitudes, « versions intermédiaires » (Weller, 2018 : 60) et « retournements » (*ibid.* : 58) qui en jalonnent le processus de fabrication. J.-M. Weller restitue « le cheminement, réglé à l'avance par la procédure pénale, mais néanmoins imprévisible quant aux multiples récits possibles qu[e le juge de proximité] va pouvoir imaginer et qu'il lui faut à la fois stabiliser et purifier » (*ibid.* : 58). Du point de vue du juge, la confection d'« un récit qui se tient » relatif à « ce qui s'est passé » suppose d'envisager une pluralité de scénarii qui sont progressivement évalués au cours d'un va-et-vient entre les pièces du dossier. Le travail narratif passe par différentes étapes de progression et d'effacement (*ibid.* : 63) dans un jeu entre les éléments factuels et les textes juridiques applicables.

Si la plupart des dispositifs algorithmiques de traitement et d'analyse des décisions de justice et les discours de leurs concepteurs reposent sur une conception « juridico-analytique » du jugement, *il me semble indispensable, pour échapper aux pièges analytiques mentionnés, de considérer les jugements juridictionnels en tant qu'ils sont à la fois des produits finis et des processus en construction*. Cela suppose de faire parler et d'observer les juges au travail, afin de capter le processus de formation des jugements juridictionnels. Le travail ethnographique très documenté de J.-M. Weller constitue à cet égard une référence.

3.3. Intégrer les acquis de la sociologie pragmatique : jugements ordinaires et jugements de droit

La sociologie pragmatique analyse les processus sociaux de jugement en identifiant les épreuves lors desquelles la grandeur des êtres est réévaluée de même que la pluralité des principes de justice à l'aune desquels cette évaluation est conduite.

3.2.1. Spécifier les jugements de droit

Dans cette acception, « le jugement constitue, en toute généralité, une activité critique d'une situation, d'une relation ou d'une action ressentie comme injuste, c'est-à-dire décevante au regard d'une convention préalable, d'une attente tenue pour légitime » (Delpuech, Dumoulin, Galembert, 2014 : 76). Dans cette perspective, Laurent Thévenot (1992 : 1279-1299) a spécifié et typifié la catégorie des « jugements de droit » en les resituant par rapport aux jugements ordinaires. Ordinaires ou de droit, « les jugements sont des actions » au sens où ils recouvrent « différentes façons selon lesquelles les acteurs ordinaires saisissent les actions des autres à partir de repères [...] qui servent de référence à l'ajustement

des conduites » (*ibid.* : 1282-1283). Ils ont en commun de résulter de multiples opérations : qualification, administration de la preuve, clôture, révision.

Les dispositifs judiciaires de règlement des litiges se distinguent toutefois des épreuves de jugement ordinaire en ce qu'ils mettent en jeu des principes de justice davantage procéduraux que substantiels. Le caractère hautement formel et procéduralisé du droit leur confère une forte spécificité. Concernant le juge, le type d'action qui convient est nécessairement une décision sur le cas puisque l'article 4 du Code civil l'oblige à juger, sauf à commettre un déni de justice passible de poursuites.

3.2.2. *Résultat du jugement et jugement en train de se faire*

Le jugement juridictionnel apparaît dès lors comme le résultat d'un processus décisionnel qui comporte une dimension narrative ; matérialisé par un écrit (le texte de l'arrêt ou du jugement du tribunal), lequel contient l'énoncé ayant force de chose jugée (le dispositif selon la terminologie juridique) et sa justification (les attendus). Le fait de rapprocher le jugement juridictionnel d'une définition plus large du jugement comme « ce qui précède l'action qui convient » ouvre le champ et laisse la place pour prendre en compte le processus qui se réalise *au cours de l'activité pratique de travail* des juges. Cela permet de pointer « un genre professionnel propre à toute la magistrature » (*ibid.* : 66), le « devoir d'hésitation », bien décrit par J.-M. Weller (2011) à propos des juges de proximité. Il conviendra d'observer comment le juge mobilise (ou pas) les dispositifs algorithmiques disponibles.

Le parallèle avec le cas de l'expertise judiciaire et de la preuve est suggestif. Bien des travaux analysent les preuves comme un déjà là. Existantes d'emblée, il suffirait de les apporter devant le prétoire pour les faire valoir pendant le procès. *A contrario*, dans la tradition des travaux de Lynch et McNally (2005), Jasanoff (1995) et Dupret (2001, 2005, 2006) notamment, V.-A. Chappe, R. Juston-Morival et O. Leclerc (2022) défendent « une analyse pragmatique de l'activité probatoire », avançant « l'hypothèse que [l]a force [de la preuve] dépend du processus qui la détermine, et que ce processus peut se laisser décrire comme un "travail de la preuve" » (*ibid.* : 7). Ils entendent ici « une activité de production distribuée entre des acteurs divers interagissant entre eux, et dont les équipements — cognitifs ou matériels — sont orientés vers l'identification, la sélection, le prélèvement, la mise en forme et en relation d'éléments empiriques hétérogènes à des fins de conviction judiciaire » (*ibid.*).

De manière comparable, nous entendons le jugement comme pouvant désigner le résultat d'une activité judiciaire se matérialisant par l'écrit désigné par le terme d'ordonnance, de jugement ou d'arrêt, mais aussi comme le processus d'accomplissement de cette activité. Il s'agit d'une activité de production distribuée entre une pluralité d'acteurs en interaction et dont les équipements cognitifs et matériels sont orientés vers l'identification, la sélection, la mise en forme et en relation d'éléments empiriques et théoriques hétérogènes à des fins de clôture d'un conflit. Le terme de *decision-making process* peut recouvrir cette activité de production du jugement, dès lors qu'il est appréhendé à travers ses opérations pratiques et sa matérialité.

Les transactions susceptibles d'intervenir ne sont pas des accords qui s'inscrivent en dehors du droit, mais bien des règlements conventionnels, garantis par des contrats de droit

privé, et qui se développent en interaction avec le système juridique étatique (Serverin, Lascoumes et Lambert, 1987). La question qui se pose ne porte donc pas seulement sur ce que les algorithmes font aux jugements juridictionnels, mais plus largement sur ce qu'ils font aux processus de règlements des litiges. *A fortiori* si l'on considère que la diffusion publique de ces outils est susceptible de mettre à l'épreuve les modes de négociation et les positions d'autorité des différentes parties, dont les justiciables. Quels en seront les effets sur la négociation et les transactions produites ?

3.4. S'intéresser au travail de *decision-making* et à sa matérialité

L'attention au processus de *decision-making* se décline dans une double direction : la prise en compte de la pluralité des équipements de la décision, et donc l'agentivité distribuée entre les humains et les non-humains (objets, machines, dispositifs de calculs, etc.) d'une part, et celle de la distribution plurielle des rôles et activités concrètes au sein de la diversité d'acteurs qui sont en interaction avec le juge d'autre part. Par le terme d'agentivité distribuée (*distributed agency*), nous désignons la répartition de qui fait quoi entre les différents acteurs et l'ensemble des ressources dont ils disposent²³.

3.3.1. Les équipements pluriels de l'activité de production des jugements

Dans les théories dites de l'activité (Licoppe, 2008), les activités humaines sont dites équipées dans la mesure où elles sont considérées à travers la diversité des appuis immatériels nécessaires à leur accomplissement pratique : normes professionnelles, compétences, savoir-faire, mais aussi intuition, astuces ou bien encore « art de la prise » (Bessy et Chateauraynaud, 2014[1995]), c'est-à-dire capacités à entrer dans les plis des objets et des situations, via des ressources cognitives et sensorielles. La perspective de la cognition distribuée (Hutchins, 1995 ; Conein, 2004) nous a en effet appris que les fonctions et capacités cognitives sont « une propriété caractéristique de collectifs hétérogènes, faits de personnes et d'artefacts » (Licoppe, 2008 : 3). Les espaces, bâtiments, bureaux cadrent physiquement en même temps qu'ils encadrent l'exercice concret de l'activité. Papiers, dossiers, ordinateurs, logiciels fournissent un outillage dont la maîtrise participe de la technicité des acteurs et qui médie les activités pratiques d'écriture, de calcul, de jugement, etc. lesquelles sont distribuées entre une diversité d'actants, humains et non humains.

Considérer le jugement à travers l'activité pratique de juger et étudier la place qu'y occupent les dispositifs algorithmiques de traitement et d'analyse des décisions de justice consiste d'abord à prendre en compte de manière réaliste et détaillée le rôle des équipements actuels de cette activité : bases de données juridiques existantes proposées par les grands groupes d'édition (et qui donnent accès à des commentaires de doctrines et d'arrêts, des décisions juridictionnelles, etc.), barèmes, tables, référentiels, nomenclatures, abaques, tableurs, formulaires, guides, listes de discussion, etc. (Licoppe et Dumoulin, 2019 ; Sayn, 2014 ; Lemaire, 2017). Les faire apparaître, c'est commencer par reconstituer la réalité de

²³ Les débats sur cette notion empruntent à différentes disciplines et de nombreux auteurs y ont pris part. Pour des éléments de synthèse, voir (Engeström, 2006). Nous nous appuyons ici sur l'idée que la capacité d'agir ne repose pas seulement sur la volonté et l'intentionnalité d'individus pensants mais qu'elle réside aussi dans les fonctionnalités et plis des objets techniques, voir notamment (Latour, 1994).

cette activité et remettre en cause l'idée de la 'justice prédictive' comme révolution qui surviendrait dans un désert technologique.

La lecture des articles de presse collectés sur la 'justice prédictive' laisse poindre l'absence de la mention de ces équipements pourtant bien réels. Si les équipements actuels sont quasi invisibles dans la controverse sur l'intérêt de dispositifs algorithmiques pour l'activité judiciaire, ce n'est pas parce qu'ils sont inexistant, absents des pratiques concrètes, mais bien plutôt parce qu'ils sont ignorés, considérés comme insignifiants ou obsolètes. On peut interpréter cela comme l'envers ou le complément du discours de survalorisation de la nouveauté et des performances des dispositifs algorithmiques, auréolés d'atouts et possibilités révolutionnaires. Cela est vrai tant du côté des concepteurs et des opérateurs qui proposent ces outils, des professionnels du droit qui en débattent que des espaces comme Twitter où la question de la 'justice prédictive' est portée.

Leur propos noircit parfois un peu le tableau pour exhausser l'ampleur du défi pour la justice de même que les bienfaits de l'intelligence artificielle. Ainsi de cet ouvrage²⁴ publié par un avocat belge spécialisé en droit des technologies et animateur d'un incubateur d'information juridique soutenu par un éditeur juridique en ligne (KnopsPublishing)²⁵, qui part « du constat que la technologie a envahi tous les grands secteurs de l'activité humaine, sauf la justice. La finance, la santé, le sport et même l'art sont impactés par les avancées technologiques portées par la quatrième révolution industrielle. La justice, quant à elle, s'érige, telle une citadelle infranchissable à la technologie. » (Van den Branden, 2019 : 2). Dessiner la justice en « citadelle » assiégée, l'assimiler à un milieu imperméable et profondément rétif aux technologies permet de grandir la cause de l'intelligence artificielle, seule planche de salut pour la justice. Ce type d'affirmation à l'emporte-pièce fait l'impasse sur la pluralité des dispositifs sociotechniques qui équipent déjà la fabrique du jugement et plus largement l'activité juridictionnelle (en audience) et judiciaire (gestion de l'activité, des dossiers, échanges avec les avocats, etc.). Certes, les greffiers et magistrats utilisent des logiciels qui ne sont pas très performants, voire sont parfois assez obsolètes. Certes, les infrastructures de connexion et d'échange d'information sont limitées et communiquent mal entre elles, surtout pour les applicatifs, créant d'importants effets de segmentation, par exemple d'une juridiction à une autre. Pour autant, les dispositifs algorithmiques n'interviennent pas dans un monde judiciaire imperméable aux technologies numériques ou vierge de toute implantation préalable. Qu'il s'agisse de bases de données de jurisprudence comme Juris-data, d'applicatifs comme Cassiopée ou APPI (application des peines, probation, insertion utilisée par les SPIP) (Larminat, 2014), du casier judiciaire national, de l'utilisation de la visioconférence (Biolley, 2013 pour le cas belge ; Dumoulin et Licoppe, 2013), de la dématérialisation de certaines procédures entre le barreau et certaines juridictions, le quotidien de l'activité judiciaire est émaillé d'interactions technologiquement médiées, dont certaines sont liées à la fabrique du jugement.

²⁴ Voir mon compte rendu dans la revue *Droit et société* : <https://ds.hypotheses.org/>

²⁵ Incubateur Avocatsnet.be : <http://www.avocatsnet.be/a-propos-de-nous/>, consulté le 10 mai 2019.

L'attention à la matérialité et à l'agentivité est développée par J.-M. Weller, à partir d'une ethnographie des opérations et supports matériels de l'activité bureaucratique judiciaire et plus largement étatique, au croisement entre sociologie pragmatique, sociologie du travail et sociologie du droit. Cette approche apparaît heuristique tant elle désacralise l'activité décisionnelle et la replace dans un continuum d'activités ordinaires, mais déterminantes dans la fabrique de documents publics comme les jugements juridictionnels, dotés d'une force performative et symbolique spécifique, et à travers lesquels l'État agit. Par la place qu'elle donne à la matérialité, aux technologies et supports concrets (trombone, pochette carton, logiciels, bureaux, etc.) de l'activité, cette approche entre dans le détail des activités de constitution d'une documentation, de qualification des situations, de production d'un récit de ce qui s'est passé et d'un écrit qui met en ordre et en forme une articulation possible des faits, du droit et d'une histoire.

3.3.2. Les acteurs pluriels de l'activité de production des jugements

Le cas observé par J.M. Weller a ceci de caractéristique qu'il met en scène des juges non professionnels, qui maîtrisent quasiment seuls « une très grande part du processus du [sic] traitement des affaires » (2018 : 66). Cela contraste avec les autres administrations qu'il étudie (Agriculture, Sécurité sociale), marquées par un éclatement plus grand du pouvoir décisionnel. Or, cette configuration, courante dans les organisations judiciaires, n'est pas la seule possible. En réalité, les configurations sont variables en fonction des types de contentieux, des cultures locales ; mais dans tous les cas, les magistrats interagissent avec des greffiers en chef, greffiers et agents de greffes, systématiquement présents dans l'organisation judiciaire. Du point de vue de la sociologie des organisations comme de la sociologie du travail et de l'activité, la place de ces acteurs d'arrière-plan n'est pas négligeable, et ce, même si le discours et l'image du juge solitaire persistent, et correspondent à certains aspects bien décrits de l'activité des juges²⁶.

Le rôle des greffiers est sous-étudié par la sociologie du droit et de la justice (Darty, Froissart et Ménard, 1997 ; Bossis, 2003), variable selon les contextes²⁷ et souvent peu valorisé. Cette invisibilité peut, au moins en partie, être interprétée comme le fait qu'ils sont perçus comme effectuant « le sale boulot » des magistrats (Vigour, Cappellina, Dumoulin *et al.*, 2022 : 115-116).

Les rares analyses de la profession de greffier et des conditions concrètes et matérielles de production des actes judiciaires mettent pourtant en évidence les multiples sous-divisions du travail entre juges, greffières, agents de greffes pour la rédaction d'actes standardisés que sont les jugements (Nouiri-Mangold, 2012) ; les relations de travail sont complexifiées par une pression accrue notamment par le manque de moyens (Girault, 2018). Si « les magistrats apparaissent portés par une rationalité instrumentale qui les conduit à rechercher, avant tout, les informations nécessaires à leur prise de décision », « les greffières

²⁶ Sur les juges des enfants, voir Bastard et Mouhanna, 2010 : 79-82 ; sur les juges de proximité, voir Weller, 2018 : chap. 1.

²⁷ Pour la justice des enfants voir Bastard et Mouhanna, 2010 : 74-78 ; pour la justice familiale, voir Collectif Onze, 2013 : 35-40 et 149-154.

sont porteuses d'une rationalité formelle et matérielle, indexée sur leur fonction d'authentification, mais aussi d'un rôle d'accueil les rapprochant des agents de guichet » (Collectif Onze, 2013 : 35-40 et 149-154). Reprenant une notion foucauldienne utilisée pour décrire l'articulation entre droit et médecine, Tonya Tartour a mis en évidence le « travail de *couture institutionnelle*, produit par les travailleuses de l'écrit que sont, au tribunal, les greffières », dans la mise en œuvre du contrôle judiciaire des hospitalisations sans consentement (2020, chap. 3). En lien avec leurs homologues, les secrétaires médicales des hôpitaux psychiatriques, elles prennent en charge le nécessaire travail d'« alignement judiciaire » (Weller, 2018) et s'engagent « dans une translation des intérêts de chacun des deux mondes » (Tartour, 2021). Ce travail est concrètement opéré autour d'objets écrits comme les certificats médicaux et par l'interfaçage avec les juges des libertés et de la détention dans les décisions médicales d'hospitalisation sans consentement des patients.

De plus, les écrits moins standardisés tels que les notes d'audience font apparaître le rôle important que les greffiers et greffière jouent parfois. Chacun les prend à sa manière, sur la forme comme sur le fond — du moment qu'elle intègre certaines informations qui doivent systématiquement y figurer. « Les audiences s'apparentent à des débats au cours desquels les greffières doivent saisir bien plus que de simples informations factuelles. Parce que leurs notes sont destinées à appuyer et faciliter la décision du juge, elles peuvent également être construites pour éclairer "l'ambiance" d'une affaire ou un "climat" familial. [...] La greffière effectue une sélection des éléments qu'elle juge pertinents pour le dossier » (Nouiri-Mangold, 2012). Certains juges s'appuient exclusivement sur ces notes une fois l'audience terminée, lorsqu'il faut prendre la décision.

Les greffières sont « chargées de documenter les dossiers en rassemblant les diverses pièces, de prendre des notes durant les audiences et de rédiger les actes judiciaires tels que les jugements, ordonnances de non-conciliation, etc. » (Nouiri-Mangold, 2012). Ce qui est valable pour les greffiers et greffières l'est aussi pour les assistants et assistantes de justice : dans certaines juridictions, ils font fonction de greffier à l'audience, préparent certains jugements, par de la recherche documentaire, la rédaction d'argumentaire et/ou la préparation d'un premier jet du jugement. Le métier et les activités du greffe ont été considérablement transformés par l'informatisation des juridictions. Les écrits judiciaires ne sont plus le résultat d'une écriture manuscrite du juge, ensuite dactylographiée par sa greffière. La division du travail s'est déplacée, de même que les équipements de l'activité : ainsi de la création de supports comme les factums²⁸ à l'articulation entre le papier et l'informatique. De ce fait, il est vraisemblable que les greffiers et assistants de justice comptent parmi les utilisateurs réels des dispositifs algorithmiques de traitement et d'analyse des décisions de justice alors même qu'ils sont complètement absents de la controverse autour de la 'justice prédictive'. Il serait donc judicieux d'imaginer qu'en réalité, l'utilisation des dispositifs algorithmiques de traitement et d'analyse des décisions de justice pourrait être

²⁸ Ce document est « conçu pour faire le lien entre la décision du juge et sa formalisation informatique par la greffière. Les cases à remplir sur le formulaire correspondent à la fois aux catégories juridiques mobilisées par le magistrat et à la trame du logiciel. Les réponses possibles sont prédéfinies, ce qui a pour effet de standardiser les réponses du juge. » (Nouiri-Mangold, 2012)

expérimentée par les juges, mais aussi par les greffiers ou assistants de justice. Différentes propositions de réforme ont d'ailleurs été avancées pour créer une équipe autour du juge²⁹ ; y compris dernièrement dans le rapport Sauvé (2022) tiré des États généraux sur la justice.

