



HAL
open science

Les grandes bases de données. Une nouvelle approche de l'histoire sociale. Le système Fichoz.

Jean-Pierre Dedieu

► To cite this version:

Jean-Pierre Dedieu. Les grandes bases de données. Une nouvelle approche de l'histoire sociale. Le système Fichoz.. Revista da Faculdade de Letras HISTÓRIA, 2005, III (5), pp.99-112. halshs-00004690

HAL Id: halshs-00004690

<https://shs.hal.science/halshs-00004690>

Submitted on 22 Sep 2005

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.

Les grandes bases de données: une nouvelle approche de l'histoire sociale.

Le système Fichoz

Le paradigme qui régit l'histoire sociale a rapidement évolué au cours des quinze dernières années¹. Il reposait depuis le milieu du XIXe siècle sur l'idée que les conduites individuelles manquaient d'intérêt, que seuls importaient les groupes. Qu'on les appelât classes sociales, ordres, civilisations, royaumes ou catégories socio-professionnelles changeait finalement peu de choses: l'historien maniait de grandes masses, définies par quelques traits caractéristiques², et il se donnait pour objectif de décrire leurs rapports, le jeu de leurs alliances et de leurs affrontements. L'individu, lorsqu'il l'étudiait, ne l'intéressait pas en lui-même: il ne valait que s'il était l'expression, l'illustration de la nature d'un groupe, ou un leader qui avait su cristalliser la stratégie d'un groupe autour de sa personne. La seconde moitié du XX siècle fut marquée par une extension considérable du regard de l'historien. Il prit en compte des contraintes techniques et écologiques jusque là négligées³: la distance devint un facteur clef; les contraintes sanitaires, alimentaires et démographiques passèrent au premier plan de ses préoccupations. Il comprit l'importance des représentations, autrement dit que la manière dont les acteurs percevaient les choses avait autant d'effet que leur réalité substantielle⁴. Il rêva d'une histoire totale, qui ferait contribuer l'ensemble des champs nouveaux qu'il s'ouvrait à l'intelligibilité d'histoires régionales, lesquelles prendraient en compte, à leur tour, l'ensemble des interactions entre les différentes régions du monde pour expliquer l'histoire de chacune d'entre elles. Il imagina en un mot pouvoir faire tenir ensemble l'ensemble des domaines qui constituent une civilisation, et l'ensemble des civilisations connues saisies d'un seul regard, mieux qu'aucune autre science sociale n'avait su le faire auparavant. Ce fut la grandeur de F. Braudel et de "l'Ecole des Annales" d'avoir su formuler cette ambition avec une force inégalée⁵.

De tout cela, redisons-le, l'individu restait absent. L'observation montrait sans cesse que les contours des ensembles géographiques, culturels, sociaux que l'on manipulait étaient flous, que leurs périphéries constituaient des lieux d'échange à l'identité incertaine que l'on avait du mal à faire entrer dans ce schéma d'interactions limitées entre acteurs bien définis, car les critères utilisés pour caractériser ces derniers s'appliquaient mal sur ces marges et cette indéfinition même semblait constitutive de leur nature. La somme de ces espaces incertains finissait par représenter une part si importante du terrain observée qu'il était

¹ Les premières pages de cet article ont fait l'objet, sous une forme légèrement différente, d'une publication dans une revue interne du CNRS: "Une nouvelle approche de l'histoire sociale: les grandes bases de données", *Sciences de l'homme et de la société (CNRS)*, N° spécial "Vie de laboratoires", 2003, n° 66, p. 35-38.

² On définit volontiers dans certains cercles d'expression française cette pratique comme "durkheimienne", faisant porter la l'illustre sociologue la responsabilité d'un état de fait maintenant rejeté.

³ Braudel (Fernand), *La Méditerranée et le monde méditerranéen à l'époque de Philippe II*, Paris, Armand Colin, 1966, 2ème éd., 2 vol.

⁴ Un premier bilan, triomphant, de cette extension fut dressé dans un ouvrage collectif d'historiens presque tous français: Leroy-Ladurie (Emmanuel), éd., *Le territoire de l'historien*, Paris, Gallimard, 1977.

⁵ Braudel (Fernand), *Civilisation matérielle et capitalisme (XVe-XVIIe siècles)*, Paris, 1967, 2 vol.

difficile de les ignorer. Par ailleurs, même dans les zones centrales du modèle, si l'on rendait assez bien compte des conflits entre groupes, sur quelque plan que l'on se place - social, culturel, économique, politique - on éprouvait de grandes difficultés à conceptualiser les relations de collaboration intergroupe, autrement dit le ciment qui faisait tenir ensemble les sociétés dont on était censé rendre compte. Une fois achevée la première exploration des terrains nouvellement défrichés, on s'aperçut de même qu'il n'était pas facile de décrire les voies par lesquels ils interagissaient l'un avec l'autre.

Réapparut alors l'acteur individuel. Fortement inspirés par une évolution parallèle d'autres sciences sociales, redécouvrant par ailleurs des courants minoritaires de leur propre discipline, les historiens s'aperçurent que, loin d'être un pur produit de forces extérieures, l'acteur secrétait littéralement la société dans laquelle il vivait. On ne rejeta pas l'acquis pour autant. L'idée subsiste, fortement ancrée, que l'individu agit dans des cadres de multiples nature qui sont autant de contraintes qui pèsent sur son action. L'idée subsiste aussi que l'action individuelle s'inscrit à l'intérieur de groupes; mais on perçoit ceux-ci de façon beaucoup plus fine qu'autrefois, on les voit beaucoup plus nombreux, multiples et divers dans leur définition. Les grands acteurs collectifs qui servaient de cadre à l'interprétation antérieure ont perdu leur caractère absolu. Ils sont perçus comme des abstractions utiles à un certain niveau d'analyse, sans préjudice d'autres, plus fins⁶. Les appartenances surtout ne sont plus exclusives. La multiappartenance est la règle. Dans chacun des domaines de la vie, l'individu appartient à un collectif différent, doté de sa logique propre. Il combine ces logiques dans son action, qui en est la synthèse⁷. Celles-ci sont elles-mêmes le produit des interactions entre les acteurs qui constituent les collectifs en action *hic et nunc*. Ces individus reproduisent les conduites antérieurement marquées comme acceptables par le groupe, mais en les distordant en fonction des intérêts qu'ils ont dans d'autres groupes. Ils les font ainsi évoluer et créent de nouvelles normes⁸.

