

essai

HISTORIEN À L'ÂGE NUMÉRIQUE

Philippe Rygiel

PRESSES DE L'ENSIB

PAPERS



Historien à l'âge numérique

Philippe Rygiel

DOI : 10.4000/books.pressesenssib.6303

Éditeur : Presses de l'enssib

Lieu d'édition : Villeurbanne

Année d'édition : 2017

Date de mise en ligne : 9 avril 2019

Collection : Papiers

EAN électronique : 9782375460818



<https://books.openedition.org>

Édition imprimée

Date de publication : 1 janvier 2017

EAN (Édition imprimée) : 9791091281935

Nombre de pages : 208

Ce document vous est offert par Ecole nationale des sciences de l'information et des bibliothèques (ENSSIB)



Référence électronique

RYGIEL, Philippe. *Historien à l'âge numérique*. Nouvelle édition [en ligne]. Villeurbanne : Presses de l'enssib, 2017 (généré le 15 mars 2023). Disponible sur Internet : <<http://books.openedition.org/pressesenssib/6303>>. ISBN : 9782375460818. DOI : <https://doi.org/10.4000/books.pressesenssib.6303>.

RÉSUMÉS

Que change, pour un historien, et pour l'écriture de l'histoire, l'entrée dans l'ordre informatique ? Quels changements sont probables dans un futur proche ? Les outils de l'intellect mis à la disposition des historiens s'imposent à eux, et exigent qu'ils en connaissent le fonctionnement : l'ère des réseaux, des portails, des *big data*, de la numérisation de masse entraîne une connaissance accrue de ces techniques et une agilité intellectuelle nouvelle.

PHILIPPE RYGIEL

Philippe Rygiel, professeur d'histoire contemporaine à l'École normale supérieure de Lyon, est spécialiste de l'histoire des migrations. Il a dirigé avec Serge Noiret *Les historiens, leurs revues et Internet* publié chez Publibook en 2005.

essai

HISTORIEN À L'ÂGE
NUMÉRIQUE

Philippe Rygiel



essai

Que signifie être historien à l'âge du numérique ? N'est-ce que participer à une évolution méthodologique, comparable à celle de l'introduction naguère des statistiques pour l'analyse du temps long ? Ne s'agit-il que de profiter d'une aubaine, l'offre de ressources numérisées à destination des « chercheurs » de toute nature s'accroissant à une échelle exponentielle ? Ou, au contraire, faut-il considérer avec circonspection la masse incessamment renouvelée d'une archive de l'internet dont la définition même n'est pas stabilisée ?

L'avènement du numérique dans le monde millénaire de l'écrit change l'approche historique comme, *mutatis mutandis*, le télescope spatial a démultiplié notre regard sur la voûte céleste.

La science historique évolue non seulement en raison de la soudaine surabondance des documents et des locuteurs, mais encore par les conditions intellectuelles de sa production en réseau, ou celles, sociales et économiques, de sa diffusion. Ce livre dresse le nouveau panorama numérique du travail de l'historien et analyse les enjeux – et les risques – de l'écriture historique contemporaine.

Philippe Rygiel, professeur d'histoire contemporaine à l'École normale supérieure de Lyon, est spécialiste de l'histoire des migrations. Il a dirigé avec Serge Noiret *Les historiens, leurs revues et Internet* publié chez Publibook en 2005.

+++++

essai

HISTORIEN À L'ÂGE
NUMÉRIQUE

+++++

Philippe Rygiel

Philippe Rygiel est historien, professeur d'histoire contemporaine à l'École normale supérieure de Lyon.

Historien à l'âge numérique / de Philippe Rygiel. – Villeurbanne : Presses de l'enssib, cop. 2017. – 1 vol. (208 p.) ; 23 cm. – (Papiers ; ISSN 2114-6551)

ISBN 979-10-91281-93-5. (br.) : 25 €

Autre édition sur un support différent :

Historien à l'âge numérique / de Philippe Rygiel. – Villeurbanne : Presses de l'enssib, cop. 2017. (Papiers ; ISSN 2492-7600)

ISBN 979-10-91281-94-2 : 15 €

Dewey : 901 - 907.2

Rameau :

Historiens -- 21^e siècle

Économie numérique

Archives -- Numérisation

Portails Internet

Web 2.0

© Enssib, 2017.

Notice rédigée par la bibliothèque de l'Enssib.



+++++

SOMMAIRE

+++++

Introduction générale 7

Partie I. Archéologie d'une pratique37

Chapitre I. L'offre faite aux historiens.
Essai de description des services web offerts
aux professionnels de l'histoire contemporaine.....38

Chapitre II. Écrire et penser un site46

Chapitre III. Les sources de l'historien
à l'heure d'Internet.....55

Chapitre IV. Les sites des archives
départementales françaises..... 72

Chapitre V. Portes, portails, téléportation83

chapitre vi. Diffusion des produits
d'une recherche historique à l'heure du Web 2.088

Partie II. Réflexivités réticulaires 97

Chapitre VII. Des archives numériques
sans historiens ?98

Chapitre VIII. L'ordinateur, le réseau
et l'écriture de l'histoire.....103

Chapitre IX. L'enquête historique
à l'heure numérique.....112

Chapitre X. De quoi le Web est-il l'archive ? 129

Chapitre XI. Écriture de l'histoire
et réseaux numériques.....151

Chapitre XII. Histoire@numérique173

Chapitre XIII. Nouvelles frontières de l'historien ... 179

Bibliographie..... 193

+++++

INTRODUCTION GÉNÉRALE

+++++

La première réaction, en pareil cas [...] consiste à déclarer doctement de la technique qu'on ne maîtrise pas qu'elle est parfaitement inutile, voire funeste. Certains font de belles carrières d'épistémologues aux mains blanches sur ce type de déni : il y faut de la pugnacité et de l'inventivité, car au fur et à mesure que progressent les outils du savoir, on doit échafauder des théories de plus en plus complexes pour en contester l'efficience.

Patrick Boucheron, *Faire profession d'historien*, 2010

Les hommes vivent, depuis cinq millénaires, environnés d'écritures¹. Les ordres graphiques qui structurent l'environnement humain entretiennent des rapports étroits avec les civilisations qui leur donnent naissance. Les systèmes de signes permettant la transmission efficace d'informations à des interlocuteurs distants et souvent inconnus sont des pièces d'une instrumentation complexe. Leur naissance suppose de très érudites élaborations et mobilise souvent les plus savants esprits. Leur création est à la fois une affaire très prosaïque, parce que liée toujours aux transactions entre hommes et aux besoins immédiats des pouvoirs et une entreprise prométhéenne, dont les artisans entendent instaurer un ordre nouveau, social autant que cognitif, et souvent transformer l'homme en sa nature. Nous modernes ne faisons pas exception.

Les technologies dont il est question ici sont celles de l'intellect. Leurs mutations transforment l'ordre du pensable, de ce que l'entendement peut produire, et donc du faisable, parce qu'écrire, c'est faire surgir des formes et des relations nouvelles, c'est penser. Leurs mutations fascinent, elles sont au cœur des transformations du monde que nous habitons et particulièrement sans doute les lettrés, les érudits, sensibles aux mutations des systèmes de signes et d'écriture, qui tout à la fois structurent leurs matériaux et sont leurs outils. Leurs transformations contemporaines

1. Je remercie ici les premiers lecteurs de ce texte, Éric Guichard, Stéphane Lamassé et Élie-Benjamin Loyer, pour leurs commentaires sagaces et pertinents.

définissent pour les historiens, en tant que communauté, ou que profession, un environnement marqué de quelques radicales nouveautés qu'il est important de penser. Ce d'abord parce qu'elles éclairent les processus génératifs du monde au sein duquel nous évoluons. Les désordres financiers actuels ont partie liée à la multiplication, partiellement automatisée, en dehors de la sphère du contrôle des États, de monnaies scripturales sans référents. L'efficacité militaire et la supériorité géostratégique dépendent aujourd'hui de l'efficacité d'algorithmes et de l'architecture de systèmes d'information permettant d'atteindre quasi instantanément des cibles bien choisies [Desbois, 2015]. Nous avons, écrit Clarisse Herrenschmidt [Herrenschmidt, 2007, p. 469], changé de monde, sans y rien comprendre ni nous en apercevoir, dans le courant des années 1970 et ces transformations sont intimement liées au déploiement des écritures informatiques et au perfectionnement des écritures mathématiques qui leur sont liées.