3.3.3. *Les usages pluriels des dispositifs algorithmiques dans différentes configurations*

Que font exactement les algorithmes de traitement aux pratiques des utilisateurs ? Répondre à cette question suppose d'articuler une analyse de ces infrastructures de calcul, de la manière dont elles ont été conçues avec une attention fine à la manière dont elles sont effectivement mobilisées dans le cours des activités. Se poser la question des usages, c'est considérer que même si les dispositifs sociotechniques comportent des scripts incorporés et formulés lors de leur phase de conception (Callon, Latour et Akrich, 2006), il reste toujours une place pour des formes de négociation. L'introduction des dispositifs de 'justice prédictive' transforme potentiellement les pratiques des acteurs du droit, que ce soit la façon dont ils raisonnent sur des cas, décident et motivent leurs décisions, dans des situations concrètes et des écologies riches de ressources très variées. L'enjeu est donc de comprendre la décision judiciaire comme une activité équipée, réalisée au sein d'un agencement ou d'un dispositif combinant des ressources à la fois multiples et hétérogènes, et susceptible de mettre à l'épreuve — au sens que la sociologie pragmatique donne à cette notion (Barthe *et al.*, 2013) — leurs manières de faire. Il importe d'aller scruter leurs pratiques pour saisir la manière dont ils s'approprient les dispositifs en leur donnant un sens pour leur activité en cours et en les insérant (plus ou moins, de différentes façons) dans des pratiques déjà structurées par des manières de raisonner, de penser et de faire, des routines, des cheminements comportant d'autres ressources, etc.

Quel que soit le discours des concepteurs sur le pouvoir prédictif des algorithmes (Benbouzid et Cardon, 2018 ; Cardon, 2019), comme le montrent les premiers retours d'usages (Licoppe et Dumoulin, 2019), tout peut être remis en jeu par les raisonnements, les pratiques et manières de faire, selon qu'à un extrême, les acteurs délèguent les décisions au système, ou au contraire choisissent d'ignorer les apports de celui-ci.

Dans nos précédentes recherches sur l'expertise judiciaire, nous avons mis en évidence un positionnement stratégique des juges dans l'utilisation de la procédure d'expertise d'une part et des rapports d'expertise d'autre part — via un « art de la pioche », c'est-à-dire une capacité à dissocier les différents éléments du rapport d'expertise pour mener à bien une argumentation. On peut imaginer que certaines fonctionnalités des dispositifs algorithmiques sont appropriées, tandis que d'autres seront laissées de côté. De même, suivant les dispositifs considérés ([PréviCompute], [Anajur], etc.), ce ne sont pas exactement les mêmes possibilités qui sont offertes. Réaliser un travail de terrain de type ethnographique ancré au ras des discours, pratiques et usages concrets (Weller, 2018) est donc indispensable pour saisir des configurations spécifiques, associant plusieurs éléments caractéristiques :

²⁹ Par exemple les rapports (Guinchard, 2008), (Delmas-Goyon, 2013). Voir (Guinchard, Varinard et Debard, 2021).

- les relations entre acteurs — en particulier juges-greffiers — dans l'activité de fabrication du jugement, pratiquée de façon plus ou moins collaborative ;
- la place des équipements déjà mobilisés dans le traitement du contentieux ;
- la nature des possibilités ouvertes par les différents dispositifs algorithmiques utilisés ;
- la nature des cas examinés, plus ou moins routiniers, plus ou moins standardisables.

Penser la fabrication du jugement comme une activité équipée permet de sortir d'une perspective individualisante sur ces décisions qui les cantonnerait à un triangle composé d'un acteur, d'un cas et d'un outil. L'écologie de la décision juridictionnelle incorpore des ressources multiples dans le temps durant lequel celle-ci se déploie, que ce soit en matière d'outillage (outre les algorithmes, de multiples textes et barèmes), de possibilité d'échanger avec des collègues (délibérés collégiaux, discussions informelles, listes de discussion, etc.), ou de pratiques sédimentées (liées à la profession, à la juridiction, au service, etc. en lien avec des politiques, normes et manières de faire locales). La question de l'agentivité se pose donc dans le cadre de ce dispositif propre à la décision judiciaire : les acteurs délèguent-ils ou pas l'activité décisionnelle au système, et dans quelle mesure ? Comprendre cette distribution de l'agentivité suppose de suivre pas à pas le processus de décision dans les différents milieux dans lesquels il s'échafaude successivement. Il suppose aussi l'attention aux spécificités des divers dispositifs algorithmiques étudiés, à ce qui fait leurs points communs et leurs différences, ce qu'ils permettent de faire (modèles informatiques et bases de décisions utilisées, interfaces utilisateurs) et à qui ils s'adressent (discours sur le produit, cibles, stratégies de développement, logiques de partenariat). En effet, les différentes solutions commercialisées ne reposent pas sur la même conception/articulation des agentivités respectives du dispositif et des acteurs qui l'utilisent. Il est essentiel de sortir d'une vision réductrice de l'intelligence artificielle comme la confrontation d'un professionnel et d'un dispositif algorithmique, à partir d'un travail de terrain approfondi, basé sur des entretiens en présence, enrichis par des démonstrations à l'ordinateur sur des cas concrets, et des observations au long cours. Si l'on ne peut s'affranchir des entretiens avec chacun des acteurs concernés (magistrats, mais aussi greffiers et assistants de justice), imaginer des entretiens en petits groupes (binôme, voire trinôme, en fonction de la réalité des collectifs de travail), pourrait s'avérer très heuristique.

Cela met en évidence la distribution complexe des tâches, pouvoirs et responsabilité entre les humains et les dispositifs. Elle peut être recomposée par les usages des algorithmes, davantage partagée et même distribuée plus largement auprès de nouveaux acteurs qui interviennent dans la conception des algorithmes. Le parallèle avec les évolutions de la médecine, le développement de protocoles d'*evidence-based medicine* à destination des professionnels de santé et la reconfiguration du pouvoir décisionnel des médecins qui en découle est intéressant ici. Il s'agit de protocoles relatifs à des pathologies ou à des problèmes de santé qui visent « à promouvoir des décisions fondées sur des données de la science agrégées plutôt que sur l'expérience individuelle des médecins » (Brunier, Castel, Chessel, 2022 : 95). Ces protocoles sont élaborés au sein d'une grande diversité d'organisations (agences sanitaires, sociétés savantes, réseaux d'experts, établissements de santé, etc.) ;

parmi lesquels des acteurs privés, comme les laboratoires pharmaceutiques. Ces derniers pilotent les essais thérapeutiques à partir desquels ces protocoles sont réalisés. La décision médicale, contrainte par ces dispositifs, apparaît plus diffractée, « beaucoup plus collective et distribuée », « l'expertise individuelle des médecins n'étant plus seule, ni même dominante. » (*ibid.*96). Toutefois, les études sociologiques disponibles montrent que cette « protocolisation » ne conduit pas à « une 'déprofessionnalisation' de la médecine, c'est-à-dire à une perte de pouvoir et d'influence du personnel médical » (*ibid.*). Les médecins développent des stratégies individuelles et collectives pour participer aux processus de fabrication des protocoles dans les organisations impliquées ; jouer un rôle clef dans les espaces de travail pertinents et « garder la main ». Ils disposent également d'une grande marge de manœuvre pour utiliser ou pas ces protocoles, en fonction des situations et rapports de force qu'ils entretiennent avec les différentes catégories d'acteurs de la santé (autres métiers patients, etc.).

Il convient donc de saisir les effets des dispositifs algorithmiques comme un déplacement de manières de faire déjà équipées, plutôt que les poser d'emblée en rupture radicale. Cela nous amènera à réfléchir sur le rapport des acteurs aux dispositifs, à la redistribution des formes d'agentivité ainsi qu'aux éventuelles transformations du rapport que les acteurs entretiennent aux cas et en particulier aux précédents, via l'utilisation de dispositifs algorithmiques de traitement et d'analyse des décisions de justice.

C'est la raison pour laquelle il importe de ne pas se pencher seulement sur les décisions juridictionnelles des juges, mais bien de déployer une analyse plus vaste de la chaîne de traitement des litiges, en incorporant aussi l'amont, c'est-à-dire les décisions qui émaillent la stratégie juridique des avocats — entamer un contentieux ou rechercher un règlement amiable du litige. Ces décisions sont elles-mêmes susceptibles d'être équipées par les dispositifs algorithmiques qui s'adressent en premier lieu aux cabinets d'avocats, directions juridiques d'entreprises et compagnies d'assurance, dans l'optique de déjudiciariser certains litiges. Si tel n'était pas l'objet du présent mémoire, le projet Just-IA que nous menons avec Christian Licoppe porte précisément sur l'ensemble de cet éventail de décisions juridiques et juridictionnelles afin de cerner dans quelle mesure les algorithmes sont utilisés pour renouveler un calcul stratégique — à savoir confronter ce que ces systèmes laissent prévoir comme gain possible au tribunal, par opposition à ce qui serait obtenu en transigeant directement, compte tenu des coûts réels et perçus d'une procédure.

Conclusion :

Dans cette « réflexion en suspension », nous avons placé le projecteur sur ce que la controverse sur la 'justice prédictive' nous apprend sur la manière dont les jugements en justice sont couramment appréhendés. À partir de données relatives au discours des concepteurs d'algorithmes, à celui des médias et aux débats sur Twitter, nous avons posé quelques jalons empiriques, appelant de plus amples investigations. En lien avec nos précédents travaux sur l'expertise judiciaire, nous avons relevé que prévaut une

représentation sociale du jugement comme décision du juge, représentation que nous qualifions de décisionniste. La description de ce que permettent les dispositifs algorithmiques est adossée à cette conception, au sein de laquelle le jugement est associé principalement à une décision tranchée qui réside dans les mains du juge. Cette conception est polarisée sur l'acte décisionnel, vu surtout sous l'angle de l'adjudication et comme faisant rupture, sur le juge, perçu comme agissant seul et maîtrisant pleinement la décision et sur le caractère accessoire et peu impactant des dispositifs algorithmiques de traitement et d'analyse des décisions de justice, dont le caractère purement instrumental est constamment affirmé.

Nous avons établi, à partir des connaissances sociologiques disponibles, que cette représentation est erronée, inexacte. Reposant sur une mythification du rôle du juge, elle est en décalage avec ce que nous savons du fonctionnement concret de la justice et de la fabrique du jugement. Face à cette représentation sociale dominante, nous avons proposé une autre manière de penser les jugements en justice : très procéduralisés, ils entretiennent certains points communs avec les jugements ordinaires ; ils sont le résultat d'une activité de production, relevant d'un processus épais, c'est-à-dire non linéaire, impliquant une pluralité d'acteurs — les juges, mais aussi les greffiers, les avocats, etc. — et une diversité d'équipements — pas seulement numériques — et reposant sur une vision distribuée de l'agentivité entre acteurs humains et dispositifs sociotechniques. Nous avons également mis l'accent sur la capacité de négociation des acteurs, dans le fait de recourir ou pas aux interactions avec les dispositifs algorithmiques et l'importance de saisir les usages concrets et non seulement possibles, imaginés ou projetés. Enfin, si les juges sont des acteurs de premier plan du jugement, ils ne sont pas seuls dans la chaîne de règlement des litiges.

Tout cela suggère que pour saisir les effets des dispositifs algorithmiques sur la justice, il faut développer un cadre d'analyse qui intègre, outre le jugement comme produit fini, le processus dense au cours duquel il se forme. Le jugement comme produit fini prend ici deux formes concrètes : celle des jugements existants qui nourrissent les dispositifs algorithmiques (les jugements sources en quelque sorte) et celle des jugements qui seront produits au cours de ce processus (les jugements destinations). Prendre soin de distinguer ces différentes manifestations empiriques est essentiel pour lutter contre une conception décisionniste informe de même qu'il est indispensable de spécifier *les configurations variées de fabrique des jugements* (en fonction des domaines, de la manière dont les précédents juridiques sont conçus, etc.).

En dernier point, nous voudrions souligner que la conceptualisation du jugement proposée ici, comme produit fini et processus épais fait écho à « la finalité longue de l'acte de juger » évoquée par Paul Ricœur (1985 : 185). L'acte de juger contribue à « la paix sociale » (*ibid.* : 190), car il est un exercice de reconnaissance mutuelle, entre celui qui juge et ceux qui sont jugés. On me pardonnera la lapalissade et la répétition : le juge juge ; ainsi, il reconnaît l'autre comme sujet de droit, membre de la communauté politique. Mais il est aussi jugé. La façon dont il a accompli l'acte de jugement est évaluée, reconnue ou pas comme pertinente, fondée, acceptable, etc., en un mot juste par celui qui est jugé (Garapon, Allard, Gros, 2008 : 183). Les travaux sur la « justice procédurale » (Tyler, 1988, 2006 ; Sunshine et Tyler, 2003) montrent qu'au-delà du règlement du litige en lui-même, les citoyens et justiciables accordent

une grande importance à la manière dont les acteurs judiciaires exercent leur mandat. Ils critiquent vertement la justice comme déjà excessivement déshumanisée, distante — en recourant à des images comme « robots » et « machines », négativement connotées (Vigour, Cappellina, Dumoulin *et al.*, 2022). Ils expriment à l'inverse des attentes en matière de considération, de traitement équitable et d'attention à la singularité des individus et des situations (*ibid.*). Or, s'il est un angle mort de la controverse sur la 'justice prédictive', c'est bien celui des perceptions que les citoyens et justiciables pourraient développer de processus judiciaires équipés par des dispositifs algorithmiques de traitement et d'analyse des décisions de justice. Il ne fait pas de doute que sur ce versant aussi il y aurait matière à investiguer pour saisir dans quelle mesure la « forme de pédagogie politique [déployée] dans l'œuvre de jugement » (Rosanvallon, 2006 : 242) est affectée ou pas par la recomposition de l'exercice concret de l'activité de jugement.

Conclusion générale

À l'issue de ces réflexions en suspension, s'il est difficile de conclure, il est possible de rassembler les fils de ce travail et de dessiner ce qui a été établi et a émergé, chapitre après chapitre. De quoi la 'justice prédictive' est-elle le nom ? À cette question centrale, posée en ouverture, ce manuscrit d'HDR apporte plusieurs réponses et éclairages.

1. La 'justice prédictive' est une construction sémantique récente qui s'est durcie au cours des débats et usages. Les chapitres — en particulier 2 et 4 — ont documenté sa trajectoire. Au fil de celle-ci, cette forme lexicale a acquis un certain pouvoir de nomination dans le monde du droit et de la justice. La multiplication des publications sur le sujet, la chronologie de la montée en puissance du terme, très semblable dans les *tweets* et articles de presse étudiés, les résonances dans l'activité universitaire (colloques, etc.) atteste de sa consolidation.
2. La 'justice prédictive' est toutefois de plus en plus discutée quant à sa pertinence sur le fond. Elle a donné lieu à une controverse portant à la fois sur le choix de l'expression, son périmètre et son acception. Plusieurs points de débat ont été repérés dans le premier chapitre et creusés ensuite au long de la démonstration. J'en retiendrai deux ici.
 - Le caractère plus ou moins neuf, innovant et/ou disruptif des dispositifs englobés sous le terme de 'justice prédictive' : L'apprentissage machine, en particulier dans son versant le plus abouti, à savoir la technique de l'apprentissage profond et les réseaux de neurones, représente une nouveauté puisque ses résultats sont élaborés inductivement à partir de grandes masses de données. Leur faible explicabilité les rend plus questionnables pour des usages par les juridictions. Ils contrastent par rapport à d'autres outils d'aide à la décision juridique qui existaient antérieurement, mais n'avaient pas été beaucoup utilisés, tels les systèmes experts (qui sont construits, eux, d'après de règles préconçues et dont les résultats sont par conséquent explicables). En revanche, tous les dispositifs qui sont présentés comme étant de la 'justice prédictive' s'appuient sur de l'intelligence artificielle faible (comme le traitement en langage naturel) ou ne relèvent pas de l'apprentissage machine (mais de traitements statistiques). Ils n'ouvrent donc pas la « révolution » annoncée ; les arguments qui circulent dans la controverse sur la 'justice prédictive' ne les concernent pas toujours. Pour partie, les débats sur la 'justice prédictive' se développent de ce fait en décalage avec la réalité de ce que sont les dispositifs algorithmiques d'analyse et de traitement des décisions de justice « nouvelle génération », c'est-à-dire apparus autour de 2016, dans le sillage de l'ouverture des données publiques. Ceux-ci s'appuient majoritairement sur des solutions techniques qui existaient auparavant.
 - Le lien problématique de la 'justice prédictive' avec la justice actuarielle : Les discours et la controverse sur la 'justice prédictive' portent en grande partie sur l'enjeu de prédiction des décisions de justice suggérée par l'expression et que les termes alternatifs (jurimétrie, justice quantitative, algorithmique, simulative, etc.) s'efforcent

de tenir à distance. Or, la logique probabiliste et prédictive est à la fois point de rencontre et point de friction entre ‘justice prédictive’ et justice actuarielle — c’est-à-dire une philosophie pénale et des dispositifs pour lesquels le crime est un risque normal qu’il s’agit de gérer selon des logiques probabilistes. Dans les discours étudiés, suivant les acteurs et les contextes, elles sont rapprochées, assimilées ou au contraire dissociées. Cela brouille ou à l’inverse clarifie le rapport que la ‘justice prédictive’ entretient avec le civil et le pénal. Dans un cas, la ‘justice prédictive’ est un label critique qui désigne un ensemble de dispositifs dont la justice actuarielle fait partie. Dans un autre, elle est une formule qui qualifie des dispositifs algorithmiques d’aide à la décision, utilisant des corpus de jugements et arrêts civils qui, à ce stade, sont destinés à fournir une aide à la décision en matière civile uniquement.

3. La ‘justice prédictive’ est plus qu’un terme à la mode ou le résultat d’un effet de buzz, perceptible à travers les pics d’activité sur Twitter. Elle encapsule avec elle, et inséparablement, l’expérimentation d’un dispositif spécifique auprès de magistrats de deux cours d’appel françaises. Ce qui a été décrit comme l’échec de cette expérimentation est une épreuve qui marque un tournant dans la controverse et dans les pratiques, démontré dans le chapitre 4.
4. Cette expérimentation, et plus largement le mouvement de création de dispositifs algorithmiques de traitement et d’analyse de décisions de justice, ont été portés par une nouvelle configuration de réforme administrative qui s’est constituée à partir de la seconde moitié des années 2010, en lien avec l’adoption de la loi pour une République numérique. La ‘justice prédictive’ apparaît dans le chapitre 3 comme l’expression, dans le domaine du droit et de la justice, d’un répertoire de réforme de l’*open data* des données publiques, autour duquel s’affrontent deux coalitions de cause. La première porte une logique hétéronome, généraliste et transectorielle ; la seconde s’appuie sur le droit pour défendre la spécificité de la justice et son autonomie, notamment dans le rapport à l’État. La controverse sur la ‘justice prédictive’ s’inscrit dans cet environnement polarisé, mais sans qu’on puisse la résumer à cela. Le rôle des acteurs privés et des enjeux économiques y est remarquable.
5. Les discours sur la ‘justice prédictive’ imaginent la contribution possible des algorithmes à la fabrique des jugements. Ils esquivent toutefois la question de qui décide, comment, à partir de quelles données. Le mythe décisionniste, qui assimile le jugement à une décision qui tranche et qui réside dans les mains du seul juge, reste prégnant. Il imprègne des discours et représentations sociales en grande partie déconnectés de la manière dont fonctionnent les juridictions et dont l’activité de jugement est pratiquée. Le face-à-face du juge avec le droit, les faits et les dispositifs algorithmiques occupe toute la place ; l’environnement institutionnel et organisationnel, la collaboration avec les greffiers et greffières, assistants et assistantes de justice, les échanges avec les autres acteurs, l’appui sur différents supports et dispositifs sociotechniques existants sont éludés.
6. Le débat sur la ‘justice prédictive’ remet en question frontalement (et ouvre un cône de vue sur) les recompositions en cours autour de la fabrique des jugements juridictionnels :

ce sont aussi les fabricants des équipements du processus décisionnel (les entreprises dites de la *legaltech*, les éditeurs juridiques) qui deviennent parties prenantes de celui-ci. Cela pousse à s'interroger sur les effets, mais également sur le sens social et politique de cette plus grande distribution de l'acte de juger.