La compréhension de ces phénomènes et des règles qui les régissent apparaît aujourd'hui comme une urgence. Nous savons qu'elle est le chemin vers l'histoire totale que les pères fondateurs de la discipline sous sa forme actuelle lui ont donné pour objectif. Elle nous pose cependant un défi technique de grande ampleur.

Le paradigme ancien, fondé sur un nombre restreint d'acteurs collectifs, réduisait considérablement le nombre des interactions possibles entre eux et de leurs combinatoires. La recherche prenait très souvent l'allure du rassemblement de données statistiques sur les différents groupes mis en présence, censées les caractériser comme acteurs collectifs. Une fois ces caractères - en nombre restreint - mis en évidence, on les reliait comme on pouvait sans trop s'attarder sur les modalités pratiques de leurs interactions, que

⁶ Revel (Jacques), dir., *Jeux d'échelle. De la micro-analyse à l'expérience*, Paris, Gallimard, 1996

⁷ Boltanski (Luc) et Thévenot (Laurent), *De la justification. Les économies de la grandeur*, Paris, Gallimard, 1991.

⁸ Beckert (Jens), "Economic action and embeddedness: the problem of the structure of action", *Economic Sociology*, electronic review (Amsterdam), 1/2000, www.pscw.uve.nl/sociosite. Repris dans Beckert (Jens), *Beyond the market: the social foundations of economic efficiency*, trad. ag., Princeton, 2000.

l'on postulait plus qu'on ne les décrivait. La masse documentaire mise en oeuvre pouvait être considérable⁹. Elle présentait cependant presque toujours deux caractéristiques: sa définition préalable à la recherche, en adéquation étroite à l'objectif poursuivi, en fonction d'une hypothèse de travail préexistante, ce qui avait pour inconvénient de rendre très rigide le processus de recherche et pratiquement impossible la réutilisation des données pour une autre enquête¹⁰; son caractère anonyme: sauf cas exceptionnel, les données n'étaient jamais nominatives¹¹. L'avantage résidait, comme on s'en aperçut *a posteriori*, dans leur volume comparativement faible.

Le but que nous poursuivons aujourd'hui impose une pratique documentaire très différente. La recherche est centrée sur les acteurs, nous l'avons vu. Même en histoire institutionnelle, même en histoire économique¹². Or, un acteur ne peut se concevoir que globalement. On ne peut distinguer en lui l'*homo oeconomicus*, de l'*homo socialis*, de l'*homo religiosus*, du *paterfamilias* ou du *magister docens*: ce sont des rôles distincts, mais dans chacun d'eux il assume simultanément tous les autres, ce qui oblige à rassembler des données sur son action dans chacun de ces domaines avant d'opiner. Par ailleurs, c'est l'étude des interactions qui est prioritaire. Cela signifie que l'individu ne pouvant être compris que dans ses relations avec d'autres, tout ensemble de données le concernant doit inclure les informations pertinentes sur l'ensemble des personnes avec lesquelles il est lié jusqu'au degré pertinent pour l'étude en cours. L'on ne pose plus enfin les interactions comme découlant mécaniquement de la juxtaposition des caractères définissant les différents groupes, mais comme des stratégies souples dépendant hautement du contexte. De ce fait, les règles qui les régissent, et dont la découverte constitue l'un des objectifs principaux de la recherche, ne sont plus déduites de la situation de départ, mais inférées de l'observation de conduites effectives. Tout ceci amène à multiplier les points d'observation et à procéder sur chacun d'eux à des rassemblements de données particulièrement lourds, sans commune mesure en tout cas avec les exigences du paradigme antérieur.

D'où un goulot d'étranglement: la masse d'information nécessaire dépasse les capacités d'un chercheur isolé travaillant selon les techniques traditionnelles de la fiche manuelle, modèle encore dominant dans la profession.

⁹ Vovelle (Michel), *Piété baroque et déchristianisation en Provence au XVIIIe siècle. Les attitudes devant la mort d'après les clauses des testaments*, Paris, Plon, 1973, 700 p.

¹⁰ De gros efforts ont été faits, notamment aux Pays-Bas, pour réutiliser les données informatisées des premiers chercheurs ayant utilisé les ordinateurs. Elles ont échoué, non point tant pour des problèmes techniques - sérieux, au demeurant - qu'à cause de cette dépendance du recueil des données vis-à-vis d'une problématique très précise. Voir: Van Horik (René), "The development of a new standard: the historical data set description scheme", in: *L'ordinateur et le métier d'historien*, Bordeaux, Maison des Pays Ibériques, 1990, p. 137-140.

¹¹ Les études fondées sur les techniques de reconstitution des familles de Louis Henri utilisaient les noms pour déterminer l'appartenance à un même groupe familial; mais l'individualisation disparaissait complètement lors de l'étude proprement dite et lors de la publication des résultats. Il faut attendre les années 1980 et les travaux menés à Paris autour de Jean Pierre Bardet ou à Cambridge autour de Allan Mc Farlane, utilisent l'individu autrement que comme un simple numéro statistique. Voir comme exemple de ces nouvelles techniques: Pélissier (Jean Pierre), *Démographie, généalogie, micro-informatique*, Paris, Ed. Christians, 2ème éd., 1987 [1983].

¹² Beckert, *art. cit.*; Bonin (Hubert), Bouneau (Christophe), Caillet (Ludovic), Fernandez (Alexandre), Marzagalli (Silvia), ed., *Transnational companies - 19th - 29th centuries*, Paris, Plage, 2002.

