



HAL
open science

Encore un effort !

Jean-Pierre Dedieu

► **To cite this version:**

Jean-Pierre Dedieu. Encore un effort!. Mélanges de la Casa de Velázquez, 2017, 47-2, pp.297-301.
10.4000/mcv.7944 . halshs-01788800

HAL Id: halshs-01788800

<https://shs.hal.science/halshs-01788800>

Submitted on 9 May 2018

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.

Encore un effort!

La plupart des commentaires sur les "humanités numériques" soulignent avec enthousiasme leur immense apport dans trois domaines: la mise à disposition des sources et de la bibliographie; la diffusion des résultats; l'accès pratique à des techniques nouvelles d'analyse, essentiellement la cartographie et le dessin¹. Les contributions au présent dossier ne font pas exception à la règle.

Loin de moi l'idée de nier qu'il s'agit là de progrès immenses! Je les utilise massivement et je mesure mieux que personne l'intense accélération qu'ils apportent au travail de l'historien. J'ajouterai même un aspect rarement évoqué, à mon avis plus important encore: la mise à disposition d'un gigantesque savoir local recueilli par des entreprises collaboratives, du type Wikipedia; par des publications massives de livres rares qu'aucun éditeur ne prendrait le risque de réimprimer; ou spontanément mis en ligne par des passionnés soucieux de transmettre les connaissances qu'ils détiennent, les déposant sur le réseau sans se soucier de trouver auparavant oreille attentive à les écouter. Depuis que j'ai découvert parmi les souvenirs d'un artisan d'aujourd'hui ce qu'était cette "graine d'Avignon" (une teinture) que transportaient les navires français à la fin du XVIIIe siècle et dont ne parlait sous ce nom aucun dictionnaire technique de l'époque, je sais qu'il y a là un outil formidable pour mobiliser des savoirs dispersés rares, indispensables à l'historien.

Tout cela n'est pas une bénédiction sans mélange. On connaît les risques. Ils ne sont pas toujours suffisamment soulignés.

La publication en ligne des archives nous prive du contact direct avec les fonds. Nous consultons en fait ce que les archivistes veulent bien nous livrer. Lorsqu'on sait combien ils sont prompts à dissimuler ce qui pourrait gêner l'institution qu'ils servent - tel archiviste de l'Inquisition avait tu dans son catalogue les procès pour "sollicitation en confession" où le clergé n'était pas à son avantage -, c'est inquiétant. Pire, nous consultons ce que nous cherchons. Or, je sais par expérience la richesse de découverte que peut amener une divagation à l'inspiration du moment, dans des zones peu fréquentées de séries peu consultées (donc rarement mises en ligne), dans l'accumulation des catalogues ouverts sur la table de travail. L'affichage sur écran rend difficile ce rapprochement physique de sources diverses et ce feuillage nonchalant des liasses qui tant m'a servi.

Publier en ligne et donner à tous accès à nos travaux, le but est louable. Mais à quel prix! Soyons net, un débat entre historiens est incompréhensible sans contresens par le profane. Nous ne mesurons pas l'énorme sédimentation de savoir sur laquelle nous travaillons, nous autres spécialistes. Nous ne nous mêlerions certainement pas d'expliquer à un grutier comment conduire sa grue; parce que nous n'avons pas accumulé ce savoir qui rend intuitifs pour lui les gestes nécessaires. Il en va de même pour un débat historique. Les conséquences de tels malentendus peuvent être dramatiques. Notre collègue Pétré-Grenouilleau, mis au pilori pour ses thèses sur l'esclavage, est là pour en témoigner.

Je ne veux pas dire pour cela que nos conclusions et surtout que notre approche, faite de remise en contexte systématique, d'élargissement des perspectives, d'apports d'éléments que la mémoire, par définition sélective, écarte, ne doivent intervenir dans le débat public. Les plus grandes fiertés de ma vie de chercheur ne sont pas mes livres, ni même les étudiants que j'ai formés, mais ma participation, bien jeune encore, au remodelage de la mémoire historique de l'Espagne lors de la "transition démocratique", et l'organisation du Congrès du Vatican, en 1998, où nous avons, une

¹ Parmi beaucoup d'autres: Classen (Christoph), Kinnebrock (Susanne), Löblich (Maria), ed., "Towards Web History", *International Historical Social Research / Historische Sozialforschung*, 2012, vol. 37/4, n° 4, p. p. 97-190; Staley (David J.), *Computers, visualization and history. How new technology will transform our understanding of the past. Second edition*, New York, Sharpe, 2014 [c. 2004], XXII + 194 p.