Ce travail sur la 'justice prédictive' a soulevé en outre plusieurs défis théoriques et analytiques, sur lesquels il avance quelques propositions.

1. Il opère une distinction analytique entre jugement et décision, en appliquant au cas de la justice, des acquis tirés de la science politique générale. Il critique l'approche décisionniste encore trop souvent implicitement présente dans les discours et les analyses des jugements juridictionnels.
2. Il prolonge des travaux, principalement issus de la sociologie pragmatique, qui ont tissé des liens entre jugement ordinaire et jugement juridictionnel. Il contribue ainsi à désenclaver les jugements juridictionnels, tout en prenant au sérieux leurs spécificités : dans le rapport au droit, à l'État, à la justice comme institution, au groupe professionnel et à l'ethos des magistrats ; dans les agencements concrets, incarnés et matériels dont ils procèdent en pratique.
3. Il propose un cadre analytique qui conceptualise le jugement comme résultat d'une activité pratique équipée par une pluralité de dispositifs sociotechniques, et d'un processus épais de formation ; et qui met l'accent sur la diversité des configurations de jugement observables, à scruter de près.
4. Il dessine un programme et une méthode de travail qui posent les fondements d'une analyse du rôle, des usages et des effets des dispositifs algorithmiques de traitement et d'analyse des décisions de justice sur la fabrique des jugements.

Sur un plan méthodologique pour finir, ce travail s'est appuyé sur des procédés classiques de la discipline (analyses de traces, observations, analyses lexicales et lexicométriques, etc.) tout en laissant la place à des sources tirées du *big data* dont l'apport est spécifique (accéder à une arène du débat public comme Twitter), mais aussi plus général : accès à des informations factuelles que nous n'avions pas vues passer ailleurs ; éléments qui amènent à relire autrement certaines données obtenues via les sources plus routinières que sont les articles de presse. Par un effet de mise en abyme, nous avons mesuré ce que le tournant algorithmique, étudié dans la justice, veut dire également pour les sciences sociales.

À l'heure de poser enfin la plume, il est clair que la conclusion du travail restitué ici ne peut qu'être provisoire. Au fil des pages, des voies de consolidation ont été identifiées, des perspectives d'approfondissement dessinées. Elles ne demandent qu'à être poursuivies. La suite du programme est excitante.

Bibliographie

AAS Katja F.(2005), *Sentencing in the age of information: from Faust to Macintosh*. London, Glass House Press.

ABBOTT Andrew (1988), *The System of Professions. An Essay on the Division of Expert Labor*. Chicago, University of Chicago Press.

ABDELNOUR Sarah et BERNARD Sophie (2018), « Vers un capitalisme de plateforme ? Mobiliser le travail, contourner les régulations ». *La nouvelle revue du travail* [en ligne], n° 13, mis en ligne le 31 octobre 2018 [consulté le 1^{er} avril 2020].

ABITEBOUL Serge et DOWEK Gilles (2017), *Le Temps des algorithmes*. Paris, Le Pommier.

ACKERMANN Werner et BASTARD Benoit (1993), *Innovation et Gestion dans l'institution judiciaire*. Paris, LGDJ.

ALAUZEN Marie (2019), « Plis et Replis de l'État plateforme. Enquête sur la modernisation des services publics en France ». Thèse de sciences, technologies et sociétés, université Paris Sciences et Lettres.

ALAUZEN Marie (2021), « Splendeurs et Misères d'une start-up d'État ». *Réseaux*, n° 225, p. 121-150.

ALETRAS Nikolaos *et al.* (2016), « Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective ». *PeerJ Computer Science* [en ligne], <https://doi.org/10.7717/peerj-cs.93>

ALGAN Yann et CAZENAVE Thomas (dir.) (2016), *L'État en mode start-up*, préface d'Emmanuel Macron. Paris, éd. Eyrolles.

ALLISON Graham T. (1971), *The Essence of Decision: Explaining the Cuban Missile Crisis*, 1^{re} éd., Little Brown [Graham ALLISON et Philip D. ZELIKOW, nouvelle éd. 1999]. Cf. traduction d'un chapitre de la nouvelle édition : ALLISON Graham T. et ZELIKOW Philip D. (2000), « L'Essence de la décision. Le modèle de l'acteur rationnel ». *Cultures & Conflits*, vol. 1, n° 36, p. 1-46.

AMICELLE Anthony (2021), « Policing & Big Data. La mise en algorithmes d'une politique internationale ». *Critique internationale*, n° 92, p. 23-48.

AMOROSO Daniele et TAMBURRINI Guglielmo (2021), « The Human Control Over Autonomous Robotic Systems: What Ethical and Legal Lessons for Judicial Uses of AI? », in KRAMER Xandra *et al.*, *New Pathways to Civil Justice in Europe: Challenges of Access to Justice*. Cham, Springer International Publishing, p. 23-42.

AMRANI-MEKKI Soraya (2018), « Le Point de vue d'une universitaire. Table ronde 1 », in Ordre des avocats au Conseil d'État et à la Cour de cassation, *La justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 49-63.

AMRANI-MEKKI Soraya (2021), « Les Plateformes de résolution en ligne des différends », in DELPECH Xavier (dir.), *L'Émergence d'un droit des plateformes*. Paris, Dalloz, p. 189-203.

- ANGQIN Julia et al. (2016), « Machine Bias ». Propublica.org [en ligne], <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- ARSENE Séverine et MABI Clément (2021), « L'Action publique au prisme de la gouvernamentalité numérique ». *Réseaux*, n° 225, p. 9-22.
- ASHLEY Kevin D. (2017), *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*. Cambridge University Press.
- ASSIER-ANDRIEU Louis (1996), *Le Droit dans les sociétés humaines*. Paris, Nathan.
- AYKUT Stephan C. et NADAI Alain (2019), « Le Calcul et le Politique. Le Débat national sur la transition énergétique et la construction des choix énergétiques en France ». *Revue d'anthropologie des connaissances*, vol. 13, n° 4, p. 1009-1034.
- BACHRACH Peter et BARATZ Morton (1963), « Decisions and Non-Decisions: an Analytical Framework ». *American Political Science Review*, vol. 57, p. 641-651.
- BADOUARD Romain, MABI Clément et MONNOYER-SMITH Laurence (2016), « Le Débat et ses arènes. À propos de la matérialité des espaces de discussion ». *Questions de communication*, vol. 30, n° 2, p. 7-23.
- BALME Richard (2019), « Rationalité ». *Dictionnaire des politiques publiques*, 5^e éd., p. 519-528.
- BANCAUD Alain (1993), *La Haute Magistrature judiciaire entre politique et sacerdoce ou le culte des vertus moyennes*. Paris, LGDJ.
- BARFIELD Woodrow et PAGALLO Ugo (eds) (2018), *Research Handbook on Law and Artificial Intelligence*. Cheltenham, Northampton, Edward Elgar Publishing.
- BARRAUD, Boris (2017), « Un algorithme capable de prédire les décisions des juges : vers une robotisation de la justice ? ». *Les Cahiers de la justice*, n° 1, 121-39.
- BARTHE Emmanuel (2008), « Arrêts de cour d'appel : Légifrance décolle », mis en ligne le 28 décembre 2008 [consulté le 16 août 2022], <https://www.precisement.org/blog/Arrets-de-cour-d-appel-Legifrance.html>
- BARTHE Emmanuel (2021), « Intelligence artificielle en droit : derrière la "hype", la réalité ». Precisement.org, mis en ligne en janvier 2017, actualisé depuis [consulté le 17 octobre 2021], <https://www.precisement.org/blog/Intelligence-artificielle-en-droit-derriere-la-hype-la-realite.html>.
- BARTHE Yannick (2006). *Le Pouvoir d'indécision. La mise en politique des déchets nucléaires*. Paris, Economica.
- BARTHE Yannick, CALLON Michel et LASCOUMES Pierre (2001), *Agir dans un monde incertain. Essai sur la démocratie technique*. Paris, Seuil.
- BASTARD Benoît avec la collaboration de BOIGEOL Anne, BENECH-LE ROUX Patricia et JAMET Ludovic (2016), *Éthique et TIC. À quoi sert l'Ordre des avocats ?* Rapport de recherche, ISP-ENS Paris-Saclay.
- BASTARD Benoit et MOUHANNA Christian (2010), *L'Avenir du juge des enfants. Éduquer ou punir ?* Toulouse, éd. Érès.

- BASTARD Irène *et al.* (2017), « Facebook, pour quoi faire ? Configurations d'activités et structures relationnelles ». *Sociologie*, vol. 8, n° 1, p. 57-82.
- BASTIN Gilles et TUBARO Paola (2018). « Le Moment big data des sciences sociales ». *Revue française de sociologie*, vol. 59, n° 3, p. 375-394.
- BAUDOT, Pierre-Yves, MARREL Guillaume et NONJON Magali (2015), « Encore une révolution informatique ? *Open et Big Data* dans les organisations administratives ». *Informations sociales*, vol. 191, n° 5, p. 8-18.
- BAUM Lawrence (2009), *The Puzzle of Judicial Behavior*. Ann Arbor, University of Michigan Press.
- BAUM Lawrence (2017), « Law and Policy in Decision Making », in HOWARD Robert M. et RANDAZZO Kirk A., *Routledge Handbook of Judicial Behavior*. New York, Routledge, p. 114-131.
- BAUM Lawrence (2017), *Ideology in the Supreme Court*. Princeton, Princeton University Press.
- BAUM Lawrence (2006), *Judges and their Audiences: A Perspective on Judicial Behavior*. Princeton, Princeton University Press.
- BECCARIA Cesare (2006[1764]), *Des délits et des peines*. Paris, Flammarion.
- BECK Ulrich (2008), *La Société du risque. Sur la voie d'une autre modernité*. Paris, Flammarion.
- BEHR Valentin et MICHON Sébastien (2020), « Le Gouvernement Macron et les nouveaux technos ». *Revue française d'administration publique*, n° 175, p. 735-753.
- BELLON Anne (2018), « Gouverner l'Internet : mobilisations, expertises et bureaucraties dans la fabrique des politiques numériques (1969-2017) ». Thèse de science politique, université Paris 1.
- BELLON Anne (2021), « L'Élaboration de la loi république numérique au gouvernement ». *Réseaux*, n° 225, p. 23-53.
- BENBOUZID Bilel (2018), « Quand prédire, c'est gérer : la Police prédictive aux États-Unis ». *Réseaux*, vol. 211, p. 221-256.
- BENBOUZID Bilel et CARDON Dominique (2018), « Machines à prédire ». *Réseaux*, vol. 211, n° 5, p. 9-33.
- BENBOUZID Bilel et Cardon Dominique (2022), « Contrôler les IA. Présentation ». *Réseaux*, vol. 232-233, p. 9-26.
- BENGHOZI Pierre-Jean (2012), « L'Âge de la multitude. Entreprendre et gouverner après la révolution numérique ». *Annales des Mines – Gérer et comprendre*, n° 110, p. 70-71.
- BERGERON Henri, CASTEL Patrick et DUBUISSON-QUELLIER Sophie (2014), « Gouverner par les labels. Une comparaison des politiques de l'obésité et de la consommation durable ». *Gouvernement et action publique*, vol. 3, n° 3, p. 7-31.
- BERNABE Boris (2016), « L'Office du juge et la liturgie du juste ». *Cahiers philosophiques*, vol. 147, n° 4, p. 48-67.
- BESSY Christian et CHATEAURAYNAUD Francis (2014 [1995]), *Experts et Faussaires. Pour une sociologie de la perception*. Paris, éd. Pétra.

- BEUSCART Jean-Samuel, DAGIRAL Éric et PARASIE Sylvain (2019), *Sociologie d'Internet*. Paris, Armand Colin, 2019.
- BEYENS Kristel, SNACKEN Sonja et VAN ZYL SMIT Dirk (2013), « Truth in the Implementation of Sentencing: Belgium and Elsewhere », in DAEMS Tom, ZYL SMIT Dirk Van et SNACKEN Sonja, *European Penology?* Oxford, Hart.
- BEZÈS Philippe (2009), *Réinventer l'État. Les réformes de l'administration française (1962-2008)*. Paris, PUF, coll. « Le lien social ».
- BEZÈS Philippe, « Construire des bureaucraties webériennes à l'ère du New Public Management ? ». *Critique internationale*, 2007, n° 35, p. 9-29.
- BEZÈS Philippe et JOIN-LAMBERT Odile (2010), « Comment se font les administrations : analyser des activités administratives constituantes ». *Sociologie du travail*, vol. 52, n° 2, p. 133-150.
- BIARD Alexandre (2019), « Justice en ligne ou nouveau Far Wwww.est ? La difficile régulation des plateformes en ligne de règlement extrajudiciaire des litiges ». *Revue internationale de droit économique*, t. 23, n° 2, p. 165-191.
- BIARD Alexandre *et al.* (2021), « Introduction: The Future of Access to Justice – Beyond Science Fiction » in KRAMER Xandra *et al.* (éd.), *New Pathways to Civil Justice in Europe: Challenges of Access to Justice*. Cham, Springer International Publishing, p. 1-20.
- BILAND Émilie *et al.* (2021), *Justice et Inégalités au prisme des sciences sociales*. Rapport pour le GIP Mission de recherche Droit et Justice.
- BING Jon (2010), « Let There Be Lite: a Brief History of Legal Information Retrieval ». *European Journal of Law and Technology* 1, n° 1 (17 mars). <http://ejlt.org/article/view/15>.
- BIOLLEY Sophie de (2013), « La Vidéo-comparution en Belgique : une solution sans problème ». *Déviance et Société*, vol. 37, p. 305-321.
- BLANC Nathalie et MEKKI Mustapha (dir.) (2019), *Le Juge et le Numérique : un défi pour la justice du XXI^e siècle*. Paris, Dalloz, coll. « Thèmes et commentaires ».
- BLUMBERG Georges (2022) « Actuariat & actuaires ». *Encyclopædia Universalis* [en ligne], consulté le 22 août 2022. <https://www.universalis.fr/encyclopedie/actuariat-et-actuaire/>
- BONAFE-SCHMIDT Jean-Pierre (1993), « Alternatif », in ARNAUD André-Jean *et al.*, *Dictionnaire encyclopédique de théorie et de sociologie du droit*. Paris, LGDJ, (2^e éd.), p. 11-15.
- BORDONABA Laure et CHOUCAN Nathalie (2016), « Les Libertés du juge ». *Cahiers philosophiques*, n° 147, p. 110-123.
- BORE Louis (2018). « Mot d'accueil », in Ordre des avocats au Conseil d'État et à la Cour de cassation, *La justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 1-4.
- BORIES Serge (2003), « L'Informatisation des données judiciaires et doctrinales : une contribution à la connaissance et à la recherche juridiques ». *Documentaliste-Sciences de l'Information*, vol. 40, n° 4, p. 272-279.
- BORIES Serge (2013), « Pierre Catala et l'Informatique juridique » , in CABRILLAC Rémy, *Hommage à Pierre Catala*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 61-66.

- BORRAZ Olivier (2008). *Les politiques du risque*. Paris, Presses de Sciences po.
- BORRAZ Olivier (ed.) (2022), *La Société des organisations*. Paris, Presses de Sciences po.
- BOSSAN Jérôme et LETURMY Laurence (2019), *Comparution sur reconnaissance préalable de culpabilité - Bilans et perspectives*. Rapport Mission de recherche Droit et Justice / EPRED.
- BOSSIS Raymonde (2003), « Greffier : les incertitudes de la professionnalisation du corps des greffiers ». *Questions pénales*, vol. 16, n° 5, p. 1-4.
- BOTTINI Thomas et JULLIARD Virginie (2017), « Entre informatique et sémiotique. Les conditions techno-méthodologiques d'une analyse de controverse sur Twitter ». *Réseaux*, vol. 4, n° 204, p. 35-69.
- BOUCHER Maxime (2019), « Expertise ». *Dictionnaire des politiques publiques*, 5^e éd., p. 256-263.
- BOULLIER Dominique (2015), « Les Sciences sociales face aux traces du big data. Société, opinion ou vibrations ? ». *Revue française de science politique*, vol. 65, n° 5-6, p. 805-828.
- BOULLIER Henri *et al.* (2021), « "En attendant d'autres données" : Raison modélisatrice, incertitudes et pièges de l'abstraction ». *Terrains & travaux*, vol. 38, p. 5-22.
- BOULLIER Henri, DEMORTAIN David, et ZEEMAN Maurice (2019), « Inventing Prediction for Regulation. The Integration of Structure-Activity Relationships in the Assessment of New Chemicals at the US Environmental Protection Agency ». *Science & Technology Studies*, vol. 32, n° 4, p. 137-157.
- BOURCIER Danièle (1995), *La Décision artificielle. Le droit, la machine et l'humain*. Paris, PUF, coll. « Les voies du droit ».
- BOURCIER Danièle et DE FILIPPI Primavera (2018), « Transparence des algorithmes face à l'open data : quel statut pour les données d'apprentissage ? ». *Revue française d'administration publique*, vol. 167, n° 3, p. 525-537.
- BOURDIEU Pierre (1986), « La Force du droit ». *Actes de la recherche en sciences sociales*, n° 64, p. 3-19.
- [BOURG Louis] (2018), « Les Perspectives ouvertes par la mise à disposition du public des décisions de justice : quelle place et quelle régulation pour la justice prédictive ? État des lieux et analyse à partir de certaines des propositions du rapport de la mission sur l'open data des décisions de justice ». *Archives de philosophie du droit*, vol. 60, n° 1, p. 117-123.
- BOURREAU-DUBOIS Cécile (2019), « La Barémisation de la justice. Une approche par l'analyse économique du droit ». Rapport BETA / GIP Mission de recherche Droit et justice.
- BOURREAU-DUBOIS Cécile, DORIAT-DUBAN Myriam et RAY Jean-Claude (2006), « Caractéristiques du juge et décisions en matière de pensions alimentaires. Une étude à partir de données expérimentales ». *Revue économique*, vol. 57, n° 3, p. 563-572.
- BOUSSAGUET Laurie (2019), « Forum ». *Dictionnaire des politiques publiques*, p. 273-278.
- BOYADJIAN Julien (2016), *Analyser les opinions politiques sur internet. Enjeux théoriques et défis méthodologiques*. Paris, Dalloz.