L'informatisation et le travail collectif s'imposent. Selon quel cahier des charges¹³?

Les données doivent évidemment être nominatives. Il n'est pas nécessaire d'insister sur ce point.

Si l'on tient compte du fait que le travail collectif est une nécessité, elles doivent pouvoir être utilisées dans plusieurs types de recherches. Ceci implique qu'elles soient stockées selon une articulation qui, tout en permettant d'y accéder rapidement, respecte strictement celle des documents de façon à les lier le moins possible à une recherche particulière. Cette nécessité est d'autant plus impérieuse que la procédure d'inférence que nous décrivons ci-dessus interdit tout *a priori* sur la manière s'agenceront les différents facteurs dans le modèle final qui sera le produit de la recherche. Comment procéder pour concilier souplesse, ergonomie et neutralité? Il y a eu au moins une tentative antérieure, très poussée, qui respectait la lettre du document et se contentait d'affecter chacun de ses éléments de marqueurs qui permettaient de retrouver ce que l'on cherchait: noms propres, noms d'objets, concepts, lieux, etc.: le langage de codification des données historiques Kleio, développé en Allemagne par Manfred Thaler¹⁴. La souplesse de l'instrument était parfaite, sa neutralité forte, son ergonomie, hélas, très faible: le nombre des types de marqueurs était tel qu'il était pratiquement impossible de dominer le système. Ceci conduisit à son échec. Le mode de marquage rendait en outre difficile l'attribution claire de chaque élément à un acteur particulier, ce qui rompait avec l'exigence nominative.

La nouvelle conception du travail en histoire sociale nous dicte la solution. Il faut traiter chaque donnée comme un événement dans la vie d'un acteur. La documentation doit donc se lire comme une suite de séquences décrivant des actions effectuées ou subies par un acteur individualisé. Chacune de ces actions doit correspondre à une entrée dans une base de données contenant tous les éléments nécessaires à son interprétation: nature et description de l'action, identification de l'acteur, date, référence, éléments de contexte, etc. Le découpage de la documentation en actions est de la responsabilité du chercheur et donne prise à une marge d'interprétation. Celle-ci doit être limitée par le principe intangible qu'en aucun cas aucune intervention individuelle mentionnée par la source, pour faible qu'il soit, ne doit être laissée de côté. L'ordinateur se charge de remonter la suite des actions concernant un individu ou de regrouper les actions similaires effectuées par des individus différents, ce qui obvie les conséquences négatives de l'effet de masse.

¹³ Le système Fichoz, dont nous décrivons à partir d'ici les caractéristiques principales, a été réalisé principalement par l'auteur de ces lignes, pour répondre aux nécessités d'une recherche entamée par Didier Ozanam et menée en compagnie de nombreux autres chercheurs, dont les réflexions ont considérablement enrichi l'instrument. Mentionnons parmi d'autres, pour le rôle moteur direct qu'ils ont joué dans cet enrichissement, Juan Luis Castellano, María Victoria López Cordón, Francisco Andújar, Teresa Nava, José María Imizcoz, Michel Bertrand, Jean Marc Buigues, Monique Peronnet et, parmi nos étudiants, Lucrecia Enriquez, Thomas Glessener, Valérie Lafage, Sylvie Pinault, Aude Bullier, Olivier Drault, Dominique Goncalves, Pascal Raynaud, Isabelle Rio, Franck Saint-Avit, Marie-Pierre Lacoste, Catherine Leroy, Céline Lèbre, Marlène Pérez, Jérôme Grondin, Clara Palmiste.

¹⁴ Thaller (Manfred), *Kleio. A data base system for historical research. Version 1.1.1, b-test Version*, Göttingen, Max-Planck-Institut für Geschichte, 1987, 127 p.

Ces actions peuvent être uninominales (exemple: naissance de l'acteur) ou binominales (exemple: l'acteur (1) appuie l'acteur (2) pour l'obtention de tel poste), et traitées quel que soit leur type par un modèle unique de fiche de type:

Acteur (1) Relation Acteur (2)

dans le cas ou l'action est binominale;

---- Action Acteur (2)

dans le cas ou l'action est uninominale, séquence identique à la précédente en laissant la première position vide. Ceci permet de traiter dans le même fichier et selon le même format informatique à la fois les événements biographiques de carrière et les relations établies dans le courant de la vie, de les manipuler avec une souplesse parfaite, et de les stocker pour une utilisation postérieure, par le chercheur qui les a rassemblés ou par tout autre.

Un tel schéma s'avère extraordinairement robuste et permet de couvrir l'ensemble des cas possibles, quelle que soit la source - chronique, archives administratives ou judiciaires, correspondance, actes notariés, registres paroissiaux, littérature secondaire ou autre. Il permet notamment un dépouillement extrêmement rapide des instruments publics - registres paroissiaux, état civil, actes notariés -, qui ne sont en fin de compte que des machines à créer de relations interindividuelles¹⁵. Il en livre de façon immédiate la matière dans le contexte de la carrière vitale de chacun des acteurs¹⁶.

Il impose surtout une lecture nouvelle de la documentation. La nécessité du découpage en séquences événementielles nominatives oblige à jeter sur les sources un regard froid, discipliné, donc plus objectif. Elle impose la désagrégation de conduites auparavant perçues en bloc comme un tout en unités constituantes, ce qui met en valeur leur complexité, tout en permettant de retrouver des comportements de base qui constituent en quelque sorte les atomes de l'action sociale, atomes que la pratique précédente, beaucoup moins encadrée, offusquait sous l'implicite. La notion de "clientélisme" par exemple, vole en éclat. Nous découvrons, sous ce que l'historiographie range d'habitude sous cette étiquette, des schémas très divers, qui ont certes en commun une relation inégalitaire entre deux acteurs, mais recouvrent des réalités et obéissent à des règles différentes. Proximité familiale, origine géographique commune, services préalables qui exigent retour, soit directement en faveur du donataire, soit indirectement par un vaste jeu de compensation qui lie à tous les autres chacun des acteurs agissant dans une sphère commune, considérations d'efficacité pratique, proximité idéologique, sont tous des éléments qui peuvent

¹⁵ Relations interindividuelles dans ce cas particulièrement faciles à exploiter, car légalement prédéfinies.