L'intelligence historique est nécessaire à l'appréhension de ces transformations. Il est peu, dans nos démocraties occidentales, de sphère de la vie sociale qui ne suppose pas aujourd'hui une infrastructure numérique, ou ne soit pas associée à des écritures réticulaires, enserrée par elles. La chose est certes récente, si l'on prend pour référence la diffusion de certains outils, la mise en contact direct de masses d'utilisateurs et de l'outillage informatique, tout en plongeant ses racines très loin dans l'histoire de l'humanité. Si nous suivons là encore Clarisse Herrenschmidt, il nous faut regarder du côté de Leibniz pour certains concepts nécessaires à la vie en réseau, voire de l'invention athénienne de l'alphabet complet pour certaines des conditions de possibilité des technologies de l'intellect que nous utilisons. En somme, l'informatisation du monde, qui est aussi sa mathématisation et son adunation, est un processus, ancien déjà de plusieurs décennies, enchâssé en d'autres, plus anciens encore. Elle est prise en d'autres transformations, indissociable d'elles, plutôt que cause ou conséquence, qui touchent à la nature des pouvoirs humains, aux formes des activités économiques, aux conditions du déploiement de la pensée humaine. Leurs mutations simultanées définissent peut-être alors un changement de civilisation, dont les révolutions industrielles n'auront été qu'un premier moment, dont le néolithique est peut-être le seul précédent, qui fit succéder aux troupes nomades les cités et les empires dont les démêlés et les interactions ont nourri depuis des millénaires les chroniques et les récits des historiens, et dont les États-nations ont hérité tant de traits.

LE CONTEXTE

+++++

Ce monde, de part en part numérique, est celui qu'il nous appartient, en tant qu'historiens, de penser dans sa genèse, celui dont il faudra à nos successeurs scruter les traces, dans une large mesure numérique. Il est aussi celui depuis lequel nous pensons. De tels constats ont rencontré d'abord l'indifférence, voire la méfiance, d'une bonne partie des mondes savants, particulièrement peut-être des historiens (voir en ce volume « Écriture de l'histoire et réseaux numériques »). Le déploiement des réseaux s'est, il est vrai, accompagné de la diffusion d'une idéologie promettant la fin des intermédiations savantes et culturelles et parfois la mort des universités. L'informatique réseau est également l'un des outils permettant à la nouvelle caste de gestionnaires, qui a émergé il y a quelques décennies au sein des mondes universitaires, un contrôle accru des pratiques des acteurs, l'orientation des conduites selon des fins hétéronomes à cet univers et par le biais de l'ubiquité des ressources produites, le rêve aussi d'un devenir-entreprise d'universités affirmées en compétition sur un marché forcément mondial.

Ces transformations, qui associent mise en place de dispositifs informatiques, volonté de contrôle et érection en norme d'une raison gestionnaire sont en cours depuis plusieurs décennies [Gueissaz, 1999]. Elles sont souvent assimilées, par les enseignants-chercheurs particulièrement, à une attaque contre les fondements mêmes des institutions universitaires. Un historien peut ainsi écrire aujourd'hui, que ces évolutions ont voué les « professionnels du savoir à l'impuissance et au mensonge [...] ravalé les connaissances au rang d'une triste marchandise », posé en somme à la fois la question du statut du savoir dans nos sociétés et de la position offerte au savant, devant lequel semble s'ouvrir un avenir, qui, s'il est incontestablement numérique, n'est point forcément radieux [Granger, 2015]. La situation apparaît sans doute particulièrement sombre aux historiens, utilisateurs de matériaux rarement disponibles sous forme numérique, qui voient se multiplier, par la grâce d'Internet, textes et évocations du passé produits par de multiples acteurs, qui leur contestent le monopole du discours légitime sur l'histoire (voir en ce volume « L'ordinateur, le réseau et l'écriture de l'histoire »). Ils sont les tenants d'un savoir dont tant la validité scientifique que l'utilité sociale sont aujourd'hui facilement remises en cause. Vincent Duclert, ainsi, ouvre un paragraphe listant les facteurs qui menacent les fondements de l'histoire en France [Duclert, 2010, p. 35] par l'évocation du livre, auquel s'identifiaient traditionnellement la recherche

scientifique et l'excellence universitaire, qui est attaqué aujourd'hui par la technologie numérique.

Si la simultanéité de transformations systémiques des mondes savants et universitaires, pour beaucoup honnies ou redoutées par les autochtones, et du déploiement dans ces univers de l'informatique réseau explique, pour partie, la méfiance persistante de beaucoup envers ce que l'on nomme, sans plus de précisions, le numérique, ni le déni ni l'évitement ne sauraient efficacement tenir lieu de politique ou de ligne de conduite. L'historien ne peut être tel que s'il est profondément de son temps. Dans notre contexte, c'est, enchâssé en une tradition multiséculaire, utiliser, nous approprier, en tant que profession, toute l'instrumentation numérique, en tant qu'individu pouvoir y puiser, être en somme des scribes numériques, ce qui ne veut pas dire devenir des informaticiens. Nous en serions bien incapables. L'informatique est une science, qui n'est pas la nôtre, l'informaticien exerce un métier, qui exige une formation spécifique et longue.

L'idée d'une *Digital History*, affaire de spécialistes, qui viendrait sagement prendre place aux côtés d'une Clio aux atours pour l'essentiel inchangés ne convainc guère non plus. La question qui est aujourd'hui posée est celle de l'instrumentation de l'histoire et de l'historien. La connaissance historique – l'érudition et non nécessairement le discours sur l'Histoire – sera de plus en plus une histoire instrumentée faisant appel, tel est déjà le cas, à des jeux d'écriture informatique reposant sur des constructions mathématiques complexes. Tout historien qui a consulté un article en ligne mis à sa disposition par le biais d'un entrepôt numérique a formé une requête adressée à une base de données, déclenché l'écriture d'une page dont une vue calculée est apparue sur son écran, sans qu'il ait eu forcément conscience qu'il faut, pour que tout cela se produise, le formalisme d'une algèbre de Boole et souvent la possibilité de rechercher efficacement au sein de vastes amas de données des expressions régulières bien formées. L'efficacité de ses pratiques pourtant y gagnerait, tout autant que la compréhension de la nature, éminemment sociale et technologique, de l'activité historique, ou la réflexion sur la formation à donner à ceux qui seront les historiens de demain.

UN PARCOURS

+++++

La conjoncture que nous évoquons ici est loin d'être nouvelle, même si elle n'a longtemps été discutée qu'en d'assez étroits cercles de spécialistes

(voir dans ce volume « Écriture de l'histoire et réseaux numériques »). De fait, au cours des cinq ou six dernières décennies, l'outillage des historiens, leurs pratiques quotidiennes, les cadres au sein desquels elles prennent place et les positions qu'ils tiennent ont déjà prodigieusement changé. Les historiens de ma génération y sont sans doute particulièrement sensibles parce que le sentiment d'avoir traversé plusieurs mondes au cours de leur carrière leur vient assez vite.

Certains de nos maîtres n'ont jamais possédé ou utilisé un ordinateur, ou une adresse e-mail. Ils confiaient leurs manuscrits à taper à des prestataires extérieurs, s'informaient des publications récentes au prix de longues heures passées dans les salles des revues des grandes bibliothèques et recopiaient patiemment les fragments d'archives dont ils nourrissaient leurs œuvres. Pour ceux, et ils n'étaient pas bien nombreux à le dire ou à l'enseigner, qui considéraient l'activité historique comme une activité instrumentée et la mise en œuvre de techniques, les boîtes à fiches et les grands inventaires imprimés étaient à peu près le tout des instruments de travail de l'historien. La photocopieuse, la machine à écrire électrique ou le magnétophone portable faisaient alors figure encore d'outils nouveaux, dont la place parmi les instruments de l'historien n'allait pas de soi et dont l'impact sur les pratiques de recherche suscitait parfois de savantes réflexions.