- BOYADJIAN Julien (2017), « Les Conditions de scientificité des *Big Data* en science politique ». *Revue française de science politique*, vol. 67, n° 5, p. 919-929.
- BRAYNE Sarah et CHRISTIN Angèle (2020), « Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts ». *Social Problems*, vol. 68, n° 3, p. 608-624.
- BRAYNE Sarah (2020), *Predict and Surveil: Data, Discretion, and the Future of Policing*. Oxford, Oxford University Press.
- BRIGANT Jean-Marie (2018), « Les Risques accentués d'une justice pénale prédictive ». *Archives de philosophie du droit*, t. 60, p. 237-251.
- BRUGUES-REIX, Béatrice et PACQUETET Ashley (2018), « La Justice prédictive : un "outil" pour les professionnels du droit ». *Archives de philosophie du droit*, vol. 60, p. 279-285.
- BRUNET Pierre (2019), « Recension de *Ideology in the Supreme Court* de Lawrence Baum ». *Droit et Société*, mis en ligne le 25 avril 2019, <https://ds.hypotheses.org/5776>
- BRUNIER Sylvain, CASTEL Patrick et CHESSEL Marie-Emmanuelle (2022), « Chapitre 5. À la recherche des décideurs. Équipement, entourage et personnalisation des décisions », in Olivier Borraz (éd.), *La Société des organisations*. Paris, Presses de Sciences po, p. 93-107.
- BURRELL Jenna et FOURCADE Marion (2021), « The Society of Algorithms ». *Annual Review of Sociology*, vol. 47, n° 1, p. 213-237.
- CADIET Loïc (2017), « L'Accès à la justice. Réflexions sur la justice à l'épreuve des mutations contemporaines de l'accès à la justice ». *Recueil Dalloz*, n° 10, p. 522.
- CADIET Loïc (2018), *L'Open Data des décisions de justice. Rapport de la Mission d'étude et de préfiguration sur l'ouverture au public des décisions de justice*. Paris, ministère de la Justice.
- CALLON Michel (1981), « Pour une sociologie des controverses technologiques ». *Fundamenta Scientiae*, vol. 2, n° 3-4, p. 381-399.
- CALLON Michel (1986), « Éléments pour une sociologie de la traduction : la domestication des coquilles Saint-Jacques et des marins-pêcheurs dans la baie de Saint-Brieuc ». *L'Année sociologique*, vol. 36, p. 169-208.
- CALLON Michel, LATOUR Bruno et AKRICH Madeleine (2006), *Sociologie de la traduction. Textes fondateurs*. Paris, Presses de l'école des Mines.
- CAMUS Aurélien (2018), « La Propriété des données publiques ». *Revue française d'administration publique*, n° 167, p. 479-490.
- CAPPELLINA Bartolomeo (2014), « Quantifier la qualité de la justice entre droits de l'homme et modernisation ». *Revue française de socio-économie*, n° 19, p. 27-46.
- CARDON Dominique (2013a). « Politique des algorithmes. Les métriques du web. Présentation ». *Réseaux*, vol. 177, p. 9-21.
- CARDON Dominique (2013b), « Dans l'esprit du PageRank : Une enquête sur l'algorithme de Google ». *Réseaux*, vol. 177, p. 63-95.
- CARDON Dominique (2015), *À quoi rêvent les algorithmes : Nos vies à l'heure des big data*. Paris, Seuil, coll. « la République des idées ».
- CARDON Dominique (2019), *Culture numérique*. Paris, Presses de Sciences po.

- CARDON Dominique, COINTET Jean-Philippe et MAZIERES Antoine (2018), « La Revanche des neurones ». *Réseaux*, vol. 211, n° 5, p. 173-220.
- CARDON Dominique et PRIEUR Christophe (2016), « Comment la conversation a façonné le web », in Martin Olivier et Dagiral Éric (dir.), *L'Ordinaire d'Internet. Le web dans nos pratiques et relations sociales*. Paris, Armand Colin, p. 226-47.
- CASILI Antonio A. (2019), *En attendant les robots. Enquête sur le travail du clic*. Paris, Seuil.
- CASSAR Bertrand (2020), *La Transformation numérique du monde du droit*. Thèse pour le doctorat en droit, Université de Strasbourg.
- CASTEL Robert (1995), *Les Métamorphoses de la question sociale*. Paris, Fayard.
- CATALA Pierre (1998), *Le Droit à l'épreuve du numérique*. Paris, PUF, coll. « Droit éthique société ».
- CHABERT Alexis (2019), « Quelle influence sur la pratique professionnelle des avocats ? », in SAYN Isabelle (éd.), *Séminaire E-Juris. De l'open data des décisions de justice à l'automatisation de la prise de décisions ? Les potentialités des bases de données décisionnelles*, MSH Lyon, p. 16-21, <https://halshs.archives-ouvertes.fr/halshs-02318411/document>.
- CHADUTEAU Olivier (2018), « Justice prédictive : l'état de l'art des legaltechs », in Ordre des avocats au Conseil d'État et à la Cour de cassation, *La justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 71-77.
- CHAISEMARTIN Arnaud de (2018), « Le Point de vue de l'avocat. Table ronde 2 », in Ordre des avocats au Conseil d'État et à la Cour de cassation, *La justice prédictive*. Paris, Dalloz, p. 89-96.
- CHANIAL Philippe (2001), « Carl Schmitt (1892-1985) et Hans Kelsen (1881-1973) la querelle du droit : positivisme, décisionnisme et utilitarisme ». *Histoire raisonnée de la philosophie morale et politique*. Paris, La Découverte, p. 660-668.
- CHAPPE Vincent-Arnaud, JUSTION-MORIVAL Romain et LECLERC Olivier, « Faire preuve : pour une analyse pragmatique de l'activité probatoire », Présentation du dossier . *Droit et Société*, vol. 110, n° 1, 2022, p. 7-20.
- CHASSAGNARD-PINET Sandrine (2017), « Les Usages des algorithmes en droit : prédire ou dire le droit ? ». *Dalloz IP/IT*, n° 10, p. 506-511.
- CHELLE Elisa (2011), « Une politique de récompense dans la haute magistrature : le cas de la prime de rendement ». *Droit et Société*, vol. 78, p. 407-427.
- [CHEVALET Victor] (2018a), « L'Éthique de la justice prédictive ». *Enjeux numériques. Les Annales des Mines*, n° 3, p. 86-91.
- [CHEVALET Victor] (2018b), « Le Droit isométrique : un nouveau paradigme juridique né de la justice prédictive ». *Archives de philosophie du droit*, t. 60, p. 287-295.
- CHEVALLIER Jacques (2018), « Vers l'État plateforme ? ». *Revue française d'administration publique*, n° 167, p. 627-637.
- CHIAMBARETTO Paul, FERNANDEZ Anne-Sophie et LE ROY Frédéric (2019), « La Coopétition ou l'Art de coopérer avec ses concurrents » in LIARTE Sébastien (dir.), *Les Grands Courants en management stratégique*. Caen, EMS éditions, p. 281-312.

CHOLET Didier (2018), « La Justice prédictive et les principes fondamentaux du procès civil ». *Archives de philosophie du droit*, t. 60, p. 223-236.

CHRISTIN Angèle (2008), *Comparutions immédiates. Enquête sur une pratique judiciaire*. Paris, La Découverte.

CHRISTIN Angèle (2016) « The Hidden Story of How Metrics Are Being Used in Courtrooms and Newsrooms to Make More Decisions ». *Ethnography Matters*, mis en ligne le 20 juillet 2018 [consulté le 13 juillet 2018], <http://ethnographymatters.net/blog/2016/06/20/the-hidden-story-of-how-metrics-are-being-used-in-courtrooms-and-newsrooms-to-make-more-decisions>

CHRISTIN Angèle (2017), « Algorithms in Practice: Comparing Web Journalism and Criminal Justice ». *Big Data & Society*, vol. 4, n° 2, <https://doi.org/10.1177/2053951717718855>

CHRISTIN Angèle (2018), « The Mistrials of Algorithmic Sentencing ». *Logic. À magazine about technology*, mis en ligne le 17 janvier 2018 [consulté le 13 juillet 2018], <https://logicmag.io/03-the-mistrials-of-algorithmic-sentencing/>

CHRISTIN Angèle (2020), *Metrics at Work. Journalism and the Contested Meaning of Algorithms*. Princeton, Princeton University Press.

CLAVIER Jean-Pierre (dir.) (2020), *L'Algorithmisation de la justice*. Bruxelles, Larcier.

CLÉMENT Marc (2019), « Quelle influence sur la façon de juger ? », in SAYN Isabelle (éd.), *Séminaire E-Juris. De l'open data des décisions de justice à l'automatisation de la prise de décisions ? Les potentialités des bases de données décisionnelles*, MSH Lyon, p. 24-29. <https://halshs.archives-ouvertes.fr/halshs-02318411/document>

CLOUTIER Laurence (2014), « Inventeurs entrepreneurs », in Pierre-Marie CHAUVIN, Michel GROSSETTI et Pierre-Paul ZALIO (dir.), *Dictionnaire de l'entrepreneuriat*. Paris, Presses de Science po, p. 380-396.

CLUZEL-METAYER Lucie (2017), « La Loi pour une République numérique : l'écosystème de la donnée saisi par le droit ». *Actualité juridique. Droit administratif*, n° 6, p. 340-349.

CLUZEL-METAYER Lucie (2018), « La Construction d'un service public de la donnée ». *Revue française d'administration publique*, n° 167, p. 491-500.

CODACCIONI Vanessa, PUCCIO-DEN Deborah et ROUSSEL Violaine (2015), « Les "Bonnes Formes" du procès à l'épreuve des mobilisations politiques », Présentation du dossier. *Droit et Société*, n° 89, p. 9-15.

COLIN Nicolas et VERDIER Henri (2012), *L'Âge de la multitude. Entreprendre et gouverner après la révolution numérique*. Paris, Armand Colin, 2^e éd. 2015.

COMMAILLE Jacques (1982), *Familles sans justice, Le droit et la justice face aux transformations de la famille*. Paris, Le Centurion.

COMMAILLE Jacques (2000), *Territoires de justice*. Paris, PUF.

COMMAILLE Jacques (2007), « La Construction d'une sociologie spécialisée. Le Savoir sociologique et la sociologie juridique de Jean Carbonnier ». *L'Année sociologique*, vol. 57, n° 2, p. 275-299.

COMMAILLE Jacques (2015), *À quoi nous sert le droit ?*. Paris, Gallimard, coll. « Folio essais ».

- COMMAILLE Jacques et DUMOULIN Laurence (2009), « Heurs et Malheurs de la légalité dans les sociétés contemporaines. Une sociologie politique de la “judiciarisation” ». *L'Année sociologique*, vol. 59, n° 1, p. 63-107.
- CONEIN Bernard (2004), « Cognition distribuée, groupe social et technologie cognitive ». *Réseaux*, vol. 124, n° 2, p. 53-79.
- CONTINI Francesco et LANZARA Giovan F. (éd.) (2009), *ICT and Innovation in the Public Sector: European Studies in the Making of E-Government (Technology, work and globalization)*. Londres, New York, Palgrave Macmillan.
- CONTINI Francesco et LANZARA Giovan F. (éd.) (2013), *The Circulation of Agency in E-Justice: Interoperability and Infrastructures for European Transborder Judicial Proceedings*. Berlin, Springer Science & Business Media.
- CORDELLA Antonio (2007), « E-Government: Towards the E-Bureaucratic Form? ». *Journal of Information Technology*, vol. 22, n° 3, p. 265–274.
- CORDELLA Antonio et BONINA Carla (2012), « A Public Value Perspective for ICT Enabled Public Sector Reform: A Theoretical Reflection ». *Government Information Quarterly*, vol. 29, n° 4, p. 512–520.
- CORDELLA Antonio et HESSE Jakob (2015), « E-Government in the Making: an Actor Network Perspective ». *Transforming Government: People, Process and Policy*, vol. 9, n° 1, p. 125–104.
- COTTEREAU Alain (1987), « Justice et Injustice ordinaire sur les lieux de travail, d'après les audiences prud'homales, 1806-1866 ». *Le Mouvement social*, vol. 141, p. 25-61.
- COTTIN Marianne (2019), « Jurisprudence et Contentieux, une (r)évolution à attendre ? ». *La Semaine juridique*, supplément au n° 44-45, p. 22-25.
- COUSIN Clément (2017), « Le Débat sur le référentiel indicatif de l'indemnisation du préjudice corporel des cours d'appel à l'heure des bases de données ». *La Semaine juridique – édition générale*, n° 17, p. 830-36.
- COUTU Michel (2018), *Max Weber's Interpretive Sociology of Law*. Londres, Routledge.
- CROZE Hervé (2017), « La Factualisation du droit ». *La semaine juridique – édition générale*, n° 5, p. 174-175.
- DAGIRAL Éric (2011), « Administration électronique ». *Communications*, n° 88, p. 9-17.
- DALLARA Christina et PIANA Daniela (2015), *Networking the Rule of Law. How Change Agents Reshape the Judicial Governance in the EU*, London, Routledge.
- DANET Jean (dir.) (2013), *La réponse pénale. Dix ans de traitement des délits*, Rennes, Presses universitaires de Rennes.
- DANZIGER Shai, LEVAV Jonathan et AVNAIM-PESSO Liora (2011), « Extraneous Factors in Judicial Decisions ». *Proceedings of the National Academy of Sciences*, vol. 108, n° 17, p. 6889-6892, <https://doi.org/10.1073/pnas.1018033108>
- DARTY Franck, FROISSART Catherine et MÉNARD François (1997), *La professionnalité des greffiers*, Rapport d'étude FORS Recherche sociale / GIP mission de recherche Droit et Justice / Ministère de la Justice.

- DE MUNCK Jean (1995), « Le Pluralisme des modèles de justice », in Antoine GARAPON, Denis SALAS, *La Justice des mineurs : évolution d'un modèle*. Paris/Bruxelles, LGDJ/Bruylant.
- DEFFAINS Bruno (2018) « Concurrence des normativités : un point de vue économique ». *Revue internationale de droit économique*, t. 32, n° 3, p. 269-279.
- DEFFAINS Bruno (2019), « Le Monde du droit face à la transformation numérique ». *Pouvoirs*, n° 170, p. 43-58.
- DEFFAINS Bruno et DORIAT-DUBAN Myriam (2019), « La Médiation à travers le prisme de l'analyse économique ». *Archives de philosophie du droit*, vol. 61, n° 1, p. 3-20.
- DEFFAINS Bruno et FERREY Samuel (2010), *Agir et Juger : comment les économistes pensent le droit*. Paris, Presses d'Assas.
- DEFFERRARD Fabrice et PAPINEAU Christelle (2017), « Le Pouvoir de juridiction des algorithmes aux États-Unis : entre fantasme et réalité jurisprudentielle ». *Dalloz IP/IT*, n° 12 (14 décembre), p. 668-71.
- DEGOS Louis (2020), « Legaltech et justice prédictive : cartographie ». *Revue pratique de la prospective et de l'innovation*, n° 2, p. 2-3.
- DELAUVIGNE Valérie (2003), « Alceste, un logiciel d'analyse textuelle ». *Texte ! Textes et Cultures*, Institut Ferdinand de Saussure, [\(hal-00924168\)](#)
- DELMAS Corinne (2019), *Les Notaires en France. Des officiers de l'authentique entre héritage et modernité*. Rennes, Presses universitaires de Rennes, coll. « Des sociétés ».
- DELMAS-GOYON Pierre (2013), « "Le Juge du XXI^e siècle". Un citoyen acteur, une équipe de justice ». Rapport à la ministre de la Justice. Paris, La Documentation française.
- DELPEUCH Thierry, DUMOULIN Laurence et GALEMBERT Claire de (2014), *Sociologie du droit et de la justice*. Paris, Armand Colin, coll. U.
- DENIS Jérôme et GOËTA Samuel (2017), « Les Facettes de l'Open Data : émergence, fondements et travail en coulisses », in MENGER Pierre-Michel, PAYE Simon (dir.), *Big Data et Traçabilité numérique. Les sciences sociales face à la quantification massive des individus*. Paris, Collège de France, p. 121-138, <https://doi.org/10.4000/books.cdf.5005>.
- [DESJARDINS Valère] (2019), « La Justice dite "prédictive" : prérequis, risques et attentes – l'expérience française ». *Les Cahiers de la Justice*, n° 2, p. 269-276.
- [DESJARDINS Valère] et [BOURG Louis] (2017), « La Mémoire numérique des décisions judiciaires. L'open data des décisions de justice de l'ordre judiciaire ». *Recueil Dalloz*, p. 1483-1487.
- DESMOULINS-CANSELIER Sonia et LE MÉTAYER Daniel (2020), *Décider avec les algorithmes. Quelle place pour l'Homme, quelle place pour le droit ?*. Paris, Dalloz, coll. « Les sens du droit ».
- DEZIEL Pierre-Luc (2018), « L'Utilisation de renseignements personnels dans le contexte de la justice prédictive : le cas des outils actuariels d'évaluation des risques de récidive ». *Archives de philosophie du droit*, vol. 60, p. 253-69.
- DODIER Nicolas et BARBOT Jeanine (2016), « La Force des dispositifs ». *Annales. Histoire, Sciences sociales*, vol. 71, n° 2, p. 421-450.

DOLEZ Bernard, FRETTEL Julien et LEFEBVRE Rémi (2019), « Introduction générale. La science politique mise au défi par Emmanuel Macron », in DOLEZ Bernard, FRETTEL Julien et LEFEBVRE Rémi (dir.), *L'Entreprise Macron*, Fontaine, Presses universitaires de Grenoble, p. 9-17.

DONDERO, Bruno (2017), « Justice prédictive : la fin de l'aléa judiciaire ? ». *Recueil Dalloz Doctrine*, 532.

DOUAY Nicolas et REYS Aurélien (2016), « Twitter comme arène de débat public : le cas du Conseil de Paris et des controverses en aménagement ». *L'Information géographique*, vol. 80, n° 4, p. 76-95.

DOUILLET Anne-Cécile et DUMOULIN Laurence (2015), « Actor Network Theory and CCTV Development », in ROBERT Dominique et DUFRESNE Martin (éd.), *Actor-Network Theory and Crime Studies: Explorations in Science and Technology*. Fanham, Ashgate Publishing, p. 21-35.

DUMOULIN Laurence (2000), « L'Expertise judiciaire dans la construction du jugement : de la ressource à la contrainte ». *Droit et Société*, n° 44/45, p. 199-223.

DUMOULIN Laurence (2007), *L'Expert dans la Justice. De la genèse d'une figure à ses usages*. Paris, Economica.

DUMOULIN Laurence (2016), « Les Juges sont-ils des décideurs politiques ? Libres propos sur ce qui apparaît comme un oxymore, mais n'en est peut-être pas un », in GRANDJEAN Geoffrey et WILDEMEERSCH Jonathan (éd.), *Les Juges : décideurs politiques ? Essais sur le pouvoir politique des juges dans l'exercice de leur fonction*, Bruxelles, Bruylant, p. 79-91.

DUMOULIN Laurence *et al.* (2017), « Justice et Nouvelles Technologies. La gestation d'un nouveau modèle de justice ? » Paris, Rapport ACI, non publié.

DUMOULIN Laurence et LICOPPE Christian (2011), « Technologies, Droit et Justice : quelques éléments de mise en perspective ». *Droit et Cultures*, vol. 61, n° 1, p. 13-40.

DUMOULIN Laurence et LICOPPE Christian (2013), « Innovation, Routinisation et Gestion de l'imprévu dans les audiences par visioconférence. Comment un écran et une caméra sur un meuble à roulettes travaillent l'organisation judiciaire en France ». *Déviante & Société*, vol. 37, n° 3, p. 323-345.