¹⁶ Autre tentative remarquable dans ce domaine, selon des principes similaires, mais plus spécialisé sur les correspondances intellectuelles, celle de Pierre-Yves Beaurepaire et Dominique Taurisson, dans le cadre du CEDIM de Montpellier. Voir: Beaurepaire (Pierre Yves), Taurisson (Dominique), éd., *Les ego-documents à l'heure électronique. Nouvelles approches des espaces et réseaux relationnels - Colloque de Montpellier, 23-2/10/2002*, site: egodoc.revues.org/octobre.2002, Montpellier, 2002. Ce système fonctionne selon des principes très proches de ceux que nous proposons ici. Il permet de lier édition typographique enrichie, recherche des liens et recherche de des actions individuelles. Il reste cependant à notre avis trop proche de l'indexation traditionnelle pour constituer un véritable instrument de recherche systématique, mais il est dans son domaine une réussite remarquable.

conduire à la "clientèle"; mais par des voies et sous des formes si différentes que le concept apparaît rapidement comme dénué de pertinence sous sa forme brute générale¹⁷. L'expérience des utilisateurs est unanime sur ce point: l'utilisation de la base a profondément modifié leur regard sur la documentation et a fait surgir des données qu'il n'y percevaient pas auparavant.

La possibilité de rassembler, au sein d'une même base et sous une forme nominative, c'est-à-dire de façon directement accessible, des données récoltées par des chercheurs différents induit à soi seul un enrichissement considérable des perspectives de recherche. Chacun ne voit que la documentation qui l'intéresse, sous l'angle qui l'intéresse. Le spécialiste de l'économie n'est que médiocrement passionné par le rôle diplomatique des marchands qu'il étudie; le spécialiste d'histoire administrative oubliera volontiers que tel employé de bureau est avant tout un grand poète, et à l'inverse le spécialiste de poésie ne prendra guère en compte que l'auteur des odes bucoliques qu'il admire a passé le plus clair de sa vie enfermé dans un bureau poussiéreux à surveiller les humeurs d'un chef grincheux. L'historiographie classique tend ainsi à aplatir le caractère multidimensionnel des carrières individuelles, à les réduire à un aspect et un seul. Cela était admissible sous le règne de l'ancien paradigme. Ce ne l'est plus sous le nouveau qui, nous l'avons vu, met au centre de ses préoccupations la coïncidence sur le même individu de logiques sociales de type différent. L'unidimensionnalité reste cependant très difficile à éviter: on ne peut demander à un chercheur individuel qu'il mène des recherches sur tout avec une égale compétence. En revanche, une base de données cumulative, si elle contient des informations sur les participations d'un même acteur à des mondes différents, oblige le chercheur qui l'utilise à les prendre en compte, en les lui servant qu'il le veuille ou pas, lorsqu'il recherche des données concernant cette personne. À condition qu'elle les contienne, c'est-à-dire que quelqu'un les ait chargées. Ceci pose la question du nécessaire travail collectif, et largement collectif, indispensable au plein rendement de l'instrument.

La constitution de la base, enfin, déborde largement le cadre de la simple collecte de données et impose une réflexion de fond sur nombre de problèmes.

Le premier, et certainement le plus passionnant, porte sur la délimitation de la catégorie "acteur". Définir un individu physique comme acteur est un mouvement spontané, dont la validité a été largement confirmée par la pratique empirique, que ce soit dans la vie courante ou dans le monde de la recherche. Définir un collectif comme acteur constitue en revanche une opération d'abstraction à la fois plus complexe et plus discutable. Un collectif¹⁸ en effet se décompose en multiples acteurs physiques qui a) sont les auteurs matériels de son action, laquelle n'est jamais que l'action d'un ou de plusieurs acteurs physiques au nom du collectif; b) débattent, s'affrontent ou collaborent au cours d'un processus de prise de décision collective, affrontements et prises de parties

¹⁷ Sur la question des clientèles, l'un des meilleurs ouvrages reste: Kettering (Sharon), *Patrons, Brokers and Clients in Seventeenth Century France*, New York, Oxford University Press, 1986. Cette complexité, qui implique la coexistence de logiques différentes, explique, selon nous, l'instabilité des clientèles souvent constatée, parfois de manière déroutante, par la plupart des auteurs qui ont traité la question.

¹⁸ Pour nous exprimer en des termes purement empiriques, un groupe d'individus agissant collectivement de façon relativement stable, que leur statut juridique leur donne ou non la qualité de "personne morale".

dont la connaissance est d'une extrême importance pour celle du collectif. Le collectif en revanche transcende à l'évidence les acteurs et doit, en bien des occurrences, être traité comme un tout. La nécessité de prendre en compte les deux faces irréductibles de la question, nous a conduit à un double marquage. Nous traitons les collectifs comme des acteurs à part entière, ayant une naissance, une mort, des relations avec d'autres, et ainsi de suite; mais chaque fois que le collectif agit, l'identificateur des personnes physiques qui prennent part à cette action est noté; et chaque fois qu'un individu agit au nom d'un collectif, l'identificateur de ce collectif est noté. On s'aperçoit en procédant ainsi qu'il n'est guère d'action au cours desquelles l'individu n'agisse en fonction d'un lien, soit avec un autre individu, soit avec un collectif. L'on est conduit de la sorte à une généralisation du concept de relation: à tout instant, l'acteur n'existe que par le lien qui le lie à l'autre, qu'il soit une personne morale ou un groupe.

