



**HAL**  
open science

# La plateforme de données de santé Health data hub : Une impossible gouvernance éthique des données massives ?

Bernard Fallery

## ► To cite this version:

Bernard Fallery. La plateforme de données de santé Health data hub : Une impossible gouvernance éthique des données massives ?. *Revue Française de Gestion*, 2021. halshs-03368851

**HAL Id: halshs-03368851**

**<https://shs.hal.science/halshs-03368851v1>**

Submitted on 7 Oct 2021

**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.

# **La plateforme de données de santé *Health data hub* : Une impossible gouvernance éthique des données massives ?**

**Bernard Fallery**

Version provisoire, en révision avant-impression  
publication dans la Revue Française de Gestion, 2021

## **Résumé**

La centralisation des masses de données de santé publique pour servir des projets d'Intelligence Artificielle a été menée au pas de charge par le *Health data hub* depuis 2018, mais tous les choix effectués (techniques, juridiques, éthiques...) ont été rapidement et fortement contestés. La gouvernance de données massives et la protection des données de santé sont apparues non pas simplement comme des problèmes à résoudre, mais bien comme des questions à penser. Les boucles dialogiques et l'éthique de reliance, deux grands principes de la pensée complexe développés par Edgar Morin, permettent alors une analyse du discours des différents acteurs dans le cadre d'une recherche-action. Les résultats font apparaître trois épisodes, dans lesquels des impératifs différents se succèdent suivant le thème, l'énonciation et la temporalité de chacune des situations éthiques qui se sont concrétisées. Accepter la diversité de ces chaînes de situations éthiques ne permet certes pas d'avancer une recette pour une soit-disant « décision » éthique, mais c'est ici « travailler à penser de façon complexe » une gouvernance des données massives de santé qui devrait pouvoir s'auto-organiser devant et avec l'ensemble des collectifs concernés.

Mots clés : Health data hub, données massives, pensée complexe, dialogique, reliance

## **Abstract**

The centralization of the masses of public health data to serve Artificial Intelligence projects has been carried out at a charge by the *Health data hub* since 2018, but all the choices made (technical, legal, ethical, etc.) were quickly and strongly contested. Big data governance and health data protection emerged not just as problems to solve, but as questions to be think about. Dialogic loops and the ethics of reliance, two main principles of complex thinking developed by Edgar Morin, then allow an analysis of the discourse of the different actors within the framework of action research. The results reveal three episodes, in which different imperatives follow one another according to the theme, the enunciation and the temporality of each of the ethical situations which have materialized. Accepting the diversity of these chains of situations is certainly not putting forward recipes for a so-called ethical decisions, but it is "working to think in a complex way" about governance of massive health data which should be able to auto-organize in front of and with all the groups concerned.

Keywords : health data hub, big data, complex thinking, dialogic, reliance

## **Introduction**

Dès la remise en mars 2018 du rapport dirigé par C. Villani « Pour donner un sens à l'intelligence artificielle », le Président de la République annonce le projet *Health data hub*. En juin une mission de préfiguration est confiée à J-M Aubert (un moment directeur de la DRESS, puis retourné à la société Iqvia, Le Monde [2019 Lien1](#)). Dès octobre un rapport définit les traits

d'un système national centralisé de toutes les données de santé publique issues de l'assurance maladie, des hôpitaux, de la recherche, des dossiers patients... (figure 1). Il s'agit de « *faire de la France un leader dans l'utilisation des données de santé au service du bien commun* » (A. Buzyn, Ministre, [2018 Lien2](#)), avec un guichet unique d'accès aux données à partir duquel l'Intelligence Artificielle pourrait optimiser des services de reconnaissance des formes et de prédiction personnalisée. Voilà qui semble une question urgente, posée de façon claire comme une innovation réputée disruptive, sans doute compliquée par les risques liés à la sécurité, mais vue comme non complexe. Et pourtant...

La « résolution du problème », confiée dès novembre 2018 au GIP *Health data hub* dirigé par Stéphanie Combes, s'est rapidement heurtée au « sens du problème », devant l'ensemble de toutes les questions qui ont été soulevées (questions politiques, techniques, juridiques, économiques, sociales, éthiques... Fallery, [2020](#)). Les réactions des professionnels de santé, des professionnels du Logiciel Libre, de députés, du Conseil national des barreaux... ont été rapidement relayées par un collectif d'organisations et de personnalités regroupées sous le nom SantéNathon. Sont alors petit à petit apparus tous les risques de multiples détournements liés aux interdépendances entre la souveraineté numérique, la centralisation ou l'inter-opérabilité, le respect des données de santé définies par le règlement RGPD (CNIL [2018](#)), les communs numériques, le pouvoir des GAFAM, la société de surveillance, les règlements européens, le consentement, le secret médical... des questions finalement bien plus complexes que seulement compliquées.

Le cas apparaît en effet typique des problèmes qui ne peuvent pas « se poser », mais qui doivent « se penser ». Oppositions, paradoxes, contradictions, dépendance à l'environnement, récursivité, évolution, niveaux multiples et finalités multiples, dualité, émergence... sont ici présentes pratiquement toutes les caractéristiques relevées par E. Morin pour définir le champ de la pensée complexe : « *Je dirais d'abord que le complexe est ce qui ne peut se réduire à une description claire, à une idée simple, non plus qu'à une loi simple... La connaissance complexe, elle, vise à reconnaître ce qui lie ou relie l'objet à son contexte, le processus ou l'organisation où il s'inscrit* » (Morin, 2008, p. 182).

Dans ce cadre, notre objectif est ici triple :

- discuter dans une première partie des possibilités qu'offre la Pensée complexe pour fournir non seulement des outils de compréhension mais aussi un cadre d'action, autour de deux principes : celui de boucle dialogique (Morin, 1982) et celui d'éthique de reliance (Morin, 2004) ;
- présenter dans une deuxième partie ce cas de gouvernance de données massives de santé vécu par l'auteur-chercheur-acteur, où les différentes problématiques s'enchaînent progressivement dans l'action à partir de juin 2020 autour de deux questions successives, celle du choix de Microsoft comme intégrateur et celle du respect des données privées de santé ;
- analyser dans une troisième partie une grande partie des discours tenus tout au long de cette action de contestation du *Health data hub*, ce qui permet d'une part de caractériser trois épisodes successifs et d'autre part de valider une proposition théorique dans le cadre de la pensée complexe : Une chaîne de situations éthiques correspond à des moments de rencontre entre plusieurs impératifs apparaissant sur différentes « roues dialogiques ». La discussion et la conclusion permettent de revenir sur la question d'une « impossible » gouvernance éthique des données massives de santé dans l'IA numérique.