DUMOULIN Laurence et LICOPPE Christian (2015), « La Visioconférence comme mode de comparution des personnes détenues, une innovation "managériale" dans l'arène judiciaire ». *Droit et Société*, vol. 90, n° 2, p. 287-302.

DUMOULIN Laurence et LICOPPE Christian (2016), « Videoconferencing, New Public Management and Organizational Reform in the Judiciary ». *Policy & Internet*, vol. 8, p. 313-333.

DUMOULIN Laurence et LICOPPE Christian (2017), *Les Audiences à distance. Genèse et institutionnalisation d'une innovation dans la justice*. Paris, LGDJ-Lextenso, coll. « Droit et Société ».

DUMOULIN Laurence, LICOPPE Christian et THOENIG Jean-Claude av. la collab. de FROMENT Jean-Charles et MOUHANNA Christian (2007), « Les Technologies dans la justice : genèses et appropriations ». Rapport final, Grenoble, Pacte-Sciences po Recherche / Paris Télécoms.

DUMOULIN Laurence et VIGOUR Cécile (2019), « On a soif d'idéal ! Des attentes à la hauteur des critiques des citoyens. Un éclairage à partir d'entretiens collectifs » in AMRANI-MEKKI Soraya (dir.), *Et si on parlait du justiciable du XXI^e siècle ?*. Paris, Dalloz, coll. « thèmes et commentaires », p. 29-81.

DUNLEAVY Patrick *et al.* (2005), « New Public Management Is Dead - Long Live Digital-Era Governance ». *Journal of Public Administration Research and Theory*, p. 1-28.

DUPRET Baudoin (2001), « Le Droit en action et en contexte. Ethnométhodologie et analyse de conversation dans la recherche juridique. Présentation ». *Droit et Société*, vol. 2, n° 48, p. 343-348.

DUPRET Baudoin (2005), « Droit et Expertise dans une perspective praxéologique. Présentation ». *Droit et Société*, vol. 3, n° 61, p. 619-625.

DUPRET Baudoin (2006), *Le Jugement en action. Ethnométhodologie du droit, de la morale et de la justice en Égypte*. Genève/Le Caire, Librairie Droz/Cedej.

ÉCORMIER-NOCCA Florence et LOUIS-SIDOIS Charles (2019). « Fronde 2.0. Les députés français sur Twitter ». *Revue française de science politique*, vol. 69, n° 3, p. 481-500.

EISENSTEIN James et JACOB Herbert (1977), *Felony Justice. An Organisational Analysis of Criminal Court*. Boston, Little Brown.

ENGESTRÖM Yrjö (2006), « L'Interagentivité orientée-objet : vers une compréhension de l'intentionnalité collective dans les activités distribuées », in Jean-Marie Barbier (éd.), *Sujets, Activités, Environnements. Approches transverses*. Paris, PUF, p. 135-173.

ÉPINEUSE Harold et GARAPON Antoine (2018), « Les Défis d'une justice à l'ère numérique de "stade 3" ». *Enjeux numériques. Annales des Mines*, n° 3, p. 16-19.

EPSTEIN Lee et LINDQUIST Stefanie A. (éd.) (2017), *The Oxford Handbook of U.S. Judicial Behavior*. Oxford University Press, <https://doi.org/10.1093/oxfordhb/9780199579891.001.0001>.

EPSTEIN Renaud (2005), « Gouverner à distance. Quand l'État se retire des territoires ». *Esprit*, novembre, p. 96-111.

ERDELEZ Sandra et O'HARE Sheila (1997), « Legal Informatics: Application of Information Technology in Law ». *Annual Review of Information Science and Technology*, n° 32, p. 367-402.

EYMERI-DOUZANS Jean-Michel (2019) « Le Technocrate et le Président. Réflexions sur Emmanuel Macron et la haute administration », in DOLEZ Bernard, FRETTEL Julien et LEFEBVRE Rémi (dir.), *L'Entreprise Macron*, Fontaine, Presses universitaires de Grenoble, p. 201-216.

EYMERI-DOUZANS Jean-Michel (2021), « Chapitre 2 – Prise de décision », in JACOB Steve et SCHIFFINO Nathalie. (dir.), *Politiques publiques. Fondements et perspectives pour l'analyse de l'action publique* ; Bruxelles, Bruylant, 2021, p. 117-194.

FABRI Marco et LANGBROEK Philip M. (éd.) (2000), *The Challenge of Change for Judicial Systems: Developing a Public Administration Perspective*. International Institute of Administrative Sciences monographs. Amsterdam, IOS Press.

FAGET Jacques (2008), « La fabrique de la décision pénale. Une dialectique des asservissements et des émancipations ». *Champ pénal / Penal Field*, vol. 5, p. 1-17.

FEELEY Malcolm M. et SIMON Jonathan (1992), « The New Penology: Notes on the Emerging Strategy of Corrections and its Implications ». *Criminology*, n° 30, p. 449-474.

FEELEY Malcolm M. et SIMON Jonathan (1994), « Actuarial Justice: the Emerging New Criminal Law », in NELKEN David (éd.), *The Futures of Criminology*. London, Sage, p. 173-201.

FENWICK Mark et VERMEULEN Erik P. M. (2019), « The Lawyer of the Future as “Transaction Engineer” : Digital Technologies and the Disruption of the Legal Profession », in CORRALES Marcelo, FENWICK Mark et HAAPIO Helena, *Legal Tech, Smart Contracts and Blockchain*. Singapore, Springer, p. 253-272.

FEREY Samuel (2018), « Analyse économique du droit, big data et justice prédictive ». *Archives de philosophie du droit*, t. 60, p. 67-81.

FERY Bénédicte (2015), *Gouverner par les données ? : Pour une analyse des processus de traduction dans l'usage des systèmes d'information : Déploiement et utilisations de Cassiopée dans l'Institution pénale*. Thèse de doctorat de science politique, Versailles Saint-Quentin en Yvelines.

FLECHER Marion (2019), « Des inégalités d'accès aux inégalités de succès : enquête sur les fondateurs et fondatrices de start-up ». *Travail et Emploi*, n° 159, p. 39-68.

[FORESTIER Morgan] (2016), « L'Impartialité des juges mise à mal par l'intelligence artificielle », in *Village de la justice, Tribunes et point de vue*, mis en ligne en mars 2016 [consulté le 8 septembre 2020], www.village-justice.com/articles/impartialite-certains-juges-mise,21760.html.

FORTUIT Philippe et HAMIDOU Farah Nemira (2017), « La Mécanisation du droit peut-elle triompher de l'avocat ? ». *Revue internationale de droit économique*, t. 31, n° 4, p. 103-109.

FOUCAULT Michel (1994 [1977]), « Le Jeu de Michel Foucault ». *Dits et écrits II*. Paris, Gallimard.

FOUILLEUX Ève (2000), « Entre production et institutionnalisation des idées : la réforme de la politique agricole commune ». *Revue française de science politique*, vol. 50, n° 2, p. 277-305.

FRANCONY Jean-Marc (2018), « L'Éditorialisation des données aux bornes des API : enjeux et perspectives pour une analyse empirique ». *Les Enjeux de l'information et de la communication*, vol. 19, n° 2, p. 69-79.

FROUILLOU Leïla, PIN Clément et VAN ZANTEN Agnès (2022), « D'APB à Parcoursup : Deux conceptions de l'affectation post-bac et leurs effets sur les inégalités », in SIMIONI Melchior et STEINER Philippe (dir.), *Comment ça matche : Une sociologie de l'appariement*. Paris, Presses de Sciences po, p. 61-99.

G'SELL Florence (2018), « L'automatisation des décisions de justice, jusqu'où ? ». *Enjeux numériques*, n° 3, p. 48-52.

G'SELL Florence (dir.) (2020a), *Le Big Data et le Droit*. Paris, Dalloz.

G'SELL Florence (2020b) « Les décisions algorithmiques », in G'SELL Florence (éd.), *Le Big Data et le Droit*. Paris, Dalloz, p. 87-109.

GALANTER Marc (2013a [1974]), « “Pourquoi c'est toujours les mêmes qui s'en sortent bien ?” : Réflexions sur les limites de la transformation par le droit », trad. de UMUBYEYI

Liliane, ISRAËL Liora, « Why the “haves” come out ahead: Speculations on the Limits of Legal Change ». *Droit et Société*, n° 85, p. 575-640.

GALEMBERT Claire de (2016). « Le “Radical”, une nouvelle figure de dangerosité carcérale aux contours flous ». *Critique internationale*, n° 72, p. 53-71.

GAMET Laurent (2019), « Justice prédictive et droit social », in BLANC Nathalie et MEKKI Mustapha (dir.), *Le Juge et le Numérique : Un défi pour la justice du XXI^e siècle*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 115-119.

GANASCIA Jean-Gabriel (2019), « Peut-on contenir l’intelligence artificielle ? ». *Pouvoirs*, n° 170, p. 71-81.

GARAPON Antoine (2017), « Les Enjeux de la justice prédictive ». *La Semaine juridique*, n° 1-2, doctrine 31.

GARAPON Antoine et LASSÈGUE Jean (2018), *Justice digitale : révolution graphique et rupture anthropologique*. Paris, PUF.

GARAPON Antoine, ALLARD Julie et GROS Frédéric (textes réunis par) (2008), *Les Vertus du juge*. Paris, Dalloz.

GARENNE Shamine et MULLER Émilie (2019), « Présentation de la plateforme Médiation mise en ligne par le CNB », in SAYN Isabelle (éd.), *Séminaire E-Juris. De l’open data des décisions de justice à l’automatisation de la prise de décisions ? Les potentialités des bases de données décisionnelles*, MSH Lyon, p. 8-9, <https://halshs.archives-ouvertes.fr/halshs-02318411/document>

GARNOT Benoît (2000), « Justice, Infrajustice, Parajustice et Extrajustice dans la France d’Ancien Régime ». *Crime, Histoire et Sociétés/Crime, History and Societies*, vol. 4, n° 1, p. 103-120.

GAUDEMET Yves (2018), « Rapport de synthèse », in *Ordre des avocats au Conseil d’État et à la Cour de cassation, La Justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 109-22.

GAUTRON Virginie (2014), « La “Barémisation” et la Standardisation des réponses pénales saisies au travers d’une étude empirique de l’administration de la justice pénale », in SAYN Isabelle (dir.), *Le Droit mis en barèmes ?*. Paris, Dalloz.

GAUTRON Virginie et DUBOURG Émilie, (2014), « La Rationalisation des outils et méthodes d’évaluation : de l’approche clinique au jugement actuariel ». *Criminocorpus* [en ligne], n° 4, mis en ligne le 26 janvier 2015 [consulté le 10 août 2022].

GAYE-PALETTES Matthieu (2019), « Le Développement des outils algorithmiques prédictifs à l’épreuve de la question prioritaire de constitutionnalité ». *Annuaire international de justice constitutionnelle*, Economica/Presses universitaires d’Aix-Marseille, n° 35, p. 638-657.

GERRY-VERNIERES Stéphane (dir.) (2019), *La Barémisation de la justice*. Rapport Mission de recherche Droit et Justice / Centre de recherches juridiques, Grenoble.

GILLESPIE Tarleton (2012), « The Relevance of Algorithms », in GILLESPIE Tarleton, BOCZKOWSKI Pablo J. et FOOT Kirsten (eds) *Media Technologies: Essays on Communication, Materiality, and Society*, Cambridge, MIT Press.

GIRARD-CHANUDET Camille (2021), « Petites Mains et Cerveaux multiples de l'intelligence artificielle. Une ethnographie du travail de conception d'un outil de pseudonymisation automatique à la Cour de cassation ». Association française de science politique, Rencontres de la science politique, Groupe Normes et Justice, Nanterre.

GIRAULT Cécile (2018), « "En fait dans la justice c'est ça, mais partout, tu as des manques de personnels que tu pallies avec des béquilles" ». *Savoir/Agir*, vol. 46, n° 4, p. 107-115.

GIUDICELLI-DELAGE Geneviève (2003), « Introduction », in CADIET Loïc et RICHER Laurent (dir.), *Réforme de la justice, Réforme de l'État*. Paris, PUF, p. 139-141.

GODEFROY Lémy (2019a), « Encadrement juridique des modes algorithmiques d'analyse des décisions (Maad) », in GODEFROY Lémy, LEBARON Frédéric et [SINCLAIR] Jacques, *Comment le numérique transforme le droit et la justice*. Rapport pour le GIP Mission de recherche Droit et Justice, Paris, Université Nice Sophia Antipolis, ENS Paris-Saclay, INRIA, CNRS, p. 38-84.

GODEFROY Lémy (2019b), « Appropriation des modes algorithmiques d'analyse des décisions (Maad) par les magistrats », in GODEFROY Lémy, LEBARON Frédéric et [SINCLAIR] Jacques, *Comment le numérique transforme le droit et la justice*. Rapport pour le GIP Mission de recherche Droit et Justice, Paris, Université Nice Sophia Antipolis, ENS Paris-Saclay, INRIA, CNRS, p. 85-110.

GODEFROY Lémy, LEBARON Frédéric et [SINCLAIR] Jacques (2019), *Comment le numérique transforme le droit et la justice*. Rapport pour le GIP Mission de recherche Droit et Justice, Paris, Université de Nice-Sophia Antipolis, ENS Paris-Saclay, INRIA, CNRS.

GOËTA Samuel (2015), « Un air de famille : les trajectoires parallèles de l'open data et du big data ». *Informations sociales*, vol. 191, n° 5, p. 26-34.

GOËTA Samuel (2016), « Instaurer des données, instaurer des publics : une enquête sociologique dans les coulisses de l'open data ». Thèse de doctorat en sociologie, Télécom ParisTech.

GRESELLE-ZAÏBET Olfa, KLEBER Aurélie et DEJOUX Cécile (2018), « Le *Hackathon* en mode *Design Thinking* ou quelles modalités pour former à des compétences méthodologiques et comportementales ? ». *Management & Avenir*, n° 104, p. 151.

GRIS Christelle (2019), « La Mise en marche des "entrepreneurs de soi". Les ressorts de l'engagement local dans une "start-up politique" » in DOLEZ Bernard, FRETTEL Julien et LEFEBVRE Rémi (dir.), *L'Entreprise Macron*, Fontaine, Presses universitaires de Grenoble, p. 243-254.

GUILLAUME Henri et MACRON Emmanuel (2007), « Enseignement supérieur, recherche, innovation. Quels acteurs ? ». *Esprit*, n° 12, p. 160-187.

GUINCHARD Serge (2008), « L'Ambition raisonnée d'une justice apaisée ». Rapport à la ministre de la Justice. Paris, La Documentation française.

GUINCHARD Serge, VARINARD André et DEBARD Thierry (2021), *Institutions juridictionnelles*. Paris, Dalloz, coll. « Précis ».

GUYOMAR Mattias (2018), « Le Point de vue du juge. Table ronde 2 », in Ordre des avocats au Conseil d'État et à la Cour de cassation, *La Justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 97-101.

- HABERMAS Jürgen (1973), *La Technique et la Science comme idéologie*. Paris, Gallimard.
- HACKING Ian (1990), *The Taming of Chance*. Cambridge, Cambridge University Press.
- HALPERN Charlotte (2019), « Décision ». *Dictionnaire des politiques publiques*, 5^e éd., p. 199-206.
- HARCOURT Bernard E. (2007), *Against Prediction. Profiling, Policing, and Punishing in an Actuarial Age*. Chicago, The University of Chicago Press.
- HARCOURT Bernard E. (2011a), « Surveiller et Punir à l'âge actuariel. Généalogie et critique », part. I. *Déviance et société*, vol. 35, n° 1, p. 5-33.
- HARCOURT Bernard E. (2011b), « Surveiller et Punir à l'âge actuariel. Généalogie et critique », part. II. *Déviance et société*, vol. 35, n° 2, p. 163-194
- HASSENTEUFEL Patrick (2011). *Sociologie politique : l'action publique*. Paris, Armand Colin.
- HAWKINS Keith (2003), « Order, Rationality, and Silence: Some Reflections on Criminal Justice Decision-Making », in GELSTHORPE Loraine and PADFIELD Nicola, *Exercising Discretion: Decision-Making in the Criminal Justice System and Beyond*. Cullompton, Willan.
- HÉNIN Clément (2021), « Confier une décision vitale à une machine ». *Réseaux*, vol. 225, p. 187-213.
- HENRY Emmanuel (2021), *La Fabrique des non-problèmes. Ou comment éviter que la politique s'en mêle*. Paris, Presses de Sciences po.
- HERPIN Nicolas (1977), *L'Application de la loi. Deux poids, deux mesures*. Paris, Seuil, coll. « Sociologie ».
- HEYDEBRAND Wolf V. (1977), « The Context of Public Bureaucracies: An Organizational Analysis of Federal District Courts ». *Law and Society Review*, vol. 11, n° 5, p. 759-821.
- HILDEBRANDT Mireille (2018), « Law as Computation in the Era of Artificial Legal Intelligence: Speaking Law to the Power of Statistics ». *University Toronto Law Journal*, n° 68 (1), p. 12-35.
- HILGARTNER Stephen et BOSK Charles L. (1988), « The Rise and Fall of Social Problems: A Public Arenas Model ». *American Journal of Sociology*, vol. 94, n° 1, p. 53-78.
- HOOD Christopher (1991), « A Public Management for all Seasons? ». *Public Administration*, n° 69, p. 3-19.
- HOWARD Robert M., RANDAZZO Kirk A. (éd.) (2017), *Routledge Handbook of Judicial Behavior*. Londres, Routledge, <https://doi.org/10.4324/9781315691527>
- HUBIN Jean-Benoît, JACQUEMIN Hervé, MICHAUX Benoît (dir.) (2019), *Le Juge et l'Algorithme, juges augmentés ou justice diminuée*. Bruxelles, Larcier.
- HUSS Jean-Victor, LEGRAND Lucrèce et SENTIS Théo (2018), *Les Enjeux éthiques de la justice prédictive. Livre blanc*. Paris, Sciences po, Predictice & Wolters Kluwer.
- HUTTON Neil (1995), « Sentencing, Rationality, and Computer Technology ». *Journal of Law and Society*, vol. 22, n° 4, p. 549-570.
- HUTTON Neil (2013), « From Intuition to Database: Translating Justice ». *Theoretical Criminology*, vol. 17, n° 1, p. 109-128.

HYDE Aurore (2021), « La Justice prédictive : enjeux et perspectives », mis en ligne le 2 juillet 2019, actualisé depuis, <https://hal.archives-ouvertes.fr/hal-02149092>

IHL Olivier (2019), « Sciences de gouvernement », in BOUSSAGUET Laurie, JACQUOT Sophie et RAVINET Pauline (éd.), *Dictionnaire des politiques publiques*. Paris, Presses de Sciences po, 5^e éd., p. 563-569.

IHL Olivier et KALUSZYNSKI Martine (2002), « Pour une sociologie historique des sciences de gouvernement ». *Revue française d'administration publique*, n° 102, p. 229-243.

JACOB Robert (1995), « Le Jugement de Dieu et la Formation de la fonction de juger dans l'histoire européenne ». *Archives de philosophie du droit*, vol. 39, p. 87-104.

JACQUEMIN Zoé (2019), « Vices et Vertus de la justice prédictive en droit de la responsabilité civile », in BLANC Nathalie et MEKKI Mustapha (dir.), *Le Juge et le Numérique : Un défi pour la justice du XXI^e siècle*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 121-128.

JASANOFF Sheila (1995), *Science at the Bar. Law, Science, and Technology in America*. Cambridge, Harvard University Press.