Nous avons élargi aux objets cette notion d'acteur porteur de liens. Nous nous sommes heurtés à la difficulté de traiter des productions, en premier lieu des productions culturelles, dont la mention était indispensable à la connaissance des carrières individuelles. Nous n'avons pu résoudre la question qu'en les analysant comme des noeuds où s'entretiennent les liens entre acteurs. L'oeuvre est liée à un ou plusieurs concepteurs, à un ou plusieurs éditeurs, renouvelés à chaque réédition ou reproduction, à des critiques, des censeurs, des préfaciers, qui entretiennent avec elle des rapports particuliers. C'est la somme de ces liens, autant au moins que son contenu¹⁹, qui la définit. Or l'oeuvre est un acteur entièrement passif. Elle n'agit que par l'intermédiaire d'une personne physique ou d'un collectif qui la représente. Nous avons donc pris le parti de lui donner rang d'acteur, de la dote de son propre identifiant. Mais nous lui adjoignons, à chacune de ses apparitions, son auteur ou son éditeur scientifique du moment, qui la représente et agit en son nom.

En chargeant des données analysées finement, on rencontre au passage une foule d'institutions sur lesquelles on n'avait pas coutume de s'attarder mais que la nécessaire transparence du travail informatique rend nécessaire d'explicitier. Un fichier annexe des institutions développe ces aspects (DIEM: Dictionnaire des Institutions de l'Epoque Moderne). De même, l'examen systématique, sans exclusive ni *a priori*, des relations entre individus, et la nécessité de définir synthétiquement des types de relations pour pouvoir tout simplement manier la base, pose le problème de leur recensement et de leur classement. Si dans le domaine de l'histoire institutionnelle la mise au jour de données nouvelles est considérable, que dire de celui, beaucoup moins travaillé, des relations interpersonnelles! La base exige l'élaboration d'une syntaxe générale de ces dernières qui constitue à elle seule un programme de recherche à part entière. C'est un terrain de recherche pratiquement vierge, dont nous avons vu à quel point il est central, lui aussi, dans le nouveau paradigme. On remarquera que l'instrument, si elle rend indispensable un tel travail, le permet aussi en rendant aisément accessibles les milliers d'événements relationnels à partir desquels doivent être inférés les règles qui les régissent.

¹⁹ Le contenu d'une oeuvre n'a pas d'existence ontologique: il n'est jamais que ce que les lecteurs y lisent.

On l'aura compris, le système que nous décrivons ne saurait être un dictionnaire biographique, sinon son contraire. Ce doit être un instrument ouvert, en perpétuelle évolution, où chacun apporte et prend²⁰. La nouvelle pratique de l'histoire sociale rend impossible la fermeture d'un corpus "exhaustif". Il est bien rare qu'un chercheur extérieur trouve dans la base du neuf sur les personnages qui sont au cœur de son étude, du moins sous l'aspect sous lequel il les envisage principalement. Il y découvre en revanche systématiquement ce qu'il n'y cherchait pas: des facettes imprévues des protagonistes qu'il croyait le mieux connaître et surtout une série de connections avec d'autres dont il n'avait généralement pas la moindre idée. Nous n'infligerons pas à nos lecteurs le récit de nos souvenirs à ce propos, pour ne pas céder à la tentation du pittoresque et ne froisser aucune sensibilité. Ils nous croiront sur paroles si nous leur affirmons qu'à bien des reprises nous avons été témoin de scènes qui ne manquaient pas de piquant, les réactions allant de l'enthousiasme le plus franc au désespoir le plus profond de devoir remettre sur le chantier une étude qui paraissait achevée. Passons.

Nous avons dit pourquoi la collaboration entre chercheurs était indispensable. Nous avons mis en place, pour la rendre effective, une procédure pour le moment efficace. Nous avons élaboré le système à propos d'une recherche sur les personnels administratifs et politiques de l'Espagne du XVIII^e siècle²¹. Nous disposons à l'heure actuelle de 250 000 événements biographiques concernant quelque 43 000 personnes, pour l'essentiel des agents de la monarchie espagnole de l'époque. Le chargement des données a évidemment constitué un travail de groupe qui, d'une façon ou d'une autre, a impliqué une quarantaine de chercheurs, d'une demi-douzaine de pays. L'instrument étant arrivé à maturité, nous le diffusons en proposant à nos partenaires une mise à disposition gratuite des données déjà engrangées; ils agrègent à celles-ci celles qu'ils dépouillent pour leur propre recherche et les mettent en contrepartie à disposition de la collectivité: l'universalité du domaine d'intérêt et l'homogénéité de la mise en forme rend, redisons-le, ces données utiles à l'ensemble des utilisateurs. Nous espérons créer à terme un instrument de référence permettant à la fois l'identification des acteurs - comme le ferait un dictionnaire biographique, totalement transparent - mais encore la fourniture de séquences biographiques et relationnelles à la demande en vue de test et d'expériences.

Collaborer est plus facile à dire qu'à faire. On aura compris que, contrairement à ce que beaucoup s'imaginent, la base ne leur fournit pas toutes mâchées les informations dont ils ont besoin pour le cœur de leur étude. On aura compris

²⁰ Le dictionnaire biographique de conception classique, c'est-à-dire fermé sur lui-même, figé dans un état donné à un moment donné, n'a plus à notre avis de fonction scientifique à l'heure actuelle, pour des raisons que le lecteur déduira aisément de ce qui précède. Il garde, en revanche, sa valeur pédagogique et sa valeur politique de détermination d'un corpus de grands ancêtres de la tribu, indispensable à sa constitution identitaire. L'usage de l'histoire à cette fin reste, qu'on le déplore ou non, l'une des principales fonctions sociales de la discipline.