## **I. LE CADRE THÉORIQUE DE LA PENSÉE COMPLEXE : BOUCLE DIALOGIQUE ET ÉTHIQUE DE RELIANCE**

De la souveraineté numérique au secret médical, en passant par l'inter-opérabilité ou la société de surveillance... les interdépendances sont tellement manifestes dans le cas du projet *Health data hub* qu'elles ne peuvent pas relever d'une démarche de décomposition analytique. La mobilisation de la Pensée complexe proposée par Edgar Morin apparaît presque comme une évidence. Parmi tous les concepts développés dans les six tomes de *La Méthode*, vouloir isoler ceux de boucle dialogique et d'éthique de reliance peut sembler comme particulièrement contraire à la pensée de l'auteur (Morin 1982). Nous nous protégerons en montrant que le principe dialogique et le principe de reliance sont tellement fondamentaux dans la pensée complexe que les mettre en avant ne représente pas un trop grand cloisonnement, ni donc une trop grande trahison.

### 1.1 Le principe dialogique (Morin 1982)

L'essence du principe dialogique est plus facile à saisir en commençant par l'opposer, même de façon un peu caricaturale, au principe de dialectique. La dialectique hégélienne est une pensée logique des contradictions, thèse/antithèse, lesquelles doivent être dépassées pour conduire à une synthèse, un retour à l'unité ; la dialectique marxiste considère même que ces contradictions doivent être recherchées en tant que moteur des processus de transformation réels (Sève et Guespin-Michel, 2005). Mais au lieu de vouloir « résoudre » ainsi les contradictions par une synthèse ou une transformation, la dialogique « *assume cette tension en deux notions devant s'exclure l'une l'autre, mais qui sont indissociables en une même réalité* » (Le Moigne et Morin 2007, p. 78).

La Pensée complexe permet ainsi l'assemblage de notions antagonistes (individu/espèce/société, féminin/masculin, rationnel/irrationnel, yin/yang...) ; la logique n'est pas binaire mais « uni-duale », plusieurs logiques différentes sont liées en une unité, mais sans que la dualité ne se perde dans l'unité complexe et dynamique. Il ne s'agit pas de hiérarchiser ou de fondre les idées opposées : De même qu'en biologie chaque cellule contient le code génétique de l'organisme, le tout n'enlève pas les parties, parce que tout et parties sont reliés de façon intrinsèque et émergente. E. Morin en donne cet exemple assez clair « *Ce qui fait l'unité de la culture européenne ce n'est pas la synthèse judéo-christiano-gréco-romaine, c'est le jeu non seulement complémentaire, mais aussi concurrent et antagoniste entre ces instances qui ont chacune leur propre logique : C'est justement leur dialogique* » (Morin, 1987, p. 28).

D'un point de vue théorique le principe dialogique est au fondement de la pensée complexe (Peyron-Bonjan, 1998). Edgar Morin le montre dans un court et lumineux chapitre (Morin, 1982, p. 181-191) qui justifie le tétragramme ordre-désordre-organisation-interaction (figure 3) :

- « *La nécessité de penser ensemble, dans leur complémentarité, dans leur concurrence et dans leur antagonisme, les notions d'ordre et de désordre nous pose exactement le problème de penser la complexité de la réalité physique, biologique et humaine* » (p. 183) ;
- « *L'idée de désordre est non seulement in-éliminable de l'univers, elle est nécessaire pour le concevoir dans sa nature et dans son évolution* » (p.189) ;
- « *L'idée enrichie d'ordre est beaucoup plus riche que l'idée de déterminisme, puisqu'elle appelle les idées d'interaction et d'organisation, puisqu'elle ne peut expulser le désordre* » (p. 185) ;
- « *Avec la notion de structure, l'idée d'ordre appelle une autre idée, qui est l'idée d'organisation* (p. 184) ».



qu'il est impossible de concevoir et assurer un souverain Bien » (p. 221). L'éthique est complexe parce qu'elle est à la fois une et multiple, affrontant l'ambiguïté, la contradiction et l'incertitude, considérées comme des valeurs sociales qui génèrent des liens : « en latin *complexus, ce qui est tissé ensemble* »;

- l'éthique de reliance est un processus itératif, ne cherchant pas l'optimisation mais se régénérant sans cesse en « *travaillant à bien penser* » (p. 251), un comportement satisfaisant au sens de H. Simon (Simon 1976). « *L'éthique complexe est inévitablement modeste* » (Morin 2004, p. 250). « *Le combat essentiel de l'éthique c'est la double résistance à la cruauté du monde et à la cruauté humaine* » (p. 246), « *Essayer de réduire la cruauté humaine c'est miser sur les forces faibles de reliance* » (p. 227) ;

- L'éthique de reliance se fonde sur la conviction que tout autre est un sujet, un alter ego ; elle est une éthique de communauté, au sens de fraternité sociale, de responsabilité et de solidarité. L'éthique est ou devrait être un impératif de reliance, à l'inverse de l'exclusion : « *Relié* » est passif, « *reliant* » est participant, « *reliance* » est activant. On peut parler de *déliance* pour l'opposé de « *reliance* » (p. 269), même s'il peut y avoir de mauvaises reliesances (le fanatisme) et de bonnes déliesances (la désaliénation).

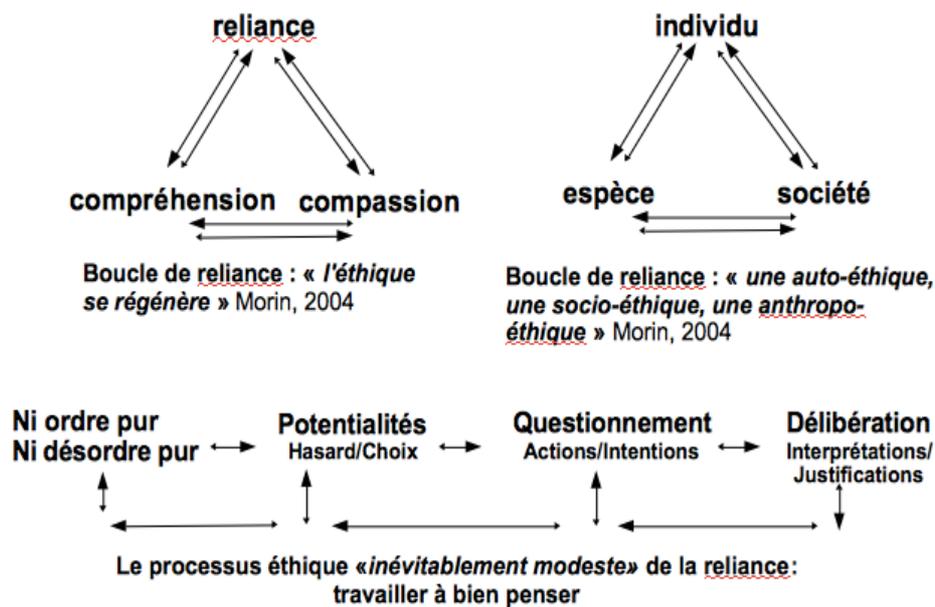


Figure 4. Le principe d'éthique de reliance : complexité, communauté, itérations, réflexivité