Certes, il existait quelques pionniers dont les déclarations retentissantes recevaient un certain écho, qui annonçaient que l'historien de demain serait programmeur ou ne serait pas, mais la formation disciplinaire que nous avons reçue ne laissait aucune place à ce que l'on appelle aujourd'hui le numérique. Nous étions de futurs hommes et femmes de lettres, qui n'avions pas plus à nous préoccuper de machines que d'images, sinon comme de potentiels, quoique peu dignes, objets d'études. Les choses ont radicalement changé en l'espace de trente ans, du moins pour ce qui est de nos pratiques professionnelles. Il n'est pas d'étudiant de licence aujourd'hui qui n'ait accès à Internet et à un ordinateur et l'appétence de beaucoup pour le son ou le flux vidéo, bien souvent préféré au texte comme support de révision, est tout à fait frappante pour qui a enseigné au cours de cette période. Pas d'historien non plus qui ne communique avec ses collègues par le biais d'Internet, compose lui-même ses textes, voire parfois les met en page au moyen d'une machine, ou n'ait rassemblé sur son disque dur quelques milliers de fichiers glanés ici ou là. Ce simple constat favorise les raccourcis faciles et les affirmations péremptoires. Tournant numérique, révolution, peut-être ou sans doute, mais à condition de ne pas oublier que nous parlons, si nous pensons en termes de mécanisation

des tâches et d'usage de l'informatique, de transformations entamées il y a plusieurs décennies, et qui ne furent ni toujours rapides ni linéaires. Le tournant numérique, vu depuis la profession historienne, n'est pas le surgissement d'une idée ou une rupture brutale, mais une histoire de moyenne durée au sein de laquelle on distinguerait facilement plusieurs époques et dont les acteurs eux-mêmes ne pouvaient guère anticiper le cours, la relecture d'anciens textes prospectifs suffit à s'en convaincre.

De cette histoire, l'auteur de ces lignes a été depuis trente ans un acteur, comme beaucoup d'autres, pas nécessairement des plus centraux, mais avec la particularité d'un parcours ménageant une pluralité de points de vue et une présence à peu près permanente au sein des débats touchant à l'instrumentation informatique de l'historien. L'avènement de la micro-informatique conduisait, à la fin des années 1980, ceux qui s'inscrivaient dans la tradition d'une histoire sociale qui ne répugnait pas à compter, à élaborer et enseigner des protocoles de recherche adaptés à une situation nouvelle. Les premières générations d'historiens qui incorporaient à leurs pratiques du traitement mécanisé de données recouraient en effet généralement à des prestataires extérieurs, qui disposaient seuls de la puissance de calcul et des compétences nécessaires. L'historien du début des années 1990 devenait lui, une fois en possession de sa machine, aussi bien l'informaticien résident que le statisticien expert d'une petite unité de production d'analyses et de discours.

La miniaturisation, en un moment peu propice aux grandes entreprises collectives, redonnait vie au travailleur isolé, renfermait l'historien en sa communauté aussi, tout en exigeant de ceux qui se lançaient dans l'aventure un assemblage assez incongru de compétences : ne pouvaient utiliser efficacement ces machines-là que ceux qui programmaient et calculaient, ce qui supposait un parcours permettant d'envisager ces tâches sans effroi. Le quantitativiste de ces années-là était un artisan historien, propriétaire de son outil de travail et maîtrisant l'ensemble du processus de production, de l'extraction de la donnée jusqu'à l'écriture finalement livrée au public. La modernité technologique donnait, pour un temps, consistance à la figure de l'artisan historien, qui hante depuis bien longtemps les pages des ouvrages d'histoire.

Elle acculturait aussi une petite communauté d'historiens à l'informatique. C'est d'elle que souvent provinrent les premiers à se pencher sur les promesses de l'informatique réseau. Les spécialistes d'histoire quantitative, qui étaient souvent aussi des spécialistes d'histoire économique et sociale, étaient nombreux parmi les membres du comité de rédaction du *Journal*

of the Association for History and Computing, dont l'auteur de ces lignes fit partie, et le Roy Rosenzweig Center for History and New-Media, place centrale dans le paysage de l'histoire numérique, porte la trace de l'inscription de son fondateur dans les perspectives de l'histoire sociale. Certains des concepteurs de bases de données du début des années 1990 sont ainsi devenus, au mitan de la décennie, producteurs et diffuseurs d'informations numériques, par le moyen d'Internet. J'ai, comme d'autres, suivi ce chemin.

En 1996, nous lançons à quelques-uns le site d'histoire sociale Clio, abrité par le serveur Barthes conçu par Éric Guichard. Ce site avait pour fonction première de mettre à la disposition des participants à un séminaire de l'École normale supérieure (ENS) les documents relatifs à celui-ci. Il devint assez vite un site savant, offrant à un public plus large, dont la correspondance nous révélait l'existence, informations et ressources relatives à l'histoire de l'immigration. Nous en fîmes également le cadre et le moyen d'expérimentations en écriture électronique savante (voir en ce volume « Écrire et penser un site »). Le site Clio permit également la naissance de la première revue électronique d'histoire savante en langue française. Il nourrit également une réflexion, tôt entamée, consacrée aux implications de l'introduction et de la généralisation de l'informatique-réseau dans la recherche historique, à laquelle contribua aussi, pour ma part, la mise sur pied et la maintenance, de 2000 à 2005, du site de la revue *Le mouvement social*. De l'atelier du travailleur isolé, nous étions passés à la *start-up*, c'est-à-dire aussi au collectif et à la recherche de dispositifs nouveaux. Nous considérions alors que si réalisations et réflexions s'inscrivaient dans une tradition savante déjà ancienne, dont les apports n'étaient en rien rendus obsolètes par ce deuxième âge des réseaux informatiques, nous intéressait alors la possibilité d'imaginer, en tant que praticien, des formes d'écriture nouvelles, plutôt que les transformations de la diffusion d'objets pour l'essentiel inchangés.

De la dernière décennie du siècle précédent jusqu'au milieu des années 2000, c'était là l'affaire de petits groupes, en contact souvent les uns avec les autres. L'atelier Internet de l'ENS en faisait partie, qui oscillait entre invention de dispositifs savants, sociologie des temps numériques et épistémologie des sciences humaines. Le Pôle informatique de recherche et d'enseignement en histoire (PIREH) de Paris 1 était lui dédié à l'élaboration de l'instrumentation numérique de l'historien, mais aussi à l'enseignement aux étudiants d'histoire des rudiments d'informatique et de statistique nécessaires à ce nouveau contexte. Nous réfléchissions dans ce cadre, avec Julien Alerini, Alain Dallo, Benjamin Deruelle, Stéphane

Lamassé en particulier, non seulement aux contenus à transmettre, mais aussi aux modes et aux outils d'enseignement les plus adaptés à cette fin. De cette réflexion naquit le projet Analyse (< <http://analyse.univ-paris1.fr/> >), qui donna naissance à une interface web facilitant l'usage du logiciel R, et par là l'analyse statistique de grands amas de données, par des étudiants de première et de deuxième année, peu à l'aise avec les lignes de commande, tout en autorisant des paramétrages fins des procédures de traitement de données. L'objectif était d'embarquer dans une machine logicielle une partie de la maîtrise procédurale nécessaire aux tâches d'analyse, de créer en somme une instrumentation adaptée à la formation d'apprenants réflexifs plus que d'opérateurs agiles, d'offrir aussi à des étudiants qui n'avaient pas tous vocation à devenir historiens, par un enseignement d'histoire au cours duquel ils étaient mis en contact avec des sources primaires, des compétences transférables en d'autres contextes. Nos institutions, et bon nombre de collègues, toléraient alors, pas toujours de bon gré, ces coupables penchants pour la technique.