²¹ Groupe PAPE, actif depuis 1989 dans le cadre de l'UMR TEMIBER (CNRS / Université de Bordeaux III). Il regroupe un réseau de chercheurs français (Bordeaux, Toulouse, Paris, Montpellier, pour l'essentiel), chiliens (Université Catholique de Santiago du Chili) et espagnols (Barcelone, Alicante, Grenade, Madrid pour l'essentiel). Il a largement bénéficié de l'appui du CNRS, notamment grâce à un PICS, et du Ministère des Affaires Étrangères et trois Actions Intégrées franco-espagnoles. Il a été transféré en janvier 2005 à l'UMR LARHRA (CNRS / Universités de Lyon / ENS), à Lyon.

également que transcrire dans la base ses propres données n'est pas une mince affaire. D'un côté, l'informatique n'admet pas l'imprécision. La mise en forme, même minimale, des données, en exige, nous l'avon vu, une compréhension totale. Par voie de conséquence il est difficile de confier un tel travail à du personnel technique, sauf de très haut niveau ou sur des sections précises, isolées et parfaitement homogènes, sous le contrôle strict du chercheur: le recueil des données, au contraire de ce que l'on pense trop souvent en sciences sociales y compris, hélas, chez les autorités qui décident des politiques de recherche, est partie intégrante du processus de recherche et nullement une tâche mécanique quasi-automatisable que l'on puisse confier à du personnel annexe. Elle joue chez nous le même rôle que la conception d'une manipulation dans les sciences dites expérimentales: le chercheur peut s'y faire aider, mais il doit garder le contrôle de tous les aspects du travail sans s'en décharger sur les épaules d'un tiers, surtout quand les données sont destinées à un usage collectif et que la valeur d'un ensemble de ce type se mesure à celle de son maillon le plus faible. La prise de conscience de la dure nécessité de mettre personnellement la main à la pâte, et de l'y mettre abondamment, continûment, a découragé plus d'une bonne volonté.

La connaissance de l'informatique est une nécessité. La consultation visuelle des données et leur chargement n'exigent que des connaissances élémentaires: toutes les tâches courantes de recherche sont en effet automatisées et accessibles par boutons figurant sur l'écran. Une batterie d'écrans de présentation permet en outre d'accéder aux données sous toutes les formes courantes. Elle permet aussi l'impression, gérée par la machine, de jeux de carrières sélectionnés. En revanche, le système ne prend pas en charge l'exploitation statistique ni les recherche très complexes: il n'est en aucun cas un système d'exploitation de données historiques, mais un (presque) pur instrument de stockage et de filtrage²². Lorsque l'on a besoin de résultats fins exploitables statistiquement, il faut d'abord sélectionner l'ensemble des fiches correspondantes, puis exporter les données nécessaires sous un logiciel de traitement, tel Excel, pour prendre un exemple bien connu; les travailler ensuite sous ce dernier logiciel. Tout cela reste à la portée du chercheur moyen, à condition qu'il ait été formé au maniement d'un logiciel de traitement en question et qu'il possède les bases de l'analyse informatique. Tant que ces disciplines indispensables n'auront pas été introduites systématiquement dans les cursus d'histoire, nous aurons là un goulot d'étranglement qui gênera le développement de systèmes tels que nous les proposons.

Nous n'insisterons pas sur le problème de la propriété intellectuelle. Il ne nous a guère gêné jusqu'à aujourd'hui, contrairement à ce que nous craignons au départ. A l'évidence, l'avantage que l'utilisateur trouve à la base compense l'inconvénient de devoir partager le fruit de son travail.

²² Le système facilite les recherches par l'intégration d'un codage systématique des événements et des relations stockées. Il permet en outre à chaque utilisateur de constituer des fichiers de codage propres, en fonction des besoins particuliers, sans interférence avec les codages effectués par d'autres chercheurs. Nous considérons que l'exploitation des données pouvant se faire dans les directions les plus variées en fonction des besoins de chaque recherche et selon des évolutions imprévisibles, la constitution d'un logiciel voulant la prendre en charge serait vouée à l'échec.

La grande question est celle de la gestion de la base. Plus elle grandit, plus elle est utile, car plus de connexions elle permet. Or l'introduction des données est nécessairement décentralisée. La première tâche du centre de gestion qui piloterait l'opération consisterait donc à récupérer les données chargées par les groupes de travail indépendants, à les valider et à les intégrer dans la base principale, puis à redistribuer les exemplaires mis à jour à l'ensemble des participants. Si, comme il est souhaitable, la base est un jour installée sur la toile mondiale et accessible à distance, cette fonction en sera transformée, mais non abolie: validation et agrégation des données resteront identiques. Le centre de gestion devrait par ailleurs assurer la formation des utilisateurs, l'évaluation de leur besoins; guider leurs attentes, vérifier leur pratique et proposer les évolutions souhaitables du système en fonction de la demande. Il devrait enfin, et c'est en cela qu'il dépasserait une pure fonction de service, coordonner ou mener les recherches nécessaires à la compréhension des mécanismes institutionnels ou relationnels nécessaires à l'interprétation des données. La création d'un tel organisme devient urgente.

Du point de vue technique, la base est tenue pour le moment sous le logiciel FileMaker.Pro, dont nous apprécions la souplesse, la puissance et l'ergonomie. Elle est transposable sous tout logiciel de base de données qui n'exige pas des champs de longueur fixe. Nous donnons en appendice, à titre indicatif, la liste des fichiers et des champs constituant le système.

Jean Pierre Dedieu, DR, UMR LARHRA, CNRS / Universités de Lyon / ENS
jean-pierre.dedieu@ish-lyon.cnrs.fr

Annexes

Liste des fichiers et des champs du système FichoZ

Le système FichoZ comprend des fichiers principaux, où sont stockés les données événementielles et relationnelles concernant les individus, et des fichiers annexes donnant des éléments de contexte. Seuls les deux premiers de la liste, marqués d'un astérisque, sont indispensables au fonctionnement du système.