D'un point de vue opérationnel nous ne trouverons donc pas ici de recettes, ni même de maximes qui permettraient de trancher un conflit éthique. Même si l'impératif de reliance conduit à formuler quelques règles (« *Notre lien consubstantiel avec la biosphère nous conduit à abandonner le rêve prométhéen de la maîtrise de la nature pour l'aspiration à la convivialité sur terre* », p. 185), ces règles sont loin d'être des programmes. La pensée complexe est une méthode pour créer du lien à partir de conjonctions (Vallat 2017) et l'application du principe éthique de reliance faite par G. Berger est éclairante (Peyron-Bonjan et Berger, 2007):

- « *Douter, questionner, problématiser, réfléchir, combiner ou croiser plusieurs regards, et plusieurs mises en scène, somme toute inventer et créer une méthode, la sienne, à tout moment : Aides-toi ; la méthode t'aidera* » ;

- « *Revenir à une sorte d'indistinction, là où les formes n'ont pas encore pris forme, là où la pluralité des formes possibles est potentielle, là où n'existent ni le désordre pur, ni l'ordre pur mais l'indistinction de l'un et de l'autre, source de toutes pluralités* » ;

- « Bref, il nous faut, en permanence, une auto-interrogation, un auto-questionnement, un auto-contrôle, une auto-réflexion, une auto-critique, une éthique et une auto-éthique pour accéder à la plénitude d'être-humain ou d'être citoyen ».

Notre responsabilité éthique reste donc celle du « travailler à bien penser » (Morin 2005), selon l'expression de Pascal, celle d'une prise de conscience et d'un travail réflexif sur la formation de nos représentations et aussi de celles des autres (Genelot 2014), dans des situations de hasard et de choix. Pour E. Morin, « travailler à bien penser, c'est à dire de penser de façon complexe » (T6, p. 70), « Aide-toi, et la complexité t'aidera » (Morin, 2005).

Sur le plan individuel c'est une « auto-éthique du questionnement » de nos comportements cognitifs et de nos intentions morales, dans des situations d'actions et d'intentions. Sur le plan collectif c'est une « socio-éthique de la délibération », dans des situations d'interprétations et de justifications (figure 4). Ces deux dimensions de l'éthique vont clairement apparaître en interaction dans le projet *Health data hub*.

## II. LA PLATEFORME DE DONNÉES DE SANTÉ HDH

En mars 2018 au Collège de France, le Président de la République annonce le projet *Health data hub*, une décision prise à la suite du rapport Villani sur l'Intelligence Artificielle remis la veille. Il s'agit de transformer le Système national des données de santé (SNDS) qui avait été créé en 2016 à partir des données médico-administratives, des données individuelles de santé issues de la carte vitale, de la tarification PMSI des établissements et des causes médicales de décès. En novembre 2018, Stéphanie Combes est nommée cheffe de projet.

Dès fin 2018, le choix de Microsoft est déjà acté comme prestataire pour la gestion des données (« en dispense de marché public », TicSanté 2019-Lien3), alors même que la définition des principes du *Health data hub* attendront juillet 2019 (dans la Loi Santé) et que ses missions ne seront définies par arrêté ministériel qu'en avril 2020 (figure1). Dès le départ la CNIL s'inquiétait de multiples finalités non explicitées et elle avait prévenu : « Au-delà d'un simple élargissement, cette évolution change la dimension même du SNDS et vise à contenir ainsi l'ensemble des données médicales donnant lieu à remboursement » (Légifrance 2019, p.1).

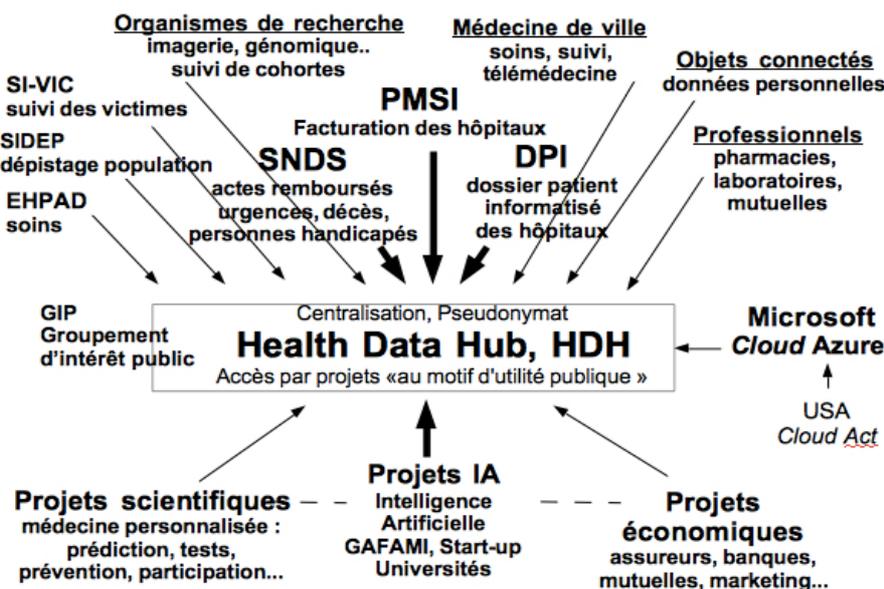


Figure 1. La plateforme HDH, l'ensemble des données médicales ouvert à l'écosystème de l'IA

## **1. Le choix contesté de Microsoft pour gérer « dans l'urgence » les données de santé françaises**

Ce sont les conditions entourant le choix de Microsoft qui vont déclencher la controverse en 2019 avec notamment une alerte lancée par le [CNLL](#), une association qui représente plus de 300 entreprises du logiciel libre. Le CNLL a vite été rejoint par [InterHop](#), collectif composé de professionnels de la santé et de l'informatique médicale, en élargissant le débat aux questions de centralisation, d'anonymat et de standards Open source. De nombreuses et diverses voix, depuis le directeur des hôpitaux de l'AP-HP jusqu'aux militants de La Quadrature du Net, expriment leur surprise devant la gestion si hâtive d'un tel projet, lequel a tout de suite été élargi aux fichiers de dépistage de la population au nom de l'état d'urgence sanitaire. La controverse devient alors une polémique, parfois assez violente de la part de certains médecins qui parlent de vol de données et de rupture du secret médical (Youtube [2020-Lien4](#)).

Le débat, devenu débat de société, débouche alors sur une première action juridique. Pour déposer une requête en urgence « référé-liberté » auprès du Conseil d'Etat, le CNLL et InterHop sont en effet rejoints en juin 2020 par un médecin (D. Sicard, ancien président du Comité national consultatif d'éthique) et un professeur en systèmes d'information (B. Fallery, auteur de cet article), par cinq syndicats (Syndicat national des journalistes, Syndicat de la médecine générale, Union française pour une médecine libre, Union générale des ingénieurs, cadres et techniciens CGT, Union fédérale CGT santé et action sociale), par deux associations (Ploss Rhône-Alpes, SoLibre) et par la société Nexedi.