Le climat changea, peu à peu, même s'il est difficile, sans mener d'enquête, de dater cela précisément. Le numérique devint, au sens propre, une affaire d'État, de grandes entreprises aussi. Des plans, des institutions virent le jour, des budgets importants apparurent, dévolus au développement de ces choses numériques désignées sous des vocables changeants, dont le choix provoquait de violents conflits. Les commissions *ad hoc* se multipliaient, consacrées à la pérennité des données des chercheurs, aux projets de numérisation de la littérature savante, à l'archivage numérique des dossiers d'étrangers détenus par l'administration. Certains des expérimentateurs de la première heure devinrent gestionnaires de projet, experts, voire comptables, ou évaluateurs, plutôt que producteurs ou expérimentateurs. J'ai participé de ce mouvement, vers la fin des années 2000, de diverses manières. J'ai contribué au projet de l'Agence nationale de la recherche Ateliers histoire et informatique (ANR ATHIS) mené par Jean-Philippe Genet, codirigé également de 2010 à 2013 le Centre d'histoire sociale du XX^e siècle, traditionnelle place forte de l'informatique historique, qui portait plusieurs projets de ce type, dont le projet européen Hope, l'Equipex Matrice et était partie prenante d'autres. Les hasards d'une carrière m'ont ensuite amené à participer à la mise sur pied du premier Massive Open Online Course (MOOC) d'histoire proposé par une université française et à coordonner une formation à distance adossée à une plate-forme informatique.

Pour qui avait traversé ces différentes époques, les transformations étaient frappantes. Équipes de grande taille, liées à de multiples institutions, plutôt que réseaux d'interconnaissance, liste de livrables à fournir en respectant un calendrier rigide, plutôt que dispositifs bricolés dans l'enthousiasme d'une idée qui surgit, négociations, parfois âpres, avec des entreprises privées plutôt que coup de main amical quand il s'agissait de disposer des matériels et outils nécessaires. L'artisan historien devenait le rouage d'entreprises assez vigoureusement sommées de s'internationaliser et fermement incitées parfois à réfléchir aux possibilités de valoriser, monétairement de préférence, l'innovation. L'intérêt de la chose était que nous retrouvions dans le cadre de ces transformations des espaces de dialogue et de collaboration avec des acteurs, venus des mathématiques, des sciences physiques ou de l'informatique. Cela se traduisait par exemple par la naissance d'un séminaire consacré à l'analyse des données organisé conjointement par les historiens du PIREH et une équipe de mathématiciens, par la participation d'une équipe d'historiens aux journées mathématiques appliquées et sciences sociales de 2012 (Modèles et apprentissage en sciences humaines et sociales, MASHS).

S'il n'est pas possible de dériver de ce parcours individuel une loi générale du développement du numérique historien, son rappel a l'intérêt, outre sa valeur d'exemple, de souligner la diversité des points de vue qui s'expriment au travers de ce recueil. Les jugements et les propos des praticiens évoquant les transformations de l'histoire à l'âge numérique sont souvent éminemment situés, parce que cet environnement définit et recompose en permanence des places et des rôles aux traits accusés, auxquels sont associés des intérêts spécifiques, parce qu'aussi il est aujourd'hui tellement vaste que personne n'est capable de l'embrasser en sa totalité et tend à juger des évolutions générales d'après ce qu'il sait du territoire qu'il arpente depuis longtemps.

Si ce parcours détermine les temps et les lieux d'où le regard s'exerce, il n'en définit pas les propriétés. L'histoire, telle que les historiens l'exercent aujourd'hui, est prodigieusement diverse, contours et frontières de l'univers historique incertains et changeants. À lire certains textes récents, l'histoire semble parfois en voie de dissolution dans l'idéologie ou la littérature. Les romanciers se prennent à faire des romans historiques et les historiens des histoires romanesques, cependant que d'aucuns affirment que les récits historiques ne sont jamais que des fables permettant la construction de mythologies modernes. L'offensive n'est pas neuve si les vocables employés parfois le sont.

Le succès de ces thèmes ne doit pas cependant faire oublier qu'existent au sein de cet univers d'autres traditions, celle en particulier, ancienne, d'une histoire savante et érudite, plus intéressée par Mabillon que par Voltaire, qui veuille bien admettre sans doute que le savoir soit toujours situé, orienté par des intérêts, en somme rapporté à un *conatus* et fragile, mais qui tient l'approche historique pour un mode de connaissance, qui suppose l'application réglée de techniques à une matérialité, ouvrant la possibilité d'un accord ou d'un dialogue entre praticiens. De cela dérive sans doute l'intérêt marqué de ses tenants pour les transformations des conditions de production du savoir historique. En témoignent les textes élaborés au long de ce parcours et réunis ici à l'invitation d'Éric Guichard.

DES TEXTES LE LONG DU CHEMIN

+++++

Ces derniers forment un ensemble daté parfois en ses parties et hétérogène. Cette diversité provient d'abord de leur étalement dans le temps. Le plus ancien des textes présentés ici date du milieu des années 1990, le dernier de 2014. Vingt ans les séparent. Les préoccupations dont ils se font l'écho, la familiarité du public auquel ils s'adressent avec les objets évoqués, la position de l'auteur dans le champ également, ont, au cours de cette période, considérablement varié. Or ces textes portent tous clairement la marque du contexte de leur production. Cela nous a conduits à présenter chacun au moyen d'un court chapeau introductif. Ils répondaient tous, soit à une commande, soit à la logique d'une collaboration ou d'un projet. Cela en faisait sinon des textes de circonstances, du moins des textes dictés par un contexte et parfois une intention, autant savante que stratégique.

De fait ils n'avaient pas tous la même fonction. Certains étaient directement liés à des projets numériques alors en cours et parfois à la nécessité pratique de les défendre, ce qui leur donnait souvent le double aspect d'un plaidoyer *pro-domo* et d'un récit de pratiques. D'autres constituent des réponses à des demandes, assez fréquentes, d'interventions évoquant de manière synthétique les transformations associées à la diffusion des pratiques informatiques et réseaux au sein des institutions historiennes. Quelques textes enfin rendent compte d'une recherche ou d'une réflexion sur certains aspects des pratiques informatiques dans le domaine de l'histoire, particulièrement de l'histoire contemporaine. Ils furent généralement produits dans le cadre d'échanges entre gens de métier engagés

dans la production de dispositifs numériques. Ils n'évitent pas de ce fait une certaine technicité.

De cette diversité de circonstances et de fonctions s'ensuit en effet que, à leur conception du moins, ces textes étaient destinés à des lecteurs distincts, au sein desquels se mêlaient, en proportions inégales selon les cas, autochtones des espaces numériques et indigènes des mondes historiens. Conçus pour un public dont la familiarité avec le numérique était incertaine, les textes évoquant de manière globale l'histoire à l'âge numérique évitaient toute référence à des procédures techniques ou à des objets mathématiques, tout en tentant d'évaluer les enjeux des transformations en cours pour les praticiens d'un domaine souvent incités par leur formation ou leurs goûts à éviter les dispositifs émergents, ou à dénier à ceux-ci la possibilité de s'accompagner de ruptures méthodologiques ou épistémologiques. Les textes produits en de telles configurations tendaient à résoudre les difficultés d'énonciation propres à un tel contexte par l'examen d'un corpus, ou un recours à l'abstraction empruntant souvent à des références extérieures au champ. Il m'a semblé que s'ils pouvaient de ce fait être parfois difficiles d'accès, cette indépendance au regard des particularités de tel ou tel outil populaire à l'époque était aussi ce qui assurait leur lisibilité dans un autre contexte que celui de leur rédaction.