JASANOFF Sheila et LECLERC Olivier (2013), *Le Droit et la Science en action*. Paris, Dalloz.

JEAMMAUD Antoine (2004), « Judiciarisation/Déjudiciarisation », in CADIET Loïc (dir.), *Dictionnaire de la justice*. Paris, PUF, p. 675-679.

JEAN Jean-Paul (2020), « Du chiffre et du sens en droit pénal. À propos de la sévérité des juges affamés », in ROUSSEAU François, GRUNVALD Sylvie et ROUSSEL Gildas (dir.), *Mélanges en l'honneur de Jean Danet*. Paris, Dalloz, p. 403-414.

JEANDIDIER Bruno, RAY Jean-Claude et MANSUY Julie (2020), « Analyses quantitatives de décisions de justice en matière de prestation compensatoire dans une perspective de justice prédictive ». *Working Papers of BETA*. Strasbourg, Bureau d'économie théorique et appliquée, Université de Strasbourg, <https://ideas.repec.org/p/ulp/sbbeta/2020-05.html>

JEANNOT Gilles (2020), « Vie et mort de l'État plateforme ». *Revue française d'administration publique*, n° 173, p. 165-179.

JEULAND Emmanuel (2017), « Justice prédictive : de la factualisation au droit potentiel ». *Revue pratique de la prospective et de l'innovation*, n° 2, p. 9-12.

JEULAND Emmanuel (2019), « Justice numérique, justice inique ? ». *Les Cahiers de la Justice*, n° 2, p. 193-199.

JOBARD Fabien et NÉVANEN Sophie (2007), « La Couleur du jugement. Discriminations dans les décisions judiciaires en matière d'infractions à agents de la force publique (1965-2005) ». *Revue française de sociologie*, vol. 48, n° 2, p. 243-272.

JOBERT Bruno (1992), « Représentations sociales, Controverses et Débats dans la conduite des politiques publiques ». *Revue française de science politique*, vol. 42, n° 2, p. 219-234.

JONCKHEERE Alexia (2013), *(Dés)équilibres. L'informatisation du travail social en justice*. Bruxelles, Larcier.

JOUËT Josiane (2000), « Retour critique sur la sociologie des usages ». *Réseaux*, 18, 100, 487-521.

- JUSTON MORIVAL Romain (2018), « Des maux et des chiffres. L'évaluation des incapacités en médecine légale du vivant ». *Sciences sociales et Santé*, vol. 36, n° 4, p. 41-64.
- JUSTON MORIVAL Romain (2021b), « Autonomie des juges ou automatisme des jugements ? La qualification judiciaire à l'épreuve des expertises médico-légales ». *Sociologie*, vol. 12, n° 4, p. 333-349.
- JUSTON MORIVAL Romain (2021a), *Médecins légistes. Une enquête sociologique*. Paris, Presses de Sciences po.
- JUSTON-MORIVAL Romain (2016), « Comment une tache de sang devient-elle une preuve ? Ingrédients et recettes des preuves médico-légales ». *Droit et Société*, vol. 93, n° 2, p. 395-416.
- KAMINSKI Dan (2015), *Condamner. Une analyse des pratiques pénales*. Toulouse, éd. Érès.
- KATZ Daniel Martin (2013), « Quantitative Legal Prediction – or – How I Learned to Stop Worrying and Start Preparing for the Data Driven Future of the Legal Services Industry? ». *Emory Law Journal*, vol. 62, mis en ligne le 11 décembre 2012, actualisé en 2013 [consulté le 28 juin 2019], https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2187752
- KATZ Daniel Martin, BOMMARITO Michael J. II et BLACKMAN Josh (2017), « A general approach for predicting the behavior of the Supreme Court of the United States ». *PLoS ONE* 12(4), <https://doi.org/10.1371/journal.pone.0174698>
- KERN Christoph A. (2017), « Le Prix des services juridiques ». *Revue internationale de droit économique*, t. 31, n° 4, p. 45-59.
- KESTENARE Emmanuel (2018), « Justice prédictive et protection juridique : quel apport dans notre relation client ? ». *Archives de philosophie du droit*, t. 60, p. 271-278.
- KINGDON John W. (1984), *Agendas, Alternatives and Public Policies*. Boston, Little Brown and Co [2^e éd. : New York, Longman, 1995].
- KIRAT Thierry (2012), *Économie du droit*. Paris, La Découverte.
- KOMTER Martha L. (2012), « Conversation Analysis in the Courtroom », in SIDNELL Jack, STIVERS Tanya (éd.), *The Handbook of Conversation Analysis*. Hoboken, Wiley-Blackwell, p. 612-629.
- LABORIER Pascale et TRÖM Danny (2003), « Introduction », in LABORIER Pascale et TRÖM Danny (dir.), *Historicités de l'action publique*. Paris, PUF, coll. « Curapp ».
- LACOUR Stéphanie et PIANA Daniela (2019), « Faites entrer les algorithmes ! Regards critiques sur la "justice prédictive" ». *Cités*, vol. 80, n° 4 (30 décembre), p. 47-60.
- LAMY Erwan (2014), « Chercheurs créateurs », in Pierre-Marie CHAUVIN, Michel GROSSETTI et Pierre-Paul ZALIO (dir.), *Dictionnaire de l'entrepreneuriat*, Paris, Presses de Science po, p. 101-113.
- LAMY Jérôme (2017), « Controverses et STS : stop ou encore ? ». *Zilsel*, vol. 2, n° 2, p. 123-130.
- LARA Miguel et LOCKWOOD Kate (2016), « Hackathons as Community-Based Learning: a Case Study ». *TechTrends*, vol. 60, n° 5, p. 486-495.

- LARMINAT Xavier de (2013), « La Technologie de mise à distance des condamnés en France. La centralisation informatique des données socio-judiciaires ». *Déviance et Société*, vol. 37, p. 359-373.
- LARRÈGUE Julien (2019), « Gouverner l'institution judiciaire par les nombres : une généalogie de la "justice prédictive" ». *Droit pénal*, n° 4, p. 8-11.
- LASCOURMES Pierre (2019), « Controverse » in Laurie Boussaguet (éd.), *Dictionnaire des politiques publiques*. Paris, Presses de Sciences po, 5^e éd., p. 170-177.
- LASCOURMES Pierre et LE GALÈS Patrick (2004), « L'Action publique saisie par ses instruments », in LASCOURMES Pierre et LE GALÈS Patrick (dir.), *Gouverner par les instruments*. Paris, Presses de Sciences po, p. 11-44.
- LATOUBRUNO Bruno (1994), « On Technical Mediation ». *Common Knowledge*, vol. 3, n° 2, p. 29-64.
- Le Collectif Onze (2013), *Au tribunal des couples. Enquête sur des affaires familiales*. Paris, Odile Jacob.
- LEBARON Frédéric (2019), « Réflexion sociologique sur les outils numériques d'analyse mathématique du droit », in GODEFROY Lémy, LEBARON Frédéric et [SINCLAIR] Jacques, *Comment le numérique transforme le droit et la justice*. Rapport pour le GIP Mission de recherche Droit et Justice, Paris, Université Nice Sophia Antipolis, ENS Paris-Saclay, INRIA, CNRS, p. 111-144.
- LEBRETON-DERRIEN Sylvie (2018), « La Justice prédictive. Introduction à une justice "simplement" virtuelle ». *Archives de philosophie du droit*, t. 60, p. 3-21.
- LEMIEUX Cyril (2007), « À quoi sert l'analyse des controverses ? ». *Mil neuf cent. Revue d'histoire intellectuelle*, n° 25, p. 191-212.
- LENOIR Rémy (1996), « Processus pénal et Mise en détention provisoire ». *Droit et Société*, vol. 30-31, p. 357-365.
- LÉONARD Thomas (2010), « Ces papiers qui font le jugement. Inégalités entre Français et étrangers en comparution immédiate ». *Champ pénal / Penal Field*, vol. 7 [en ligne].
- LÉONARD Thomas et MORITZ Marcel (2020), « Outils de "justice prédictive". Enjeux et cartographie sociologique des professionnels concernés ». GIP Mission de recherche Droit et Justice/CERAPS Université Lille 2
- Les Cahiers de la justice* (2015), « Des juges sous influence ». Dossier spécial, n° 4.
- LEVY René (1987), *Du suspect au coupable, le travail de police judiciaire*. Genève / Paris, Médecine et hygiène / Méridiens Klincksieck.
- LICOPPE Christian et DUMOULIN Laurence (2019), « Le Travail des juges et les algorithmes de traitement de la jurisprudence. Premières analyses d'une expérimentation de "justice prédictive" en France ». *Droit et Société*, n° 103, p. 535-554.
- LICOPPE Christian (2008), « Dans le "carré de l'activité" : perspectives internationales sur le travail et l'activité ». *Sociologie du travail*, vol. 50, n° 3, p. 287-302.

- LOUVARIS Antoine (2018), « La Justice prédictive entre être et devoir-être », in *Ordre des avocats au Conseil d'État et à la Cour de cassation, La Justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 23-38.
- LOUVEL Bertrand (2018), « Communication à l'occasion du colloque pour le bicentenaire de l'ordre des avocats au Conseil d'État et à la Cour de Cassation », in *Ordre des avocats au Conseil d'État et à la Cour de cassation, La Justice prédictive*. Paris, Dalloz, p. 15-18.
- LOWI Theodor J. (1972), « Four Systems of Policy, Politics, and Choice ». *Public Administration Review*, vol. 32, n° 4, p. 298-310.
- LYNCH Michael et McNally Ruth (2005), « “Science”, “Sens commun” et Preuve ADN : une controverse judiciaire à propos de la compréhension publique de la science ». *Droit et Société*, vol. 61, n° 3, p. 55-681.
- MABI Clément (2015), « La Plateforme “data.gouv.fr” ou l'Open Data à la française ». *Informations sociales*, vol. 5, n° 191, p. 52-59.
- MAILLARD Jacques de, et KÜBLER Daniel (2019), *Analyser les politiques publiques*, Presses universitaires de Grenoble.
- MALABAT Valérie (2019), « Justice prédictive et droit pénal substantiel », in BLANC Nathalie et MEKKI Mustapha (dir.), *Le Juge et le Numérique : Un défi pour la justice du XXI^e siècle*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 105-114.
- MALAPERT Baptiste (2020), « Commentaire de la décision State v. Loomis de la Cour suprême du Wisconsin : le droit à un procès équitable à l'épreuve de l'opacité algorithmique dans les systèmes de justice américain et européen », mis en ligne le 26 février 2020 [consulté le 23 août 2022], <https://blogs.parisnanterre.fr/article/commentaire-de-la-decision-state-v-loomis-de-la-cour-supreme-du-wisconsin-le-droit-un-0>
- MARCH James G. et OLSEN Johan P. (1976), *Ambiguity and Choice in Organizations*. Bergen, Universitetsforlaget.
- MARGETTS Helen (2003), « Electronic Government: A Revolution in Public Administration? », in PETERS B. Guy, PIERRE Jon (éd.), *Handbook of Public Administration*, London, Sage Publications, p. 366–377.
- MARGETTS Helen et DUNLEAVY Patrick (2013), « The Second Wave of Digital-Era Governance: a Quasi-Paradigm for Government on the Web ». *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 371.
- MARIN Jean-Claude (2018), « Communication à l'occasion du colloque pour le bicentenaire de l'ordre des avocats au Conseil d'État et à la Cour de cassation », in *Ordre des avocats au Conseil d'État et à la Cour de cassation, La Justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 19-22.
- MARONIKOLAKIS Antonios *et al.* (2020), « Analyzing Political Parody in Social Media », in *Association for Computational Linguistics, Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* [en ligne], p. 4373-4384, <https://doi.org/10.18653/v1/2020.acl-main.403>
- MARQUET Clément (2016), « Faire du smartphone un instrument de la relation de service ? Handicap, mobilité et infrastructure d'accessibilité ». *Réseaux*, n° 200.

MARTIGNY Vincent et STRUDEL Sylvie (2019), « De la campagne à la présidence Macron : vers la fin du “en même temps” ? » in DOLEZ Bernard, FRETTEL Julien et LEFEBVRE Rémi (dir.), *L'Entreprise Macron*, Fontaine, Presses universitaires de Grenoble, p. 127-138.

MARY Philippe (2001), « Pénalité et Gestion des risques : vers une justice “actuarielle” en Europe ? ». *Déviance et Société*, vol. 25, n° 1, p. 33-51.

MATHIS Bruno (2018), « Rapport Cadet : une vision frileuse de l'open data des décisions de justice ». *La Semaine juridique entreprise et affaires*, n° 7, p. 1086.

MATHIS Bruno (2019), « Open Data des décisions de justice : un projet à moyens constants », mis en ligne le 12 décembre 2019 [consulté le 16 avril 2021], <https://www.village-justice.com/articles/open-data-des-decisions-justice-projet-moyens-constants,33215.html>.

MAVEETY Nancy (éd.) (2003), *The Pioneers of Judicial Behavior*. Chicago, The University of Michigan Press.

MEDVEDEVA Masha, VOLS Michel et WIELING, Martijn (2019), « Using Machine Learning to Predict Decisions of the European Court of Human Rights ». *Artificial Intelligence and Law*, vol. 27, n° 2, <https://link.springer.com/article/10.1007/s10506-019-09255-y>

MEMMI Daniel (1990), « Connexionnisme, Intelligence artificielle, et Modélisation cognitive ». *Intellectica. Revue de l'Association pour la recherche cognitive*, n° 9-10, p. 41-79.

MERTON Robert K. (1968), « The Self-Fulfilling Prophecy », in *Social Theory and Social Structure*. New York, The Free Press, p. 475-490.

MICHAUT Françoise (2000), *La Recherche d'un nouveau paradigme de la décision judiciaire à travers un siècle de doctrine américaine*. Paris, L'Harmattan.

MILBURN Philipp (2003), « L'Avocat et l'Algorithme ». *Les Cahiers de la justice*, n° 3, p. 499-510.

MILLARD Éric (2014), « Réalisme scandinave, réalisme américain ». *Revus* [en ligne], vol. 24, mis en ligne le 22 août 2014 [consulté le 19 octobre 2022] <http://journals.openedition.org.iepnomade-2.grenet.fr/revus/2915>

MINTZBERG Henry (1982), *Structure et dynamique des organisations*. Paris, Éditions d'organisation.

[MOLINE Laurent] (2017). « La Justice prédictive, version moderne de la boule de cristal ». *RTD civil*, n° 2, p. 527-28.

MOLINIÉ François (2018), « Le Point de vue d'un avocat. Table ronde 1 », in *Ordre des avocats au Conseil d'État et à la Cour de cassation, La Justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 64-70.

[MONTA Géraud] (2018), « Du droit saisi par l'IA au droit saisissant l'IA, éléments de réflexion ». *Archives de philosophie du droit*, t. 60, p. 103-116.

[MONTA Géraud] et [SINCLAIR] Jacques (2017), « L'Intelligence artificielle au service de la valorisation du patrimoine jurisprudentiel ». *Dalloz IP/IT*, n° 10, p. 500-505.

MOUHANNA Christian et BASTARD Benoit (2010). « Procureurs et Substituts : l'évolution du système de production des décisions pénales ». *Droit et Société*, vol. 74, n° 1, p. 35-53.

- MOUILLARD Agnès (2018), « Présentation générale de la table ronde 2 : L'office du juge », in Ordre des avocats au Conseil d'État et à la Cour de cassation, *La Justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 81-84.
- MOURIESSE Élise (2018), « Quelle transparence pour les algorithmes de justice prédictive ? ». *Archives de philosophie du droit*, t. 60, p. 125-145.
- MULLER Pierre (2015), *La Société de l'efficacité globale*. Paris, PUF.
- MULLER Pierre et SUREL Yves (1998), *L'Analyse des politiques publiques*. Paris, Montchrestien, coll. « Clefs ».
- NAGEL Stuart S. (1963), « Applying Correlation Analysis to Case Prediction ». *Texas Law Review*, n° 42, p. 1006.
- NAY Olivier et SMITH Andy (2002), « Les Intermédiaires en politique. Médiation et jeux d'institutions », in NAY Olivier et SMITH Andy (dir.), *Le Gouvernement du compromis : courtiers et généralistes dans l'action politique*. Paris, Economica, p. 1-21.
- NEGRON Éric, BORIES Serge Charles et DA SILVA Jérémias (2017), « À la recherche d'une justice perdue ? » *La Semaine juridique*, n° 26, p. 1246-1249.
- NEUMAN Russel *et al.* (2014), « The Dynamics of Public Attention: Agenda-Setting Theory Meets Big Data ». *Journal of Communication*, vol. 64, p. 193-214.
- NEVEU Érik (2015), *Sociologie politique des problèmes publics*. Paris, Armand Colin.
- NEYLAND Daniel (2019), *The Everyday Life of an Algorithm*. Londres, Palgrave Macmillan.
- NOUIRI-MANGOLD Sabrina (2012), « Se distinguer dans un espace standardisé. Le cas des greffières aux affaires familiales », in COTON Christel, PROTEAU Laurence, *Les Paradoxes de l'écriture. Sociologie des écrits professionnels dans les institutions d'encadrement*. Rennes, Presses universitaires de Rennes, coll. « Le sens social », p. 105-125.
- O'NEIL Cathy (2016), *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, Crown Publishers.
- OFFERLE Michel (2019), « "Les Patrons" ou "des patrons" avec Emmanuel Macron. Capitaux entrepreneuriaux et capital politique », in DOLEZ Bernard, FRETTEL Julien et LEFEBVRE Rémi (dir.), *L'Entreprise Macron*, Fontaine, Presses universitaires de Grenoble, p. 79-92.
- OLIVESI Aurélie et HUBÉ Nicolas (2016), Présentation du dossier « Des vraies gens aux followers ». *Politiques de communication*, n° 6, p. 5-17.
- Ordre des avocats au Conseil d'État et à la Cour de cassation (2018), *La Justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires ».
- OST François (2007), « Jupiter, Hercule, Hermès : trois modèles du juge », in *Dire le droit, faire justice*. Bruxelles, Bruylant, p. 33-60 [1^{re} éd. 1991].
- OWENS Boyce Robert (2016), « Judicial Decision Making as Knowledge Work ». *Law & Social Inquiry*, vol. 41, n° 2, 502-21.
- PALIWALA Abdul (2010), « A History of Legal Informatics: An Introduction to the Special Issue ». *European Journal of Law and Technology*, vol. 1, n° 1.