Le Conseil d'Etat examine cette première demande de suspension immédiate du *Health data hub*. Sur le fond la juge reporte la question au niveau « *des requêtes pendantes devant le Tribunal de l'Union européenne* ». Mais sur la question de l'urgence, l'ordonnance rendue stipule notamment que dans un délai de quinze jours « *La Plateforme des données de santé fournira à la CNIL tous éléments relatifs aux procédés de pseudonymisation utilisés, propres à permettre à celle-ci de vérifier que les mesures prises assurent une protection suffisante des données de santé traitées...* » (Conseil d'Etat, [19 juin 2020](#) article 1).

## **2. La question de la protection des données de santé, reconnue comme complexe**

C'est en fait la question du respect des données privées de santé qui va élargir le débat (figure 2). En juillet 2020 se produit en effet ce que certains appellent un véritable *big bang* réglementaire. Suite à une deuxième plainte de l'avocat autrichien Max Schrems sur les données européennes gérées par Facebook, la Cour de Justice de l'Union Européenne invalide un accord dit *Privacy Shield* (accord général de 2016, gérant le transfert de données entre l'Europe et les États-Unis). Cet accord ne permettait pas une protection des règles européennes du RGPD au vu du droit des États-Unis : D'une part un juge étasunien peut exiger toutes les données personnelles gérées par une société étasunienne dans le monde (*Cloud Act*) et d'autre part les programmes de surveillance (*FISA Act, Executive Order 12333*) permettent un accès aux données en transit vers les USA sans surveillance judiciaire ni droit de recours. S'appuyant alors sur cette décision de la Cour de justice de l'union européenne, le collectif des requérants SantéNathon s'élargit encore à trois autres associations et il dépose deux nouveaux recours successifs au Conseil d'État en urgence référé-liberté.

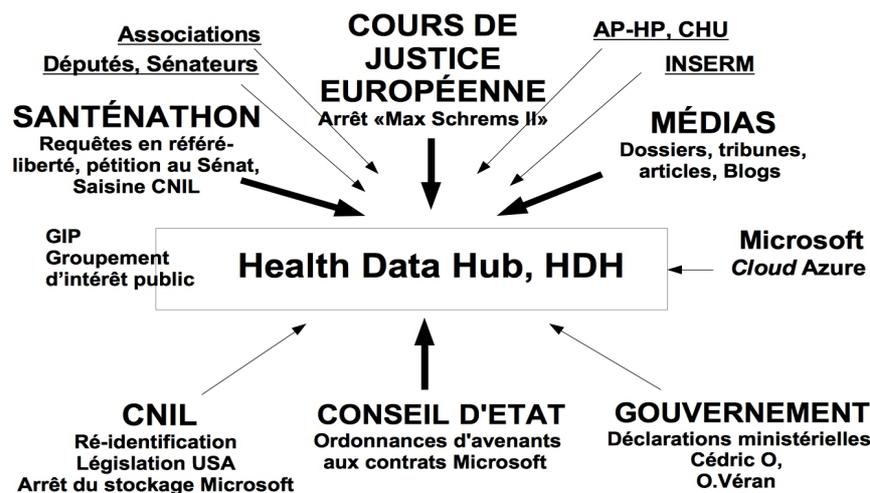


Figure 2. Actions juridiques et médiatiques, la question de Microsoft et celle des données personnelles

Le Conseil d'État accepte à nouveau de juger ces requêtes demandant un moratoire sur le *Health data hub*. Juste avant le début de l'audience, la CNIL publie son avis demandant un arrêt du stockage des données chez Microsoft (Mediapart [2020-Lien5](#)). Le 14 octobre, le Conseil d'Etat reconnaît que Microsoft ne protège pas les données de santé des français contre les intrusions du gouvernement des USA et il indique au ministère être dans l'attente de garanties supplémentaires et d'une autre solution évoquée par la CNIL. Même si la juge n'a pas voulu considérer de menaces « graves et imminentes » justifiant de mettre immédiatement un terme à la plateforme, le choix de Microsoft comme hébergeur paraît désormais bien compromis. Les ministres de la santé ou du numérique le reconnaissent mais veulent s'accorder un délai de deux ans, le collectif SantéNathon proteste et lance une pétition pour soutenir la demande d'une commission d'enquête au Sénat, une saisine à la CNIL et des tribunes dans la presse (Libération [2020-Lien6](#)).

Un autre front est ouvert en février 2021 par un collectif d'associations de patients et de syndicats de médecins, pour demander cette fois l'annulation du partenariat entre le gouvernement et la société Doctolib qui fait gérer par Amazon Web Services toutes les données recueillies lors des rendez-vous de vaccination. Le 12 mars, la juge des référés écarte cette demande, en estimant que le niveau de protection n'est pas « *manifestement insuffisant* » compte tenu de la nature des données en cause.

### 3. LES RÉSULTATS : UNE ANALYSE « EN COMPLEXITÉ » DE TROIS ÉPISODES DE DIFFICULTÉS ÉTHIQUES

L'analyse faite ici du cas *Health data hub* peut être qualifiée de recherche-action, dans la mesure où l'auteur a participé au travail et aux actions du collectif SantéNathon de mai à décembre 2020. On montre comment une « simple polémique » de départ sur l'urgence s'est transformée en trois questions « d'éthique complexe » dans lesquelles des controverses et des actions amènent les différents acteurs à travailler à bien penser. On qualifie ces trois questions autour des impératifs de souveraineté, des impératifs de surveillance et des impératifs de l'état de droit, suivant la définition que donne E. Morin des difficultés éthiques « *qui sont moins dans une insuffisance que dans un excès d'impératifs* » (Morin, T6, p. 53).

### **Méthodologie d'analyse de discours**

Les résultats présentés sont issus de l'analyse de documents internes au collectif SantéNathon (courriels, compte-rendus de réunions, rédaction de mémoires et de pages Web), d'interventions à la Ligue des Droits de l'Homme (enregistrements de deux heures de débat) et de documents externes publiés par les différents acteurs (dossiers et articles de presse, interviews, blogs et sites Web, ordonnances de justice, questions-réponses au Sénat...).

L'ensemble des documents représentent plus de mille pages et a fait l'objet d'une analyse de discours, selon la définition qu'en donne A. Seignour (2011) : Une « grille de lecture » permettant de faire apparaître des indicateurs énonciatifs (modalisation, verbes), des indicateurs référentiels (champs sémantiques, arguments), des indicateurs organisationnels (séquences, chronologie) pour finalement qualifier le *fonctionnement global* d'un discours (thème, thèses, présupposés, système d'énonciation). Pour ne pas surcharger la présentation, les verbatims ne sont pas recopiés dans le texte, mais une liste en webographie présente des exemples de liens actifs et les mots-clés indicateurs sont présentés ci-dessous.

Le codage a été réalisé à la main sous la forme de surlignages et d'annotations par une soixantaine de mots-clés, puis ces indicateurs ont pu être regroupés en trois grandes thématiques (la souveraineté, la surveillance, le droit) se développant à chaque fois suivant trois systèmes d'énonciation (la polémique, la controverse, l'action) et trois horizons de justification (l'urgence, le potentiel, l'essentiel).