La demande à laquelle répondaient ces textes fut en outre plusieurs fois formulée en termes relativement similaires : qu'est-ce que le numérique change pour l'histoire et l'historien ? Cela avait un double inconvénient. Il fallait, tout en mettant en avant le nouveau ou le possible, éviter un discours a-historique opposant un avant et un après par essence distincts et chacun stables, alors même que nous étions pris dans des processus enchevêtrés, dont l'origine était parfois ancienne et les futurs incertains. Cela incitait de plus à évoquer les activités numériques de l'historien, ou l'environnement numérique de son travail, sous la forme d'un tout indifférencié, quand bien même l'ubiquité croissante des dispositifs numériques conduit à une différenciation des pratiques et à l'apparition d'enjeux et de controverses locales dont les termes ne sont pas identiques ni superposables. Il est possible de dire qu'il existe aujourd'hui, et depuis d'ailleurs déjà longtemps, une informatique de recherche, une informatique-réseau pédagogique, et une informatique documentaire, pour ne prendre que quelques exemples. Autour d'outils, de configurations d'acteurs et d'enjeux distincts, se structurent des champs qui ont chacun leur logique propre et ne mettent pas en jeu, ni les mêmes compétences, ni les mêmes intérêts, ni les mêmes acteurs. Traiter du numérique comme d'un tout qui

serait essentiellement distinct de la sphère historienne tend à empêcher la compréhension des situations concrètes autant que celle des positions des acteurs engagés dans ces différents domaines, qui peuvent sembler agir à fronts renversés parfois, si l'on comprend le numérique comme une entité à laquelle on serait supposé soit s'opposer soit s'abandonner et non comme la texture même des environnements savants contemporains.

La mise à disposition aujourd'hui d'un tel ensemble exigeait quelques décisions éditoriales, afin d'en assurer la lisibilité et la pertinence. J'ai considéré, choix discutable, qu'actualiser les textes réunis ici n'avait pas beaucoup de sens, c'est-à-dire par exemple y signaler les nouveautés techniques intervenues depuis leur rédaction, ôter au contraire les références à des dispositifs obsolètes, voire en replacer les propositions dans le contexte des débats contemporains plutôt que dans celui de leur conception. Deux raisons m'ont conduit à ce choix. D'une part, cela n'aurait guère fait que retarder, et de peu de temps, l'obsolescence de ces passages, d'autre part, il me semblait que ces textes souvent avaient pour valeur première de porter trace de différents moments des rapports d'une partie de la communauté historienne aux dispositifs informatiques, caractérisés chacun par une configuration particulière.

Par contre, dans un souci de lisibilité, les références à des ressources électroniques ont été mises à jour aussi souvent que possible. Les textes présentés ici, d'âge divers, incluaient des renvois à des ressources dont certaines ont depuis disparu, cependant que d'autres se modifiaient considérablement. Les références ont été vérifiées et mises à jour lorsque les objets mentionnés avaient peu varié et étaient encore accessibles. Dans le cas de ressources n'étant plus aujourd'hui en ligne, nous avons indiqué les coordonnées de celles-ci à l'époque de leur dernière consultation, ce qui permet dans certains cas de les retrouver par le biais des archives du net, particulièrement d'Internet Archive. Nous avons procédé de même dans le cas de ressources dont le contenu actuel n'était plus pertinent au regard de l'analyse proposée.

Figés en leur contenu informatif et leur argumentaire, ces textes n'en ont pas moins été réécrits pour partie. L'honnêteté m'impose de dire que cela fut fait à la demande expresse d'Éric Guichard, qui exerçait là légitimement son regard critique. Il considérait que ces textes étaient écrits en une langue dont la complexité rebutait et les enjolivures distraient, cependant que les allusions savantes ne risquaient guère de dérouter le lecteur, puisqu'elles étaient, selon toute vraisemblance, indéchiffrables. Il est vrai que l'habitude de produire des textes destinés à un public de

quelques dizaines de spécialistes, qu'il serait possible de nommer presque tous, favorise une écriture nourrie d'implicites et n'encourage pas à simplifier des formulations complexes, confiant que l'on est en la capacité de ses lecteurs à décoder ce qui doit l'être, puisque aussi bien l'on est de connivence avec eux. J'ai donc repris ces textes en explicitant certaines allusions, en mettant au jour des références implicites. Je ne l'ai pas toujours fait parce qu'une telle exigence, systématiquement tenue, imposerait une régression à l'infini et des notes bien longues. La syntaxe utilisée a été aussi simplifiée en certains passages, évitant d'inutiles obscurités, et quelques intertitres ajoutés. Il m'a semblé là encore cependant que la complexité de la phrase découlait parfois de la difficulté de la matière et ne pouvait être évitée. J'espère avoir ainsi facilité l'accès à ces textes sans avoir sacrifié le mouvement qui les animait.

Il fallait plus que l'amitié née d'un long compagnonnage numérique pour convaincre un auteur, d'abord réticent, d'entreprendre ce travail. Deux arguments ont particulièrement porté. Le premier est qu'ayant souvent insisté sur l'idée que ce que l'on nomme le numérique ne permettait pas de définir deux mondes aux propriétés essentiellement différentes, mais, à hauteur d'acteur et donc d'action, désignait un processus inachevé de moyenne durée remodelant en permanence le monde de l'historien et ce depuis plusieurs décennies, j'aurais eu mauvaise grâce à ne pas vouloir documenter cette historicité en réunissant les traces d'un parcours. Le second était que ces textes, pour partie du moins, conservaient une pertinence.

DES DÉBATS NÉCESSAIRES

+++++

Ce souci d'archives n'est en effet pas pur réflexe d'historien. L'historicité même du phénomène est l'un des éléments à prendre en compte lorsque l'on veut penser les transformations d'un métier et prendre part aux débats qu'elles doivent nécessairement susciter, c'est-à-dire aussi orienter sa conduite en tant que praticien et que membre d'une profession. La conjoncture numérique crée en effet pour l'historien un environnement dont les ressorts ne lui sont pas spontanément connus, parce que jouant en des champs, ceux de la science, de la technologie, de la grande industrie, dont il ne participe pas et dont il est peu familier. Les mutations de son environnement professionnel, de ses outils sont alors facilement perçues comme dérivant de facteurs exogènes et parfois incompréhensibles. Cela d'autant plus qu'en matière numérique, une certaine opacité est de

règle. Des décisions importantes résultent souvent d'interactions, à l'abri de portes solides, d'acteurs puissants aptes à dissimuler les enjeux des évolutions en cours sous des monceaux de discours d'accompagnement experts, souvent produits par des universitaires stipendiés [Chester, 2007, p. 85]. Les textes présentés ici rendent compte de ce moment et de cette position, par leur visée même. L'enjeu de ceux-ci n'était pas de produire une épistémologie, mais d'élaborer des représentations manipulables d'un monde changeant, vu depuis le territoire de l'historien, et plus particulièrement parfois de la position, pas toujours la même, du praticien, afin de s'y orienter et d'y agir de manière pertinente. En ce sens, ils consistent en l'élaboration d'outils réemployables mais aussi en contribution à des débats – et c'est à dessein qu'est ici employé le pluriel – toujours ouverts.

L'une des idées défendues en la plupart de ces textes, et d'autres, est en effet que les transformations en cours touchaient simultanément à tous les aspects d'un métier, c'est-à-dire tant aux conditions de production du savoir historique, qu'à celle de sa diffusion et de son enseignement. Une conversation générale sur le numérique et ses effets, surtout tenue entre spécialistes attirés, a donc peu de sens. À l'inverse, la prise en compte de l'horizon numérique dans la plupart des débats agitant la profession est nécessaire. Il s'ensuit de cela que nul n'est tenu d'intervenir au sein de débats relatifs aux changements accompagnant le développement du numérique qui ne touchent que de loin ses champs d'intervention ou de compétence. J'ai tenté autant que possible de m'en tenir à cette règle. Cela ne signifie pas mésestimer l'importance des débats auxquels on ne prend pas part. Deux ainsi, sur lesquels je suis peu revenu, me semblent aujourd'hui d'une importance particulière.