Fichiers principaux

**Fichier Actoz_Diccionario_actores*

Contient la liste de tous les identificateurs d'acteur utilisés dans la base, informatiquement lié au fichier Actoz_acontecimiento.. Sert à la fois à guider l'utilisateur dans le choix des identificateur nouveaux et de pivot pour le montage des carrières individuelles. Porte la donnée concernant le sexe des individus, seul caractère stable tout au long de leur carrière²³.

**Fichier Actoz_acontecimientos*

C'est le fichier principal du système. Il contient l'ensemble des évènements uni- et binominaux à raison d'un par registre informatique. Chaque registre comprend, le nom du ou des acteurs, un numéro d'identification de celui-ci ou ceux-ci²⁴, la date de début et de fin de l'évènement, la liste des sources utilisées, des remarques et la mentions d'évènements annexes. Il contient également un code, inséré par l'administrateur, qui permet de retrouver aisément tous les évènements d'une même classe (voir ci-dessous le paragraphe sur le fichier Codipost). Un champ donne l'âge de l'acteur ou des acteurs concernés à la date de l'évènement. Celui-ci est automatiquement calculé par l'ordinateur si l'on possède pour le personnage une année de naissance et pour l'évènement le millésime.

Fichier Actoz_documentos

Lors du dépouillement des actes publics (protocoles notariaux, registres paroissiaux) ou de pièces de correspondance (lettres, billets, etc.), Actoz_documentos reçoit le texte ou un résumé de l'acte ou de la lettre. Informatiquement lié à Actoz_acontecimientos, il permet de faire apparaître le contenu de du document dépouillé sur une moitié de l'écran et, en regard sur l'autre moitié, la liste des relations qu'il engendre. Il permet aussi de sélectionner automatiquement dans Actoz_acontecimientos l'ensemble des registres issus d'un même document public ou d'une même pièce de correspondance.

Fichier Actoz_datos_actores

²³ Sauf exceptions qui au XVIIIe siècle sont rarissimes.

²⁴ Ce numéro d'identification, attribué par le chercheur la première fois qu'il rencontre l'individu et qui suit celui-ci à jamais dans la base, permet de résoudre de façon très économique le problème des homonymies et des variations onomastiques.

Stocke au vol les données individuelles non relationnelles mentionnées dans les actes publics ou les lettres. Son utilisation accélère considérablement le dépouillement de ces pièces.

Fichier Actoz_importación

Fichier de même structure que Actoz_acontecimientos, qui permet d'importer dans la base principale de façon ergonomique les fichiers constitués par des collaborateurs extérieurs.

Fichiers de codification

Ces fichiers sont à la disposition des utilisateurs pour procéder aux codifications qui leurs semblent nécessaires pour leur recherche. Liés aux fichiers principaux, ils permettent d'accéder directement aux données que ceux-ci contiennent. Un jeu de fichiers vides est créé par copie pour chacun des utilisateurs. Chaque jeu est automatiquement branché sur les fichiers principaux dès qu'on donne aux fichiers qui le composent leur nom canonique. Pour débrancher les fichiers et laisser la place à d'autres, il suffit de transformer leur nom en leur donnant, par exemple, celui de leur utilisateur.

Actoz_codif_acontecimientos

Codifie chacun des champs du fichier principal Actoz_acontecimientos.

Actoz_codif_datos_actores

Codifie chacun des champs du fichier principal Actoz_datos_actores.

Actoz_codif_documentos

Codifie chacun des champs du fichier principal Actoz_documentos.

Actoz_codif_marcadores

Permet de marquer des événements mentionnés dans l'un des fichiers principaux, quel qu'il soit, pour son usage ultérieur.

Fichiers de données annexes

Les fichiers de données annexes contiennent des renseignements sur le contexte qui permettent d'exploiter utilement les données événementielles et relationnelles contenues dans la base. Ils sont le produit d'une recherche élaborée de grande ampleur, et ouverts à toute évolution exigée par la nature des données nouvellement engrangées.

Atlasesp

Fichier File Maker donnant, pour chacun des lieux-dits d'Espagne²⁵, à la date de 1849 environ, province, commune, population, diocèse, localisation, entre autres éléments. Cet atlas informatisé constitue un très puissant instrument d'identification des lieux figurant dans la documentation²⁶.

Cataom

²⁵ Il reste à charger environ les deux tiers des lieux dits de Galice, mais pour le reste de l'Espagne le travail est complet.

²⁶ Basé sur le contenu de Madoz (Pascual), *Diccionario geográfico-estadístico-histórico de España y sus posesiones de Ultramar*, Madrid, 1848-1850, 16 t., il a été réalisé par une équipe de vacataires payés par la Région Aquitaine placée sous la direction de fait de Clara Palmiste.

Fichier FileMaker. Index collectif des catalogues d'information de pureté de sang. Donne pour chaque information la liste des noms de famille du prétendant, de ses parents, grands-parents et arrière-grands-parents; ainsi que l'ensemble des lieux de naissance de ces personnes. Stocke ainsi le contenu de tous les catalogues qui passent entre les mains de l'équipe. Le fichier contient à l'heure actuelles les données de tous les catalogues publiés d'informations de pureté de sang des ordres militaires, des chanoines de la cathédrale de Séville, des gardes-marines et des procès d'hidalguía de la Chancellerie de Valladolid²⁷. Il constitue un puissant instrument de recherche, car il permet de localiser les données généalogiques concernant non seulement le prétendant, mais encore les ascendants, souvent présents dans la base, bien que n'ayant jamais fait eux-mêmes d'information²⁸.

Codipost

Fichier Word. Contient la liste des codes codant les événements de Fichoz_acontecimientos. Les relations interpersonnelles sont codées d'un mot indiquant leur nature, selon un système encore provisoire. Les événements unipersonnels, en revanche, ont fait l'objet d'une codification hiérarchique arborée qui permet de retrouver n'importe lequel d'entre eux selon sa classe. En particulier, la codification des postes administratif a impliqué la reconstitution de l'organigramme complet des administrations royale, seigneuriales et municipales du XVIIIe siècle dans la partie européenne de la Monarchie espagnole²⁹.