## **1. La difficulté éthique des impératifs de souveraineté, l'épisode GAFAM**

Entre l'objectif affiché de «... *faire de la France un leader dans l'utilisation des données de santé* » (A. Buzyn, Ministre, 2018) et le contrat avec Microsoft avec qui « *on a préféré aller vite, pour ne pas prendre de retard et pénaliser la France* » (S. Combes, Directrice HDH, [2020-Lien7](#)), les impératifs contradictoires de souveraineté se concrétisent dans un « épisode GAFAM » :

- un épisode, parce que l'analyse des discours montre que cette thématique a culminé au début de l'année 2020, avec ses mots-clés caractéristiques : Microsoft, Gafam, urgence, marchandisation de la santé, Logiciel libre, communs numériques, hébergement, centralisation, marchés publics, anonymat, risques, certifications, clés de chiffrement...
- un épisode où l'on peut voir que le système d'énonciation évolue entre la polémique (mots-clés : conflits d'intérêt, favoritisme, urgence, captivité numérique, hold-up sur les données de santé, réalité travestie, données en otage...), la controverse argumentée (mots-clés : interopérabilité, pseudonymat, clés de chiffrement, standardisations *Open source*, *Federated Learning*, approches décentralisatrices ...) et l'action (tribunes et pétitions dans la presse, saisine du ministère, [questions-réponses](#) au Parlement, et finalement première requête en référé-liberté au Conseil d'Etat) ;
- un épisode où l'urgence (toujours évoquée comme justification du contrat Microsoft) s'oppose au potentiel, défendu par les tenants des communs numériques en démontrant les alternatives actuelles et futures d'un écosystème autour du Logiciel libre (exemple du projet [eHop Lien8](#), ou des entrepôts de données de santé de l'APHP, de Marseille, de Toulouse, de Lille...)

## **2. La difficulté éthique des impératifs de surveillance, l'épisode COVID**

Dans la lutte contre l'épidémie de Covid-19 c'est l'application StopCovid qui a tout de suite soulevé la question de la surveillance des contacts (Rodhain 2020). Mais cet arbre StopCovid a caché la forêt *Health data hub* (Fallery 2020), car la forte polémique StopCovid a en

fait occulté le [décret du 21 avril 2020](#). C'est pourtant ce décret qui élargit par surprise la plateforme de données HDH : non seulement les fichiers Contacts Covid et SI-Dépistage ont été intégrés, mais « dans la foulée » les données venant aussi des pharmacies, des applications mobiles de santé ou télémédecine, des laboratoires de ville, des Samu, des Ehpad, des enquêtes de vécu, des centres de santé, et même du fichier très contesté de suivi des victimes...

Les impératifs contradictoires de surveillance qualifient ce qu'on peut appeler « l'épisode COVID » :

- un épisode notamment déclenché par la réaction sévère de la Cnil, qui note dès le 23 avril en soulignant avec des guillemets : *A titre d'exemple, le projet mentionne, sans plus de détail, la remontée possible de données issues du SNDS ou de «données de pharmacie»* ;
- un épisode où l'on peut voir que le système d'énonciation évolue là aussi entre la polémique (mots-clés : StopCovid, mouchards, *Big brother* médical, urgence, affront à la recherche médicale, glissement de la santé vers la sécurité...), la controverse scientifique (mots-clés : interconnexion de fichiers, liberté-sécurité, données sensibles, surveillance, consentement, secret médical, datification, ré-identification, finalités non explicitées...) et l'action (groupe de travail Santé à la LDH, demande de commission d'enquête parlementaire, deuxième enquête en référé-liberté au Conseil d'Etat) ;
- un épisode où l'urgence, toujours justifiée par la crise sanitaire, s'oppose à l'essentiel, défendu par les tenants d'une médecine « *qui n'a pas attendu le HDH pour faire de la recherche médicale* ».

### **3. La difficulté éthique des impératifs de l'état de droit, l'épisode EUROPE**

L'onde de choc juridique du 16 juillet 2020, suite à l'invalidation par la Cour de Justice de l'Union Européenne ([Curia-Europe 2020](#)) de l'accord dit *Privacy Shield* qui gérait le transfert de données entre l'Europe et les États-Unis, constitue « l'épisode EUROPE » dans les discours sur la plateforme HDH :

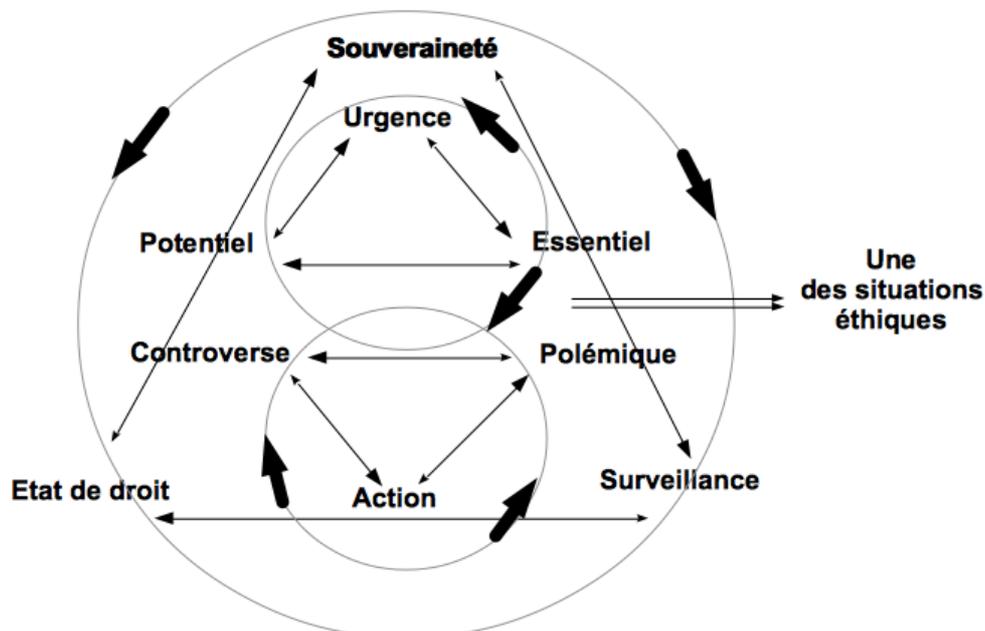
- en juin le Conseil d'Etat avait pu s'en remettre aux « *requêtes pendantes devant le Tribunal de l'Union européenne* », mais en octobre la juge doit reconnaître, suite à la troisième requête du collectif SantéNathon, l'existence d'un risque de transfert des données du HDH vers le gouvernement des USA. Elle demande au ministère des garanties supplémentaires, en attente d'une solution pérenne ;
- dans les discours l'épisode EUROPE voit encore s'alterner l'énonciation polémique (mots-clés : *Cloud Act*, Max Schremes, *Privacy Shield*, urgence, protection, *Foreign Intelligence Surveillance Act*, menaces de Facebook, données bradées, souveraineté européenne...), la controverse scientifique (mots-clés : RGDP, éthique européenne de l'IA, principe de finalités, principe de proportionnalité, principe de recours...) et l'action (nouvel avis CNIL le [29 octobre](#), lettre du ministre de la Santé le [19 novembre](#), saisine à la CNIL par SantéNathon le [7 décembre](#)) ;
- enfin dans cet épisode EUROPE, l'urgence ne pouvant quand même pas être évoquée par le ministère face au respect fondamental de l'État de droit, la polémique se renverse totalement. Le Ministère veut s'accorder un délai de deux ans pour se mettre en conformité, la contestation dénonce une manœuvre dilatoire : pour avoir le temps de négocier d'autres accords transnationaux ou pour espérer que la réversibilité devienne définitivement impossible.