Le premier touche à l'enseignement de l'histoire, plus précisément à l'histoire que nous enseignons. Notre nature est aujourd'hui numérique. Des dispositifs techniques, qui incorporent des règles juridiques, des normes sociales, autant que des procédures mathématiques ou des stratégies économiques déterminent le cadre de notre expérience sociale. Nous enseignons une histoire pourtant, que ce soit à nos étudiants ou bien aux élèves du secondaire, qui généralement ignore assez superbement tant le droit que la technique ou la science et ne prête aux faits économiques et sociaux qu'une attention polie mais passagère, tendance que la dernière mouture des programmes d'histoire destinés aux classes du secondaire ne fait que renforcer. Si l'histoire que l'on enseigne a fonction de dire quelque chose du monde et non seulement de satisfaire une curiosité d'antiquaire ou de nourrir la nostalgie, bien vivante, de temps plus simples, qui ne

l'étaient pourtant guère, il nous faut alors, *a minima*, poser fermement que l'enseignement de l'histoire n'a pas de sens à être celui d'un canon immuable au sein duquel figurent de droit immémorial les mêmes figures et les mêmes vignettes. Cela d'ailleurs n'a jamais été le cas. Je doute qu'un bachelier d'aujourd'hui ait jamais entendu parler du tambour Bara ou de Jeanne la Huchette, ni qu'il y ait là matière à déploration, ces héros indomptables ayant bien mérité un long repos.

La question prend une acuité particulière lorsque nous devons nous pencher sur les formations plus spécifiquement destinées aux futurs historiens. Avertissements et lamentations se sont en ce domaine multipliés depuis plusieurs décennies, les expérimentations et les réalisations aussi. Il semble pourtant que le débat soit destiné à demeurer toujours ouvert. La temporalité dans laquelle nous sommes pris semble suffire à l'expliquer. Puisque l'avènement du numérique n'est pas un événement mais un processus systémique, long et inachevé, qui affecte aussi bien les mondes savants que les sociétés qui les abritent, nul ne peut être sûr des traits du monde qui vient et encore moins des formes sociales et techniques que prendra une activité savante, inconnaissable en son futur par cela même qu'elle est, de fait, de moins en moins constituée en un domaine autonome. Nous savons déjà que les ordinateurs de demain ou d'après-demain seront pour certains quantiques, pour certains des robots, intégreront sans doute des éléments d'intelligence artificielle. Nous ne savons pas quel monde les incorporera, quelle forme y prendra l'activité savante, mais soupçonnons que cet univers ressemblera peu à celui que nous connaissons.

Prisonnier de plus de nos propres pratiques, nous ne voyons pas toujours que ceux qui auront à mettre en œuvre une compétence d'historien ne seront pas forcément, au cours des prochaines décennies, d'abord ou surtout des auteurs de livres et d'articles. Tant les objets savants – atlas, bases de données, reconstitutions d'environnements disparus, encyclopédies locales, simulations – que les moyens de l'exposition de la connaissance historique se modifient sous nos yeux. Les expositions, les reconstitutions, les spectacles, les jeux, le film sont aujourd'hui déjà d'importants vecteurs de diffusion de connaissances et de contenus historiques, les plus utilisés par une partie d'un public dont l'appétit pour la connaissance historique n'en paraît pas pour autant moindre qu'autrefois, même si ses formes diffèrent. Quelle importance ces productions auront-elles dans l'activité des historiens professionnels, comment leur activité s'articulera-t-elle à celle d'autres professionnels provenant d'univers tout à fait différents, qui aura demain besoin d'une culture historique en complément

de la maîtrise d'un autre métier ? Aucune de ces questions ne semble avoir de réponse évidente. Il en résulte que nous ne pouvons enseigner les tâches et les tours de main des futurs historiens parce que nous ne les connaissons pas et tout apprentissage qui soit purement procédural – comment faire telle chose avec tel logiciel – a de bonnes chances de doter ceux qui le reçoivent de compétences, certes souvent utiles dans l'immédiat, mais rapidement obsolètes, parce que les logiciels vivent et meurent, les formats changent, les interfaces évoluent et les métiers surtout se transforment.

D'autre part, du fait des transformations incessantes de l'environnement numérique, les compétences pertinentes, de même d'ailleurs que le coût de leur acquisition, tendent à varier dans le temps, sans même parler de l'évolution des normes de la discipline, réelles au cours de la période, sous l'effet de dynamiques parfois autres. La réalisation d'un dispositif numérique pertinent ainsi, d'un site diffusant les premiers résultats d'une recherche en cours ou permettant les débats au sein d'un milieu de spécialistes, est devenue incomparablement plus aisée aujourd'hui qu'il y a quelques années. La transformation des conditions de l'apprentissage des langages et outils informatiques par la multiplication des ressources réseaux et la généralisation de la programmation par assemblage de briques logicielles rendent accessibles à un étudiant persévérant la création de dispositifs qui exigeaient autrefois l'intervention d'un programmeur expérimenté. Créer, pour ne prendre que quelques exemples, un wiki ou un forum, un sondage en ligne, n'est aujourd'hui guère plus difficile qu'écrire un blog et peut se faire sans avoir jamais vu une ligne de code ni connaître les principes qui régissent ces objets.

À cela s'ajoutent les effets de l'ubiquité croissante des dispositifs numériques. L'histoire aujourd'hui, en tant que pratique et que produit, est déjà numérique, tant par les conditions de sa production que par les moyens de sa diffusion. Elle est de plus particulièrement diverse par les matériaux qu'elle emploie et les modes de construction des objets savants que ses praticiens mettent en œuvre, ce qui favorise la multiplication de dispositifs *ad hoc*. De plus, les environnements numériques auxquels sont et seront confrontés les historiens sont prodigieusement divers, selon les tâches qu'ils ont à accomplir, mais aussi sous l'effet d'une spécialisation croissante. Participer à l'élaboration d'un MOOC en tant qu'enseignant, concevoir un blog ou un site permettant de rendre compte d'un chantier de recherches, analyser des données volumineuses, ou prendre part à la numérisation d'un corpus sont des tâches que déjà beaucoup d'historiens

ont eues à accomplir, qui à chaque fois confrontent à des dispositifs logiciels, mais exigent des compétences tout à fait différentes, parce que l'introduction du numérique dans chacune ne suffit pas à faire de l'instrumentation de la recherche, de la documentation et de la pédagogie des réalités de même nature.

De la même façon, pour ne s'en tenir qu'à l'instrumentation informatique de la recherche, travailler à l'élaboration d'un système semi-automatisé de description d'archives vidéo, approprié à l'utilisateur, parmi lesquels les historiens, au meilleur mode possible de numérisation des plans parisiens anciens ou aux systèmes de visualisation et d'analyse des données les plus pertinents pour les historiens supposent l'insertion dans des environnements informatiques radicalement différents, que ne peuplent pas les seuls historiens et qui tous requièrent une formation spécifique, pour le coup assez lourde et des compétences différentes, qu'il est assez illusoire d'imaginer réunies en un seul individu. De ce fait, parler d'une histoire numérique qui viendrait en appui d'une histoire qui pourrait paisiblement ne pas l'être a peu de sens.

Ces interrogations ne sont pas propres au champ de l'histoire, elles rejoignent celles plus générales qui touchent à la définition et à l'enseignement d'une culture numérique, ou à l'élaboration d'un humanisme numérique. Posés en ces termes très généraux, ce sont des débats en lesquels je ne suis jamais entré. Les humanités, dont j'ai bien du mal à percevoir les contours, sont pour moi quelque chose de bien trop vaste pour ne pas gommer une identité professionnelle construite en référence à une tradition disciplinaire que définit, entre autres choses, un rapport particulier de l'historien à ses matériaux, que l'effort réflexif imposé par le développement de l'informatique historique a permis de mettre en évidence, il y a déjà longtemps [Genet, 1986]. Il est fort rare en ce champ que le chercheur travaille des données déjà constituées par des opérateurs savants, dont les pratiques soient bien documentées, les catégories compatibles avec les schémas conceptuels des chercheurs et la fiabilité éprouvée. Plus rare encore qu'il dispose de relevés de mesures physiques obtenus par le moyen de capteurs aux propriétés connues, ce qui peut être le cas de certains géographes, voire aujourd'hui de sociologues ou d'économistes examinant les flux qui traversent les villes connectées.