quarantaine d'historiens de l'inquisition, fermement expliqué à l'Eglise romaine que l'Inquisition n'était pas une imposition des souverains laïques au pape, mais une invention de la papauté². Je ne peux oublier ni ce prêtre qui prit un jour comme thème de son sermon une longue interview dans *El País* de Jaime Contreras³, alors simple doctorant; ni ce théologien qui me dit privément, à l'issue de la rencontre de Rome, que nous obligeons l'Eglise à revoir son ecclésiologie; ni les multiples conférences que j'ai improvisées dans des trains ou des autocars, sur toutes sortes de thèmes historiques en relation avec l'actualité ou les lieux que nous traversons, à la demande de mes compagnons de voyage. Mais cette communication doit se faire sur un mode qui n'est pas celui du débat entre nous. Je ne suis pas sûr qu'ouvrir un accès illimité et sans médiation à des informations dont la difficulté d'accès ne réside pas seulement dans leur enfermement physique, mais aussi dans des obstacles cognitifs qu'il serait naïf de nier, soit toujours sans péril. En tout cas, cette ouverture inévitable, nous impose des devoirs immenses qu'il convient de souligner: l'accompagner, d'une part, des éléments nécessaires à sa compréhension; diffuser d'autre part, massivement, en tous lieux et par les mêmes moyens, les spécificités de notre approche, qui donnent les clefs d'une compréhension du passé détachée des mouvements les plus immédiats de la mémoire.

Ce sont là de graves soucis; mais il en est un autre qui me taraude et parfois me désespère. Près d'un demi-siècle après les mots fameux d'Emmanuel Leroy Ladurie sur l'historien qui sera programmeur ou ne sera pas, nous restons très loin du compte. Ce n'est pas à la mise à disposition de la documentation ni à la diffusion des résultats qu'il se référait, pour la bonne et simple raison que personne alors n'avait idée de ce qu'allaient être la mise en ligne des documents et la Toile Il voulait dire que l'ordinateur est un outil d'analyse des données qu'il faut placer au coeur même du travail historique, un outil de production des connaissances historiques. Je le sais d'autant mieux que j'ai vu l'informatique des lèvres de Michel Demonet, l'ingénieur qui programmait pour lui. C'est avec Michel que j'ai appris à parler Fortran. Nous étions quelques-uns en France à l'époque, André Zysberg, entre autres, dont je fus un temps proche; en Angleterre, un groupe plus nombreux autour de la revue *History and Computing*. Certains parmi nous prirent la chose au sérieux et comprirent que l'informatique - c'est à dessein que je n'utilise pas ici le mot "numérique" - exige un apprentissage semblable à celui qu'impose le maniement de l'écrit ou la conduite d'un bolide de course. Notre récompense fut grande. C'est une analyse factorielle fondée sur la réduction de sept mille procès d'inquisition en sept mille objets composés d'une trentaine de caractères chacun, qui m'a donné la clef de ma thèse. Je saisisais tout d'un coup un panorama immense que personne n'avait contemplé avant moi. Je touchai ce jour-là le bonheur absolu.

Or, qu'observons-nous aujourd'hui? Des mises en ligne de résultats, oui; des mises en ligne de documents, oui. Une intégration de l'outil informatique à l'acte même d'extraire des connaissances nouvelles de la documentation, non. Ou si peu. Certes on utilise des outils informatiques d'analyse. On fait des courbes sous Excel, on cartographie - enfin, on commande des cartes à des ingénieurs - on dessine des graphes sous Pajek - oui, mais Pajek est bien compliqué à mettre en oeuvre, trouvons plus simple, beaucoup moins puissant, certes, mais plus simple -... Je sors cafardeux de tous les séminaires, groupes de réflexion et réunions de travail auxquels je participe. Quand je me rends compte qu'une vulgaire "table pivot" est accueillie par de jeunes doctorants comme une révélation; qu'ils considèrent l'exportation d'un fichier tabulé, puis sa réimportation sous un autre logiciel, comme une opération d'une suprême difficulté; que je les vois peiner sur des tableaux Excel de plusieurs centaines de colonnes (authentique) parce qu'on leur a mis dans la tête que c'est la seule "base de données" (sic; Excel n'a jamais prétendu être une base de données) à leur portée, je me sens envahi d'une sainte fureur. Même les mises en ligne de documents pâtissent trop souvent de l'ignorance où sont leurs auteurs des règles élémentaires du traitement informatique.