### **4. DISCUSSION SUR UNE « IMPOSSIBLE » GOUVERNANCE ÉTHIQUE DES DONNÉES MASSIVES**

L'analyse de discours a permis de mettre ici en avant au moins trois boucles dialogiques à

la fois concurrentes, complémentaires et dynamiques, suivant ce que E. Morin appelle des impératifs (Morin, T6, p. 53) : une boucle sur les grands thèmes soulevés (ici les impératifs de souveraineté, de surveillance et d'État de droit), ensuite sur le système d'énonciation (ici les impératifs de polémique, de controverse et d'action) et enfin sur la temporalité (ici les impératifs d'urgence, de potentiel et d'essentiel). Nous pouvons alors avancer (figure 5) que ces boucles sont comme des « roues dialogiques » qui tourneraient chacune au cours du temps, dans des sens différents et à des allures différentes :

- à un instant déterminé, un impératif peut prévaloir sur une des boucles, à l'instant suivant un autre impératif peut se trouver en position dominante : Dans l'épisode COVID par exemple, un alignement provisoire s'est caractérisé à un moment par une rencontre du thème Surveillance, de la temporalité Essentiel et de l'énonciation Polémique ;
- ceci permet de généraliser en proposant la définition de situations éthiques. Pour J. Girin la situation de gestion « peut être vue comme une réaction collective à un impératif », elle réunit des acteurs (participants et autres) qui « doivent accomplir dans un temps déterminé, une action collective conduisant à un résultat soumis à un jugement externe » (Girin, 1990, p. 2). Une « situation éthique » est alors une situation d'action collective dans laquelle des principes de niveau moral se déclinent en comportements de niveau éthique et qui correspond à un moment de rencontre d'impératifs concurrents apparaissant sur plusieurs boucles dialogiques. Ces situations éthiques se succèdent dans le temps et forment un processus itératif, ne cherchant pas l'optimisation mais se régénérant en permettant de « travailler à bien penser ».



**Figure 5. Les « situations éthiques » : une pluralité d'impératifs, sur différentes boucles dialogiques**

Cet alignement provisoire de différents impératifs à la fois complémentaires et antagonistes est justifié à plusieurs reprises par E. Morin :

- « Dans la tri-finalité Liberté-Egalité-Fraternité, nous devrions, selon les périodes, donner la priorité à l'une, sans oublier les deux autres » (T6, p. 67)
- « L'auto-éthique, la socio-éthique et l'anthropo-éthique sont à la fois complémentaires, concurrentes et antagonistes. Il faut en chaque occasion établir une priorité et effectuer un choix qui est un pari » (T6, p. 69)
- « A force de sacrifier l'essentiel pour l'urgence, on finit par oublier l'urgence de l'essentiel » (T6, p. 58).

En cherchant à suivre deux grands principes de la pensée complexe, l'analyse des discours sur le *Health data hub* a permis la définition d'un processus composé d'une succession de situations éthiques. Ces résultats montrent que la pensée complexe, en proposant la reliance plutôt que l'exclusion et en proposant la dialogique plutôt que le dilemme, permet une élucidation d'impératifs concurrents. Il faut alors discuter les apports possibles dans le champ des sciences de gestion et les recommandations possibles dans le champ des données massives.

## 1. Un apport théorique possible de la pensée complexe dans le champ de la décision

Travailler à « penser de façon complexe » dans le cas du *Health data hub* a montré que l'on doit reconnaître la diversité des impératifs éthiques (notamment sur les enjeux et sur la temporalité) et que l'on doit reconnaître des épisodes dans un conflit éthique (l'action permettant notamment d'enclencher des polémiques et des controverses). Si l'on accepte alors de considérer que la prétendue « décision » éthique n'est pas un simple choix de solutions mais une chaîne de situations confrontant différentes réponses, on voit que la pensée complexe peut offrir des perspectives croisant la rationalité procédurale (Simon [1976](#)) et la rationalité située (Weick et al. [2005](#)).

La vision comportementale de la rationalité « procédurale » développée par Herbert Simon s'est constituée autour du *satisficing* (le principe du seuil de satisfaction), des processus cognitifs heuristiques (les boucles récursives intelligence-modélisation-choix) et du rôle du contexte organisationnel (les contraintes et limites). La vision collective de la rationalité « située » développée par Karl Weick s'est constituée autour du *sensemaking* (l'action est préalable à l'émergence de sens), du *mindfulness* (la prise de conscience des situations est collective), de l'*organizing* (le rôle de la communication dans les situations de crise) et de l'improvisation organisationnelle (la métaphore de l'orchestre de jazz et les stratégies chemin-faisant).

Dans l'analyse d'une chaîne de situations éthiques, la pensée complexe devient alors une méthode pour aider à penser, sans détacher le flux de l'action et l'incomplétude de la raison, sans détacher l'ordre et le désordre, sans détacher la cognition et la communication, sans détacher la connaissance et l'éthique, et donc au final sans détacher les potentialités, les intentions et les délibérations. Dans la décision stratégique ce travail a été initié par C. Martinet ([2006](#)) et dans l'apprentissage organisationnel par D. Vallat ([2017](#)).