La donnée de l'historien est généralement fabriquée par celui-ci, au moyen souvent de jugements experts complexes, impliquant une modélisation de l'information pertinente, à partir des débris de mondes sociaux disparus, qui n'avaient pas pour fonction première de servir de support

à une entreprise de connaissance et dont la production obéissait à des règles parfois inconnues ou dont la reconstitution exige un long travail d'enquête. Le manque, l'incertitude, l'imprécision aussi sont ici la norme en même temps que le détournement d'usage et le doute systématique. Il est rare de plus qu'existe un double numérique des matériaux ayant permis la production des données d'une enquête, et cela demeurera vrai longtemps. Les évolutions de la discipline dérivant à la fois de mutations des schémas conceptuels mis en œuvre et de mise en exploitation de matériaux nouveaux, il est peu fréquent de plus, même si le cas peut se produire, que les données élaborées dans le cas d'une enquête soient réutilisées, quand bien même elles pourraient l'être. Il s'ensuit que l'acquisition de la donnée est pour l'historien une opération coûteuse dont certaines phases ne peuvent que difficilement être automatisées.

L'information obtenue n'est de plus utilisable que si est conservé le lien aux matériaux qui en ont permis la production et conservée la mémoire des opérations effectuées, ce qui crée des contraintes fortes en matière de gestion de l'information. Il faut en particulier prévoir des mécanismes de rétroaction sur les données au fur et à mesure des progrès de l'enquête. Les données obtenues de plus, en raison d'effets de seuil, ou d'autres de leurs propriétés, ne se prêtent pas toujours à des modes d'analyse standardisés élaborés en d'autres contextes. Si nous n'en concluons pas que chacun doit souffler sa propre forge, l'idée d'une instrumentation à la fois radicalement distincte de celle des sciences dites exactes – alors même par exemple que les problématiques de l'e-biologie ne sont pas si différentes [Roux, 2009] – et commune à tout ce qui n'en fait pas partie – du géomaticien au spécialiste de la science des textes ou de la visualisation de données – convainc peu. Il gomme la nécessité d'une appropriation en contexte, qui est pour partie un contexte disciplinaire [McCrank, 2001], tout autant que la parenté des pratiques de modélisation quel que soit l'univers au sein desquels elles se déploient.

L'expression d'humanités numériques, utile dès lors qu'il s'agit de fédérer les demandes d'acteurs engagés dans la construction d'infrastructures numériques dont historiquement ne disposaient pas de nombreuses communautés savantes, paraît par contre trop étroite pour le praticien qui doit penser et agir dans un monde numérique qui engage à renouer le commerce entre tous les littérateurs : historiens, orateurs, mathématiciens et physiciens, tous ces manieurs de symboles que Colbert, aux dires du vieux Fontenelle, souhaitait réunir en une seule académie.

Posons cependant que, quelles que soient les positions des uns et des autres, tous les acteurs des mondes érudits impliqués dans les chantiers numériques font un même constat. Il faudra aux praticiens futurs une culture leur permettant de travailler et d'agir dans un monde saturé plus encore que le nôtre par les dispositifs numériques. Il n'est pas certain de plus, et les historiens ne font pas exception à la règle, qu'un accord règne ou soit possible sur les contours de celle-ci. Cela tient à ce que la tâche est de fait difficile pour les raisons évoquées plus haut, mais renvoie aussi à des divergences tactiques ou professionnelles liées souvent aux positions institutionnelles des uns ou des autres, plus ou moins attachés à l'autonomie d'une discipline ou d'un champ, aux probabilités d'une alliance avec tel ou tel type de partenaires, choix dont les conséquences ne sont pas neutres en termes de postes, de budgets ou de maquettes d'enseignement, et d'autant plus avidement scrutées que la position au sein des institutions universitaires – elles-mêmes bien mal en point – des lettres, des sciences sociales et des disciplines érudites n'est aujourd'hui guère solide.

Si j'ai en ce domaine quelques convictions, elles ne font pas exception et sont le produit d'un parcours. La première est qu'au regard de l'extrême volatilité de l'environnement professionnel qui est le nôtre, la définition d'un cadre et d'un canon pérenne qui soit une liste d'outils n'est guère possible. Ni l'apprentissage du Common Business Oriented Language (COBOL) ni celui du maniement des feuilles semi-log ou du pied à coulisse n'ont aujourd'hui de sens, alors que d'honnêtes pédagogues y consacraient autrefois de longues séances en ayant le sentiment justifié de faire œuvre utile.

La forte technicité de certaines tâches et la tendance de longue durée à l'hyperspécialisation (à chaque historien un domaine, un type de sources, un terrain, voire une question ou un type d'intervention) ont peu de chances de s'atténuer dans un futur proche, empêchant d'imaginer former un historien qui serait à sa sortie de l'université paré à toutes les éventualités et apte à mettre en œuvre immédiatement toutes les techniques à disposition, effet aussi d'un contexte qui est à la fois marqué par l'approfondissement de la division du travail au sein des mondes savants et l'intégration accrue de ceux-ci à d'autres sphères sociales (voir en ce volume « L'enquête historique à l'ère numérique »). Dans le même temps, les réformes successives des études doctorales et la brutale concurrence à laquelle sont aujourd'hui exposés les jeunes historiens impliquent une maîtrise immédiate des outils les plus adaptés à la tâche, tout en les privant du temps nécessaire à des apprentissages complexes. La thèse en

change en nos disciplines de nature, qui n'est plus l'occasion d'un apprentissage, mais la monstration, qui tend à devoir être plusieurs fois répétée, de ce que l'on sait déjà faire.

Il en découle que la formation doit être, pour partie au moins, déportée vers l'amont et tout à la fois donner à chacun un socle solide qui permettra dans le temps long d'un parcours des appropriations et des apprentissages, offrira une compréhension des outillages disponibles et autorisera la différenciation des parcours. Disons, pour prendre un exemple, qu'un futur historien probablement amené à manier des données sérielles n'a pas besoin pour achever sa thèse de maîtriser à la fois les techniques d'analyse multivariées, la régression logistique, l'analyse réseau ou la modélisation des processus markoviens, mais qu'il est utile qu'il connaisse l'existence de ces techniques, leurs contextes d'usage et dispose des prérequis permettant de les aborder si besoin, autant que d'un accès à un lieu l'y autorisant.

Cela conduit me semble-t-il à prôner, autant qu'elle est possible, une formation initiale qui fasse une large place à l'abstraction, qu'elle soit méthodologique – parce que ce contexte transforme l'historien en stratège en permanence contraint à optimiser ses apprentissages et son temps de travail, ce qui suppose une forte réflexivité –, mais aussi à des rudiments de mathématique et d'informatique, parce que ce sont là les gages de futurs apprentissages possibles, mais aussi la condition de possibilité d'échanges avec les professionnels du traitement de l'information (voir en ce volume les « Nouvelles frontières de l'historien »).

Un tel programme est difficile à mettre en œuvre au sein de nos institutions, particulièrement en temps de pénurie, qui n'abritent ni nécessairement les ressources humaines, ni non plus toujours un nombre suffisant d'étudiants en besoin des mêmes apprentissages, alors même que le temps de formation est simultanément un temps de production et doit ménager apprentissages fondamentaux et possibilité de l'appropriation d'une palette différenciée d'outils immédiatement mobilisables dans le cadre des premiers travaux de recherche. Un département d'histoire n'a guère de raisons d'ouvrir un atelier de formation à la cartographie assistée par ordinateur pour le bénéfice des deux ou trois étudiants ayant besoin d'acquérir cette compétence, non plus qu'un séminaire dédié à la visualisation de données. L'autoformation, longtemps la règle, atteint cependant aujourd'hui ses limites. Grande dévoreuse de temps, elle a de plus le défaut de ne guère permettre une socialisation centrée sur l'usage des outils, souvent particulièrement efficace, nécessaire aussi, parce qu'il est douteux que le contexte permette à beaucoup d'être des intellectuels

isolés peaufinant leur œuvre dans une solitude paisible. Il n'y a guère d'autre perspective qu'une mutualisation, éventuellement d'ailleurs assistée par le réseau, entre disciplines et entre établissements, qui suppose la possibilité d'une circulation assez libre des étudiants en cours de formation entre départements d'une même institution et entre établissements. La chose n'est pas tout à fait simple, pas impossible cependant et sans doute nécessaire.