Diemart

Fichier FileMaker. Dictionnaire Institutionnel de l'Espagne Moderne. Contient la définition de tous les termes techniques, concepts, institutions et événements marquants rencontrés lors de la constitution de la base et tente d'en donner une

²⁷ Cadenas Vicent (Vicente), [Extracto de las informaciones de los] *Caballeros de Santiago. Siglo XVIII*, Madrid, Hidalguía, 1977-1980, 5 t., (1701-1777); Cárdenas Píera (Emilio), *Caballeros de la Orden de Santiago. Siglo XVIII*, Madrid, 1994 - 1995, 2 t. (1789-1799); Cadenas Vicent (Vicente), *Caballeros de la orden de Santiago que efectuaron sus pruebas de ingreso durante el siglo XIX*, Madrid, Hidalguía, 1958; Cadenas y Vicent (Vicente), *Caballeros de la orden de Alcántara que efectuaron sus pruebas de ingreso durante el siglo XVIII*, Madrid, Hidalguía, 1991-1992, 2 t.; Cadenas Vicent (Vicente), *Caballeros de la orden de Alcántara que efectuaron sus pruebas de ingreso durante el siglo XIX*, Madrid, Hidalguía, 1956; Cadenas Vicent (Vicente), *Caballeros de la orden de Calatrava que efectuaron sus pruebas de ingreso durante el siglo XVIII*, Madrid, Hidalguía, 1986-1987, 4 t.; Cadenas Vicent (Vicente), *Caballeros de la orden de Calatrava que efectuaron sus pruebas de ingreso durante el siglo XIX*, Madrid, Hidalguía, 1976; Cadenas Vicent (Vicente), *Caballeros de Montesa que efectuaron sus pruebas de ingreso durante el siglo XIX*, Madrid, Imprenta Maestre, 1957; Cadenas Vicent (Vicente), *Extracto de los expedientes de la orden de Carlos III, 1771-1847*, Madrid, Hidalguía, 1979-1988, 13 t.; Cadenas Vicent (Vicente), *Indice de apellidos probados de la Orden de Carlos III. Antecedentes que se conservan en sus expedientes*, Madrid, Hidalguía, 1965; Cadenas Vicent (Vicente), *Pleitos de hidalguía que se conservan en el archivo de la Real Chancillería de Valladolid. Extracto de sus expedientes. Siglo XVIII, extractados y papelados...*, Madrid, 1981-1996, les dix-huit premiers tomes; Salazar Mir (Adolfo de), *Los expedientes de limpieza de sangre de la catedral de Sevilla (Genealogías)*, Madrid, Hidalguía, 1995 - 1998, 3 t.; Valgoma (Dalmiro), Finestrat (Barón de), *Real compañía de guardias marinas y colegio naval. Catálogo de pruebas de aspirantes*, Madrid, Instituto histórico de Marina, 1943-1956, 7 vol.

²⁸ La saisie des données a été réalisée, en grande partie bénévolement, par Sophie Dedieu.

²⁹ Le travail est très avancé en ce qui concerne l'administration royale, encore en cours dans les deux autres domaines.

définition dynamique fonctionnelle. Encore en cours de constitution, il distingue soigneusement ce qui est certitude avérée, des commentaires de l'auteur de l'entrée sur le rôle ou le poids de tel ou tel élément. Ce fichier constituera à terme un dictionnaire institutionnel indépendant.

Diembibl

Fichier Word. Contient les références des sources utilisées pour constituer Diemart.

Diemcro

Fichier Word. Contient les dates précises d'un ensemble d'événements remarquables (batailles, décès de personnages connus, etc.) utilisés par les sources pour dater relativement des événements concernant les individus. Facilite grandement le dépouillement des dossiers, spécialement des relations de mérites. En voie de constitution, au gré des besoins.

Juris

Fichier FileMaker donnant à la date de 1789 la liste complète des villes (*villas*) et cités (*ciudades*) de la partie européenne de la Monarchie espagnole et, pour chacune, son statut juridictionnel et le nom de son seigneur (lorsqu'il est connu)³⁰.

Fichier généalogique

Fichoz_genealogía

Fichier tenu sous Heredis_2001, logiciel commercial de généalogie. Les relations de parenté n'ont pu être informatisées comme les autres dans le fichier principal Fichoz_acontecimientos, étant donné que l'établissement d'une seule relation en modifie des centaines d'autres. Nous utilisons donc un logiciel généalogique du commerce, de bonne qualité au demeurant, indépendant du reste de la base. Chaque individu y est affecté du numéro d'identification qui l'individualise dans la base principale, ce qui permettra à terme de lier informatiquement les deux fichiers.

Manuels

Fichoz_codipost

Fichier Word contenant la liste des codes insérés dans la base par les administrateurs.

Fichoz_diario

Fichier Word indiquant la chronologie des événements marquants de l'histoire de la base.

Fichoz_descriptivo

Fichier Word décrivant dans les détails chacun des champs et scripts de la base. Ce fichier constitue le manuel de référence de l'utilisateur averti.

Fichoz_matriz_general

³⁰ Fichier constitué par Catherine Leroy sur la base du *Nomenclator o diccionario de las ciudades, villas, lugares, aldeas, granjas, cotos redondos, cortijos y despoblados de España y sus islas adyacentes, con expresión de la provincia, partido y término a que pertenecen, y la clase de justicias que hay en ellas*, Madrid, 1789.

Tableau croisé sous Excel contenant la liste complète des champs, formats d'écran et scripts³¹ dont dispose le système. Donne pour chacun d'eux la listes de champs, formats d'écran et scripts que chacun d'eux utilise. Ce fichier constitue en quelque sorte le plan général technique du système Fichoz dans son ensemble.

³¹ Script (en espagnol: "guión"): désignation des programmes écrits sous le langage de programmation de FileMaker permettant d'assurer automatiquement les routines les plus courantes de gestion des données.