## 2. Une recommandation possible pour la gouvernance éthique de données massives

Les données massives sont devenues une nécessité pour la reconnaissance des formes dans l'IA (Pellegrini [2017](#)) et un champ de recherche pour de nombreuses disciplines autres que l'informatique : sur la société des traces (Quessada [2010](#)), sur la science sans théories (Boyd et Crawford [2011](#)), sur la gouvernementalité algorithmique (Rouvroy et Berns [2013](#))... Quant à la gouvernance de ces données massives, c'est alors la nécessité d'organiser les rôles, les responsabilités et les processus, qui pose un ensemble de questions : Pour garantir aussi bien la gestion des données (capture/rationalisation/stockage, mais aussi stratégies/finalités/risques), que pour garantir leur partage (confiance/accès/décision). Or dans l'écosystème des données massives de santé la liste des impératifs concurrents est presque interminable :

- impératifs concurrents des nombreux acteurs impliqués sur un marché des données de santé qui cherche à se constituer : les acteurs concernés (patients, leurs associations et leurs espaces d'échanges, médecins, pharmaciens et laboratoires de ville...), les acteurs publics (assurance maladie, centres hospitaliers, ministères, organismes de recherche, CNIL, *Health data hub*...) et les sociétés privées très variées (hébergeurs de données, consortiums sur les ontologies de données et les standardisations, plates-formes de

services médicaux, mutuelles banques et assureurs, fabricants de dispositifs connectés, sociétés de conseil... et tout l'écosystème de l'IA) ;

- impératifs concurrents des finalités impliquées dans la réutilisation des données de santé, aussi bien pour des usages nobles mais encore risqués (de la médecine personnalisée « précise, prédictive, préventive et participative »... à la santé connectée), que pour des usages moins nobles mais déjà en place (un contrôle très intrusif des comportements sanitaires, de la relation patient-médecin... et au final un contrôle de « bonnes pratiques » sociales) ;

- impératifs concurrents des situations économiques, avec d'un côté un travail très qualifié de saisie des données largement financé par le service public (travail chronophage sur le terrain en sous-effectif et sous-financement, mais avec ici libre circulation des données ouvertes) et de l'autre côté un travail d'exploitation et valorisation des données largement privatisé (travail à forte valeur ajoutée, mais relevant ici de la propriété intellectuelle et du droit des brevets).

Avec un tel niveau de variété possible des états du système et avec une grande faiblesse de variables de commande, la tentation est grande non pas de travailler à penser de façon complexe, mais de travailler par simplification : soit par délégation, soit par régulation.

- La gouvernance par délégation revient en fait à confier la réflexion éthique aux « personnes autorisées » que constituent les praticiens de l'IA et des *Big data*. L'éthique numérique relève alors de leurs propres chartes de déontologie : serment d'Hippocrate pour *Data Scientist*, serment *Holberton-Turing*, *Déclaration de Montréal*, préconisations *Allistene*, somment *AI4People*, bonnes pratiques *AI HLEG*... sans compter les chartes spécifiques de *Google*, *Facebook*, *Microsoft*... On pourrait parler d'une gouvernance par consentement, mais malgré tous ces engagements les révélations du *Wall Street Journal* en 2019 ont révélé le projet *Nightingale*, un moyen pour Google de mettre la main sur les données médicales de dix millions d'Américains ;

- La gouvernance par régulation revient à confier la certification des pratiques sur les données à des « tiers de confiance » légitimes, qui pourraient labelliser une conduite éthique. On pourrait parler d'une gouvernance par le contrôle, c'est la voie envisagée par le Comité consultatif national d'éthique : « *Il est essentiel que le titulaire des données soit informé des modalités par lesquelles l'autorité de contrôle assure sa fonction de tiers de confiance* » (CCNE [2019](#)), une voie promue par exemple par J. Béranger ([2017](#)) pour « *accompagner la mise en conformité des organisations tout en déléguant une partie de la charge du contrôle aux organismes émetteurs des labels* ».

Mais c'est à une troisième forme de gouvernance que nous invite la pensée complexe, la gouvernance par auto-organisation. Historiquement l'auto-organisation est la forme de gouvernance qui a permis à Internet de se développer (Fallery et Rodhain [2010](#)) : Des règles élaborées par les destinataires eux-mêmes de ces règles, par ajustements et réflexivité dans de nombreux collectifs de niveaux à la fois variés et imbriqués tels que l'ISOC, l'IETF ou le W3C. C'est aussi une des formes de gouvernance de l'Open Source : la véritable constitution du projet *Debian*, les règles de la communauté *Apache* interdisant notamment les dictateurs bienveillants... C'est aujourd'hui la proposition d'une ontologie de la gouvernance « modulaire » des plateformes Internet, un réseau en couches à partir de composants informatiques polyvalents, portables et interopérables (Schneider et al. [2021](#)) : cercles, jurys, jugements majoritaires, élections, pétitions, révocations, jetons, permissions... s'appuyant sur le paradigme IAD de gouvernance polycentrique des biens communs (Ostrom, [2009](#)). On pourrait alors parler d'une gouvernance par l'éthique complexe.

## CONCLUSION

Bien que nous ayons été personnellement engagé dans cette action de contestation du

projet *Health data hub*, nous avons ensuite cherché, en s'appuyant sur la rigueur de l'analyse de discours, à « travailler pour penser complexe » et à garder une modeste volonté d'éthique de reliance.

Cette reliance, qui doit affronter l'ambiguïté, la contradiction et l'incertitude, devient indispensable devant au moins trois risques liés aux *big data* et à l'Intelligence Artificielle : le risque d'une perte générale de la connaissance au profit de la simple reconnaissance de signaux, le risque d'une perte de la délibération ou du libre arbitre, enfin le risque d'une perte de la responsabilité par l'automatisation du travail.

Cette reliance, qui doit être à la fois une auto-éthique du questionnement et une socio-éthique de la délibération, devient nécessaire car l'omniprésence du Réseau concerne des individus qui sont aussi des citoyens. Les conflits, déviances, addictions, dilemmes, paradoxes... liés aux technologies numériques (Stiegler [2007](#)) associent en effet aujourd'hui le comportement des individus (l'action devenant de plus en plus une action numérique, Rosati [2012](#)) avec le comportement de ces technologies (non seulement ambivalentes, mais devenant de plus en plus autonomes).

En gardant cette volonté d'éthique de reliance, il faut donc garder la volonté de repenser le mode de gouvernance centralisée du projet actuel *Health data hub*. Une telle plateforme pourrait devenir une instance qualifiée pour la régulation de l'accès aux données par l'écosystème de l'IA. Mais les instances de production des données de santé et les instances d'hébergement pourraient être différentes, à la fois décentralisées et intégrées, pour assurer la qualité et la sécurité grâce aux compétences des praticiens qui ne sont heureusement pas seulement des « *data scientist* » et pour assurer la confiance de patients qui ne sont heureusement pas seulement des « *data* ». La gouvernance particulièrement sensible des données de santé, alors reconnues comme biens communs, pourrait devenir un exemple pour la renaissance d'une culture numérique démocratique.