- PASQUALE Frank (2015). *The Black Box Society: the Secret Algorithms that Control Money and Information*. Cambridge, Harvard University Press.
- PASQUALE Frank et CASHWELL Glyn (2018), « Prediction, Persuasion, and the Jurisprudence of Behaviourism ». *University of Toronto Law Journal*, vol. 68, n° 1, p. 63-81.
- PAWSON Ray (2021), « The Coronavirus Response: Boxed in by Models ». *Evaluation*, vol. 27, n° 2, p. 149-167.
- PÉRÉ Dominique (2018), « Le Point de vue de l'éditeur. Table ronde 2 », in *Ordre des avocats au Conseil d'État et à la Cour de cassation, La Justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 85-88.
- PHILIPPE Arnaud (2022), *La Fabrique des jugements. Comment sont déterminées les sanctions pénales*. Paris, La découverte.
- [PIERRE Arthur], [LAVAUDAN Georges] et [MARIANO Nicolas] (2016), « Les Enjeux du projet Open Law *Le droit ouvert ». *I2D - Information, données & documents*, vol. 53, n° 1, p. 23-25.
- PIERSON Paul (2000), « Increasing Returns, Path Dependence, and the Study of Politics ». *The American Political Science Review*, vol. 94, n° 2, p. 251–267.
- [PLACE Simon] (2014), « Le Référentiel indicatif régional d'indemnisation du préjudice corporel », in SAYN Isabelle (dir.), *Le Droit mis en barèmes ?*. Paris, Dalloz.
- POLLITT Christopher et BOUCKAERT Geert (2004), *Public Management Reform: A Comparative Analysis*. Oxford, Oxford University Press, 2^e éd.
- PORTER Theodore M. (2020[1995]), *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, Princeton University Press.
- PRATES Fernanda (2011), « Le Jugement comme pratique sociale : une réflexion sur la justice pénale brésilienne ». *Champ pénal / Penal Field*, vol. 8 [en ligne].
- PROULX Serge (2015), « La Sociologie des usages, et après ? ». *Revue française des sciences de l'information et de la communication* [en ligne], n° 6, mis en ligne le 23 janvier 2015 [consulté le 14 octobre 2022].
- QUIJOUX Maxime et SAINT-MARTIN Arnaud (2020), « Start-up : avènement d'un mot d'ordre ». *Savoir/Agir*, n° 51, p. 15-22.
- RAVINET Pauline et PALIER Bruno (2015), « "Global, es-tu là ?" Réflexions sur une catégorie fantomatique », in BOUSSAGUET Laurie, JACQUOT Sophie et RAVINET Pauline (éd.), *Une French touch dans l'analyse des politiques publiques ?*. Paris, Presses de Sciences po, p. 27-55.
- REILING Dory (2021), « Court Information Technology: Hypes, Hopes and Dreams », in KRAMER Xandra et al. (éd.), *New Pathways to Civil Justice in Europe: Challenges of Access to Justice*. Cham, Springer International Publishing, p. 43-60, https://doi.org/10.1007/978-3-030-66637-8_3.
- REVERDY Thomas (2021), *Antimanuel de management de projet*. Paris, éd. Dunod.
- REYNAUD Jean-Daniel (1988), « Les Régulations dans les organisations : régulation de contrôle et régulation autonome ». *Revue française de sociologie*, p. 5-18.

RHATTAT Rachid (2019), « L'Incubateur du Barreau de Paris : un dispositif d'accompagnement entrepreneurial dédié à l'innovation juridique ». *Entreprendre & Innover*, vol. 40, n° 1, p. 57-58.

RICCEUR Paul (1985), « L'Acte de juger ». *Le Juste*. Paris, éd. Esprit, p. 185-192.

RIEDER Bernard et SMYRNAIOS Nikos (2012), « Pluralisme et Infomédiation sociale de l'actualité : Le cas de Twitter ». *Réseaux*, vol. 176, n° 6, p. 105-139.

ROBERT Philippe, FAUGERON Claude et KELLENS Georges (1972), *Les Attitudes des juges à propos des prises de décision. Prérecherche exploratoire*. Paris, ministère de la Justice, service d'études pénales et criminologiques.

ROSANVALLON Pierre (2006), *La Contre Démocratie : la Politique à l'âge de la défiance*. Paris, Seuil.

ROSENTHAL Claude (2017), « Mettre en scène des technologies dans un salon de l'innovation. Le cas Vivatech 2017 ». *Occasional paper CEM*, n° 50, <http://cems.ehess.fr/index.php?4183>.

ROTTIER Edouard (2018), « La Justice prédictive et l'Acte de juger ». *Archives de philosophie du droit*, vol. 60, p. 189-93.

ROUBAN Luc (2018), *Le Paradoxe du macronisme*. Paris, Presses de Sciences po.

ROUVROY Antoinette et BERNS Thomas (2013), « Gouvernamentalité algorithmique et Perspectives d'émancipation ». *Réseaux*, vol. 177, n° 1, p. 163-196.

SAAS Claire, LORVELLEC Soizic et GAUTRON Virginie (2013), « Les Sanctions pénales : une nouvelle distribution », in DANET Jean (dir.), *La Réponse pénale. Dix ans de traitement des délits*. Rennes, Presses universitaires de Rennes, p. 159-210.

SABATIER Paul A. et JENKINS-SMITH Hank C. (1999), « The Advocacy Coalition Framework, an Assessment », in SABATIER Paul A. (éd.), *Theories of the Policy Process*, Boulder, Westview Press, p. 117-166.

SABATIER Paul A. (2019), « Advocacy Coalition Framework », in BOUSSAGUET Laurie, JACQUOT Sophie et RAVINET Pauline (dir.), *Dictionnaire des politiques publiques*. Paris, Presses de Sciences po, p. 46-53.

SABATIER Paul A. et JENKINS-SMITH Hank C. (éds) (1993), *Policy Change and Learning. An Advocacy Coalition Approach*, Boulder, Westview Press.

SADIN Éric (2018), *L'Intelligence artificielle ou l'Enjeu du siècle : Anatomie d'un antihumanisme radical*. Paris, L'échappée.

SAINT-MARTIN Denis (2014), « Expertise ». *Dictionnaire des politiques publiques*, 4^e éd., p. 266-274.

SANTISO Carlos (2020), « Gilles Babinet, *Refondre les politiques publiques avec le numérique*, Dunod, 2020 ». *Revue française d'administration publique*, vol. 177, n° 1, p. 249-252.

SAPIR Yoav (2008), « Against Prevention ? A Response to Harcourt's *Against Prediction* on Actuarial and Clinical Predictions and the Faults of Incapacitation ». *Law & Social Inquiry*, vol. 33, n° 1, p. 253-264.

SAUVÉ Jean-Marc (2018), « Communication à l'occasion du colloque pour le bicentenaire de l'Ordre des avocats au Conseil d'État et à la Cour de cassation », in *Ordre des avocats au Conseil d'État et à la Cour de cassation, La Justice prédictive*. Paris, Dalloz, coll. « Thèmes et commentaires », p. 7-13.

SAUVE Jean-Marc (2022), *Rendre justice aux citoyens*. Rapport du Comité des États généraux de la justice (octobre 2021-avril 2022). Paris, La Documentation française.

SAYN Isabelle (2019), *Les Barèmes (et autres outils techniques d'aide à la décision) dans le fonctionnement de la justice*. Rapport CERCRID / GIP Mission de recherche Droit et Justice.

SAYN Isabelle (2019c), « Connaître la production des juridictions ou prédire les décisions de justice ? ». *Les Cahiers de la Justice*, vol. 2, n° 2, p. 229-242.

SAYN Isabelle (2020), « La Banalisation des outils d'aide à la décision : une dynamique de transformation sans réforme ? ». *La Revue juridique Themis*, n° 54, p. 337-362, https://ssl.editionsthemis.com/uploaded/revue/article/21311_12-Sayn.pdf

SAYN Isabelle (dir.) (2014), *Le Droit mis en barèmes ?*. Paris : Dalloz, coll. « Thèmes et commentaires ».

SAYN Isabelle (éd.) (2019a), *Séminaire E-Juris. De l'open data des décisions de justice à l'automatisation de la prise de décisions ? Les potentialités des bases de données décisionnelles*, MSH Lyon, <https://halshs.archives-ouvertes.fr/halshs-02318411/document>

SAYN Isabelle (éd.) (2019b), « État du droit, ou Quel régime juridique des décisions individuelles automatiques ? », in SAYN Isabelle (éd.), *Séminaire E-Juris. De l'open data des décisions de justice à l'automatisation de la prise de décisions ? Les potentialités des bases de données décisionnelles*, MSH Lyon, <https://halshs.archives-ouvertes.fr/halshs-02318411/document>

SBRICCOLI Mario (2007), « Histoire sociale, dimension juridique : l'historiographie italienne récente du crime et de la justice criminelle ». *Crime, Histoire & Sociétés*, vol. 11, n° 2, mis en ligne le 1^{er} mars 2011.

SEGAL Jeffrey A. (1984), « Predicting Supreme Court Cases Probabilistically: The Search and Seizure Cases, 1962–1981 ». *American Political Science Review*, vol. 78, n° 4, p. 891-900.

SEGAL Jeffrey A. (2011), « Judicial Behavior », in GOODIN Robert (éd.), *The Oxford Handbook of Political Science* [en ligne].

SEGAL Jeffrey A. et SPAETH Harold J. (2002), *The Supreme Court and the Attitudinal Model Revisited*. Cambridge, Cambridge University Press.

SERVERIN Évelyne (1993), « Juridiction et Jurisprudence : deux aspects des activités de justice ». *Droit et Société*, n° 25, p. 339-349.

SERVERIN Évelyne (2018), « De l'Informatique juridique aux services de justice prédictive, la longue route de l'accès du public aux décisions de justice dématérialisées ». *Archives de philosophie du droit*, t. 60, p. 23-47.

SERVERIN Évelyne, LASCOUMES Pierre et LAMBERT Thierry (1987), *Transactions et Pratiques transactionnelles. Sujets et objets des transactions dans les relations conflictuelles de droit privé et de droit public*. Paris, Economica.

SFEZ Lucien (1992), *Critique de la décision*. Paris, Presses de Sciences po, 4^e éd.

SIMIONI Melchior et STEINER Philippe (2022). *Comment ça matche : Une sociologie de l'appariement*. Paris, Presses de Sciences po.

SMYRNAIOS Nikos et RATINAUD Pierre (2013), « Comment articuler analyse des réseaux et des discours sur Twitter ». *Tic & Société*, 7(2).

SNACKEN Sonja, VAN ZYL SMIT Dirk et BEYENS Kristel (2013), « Sentencing », in BODYGENDROT Sophie et al., *The Routledge Handbook of European Criminology*. London, Routledge.

STAR Suzan L. et RUHLER Karen (1996), « Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces ». *Information Systems Research*, vol. 7, n° 1, p. 111-134.

STARR Sonja B. (2014), « Evidence-Based Sentencing and the Scientific Rationalization of Discrimination ». *Stanford Law Review*, vol. 66, n° 4, p. 803-872.

STEINER Philippe (2022), « De corps à corps. L'appariement dans la transplantation rénale », in SIMIONI Melchior et STEINER Philippe (dir.), *Comment ça matche. Une sociologie de l'appariement*. Paris, Presses de Sciences po, p. 295-322.

STONE SWEET Alec (2000), *Governing with Judges. Constitutional Politics in Europe*. Oxford / New York, Oxford University Press.

[STRIPPOLI Thibault] (2017), « La Justice à l'épreuve de sa prédictibilité ». *Actualités juridiques Pénal*, 334.

SULEA Octavia-Maria et al. (2017), « Predicting the Law Area and Decisions of French Supreme Court Cases » [en ligne], <https://arxiv.org/pdf/1708.01681.pdf>.

SUNSHINE Jason et TYLER Tom R. (2003), « The role of Procedural Justice and Legitimacy in Shaping Public Support for Policing ». *Law and Society Review*, vol. 37, n° 3, p. 513-548.

SUNSTEIN Cass R. (éd.) (2001), *Behavioral Law and Economics*. Cambridge, Cambridge University Press.

SUSSKIND Richard et SUSSKIND Daniel (2015), *L'Avenir des professions libérales : Comment les nouvelles technologies vont transformer le travail des experts humains*. Oxford, Oxford University Press.

TAGNY-NGOMPÉ Gildas et al. (2019), « Detecting Sections and Entities in Court Decisions Using HMM and CRF Graphical Models », in PINAUD Bruno et al. (éd.), *Advances in Knowledge Discovery and Management*. Cham, Springer International Publishing, p. 61-86.

TAGNY-NGOMPÉ Gildas et al. (2020), « Identification of Judicial Outcomes in Judgments: A Generalized Gini-PLS Approach ». *Stats*, n° 4, p. 427-443.

TARTOUR Tonya (2020), *L'Administration du désordre : gouverner l'hôpital psychiatrique depuis les années 1980*. Thèse de doctorat en sociologie, IEP de Paris / CSO.

TARTOUR Tonya (2021), « De la couture institutionnelle. Fabrique et usage des certificats médicaux dans le cadre du contrôle judiciaire des hospitalisations sans consentement en psychiatrie ». Séminaire jeunes chercheur·euse·s, *Handicap(s) et Sociétés*. Paris, [hal-03522655](https://hal.archives-ouvertes.fr/hal-03522655)

- TATA Cyrus (1997), « Conceptions and Representations of the Sentencing Decision Process ». *Journal of Law & Society*, vol. 27, p. 395-420.
- TATE C. Neal, (1983), « The Methodology of Judicial Behavior Research: A Review and Critique ». *Political Behavior*, vol. 5, n° 1, p. 51-82.
- TERESI Laurent (2016), « L'Open Data et le Droit de l'Union européenne ». *Actualité juridique. Droit administratif*, p. 87-91.
- THERY Irène (2001), *Le Démariage. Justice et vie privée*. Paris, Odile Jacob, coll. « Opus », 2^e éd.
- THÉVENOT Laurent (1986), « Les investissements de forme », in THÉVENOT Laurent (dir.). *Conventions économiques*. Paris, PUF (Cahiers du Centre d'études de l'emploi).
- THÉVENOT Laurent (1992), « Jugements ordinaires et Jugement de droit ». *Annales. Économies, Sociétés, Civilisations*, vol. 47, n° 6.
- THIBOUT, Charles (2019), « La Compétition mondiale de l'intelligence artificielle ». *Pouvoirs*, n° 170, vol. 3, p. 131-142.
- THIEULIN Benoît (2018), « Gouverner à l'heure de la révolution des pouvoirs ». *Pouvoirs*, n° 164, p. 19-31.
- THOENIG Jean-Claude (1985), « L'Analyse des politiques publiques », in GRAWITZ Madeleine et LECA Jean (dir.). *Traité de science politique*, vol. 4. Paris, PUF.
- TIMSIT Gérard (1993), *Les Figures du jugement*. Paris, PUF, coll. « Les voies du droit ».
- TOPALOV Christian (1999). *Laboratoires du nouveau siècle. La nébuleuse réformatrice et ses réseaux en France (1880-1914)*. Paris, Éditions de l'EHESS.
- TULKENS Françoise (1990), « La Sécurité juridique : un idéal à reconsidérer ». *Revue interdisciplinaire d'études juridiques*, 24, p. 25-42.
- TYLER Tom R. (1988), « What is Procedural Justice? Criteria Used by Citizens to Assess the Fairness of Legal Procedures ». *Law and Society Review*, vol. 22, n° 1, p. 103-134.
- TYLER Tom R. (2006), *Why People Obey the Law*. Princeton, Princeton University Press.
- UBALDI Barbara (2013), « L'Administration électronique, support de l'innovation dans les services publics ». *Revue française d'administration publique*, p. 449-464.
- VAN DEN BRANDEN Adrien (2019), *Les robots à l'assaut de la justice : l'intelligence artificielle au service des justiciables*. Bruxelles, Bruylant.
- VANHAMME Françoise et BEYENS Kristel (2007), « La recherche en sentencing : un survol contextualisé ». *Déviance et Société*, vol. 31, n° 2, p. 199-228.
- VANNEUVILLE Rachel (2008), « Sécuriser le droit pour mieux gouverner les conduites : les enjeux sociopolitiques de la promotion contemporaine de la sécurité juridique ». *Revue Gouvernance*, vol. 5, n° 2.
- VAUCHEZ Antoine (2004), *L'Institution judiciaire remotivée. Le processus d'institutionnalisation d'une « nouvelle justice » en Italie (1964-2000)*. Paris, LGDJ.

VAUCHEZ Antoine (2013), *L'Union par le droit. L'invention d'un programme institutionnel pour l'Europe*. Paris, Presses de Sciences po.

VAUCHEZ Antoine (2016), « Une République des avocats fonctionnaires ? Enquête aux frontières des espaces politico-administratifs et du barreau d'affaires ». *Les Cahiers de la Justice*, vol. 1, n° 1, p. 117-127.

VERHEIJ Bart (2020), « Artificial Intelligence as Law: Presidential Address to the Seventeenth International Conference on Artificial Intelligence and Law ». *Artificial Intelligence and Law*, vol. 28, n° 2, p. 181-206. <https://doi.org/10.1007/s10506-020-09266-0>

VERMEYS Nicolas (2021), « The Computer as the Court: How Will Artificial Intelligence Affect Judicial Processes? », in KRAMER Xandra et al. (éd.), *New Pathways to Civil Justice in Europe: Challenges of Access to Justice*. Cham, Springer International Publishing, p. 61-80, https://doi.org/10.1007/978-3-030-66637-8_4

[VIAL Thomas] (2018), « Quel avenir pour la “justice prédictive” ? Enjeux et limites des algorithmes d'anticipation des décisions de justice ». *La Semaine juridique – édition générale*, n° 7, p. 190.

[VIAL Thomas] (2020), *L'Intelligence artificielle en procès. Plaidoyer pour une réglementation internationale et européenne*. Bruxelles, Bruylant.

VIGOUR Cécile (2015), « Professions in Policy and Knowledge Transfer: Adaptations of Lean Management, and Jurisdictional Conflict in a Reform of the French Public Service ». *International Journal of Sociology*, vol. 45, n° 2, p. 112-132.

VIGOUR Cécile (2019), *La Justice à l'épreuve de la nouvelle gestion publique. Sociologie de la gestionnarisation des organisations publiques*. Mémoire d'HDR en science politique, IEP de Paris.

VIGOUR Cécile et al. (2022), *La justice en examen. Attentes et expériences citoyennes*. Paris, PUF, coll. « Le Lien social ».

VILLANI Cédric (2019), « Donner un sens à l'intelligence artificielle. Pour une stratégie nationale et européenne ». Rapport au Premier ministre, https://www.aiforhumanity.fr/pdfs/9782111457089_Rapport_Villani_accessible.pdf

WEINSHALL-MARGEL Keren et SHAPARD John (2011), « Overlooked Factors in the Analysis of Parole Decisions, Israeli Courts Research Division, The Supreme Court, Jerusalem and Federal Judicial Center, Washington ». *PNAS*, vol. 108, n° 42.

WELLER Jean-Marc (1999), *L'État au guichet. Sociologie cognitive du travail et modernisation administrative des services publics*. Paris, Desclée de Brouwer.

WELLER Jean-Marc (2011), « Comment décrire ce qu'on ne voit pas ? Le devoir d'hésitation des juges de proximité au travail ». *Sociologie du travail*, vol. 53, n° 3, p. 349-368.

WELLER Jean-Marc (2018), *Fabriquer des actes d'État. Une ethnographie du travail bureaucratique*. Paris, Economica.

WEST Darrell M. (2005), *Digital Government: Technology and Public Sector Performance*. Princeton, Princeton University Press.

YOLKA Philippe (2016), « Open data : “L’ouverture, c’est l’aventure” ». *Actualité juridique. Droit administratif*, p. 79.