2 Publié tardivement: Borromeo (Agostino), ed., *L'Inquisizione. Atti del Simposio internazionale Città del Vaticano, 29-31 ottobre 1998*, Città del Vaticano, Biblioteca Apostolica Vaticana, 2003, 790 p.

3 "Análisis de una estadística de 50.000 sentencias de la Inquisición", *El País*, 1/11/1978 (http://elpais.com/diario/1978/11/01/cultura/278722801_850215.html, consulté le 8 mars 2017)

Car entre la donnée et les instruments d'analyse, il y a la base de données. Une base de données n'est pas un simple réservoir où l'on accumule des informations. C'est un espace où cette information est transformée en données en fonction de l'utilisation qui va en être faite. Données, c'est-à-dire un ensemble d'unités homogènes décrivant chacune un individu objet du traitement futur sous la forme d'une collection de caractères dont chacun représente l'une des dimensions de l'analyse envisagée. On peut simplifier le travail en constituant des unités massives sommairement décrites, qui ne permettront pas une analyse par la machine, mais autoriseront la constitution de sous ensembles qui faciliteront l'analyse intuitive du chercheur, sur un mode traditionnel; un fichier manuel amélioré, qui fait gagner du temps mais n'apporte rien de plus. On peut aussi simplifier l'objet en limitant le nombre des caractères et le nombre des valeurs qu'elles peuvent prendre, mais là on ne fait plus de la recherche, car on a toutes les chances de retrouver à la sortie les choix de départ. La donnée historique doit rendre compte de toute l'information, et constituer en même temps l'atome de base que la machine manipulera pour livrer des conclusions. Des conclusions qui, bien entendu, auront été préparées par un paramétrage adéquat de l'instrument, effectué par le chercheur, certainement pas par un ingénieur, car ce paramétrage dont dépend le résultat est un acte fondamental de la recherche. La machine est stupide. Elle ne peut réaliser qu'un nombre extrêmement limité d'opérations (ce n'est pas moi qui le dit, c'est Turing, le père de toute informatique), qui combinées en algorithmes complexes arrivent à des résultats extrêmement raffinés. Le problème, c'est que ces algorithmes il faut les écrire. Que le chercheur (voir ci-dessus) doit les écrire. Et qu'il lui faut pour cela une double compétence. Il sera programmeur, ou il ne sera pas. Nous sommes loin du compte.

C'est une illusion d'imaginer que l'informatique utilisée dans la recherche se simplifie au cours du temps. On crée chaque jour des collections d'algorithmes nouveaux, qui sont comme autant de prêts-à-porter répondant à des tâches prédéfinies sur des objets prédéfinis. Or en recherche, il n'existe ni tâches prédéfinies, ni objets prédéfinis. Même les documents apparemment les plus structurés révèlent à l'examen un degré impressionnant de variabilité. D'où la nécessité de procéder au cas par cas, tout au plus guidé par des modèles de conduite issus de l'expérience, certainement pas par des recettes mécaniquement applicables. Bien difficile tout cela. D'où le silence assourdissant des ouvrages à la gloire des humanités numériques sur des questions aussi fondamentales que la constitution des bases de données dans le contexte révolutionnaire de l'apparition d'un instrument d'une puissance proprement gigantesque, qui reste ainsi sous-employé⁴. La situation n'est pas propre aux sciences sociales. Quelques contacts avec des physiciens m'ont éclairé à ce propos. Tout au plus, eux-mêmes, comme les sociologues et les économistes, en restent-ils à exploiter avec des outils nouveaux des bases constituées comme elles l'étaient à l'ère du crayon et du papier. Ce qui n'est pas le but de la manœuvre.

Que faire? Je ne vois qu'une solution. Introduire la programmation, même en dehors de tout contexte disciplinaire, dans le cursus des étudiants, au moins dans un premier temps de ceux qui se destinent à la recherche, comme matière obligatoire et importante. Non pas sous forme de conférences sur le numérique, mais en les obligeant par un apprentissage concret, soutenu et étalé dans le temps, à transformer une masse d'informations brutes en une collection de données directement utilisables pour obtenir un résultat planifié d'avance. Le reste suivra naturellement. Doté d'un outillage de base, le chercheur l'adaptera lui-même aux besoins spécifiques de son champ de recherche.

Jean Pierre Dedieu
CNRS (émérite) / Framespa Toulouse / IAO Lyon

⁴ Par exemple: Burdick (Anne), Drucker (Johanna), Lunefeld (Peter), Presner (Todd), Schapp (Jeffrey), *Digital humanities*, Cambridge (Massachusetts), Massachusetts Institute of Technology, 2012, 141 p.