## Bibliographie

Béranger J. (2017). Vers une régulation éthique de la gouvernance numérique médicale. Communication au 7ème colloque Santé, Marseille. [En ligne](#)

Bolle de Bal M. (2009). Éthique de reliance, éthique de la reliance : Une vision duelle illustrée par Edgar Morin et Michel Maffesoli. *Nouvelle revue de psychosociologie*, 8(2), 187-198. [En ligne](#)

Boyd D., Crawford K. (2011). Six provocations à propos des Big Data. Symposium Une décennie avec Internet, article *Oxford Internet Institute*, traduction. [En ligne](#)

Eynard J. (2012). L'éthique à l'épreuve des nouvelles particularités et fonctions des informations personnelles. *Éthique publique*, vol. 14, n° 2. [En ligne](#)

Fallery B. (2020). Données de santé : L'arbre StopCovid qui cache la forêt Health data hub. Revue en ligne *The Conversation*, juin 2020. [En ligne](#)

Fallery B., Rodhain F. (2010). Fondements théoriques pour une régulation d'Internet : La légitimation faible et la réflexivité forte. *Systèmes d'Information et Management*, 15(3), 41-70 [En ligne](#)

Gaillard J., Coenen R., Friehe F., Hardy G. (2011). Vers une neuro-éco-systémique: Manifeste pour l'urgence d'un changement. *Thérapie Familiale*, vol. 32 (1). [En ligne](#)

Genelot D. (2014). Responsabilité éthique dans l'action en complexité. *Revue de l'AIEMS*,

Réseau Intelligence de la complexité. [En ligne](#)

Girin J. (1990). L'analyse empirique des situations de gestion. In *Épistémologies et Sciences de gestion*, Martinet A.-C. (coord.), *Economica* [En ligne](#)

Quessada D. (2010). De la sousveillance : La surveillance globale, un nouveau mode de gouvernementalité. *Multitudes*, 40(1), 54-59 [En ligne](#)

Le Moigne J-L (2008). Edgar Morin, le génie de la Reliance. *Synergie Monde*, n°4, 177-184. [En ligne](#)

Le Moigne J.-L., Morin E. (dir.) (2007). *Intelligence de la complexité : épistémique et pragmatique*. Editions de l'Aube, Aix-en Provence

Martinet, A. (2006). Stratégie et pensée complexe. *Revue française de gestion*, 1(1), 31-45. [En ligne](#)

Morin E. (1982). *Sciences avec conscience*, nouvelle édition 1990. Points, Edition du Seuil, Paris

Morin E. (1987). *Penser l'Europe*. Gallimard, Paris

Morin E. (2004). *La Méthode T.6 Éthique*. Points, Edition du Seuil, Paris

Morin E. (2005). La pensée complexe comme méthode pour une organisation apprenante « Aide-toi, et la complexité t'aidera ». *Cahier SOL* n°4 1-7 [En ligne](#)

Morin E. (2008). *Mon chemin*. Entretiens avec Djénane Karech Tager. Fayard, Paris

Ostrom E. (2009). Par-delà les marchés et les états, la gouvernance polycentrique des systèmes économiques complexes. Conférence Stockholm, Fondation Nobel. [En ligne](#)

Pellegrini F. (2018), Intelligence artificielle, mégadonnées et gouvernance. *Revue Lamy Droit de l'immatériel*, 56-59. [En ligne](#)

Peyron-Bonjan C. (1998). Complexité. La pensée d'Edgar Morin réinterrogée. Colloque Unesco. [En ligne](#)

Peyron-Bonjan C., Berger G., (2007). Complexité, in *Vocabulaire de la complexité*, Post-scriptum à La Méthode d'Edgar Morin, de M. Mukungu Kakangu. L'Harmattan, Paris. [En ligne](#)

Rodhain F. (2020). Fact check : L'application StopCovid contient-elle un mouchard ? Revue en ligne *The Conversation*, juillet 2020. [En ligne](#)

Rouvroy A., Berns T. (2013). Gouvernamentalité algorithmique et perspectives d'émancipation. *Réseaux* n°177, 163-196. [En ligne](#)

Rosati M. (2012). Une éthique appliquée ? Considérations pour une éthique du numérique. *Éthique publique*, vol. 14, n° 2 [En ligne](#)

Sánchez A. (2011). Féminin-masculin : Une dialogique inachevée. *Hermès, La Revue*, 60 (2), 217-222. [En ligne](#)

Schneider N. et al. (2021). Modular Politics : Toward a Governance Layer for Online Communities. *Proceedings ACM Human-Computer Interactions*, 5, article 16. [En ligne](#)

Seignour A. (2011). Méthode d'analyse des discours : L'exemple de l'allocation d'un dirigeant d'entreprise publique. *Revue française de gestion*, 211(2), 29-45. [En ligne](#)

Sève L., Guespin-Michel J. (2005). Émergence, complexité et dialectique : Sur les systèmes dynamiques non linéaires. Odile Jacob, Paris. [Google.books](#)

Simon H. (1976). De la rationalité substantive à la rationalité procédurale. *Revue Piste* n°3, Traduction [en ligne](#)

Stiegler B. (2007). Questions de pharmacologie générale. Il n'y a pas de simple pharmakon. *Psychotropes*, vol. 13(3-4), 27-54. [En ligne](#)

Vallat D. (2017). Manager les connaissances dans un environnement complexe : questions éthiques, épistémologiques et stratégiques. HDR Université Jean Moulin. [En ligne](#)

K. Weick, K. Sutcliffe, D. Obstfeld (2005), Organizing and the Process of Sensemaking. *Organization Science* 16(4), 409-421. [En ligne](#)

### Webographie

[Lien1](#) [www.LeMonde.fr](http://www.LeMonde.fr) 2019/12/24. Mots-clés : Conflit d'intérêts au cœur de la nouvelle plateforme

[Lien2](#) [www.Solidarites-sante.gouv.fr](http://www.Solidarites-sante.gouv.fr) 12/10/18. Mots-clés : Agnès Buzyn annonce la création d'un *Health data hub*

[Lien3](#) [www.Ticsante.com](http://www.Ticsante.com) 27/12/2019. Mots-clés : Microsoft prestataire du *Health data hub* : un choix d'opportunité pour « aller vite »

[Lien4](#) *Youtube* [05/4/20](#). Mots-clés : *Health data hub* = Vol des données de santé au profit de Microsoft ? Fin du secret médical?

[Lien5](#) [www.Mediapart.fr](http://www.Mediapart.fr) [09/10/2020](#) Mots-clés : La Cnil demande l'arrêt du stockage de nos données de santé par Microsoft

[Lien6](#) [www.Libération.fr](http://www.Libération.fr) [14/12/2020](#) Mots-clés : *StopHealthDataHub* : les données de santé en otage chez Microsoft

[Lien7](#) *Youtube* 19/12/19. Mots clés : *Health data hub*, plateforme de la discorde ou de la concorde ? Avec S. Combes, Directrice HDH

CCNE Comité consultatif national d'éthique (2019). Données massives et santé, une nouvelle approche des enjeux éthiques. Rapport 94 pages [En ligne](#)

CIGREF (2014). Gouvernance du Numérique, création de valeur, maîtrise des risques, allocation des ressources. Rapport 42 pages. [En ligne](#)

CJUE Cour de justice de l'Union européenne (2020). Communiqué n° 91/20, *Curia-Europe*. [En ligne](#)

CNIL (2018). Qu'est-ce qu'une donnée de santé ? [En ligne](#).

CNIL (2019). Avis sur un projet de loi relatif à l'organisation et à la transformation du système de santé. *Légifrance Délibération* 2019-008, 31/12/19. [En ligne](#)