ZAMBRANO Guillaume (2015), « Précédents et Prédictionns jurisprudentielles à l’ère des big data », document de recherche, <https://hal.archives-ouvertes.fr/hal-01496098>

Table des graphiques, encadrés et tableaux

Graphique 2.1. Dendrogramme représentant le résultat de la classification de Reinert. Corpus d'articles de presse « prédictif.ve » (2010-2020).....	87
Graphique 2.2. Distribution de la variable année au sein de chacune des classes. Corpus d'articles de presse (n=1 622), 2010-2020.....	89
Graphique 2.3. 15 mots les + représentatifs de la classe 2	90
Graphique 2.4. Nuage des 100 mots les plus significatifs de la classe 2	90
Graphique 2.5. Variable année de la classe 2	90
Encadré 2.1. Les 5 segments de texte les plus caractéristiques de la classe 2	92
Graphique 2.6. Nuage représentant les 100 mots les plus significatifs de la classe 3.....	95
Encadré 2.2. Les cinq segments de texte les plus caractéristiques de la classe 3. Corpus d'articles de presse, mot-clef « prédictif.ve » 2010-2020	96
Graphique 2.7. Nuage représentant les 100 mots les plus significatifs de la classe 4.....	96
Encadré 2.3. Les cinq segments de texte les plus caractéristiques de la classe 4	97
Graphique 2.8. Distribution annuelle du nombre d'articles de presse 'justice prédictive'. Corpus 2010-2020	100
Graphique 2.9. Distribution annuelle du nombre d'articles de presse 'justice prédictive'. Corpus 2010-2020 selon le type de publication.....	100
Graphique 2.10. Nombre de tweets émis par année sur la période 2010-2020	101
Graphique 2.11. Nombre de tweets émis par mois sur la période 2010-2020.....	102
Tableau 2.1. La 'justice prédictive' sur Twitter, de 2010 à 2020 : Deux modèles à l'origine des pics d'intensité	105
Tableau 2.2. Les 20 termes les plus fréquents au sein du corpus de tweets 2010-20	108
Tableau 2.3. Les 20 termes les plus fréquents au sein du corpus de presse 2010-20.....	108
Graphique 2.12. Nuage des 100 mots les plus fréquents du corpus de tweets 2010-20.....	108
Graphique 2.13. Nuage des 100 mots les plus fréquents du corpus de presse 2010-20	108
Graphique 2.14. Dendrogramme représentant le résultat de la classification de Reinert ...	111
Graphique 2.15. Projection factorielle des classes identifiées. Corpus 2010-2020.....	112
Graphique 2.16. Les mots les plus significatifs de la classe 2	113
Graphique 2.17. Nuage de mots représentant les 100 mots les plus significatifs de la classe 2	113
Graphique 2.18. Analyse de graphe du mot tester.....	113

Graphique 2.19. Les mots les plus significatifs de la classe 4	114
Graphique 2.20. Nuage de mots représentant les 100 mots les plus significatifs de la classe 4	114
Graphique 2.21. Analyse de graphe du mot expérimentation	114
Encadré 2.4. Les 5 tweets les plus caractéristiques de la classe 2.....	115
Encadré 2.5. Les 5 tweets les plus caractéristiques de la classe 4.....	115
Tableau 2.4. Récapitulatif des énoncés médiatiques dont les mots sont présents dans les mots les significatifs des classes 2 et 4.....	116
Graphique 2.22. Les mots les plus significatifs de la classe 8	117
Graphique 2.23. Nuage de mots représentant les 100 mots les plus significatifs de la classe 8	117
Graphique 2.24. Distribution de la variable genre au sein de chacune des classes	118
Encadré 2.6. Les 5 tweets les plus caractéristiques de la classe 8.....	118
Graphique 2.25. Distribution de la variable #justicepredictive au sein de chacune des classes	119
Graphique 2.26. Distribution de la variable presencementions au sein de chacune des classes	119
Graphique 2.27. Distribution de la variable retweets au sein de chacune des classes	120
Graphique 2.28. Distribution de la variable favorites au sein de chacune des classes	120
Graphique 2.29. La distribution des tweets émis par [Anajur] au sein des classes.....	121
Graphique 2.30. La distribution des tweets émis par [PréviCompute] au sein des classes ..	121
Graphique 2.31. Les mots les plus significatifs de la classe 6	122
Graphique 2.32. Nuage de mots représentant les 100 mots les plus significatifs de la classe 6	122
Encadré 2.7. Les 5 tweets les plus caractéristiques de la classe 6.....	122
Graphique 2.33. Les mots les plus significatifs de la classe 7	123
Graphique 2.34. Nuage de mots représentant les 100 mots les plus significatifs de la classe 7	123
Encadré 2.8. Les 5 tweets les plus caractéristiques de la classe 7.....	124
Graphique 2.35. Distribution de la variable réponse au sein de chacune des classes	124
Graphique 2.36. Distribution de la variable replies au sein de chacune des classes.....	125
Graphique 2.37. Les mots les plus significatifs de la classe 1	126
Graphique 2.38. Nuage de mots représentant les 100 mots les plus significatifs de la classe 1	126
Encadré 2.9. Les articles médiatiques de la classe 1.....	126

Encadré 2.10. Les 5 tweets les plus caractéristiques de la classe 1.....	127
Graphique 2.39. Les mots les plus significatifs de la classe 3	128
Graphique 2.40. Nuage de mots représentant les mots les plus significatifs de la classe 3 .	128
Encadré 2.11. L'article viral au cœur de la classe 3.....	128
Encadré 2.12. Les 5 tweets les plus caractéristiques de la classe 3.....	128
Graphique 2.41. Les mots les plus significatifs de la classe 9	129
Graphique 2.42. Nuage de mots représentant les 100 mots les plus significatifs de la classe 9	129
Encadré 2.13. Les 5 tweets les plus caractéristiques de la classe 3.....	129
Encadré 2.14. Les articles de presse de la classe 3	130
Graphique 2.43. Les mots les plus significatifs de la classe 5	130
Graphique 2.44. Nuage de mots représentant les 100 mots les plus significatifs de la classe 5	130
Graphique 2.45. Distribution de la variable genre au sein de chacune des classes	131
Encadré 2.15. Les 5 tweets les plus caractéristiques de la classe 5.....	132
Graphique 2.46. Distribution de la variable #ia au sein de chacune des classes.....	132
Graphique 2.47. Distribution de la variable #legaltech au sein de chacune des classes.....	133
Tableau 2.5. La 'justice prédictive' sur Twitter, de 2010 à 2020 : pics d'intensité et classes de discours	135
Graphique 3.1. Les anglicismes dans le menu du site de l'association Open Law (bandeau en haut de page)	162
Graphique 3.2. : Les relations entre les différentes structures d'action collective autour de l'informatique juridique puis de l'open data du droit	163
Graphique 3.3. Les acteurs de la coalition particulariste de l'open data	182
Tableau 3.1. Les croyances des acteurs de la coalition particulariste de l'open data.....	182
Graphique 3.4. Les acteurs de la coalition généraliste de l'open data	187
Tableau 3.2. Les croyances des acteurs de la coalition généraliste de l'open data	187
Tableau 4.1. Composition du jury Open Case Law 2016.....	198
Graphique 4.1. Distribution mensuelle de la publication des articles de presse du sous-corpus 'justice prédictive', 2010-2020.....	202
Graphique 4.2. Nombre de tweets émis par jour au cours des mois de septembre, octobre et novembre 2017.	203
Graphique 4.3. Distribution de la variable antepost9oct au sein de chacune des classes, articles de presse, sous corpus 'justice prédictive' 2010-2020	204

Graphique 4.4. Evolution chronologique des surreprésentations des classes de discours dans le sous corpus 'justice prédictive' d'articles de presse	205
Graphique 4.5. Les mots les plus significatifs et le poids de la classe 6 du sous-corpus d'articles de presse (2010-2020) et des classes 2 et 4 du sous-corpus de tweets (2010-2020)	206
Graphique 4.6. La distribution des tweets émis par [PréviCompute] au sein des classes. ...	207
Graphique 4.7. La distribution des tweets émis par [Anajur] au sein des classes.	208
Encadré 4.1. L'analyse prédictive par [PréviCompute]	209
Graphique 4.8. Corpus de tweets 'justice prédictive' 2010-2019.....	218
Encadré 4.2. Échange de <i>tweets</i> quant à l'enjeu de définition de la 'justice prédictive'.	220
Encadré 4.3. Les 5 <i>tweets</i> les plus caractéristiques de la classe 5, corpus de <i>tweets</i> 'justice prédictive', 2010-2020	222
Graphique 5.1. La classe 5, poids, mots représentatifs. Corpus d'articles de presse 'justice prédictive', 2010-2020	237
Graphique 5.2. Nuage de mots représentant les 100 mots les plus significatifs de la classe 5. Corpus d'articles de presse 'justice prédictive', 2010-2020	237
Graphique 5.3. Graphe du mot calculer au sein de la classe 5. Corpus d'articles de presse 'justice prédictive', 2010-2020.....	238
Encadré 5.1. Les 5 segments de texte les plus caractéristiques de la classe 5. Corpus d'articles de presse 'justice prédictive', 2010-2020	238
Graphique 5.4. Les jugements entre deux pôles	241
Graphique 5.6. Analyse de graphe du mot jugement. Tweets 'justice prédictive', 2010-2020	244
Graphique 5.7. Analyse de graphe du mot décision. Tweets 'justice prédictive', 2010-2020	244
Graphique 5.8. Analyse de graphe du mot procès. Tweets 'justice prédictive', 2010-2020.	245

Table des matières

Remerciements	1
Table des sigles.....	5
Introduction générale.....	7
I / La « société des algorithmes »	7
1.1. <i>Big data</i> et algorithmes	7
1.2. <i>E-gouvernement, open data</i> et modèles prédictifs dans l'action publique	8
II / La justice dans la société algorithmique	10
2.1. 'justice prédictive' et intelligence artificielle.....	10
2.2. La 'justice prédictive' à la croisée des chemins.....	10
2.3. Prédiction et justice actuarielle.....	16
III / Ce que nous disent les controverses : question de recherche et problématique	19
3.1. La 'justice prédictive', un objet carrefour	19
3.2. Controverses sur la 'justice prédictive'	20
3.3. L'analyse de dispositifs algorithmiques de traitement des décisions de justice.....	22
3.4. Justice pénale ou civile ?	23
3.5. Deux enjeux de connaissance.....	24
IV / Des terrains à géométrie et granulométrie variables.....	28
4.1. Articles de presse	28
4.2. <i>Tweets</i>	29
4.3. Archives et traces	30
4.4. Entretiens semi-directifs et observations.....	30
4.5. Anonymiser oui, mais comment ?	31
V / Annonce du plan : le contenu des chapitres	32
Chapitre 1 : Les algorithmes dans la justice, de l'effervescence des expériences à l'analyse de leurs effets.....	37
I / Une controverse sur la performance des algorithmes dans la justice.....	39
1.1. Des expériences disruptives concernant les décisions de justice ?.....	40
1.2. Ces modèles prédictifs sont-ils si neufs et performants ?.....	46
II / Une controverse autour des effets possibles des algorithmes sur le système juridique et judiciaire français	54
2.1. Peut-on prédire le jugement ?	54
2.2. Sécurité juridique accrue ou standardisation des décisions ?	57

2.3.	Déjudiciarisation des conflits vs accès à la justice.....	62
2.4.	Libération vs disparition des professionnels du droit	64
2.5.	Transformations profondes du système juridique et du rôle du droit	65
2.6.	Quelle régulation ? <i>Soft law</i> , autorégulation ou réglementation ?	69
III /	Les algorithmes à l'épreuve du terrain	71
3.1.	Que disent concepteurs et magistrats des dispositifs algorithmiques ?.....	71
3.2.	Les algorithmes à travers leurs usages.....	76
Chapitre 2 : La 'justice prédictive', controverse autour d'une notion floue.....		83
I /	Une justice a-technologique confrontée à la modernité ?	86
1.1	Variations autour du tournant prédictif.....	86
1.2.	Du pénal au civil : le double sens de la 'justice prédictive'	90
1.3.	Anticipation et délégation à des machines	93
1.4.	De quoi l'avenir sera-t-il fait ? Stupeurs et tremblements.....	94
II /	La trajectoire ascendante de l'expression 'justice prédictive' au milieu des années 2010.....	99
1.2.	Des trajectoires comparables dans la presse et les tweets : le pic de 2017	99
1.3.	Analyse de séquences et de pics d'intensité.....	102
III /	La 'justice prédictive', une innovation algorithmique qui pose question.....	107
1.4.	La rencontre du droit, des innovations et de l'université	107
1.5.	La controverse sur les algorithmes d'analyse des décisions de justice civiles	110
1.6.	La dynamique de la controverse, à travers le croisement des phases et des classes de discours.....	134
1.7.	Comparaison avec l'analyse lexicale des articles de presse.....	137
Chapitre 3 : La traduction de la politique d' <i>open data</i> dans le secteur de la justice		143
I /	La politique d' <i>open data</i> , résultat d'une nouvelle configuration de réforme	145
1.1.	Le répertoire de réforme de l' <i>open data</i> : État-plateforme, méthode agile et <i>startup Nation</i> 145	
1.2.	Une configuration de réforme favorable à la « <i>startup Nation</i> »	151
1.3.	L'expression d'une nouvelle norme publique	155
II /	Un répertoire de réforme promu par Open Law	157
2.1.	De Juriconnexion et l'Adij à Open Law	158
2.2.	Des parcours hybrides et multipositionnés : les courtiers	159
2.3.	Compétition, collaboration et hackathon : le « catéchisme » de l'innovation numérique ..	160
2.4.	Une « communauté de la donnée » pour un service public des données juridiques	162
2.5.	Le Village de la legaltech : un événement-carrefour qui incarne le rapport public-privé et les valeurs d'ouverture du droit et d'innovation par les technologies	163

III / Trois temps forts dans l' <i>open data</i> des décisions de justice.....	166
3.1. Un décret qui conforte le pouvoir des hautes juridictions et des magistrats dans la mise en œuvre de l' <i>open data</i> dans la justice	166
3.2. Une loi qui bride les usages de décisions déjà accessibles.....	169
3.3. Une circulaire qui empêche la récupération de décisions de justice en masse.....	170
IV / Coalitions de cause autour de l'alignement de la justice sur la politique d' <i>open data</i> ou l'incarnation d'un rapport sectoriel-global conflictuel	171
4.1. Deux coalitions de cause qui se constituent progressivement	173
4.2. Des <i>policy brokers</i> qui font le lien entre les coalitions de cause.....	188
4.3. À travers deux coalitions, un rapport global-sectoriel conflictuel	189
Chapitre 4 : Un nouveau répertoire de réforme en action : la fabrique d'une expérimentation de 'justice prédictive' auprès de magistrats	193
I / L'expérimentation [PréviCompute], fruit d'un concours et incarnation de la 'justice prédictive'	195
1.1. Un module de calcul et de data-visualisation né d'un hackaton	195
1.2. [PréviCompute], l'expérimentation et son « échec » au cœur de la controverse sur la 'justice prédictive'	201
1.3. [PréviCompute], point de focalisation et porte-parole de la 'justice prédictive'	206
II / L'expérimentation [PréviCompute], une épreuve pour la 'justice prédictive'	211
2.1. Des <i>startups</i> pas tellement alignées.....	211
2.2. L'opposition autour de l'expression 'justice prédictive' : [PréviCompute] vs [CLA]	216
III / Quels effets de la controverse ?	221
3.1. La 'justice prédictive' existe	221
3.3. Détricotages, ripostes, apprentissages et repositionnements.....	223
3.4. Un faisceau de réalisations visant à utiliser et rendre comparables les décisions de justice	226
Point(s) de suspension	233
Du jugement « produit fini » au jugement comme action équipée	233
I / La relation entre algorithmes et jugements : éléments empiriques	233
1.1. Une focalisation sur l'adjudication exprimée sous forme quantifiée par le juge.....	234
1.2. ... pour mieux déjudiciariser et développer le règlement amiable	239
1.3. Une vision instrumentale des dispositifs algorithmiques qui préserve la liberté du juge	241
1.4. Twitter, une arène où le jugement apparaît comme la chose du juge	242
II / Une conception décisionniste en décalage avec la réalité du fonctionnement judiciaire	246
2.1. Décision, rationalité et décisionnisme	246

2.2.	Décisionnisme, droit et justice	248
2.3.	Le décisionnisme malgré la sociologie du droit et de la justice	251
III /	Comment élaborer une analyse différente des dispositifs algorithmiques de traitement et d'analyse des décisions de justice qui ne repose pas sur cette conception décisionniste du jugement ?.....	255
3.1.	Clarifier la notion de jugement : l'épaisseur du jugement contre l'irruption de la décision	255
3.3.	Intégrer les acquis de la sociologie pragmatique : jugements ordinaires et jugements de droit	258
3.4.	S'intéresser au travail de <i>decision-making</i> et à sa matérialité.....	260
	Conclusion générale	269
	Bibliographie	273
	Table des graphiques, encadrés et tableaux.....	303
	Table des matières	307

Titre : De quoi la 'justice prédictive' est-elle le nom ? Algorithmes, décision et jugement

Mots clés : Sociologie de l'action publique, Configuration de réforme administrative, Dispositifs algorithmiques, Innovation, Intelligence artificielle, *Open data*, *Legaltech*, Magistrats, Avocats

Résumé : La 'justice prédictive' est une construction sémantique récemment apparue. Elle désigne un ensemble disparate de plateformes en ligne qui, à partir d'un grand nombre de décisions de justice, dont le traitement est assuré par différentes méthodes mathématiques, statistiques et/ou algorithmiques, simulent différentes décisions de justice possibles sur un cas donné et évaluent leur probabilité respective. Ce mémoire propose une sociologie politique de la rencontre entre ces dispositifs algorithmiques de traitement et d'analyse des décisions de justice d'une part et le monde du droit et de la justice d'autre part. Il revient sur la genèse, le cadrage et les enjeux de ces dispositifs. La montée en puissance de la 'justice prédictive' est interprétée comme la manifestation de l'émergence d'une nouvelle configuration de réforme administrative qui s'est mise en place en France depuis les années 2010 pour favoriser l'*open data* des décisions de justice et encourager

l'innovation dans le droit. Une attention particulière est accordée à la manière dont les jugements juridictionnels sont conçus dans la controverse sur la 'justice prédictive' : la vision mythifiée du jugement comme pure décision du seul juge y est encore prégnante. En contrepoint, le mémoire propose un cadre analytique qui conceptualise le jugement comme le résultat d'une activité pratique, équipée par des dispositifs sociotechniques et qui découle d'un processus de formation non linéaire et marqué par une certaine épaisseur. La diversité des configurations de jugement observables est également pointée.

Ce travail repose sur des investigations empiriques quantitatives et qualitatives, menées à partir de sources diversifiées (traces et archives, *tweets* et articles de presse, observations).

Title : What is 'justice prédictive' all about ? Algorithms, decision and judicial decision-making

Keywords : Sociology of public action, Configuration of administrative reform, Algorithmic devices, Innovation, Artificial intelligence, Legal analytics, Open data, Judges, Lawyers

Abstract : 'Predictive justice' is a term that has recently emerged. It refers to a disparate set of online platforms, which are based on a large number of court decisions, processed by mathematical, statistical and/or algorithmic methods. They simulate several possible court decisions on a given case and evaluate their respective probability. This HDR manuscript is centred upon a political sociology of the encounter between these algorithmic devices for processing and analysing rulings on the one hand and the world of law and justice on the other. It looks back at the genesis, the framing and the stakes of these devices. The rise of 'predictive justice' is interpreted as the manifestation of the emergence of a new configuration of administrative reform that has been in place in France since the 2010s, to promote open data on court decisions and encourage

innovation in the law.

Particular attention is paid to the way in which adjudication is conceived in the 'predictive justice' controversy: the mythologised view of judgement as a pure decision of the judge alone is still prevalent. As a counterpoint, the manuscript proposes an analytical framework that conceptualises ruling as the result of a practical activity equipped with socio-technical devices and stemming from a thick and non-linear process. The diversity of observable judgement configurations is also pointed out.

This work is based on quantitative and qualitative methods, carried out from diversified (traces and archives, tweets and press articles, observations).