UN TOURNANT RÉFLEXIF

+++++

Insister sur l'importance de tels débats ou de telles conversations revient à dire, les lignes qui précèdent le montrent, que l'un des effets de la conjoncture que nous traversons est l'interrogation, sous un angle, nouveau parfois, des pratiques savantes, examinées non seulement en leurs produits, leurs intentions ou leurs déterminants idéologiques et sociaux, mais en les circonstances de leur actualisation. Ce mouvement, pour partie lié à la nécessité d'un ajustement constant des pratiques à un environnement changeant et peu maîtrisé, peut également être mis en relation avec le tournant réflexif théorisé par certains auteurs de sciences sociales dès le début des années 1990.

Les sciences historiques n'ont pas été les plus touchées ou les plus précocement atteintes par ce mouvement, dont cependant les travaux issus de l'ANR ATHIS que menait Jean-Philippe Genet ont prouvé toute la fécondité, montrant, par exemple, qu'à l'élaboration d'une instrumentation nouvelle répondent, pour l'étude du texte médiéval, une redéfinition des objets d'études et un renouvellement des concepts à l'œuvre. Les textes rassemblés en ce recueil participent à leur façon de ce moment qui interroge l'activité historique en tant qu'elle est manipulation instrumentée d'une matérialité, permettant parfois ainsi d'en imaginer plusieurs futurs possibles, autorisant quelques fois l'élaboration de conjectures dont certaines me semblent toujours valides.

Si nous suivons ce fil, nous sommes amenés à considérer que la pratique historique est d'évidence une pratique sociale, ce qu'écrivait, il y a déjà longtemps, Antoine Prost [Prost, 2010], inscrite en un environnement aussi bien institutionnel et politique qu'économique ou que matériel, dont les évolutions sont déterminantes pour qui veut comprendre les dynamiques des transformations de la discipline et du métier qui lui est associé. La nature numérique en laquelle sont plongés les acteurs n'en

demeure pas moins le produit de processus politiques et sociaux et non le simple déploiement dans le réel d'une idée dont serait porteuse la nouveauté technologique.

Les injonctions ainsi adressées aux universitaires, de plus en plus sommés de s'inscrire dans des dispositifs de formation partiellement ou totalement médiatisés, ou que certains souhaiteraient fermement inviter à déposer sur un répertoire central toutes leurs productions, se comprennent en référence à des processus politiques, à des intérêts économiques, à des choix stratégiques également, ne dérivant des propriétés des dispositifs numériques qu'à titre d'accomplissement d'une virtualité, ceux-ci n'étant en l'espèce que condition de possibilité.

Du caractère social des pratiques découle leur historicité, mais aussi leur plasticité et leur relative indétermination. Si les mutations de fait de nos métiers, au cours des quelques dernières décennies, sont considérables, elles ne sont ni soudaines ni achevées, pas plus qu'elles ne peuvent se lire toutes entières dans les entrailles des machines qui trônent sur nos bureaux ou les pages de code qui en règlent le fonctionnement, d'où, pour les praticiens, à la fois des marges de manœuvre et des difficultés, les unes comme les autres changeant en fonction des contextes. Les textes rassemblés ici en témoignent. Par force, ils ne disent pas celles qui apparaissent aujourd'hui les plus actuelles ni ce qu'il serait aujourd'hui nécessaire de penser.

Le contexte semble ainsi se prêter particulièrement à une historicisation du rapport de la profession historienne au numérique et à l'instrumentation, qui permettrait de rompre avec un rapport fétichiste à celui-ci. L'autorise tant la disponibilité de sources que l'épaisseur temporelle, si même nous ne nous interrogeons qu'au rapport à l'informatique-réseau ou à son aspect le plus visible qu'est le Web. Un exemple rapidement évoqué permet d'éclairer ce point. L'écriture historique destinée à la mise en ligne est une réalité à la fois massive et ancienne. Ses formes cependant ont prodigieusement changé au cours des vingt dernières années.

Les premiers sites savants, au milieu des années 1990, sont produits par des individus ou de petits groupes. Le site Clio, dont nous pouvons retrouver trace de la forme qu'il a alors grâce à Internet Archive², fait partie en France de cette première génération. Un regard au code de ces pages permet quelques constats. Il y a une vingtaine d'années, écrire pour le Web c'est très littéralement écrire le Web, c'est-à-dire ouvrir un éditeur

2. < <http://web.archive.org/web/20000119121407/http://barthes.ens.fr/clio/indexoktxt.html> >.

de texte et taper une chaîne de caractères. Ce n'est pas coder, mais écrire un texte que l'on va baliser, exactement comme aujourd'hui on utilise l'eX-tensible Markup Language (XML) pour éditer un texte. Il faut également construire l'architecture des fichiers et celle des contenus, parce qu'il n'y a pas à l'époque de personnage qui prenne en charge ces fonctions quand on veut arriver à diffuser de manière électronique du contenu.

En termes de compétences, cela suppose de connaître un langage de description, l'HyperText Markup Language (HTML), être capable de se connecter à un serveur File Transfer Protocol (FTP), ouvrir une connexion Secure Shell (SSH), donc être capable de taper quelques lignes de commande et bien maîtriser les notions liées à la construction d'une arborescence. Cela ne suffit pas dans la seconde moitié des années 1990 à créer un contenu accessible. Il faut aussi être en contact avec un administrateur réseau ou le responsable d'un serveur convaincu que vous n'allez pas gaspiller de l'espace disque ou de la bande passante avec des fichiers qui ne sont pas vraiment de la science. C'est peut-être à l'époque le plus difficile.

Tout cela paraît un peu antique et rudimentaire, mais cela suffit déjà à faire beaucoup de choses. Il est possible de gérer des listes de diffusion, de proposer l'accès à des bases de données, même si de manière beaucoup moins souple qu'aujourd'hui. Il est aussi possible de mettre en place assez facilement des dispositifs interactifs. Il n'y a pas beaucoup cependant d'utilisateurs et la plupart découvrent un univers qui leur est tout à fait étranger. De fait, les interactions passent souvent par le mail, et l'on découvre alors qu'un site savant génère parfois beaucoup de courriers, c'est-à-dire qu'il est une plate-forme d'interaction, et parfois surtout cela, pour un collectif assez complexe en ses formes.

Faisons maintenant un saut d'une dizaine d'années. Nous sommes en 2005. Un site savant d'histoire ressemble alors souvent au site d'une revue d'histoire telle que *Le mouvement social*, ici à l'été 2003³. Nous avons désormais affaire à une institution de la recherche et non plus à un réseau informel de jeunes chercheurs. Nous sommes passés du fanzine au site attaché à une institution. Cela correspond à une évolution plus générale, qu'il faudrait une enquête pour décrire avec plus d'exactitude, mais la multiplication des sites attachés à une institution de recherche est bien l'une des nouveautés de la première moitié de la décennie écoulée.

L'évolution du design et de l'ergonomie est assez nette. Nous arrivons en quelques années à une forme très proche des normes actuelles. On

3. < http://web.archive.org/web/20030802052615fw_/http://biosoc.univ-paris1.fr >.

reconnaît la région centrale qui renferme l'information spécifique à la page, le bandeau de navigation supérieur sous forme d'un menu déroulant, les deux bandeaux latéraux réservés pour les renvois externes. Le code d'une telle page est beaucoup plus touffu et long que le précédent. Il fait appel à JavaScript, et contient une référence à des feuilles Cascading Style Sheets (CSS) externes. Cela renvoie là à l'évolution de normes et de pratiques tout à fait hétéronomes au champ historique et aux mondes savants, d'une part la dissociation accrue du contenu informationnel et de sa mise en forme, d'autre part la multiplication de scripts permettant un affichage changeant en fonction des actions du lecteur, enfin la diffusion de normes de bonnes pratiques venues de l'univers de l'écriture logicielle. Nous pourrions lier certaines de ces évolutions à des transformations techniques. L'usage accru d'éléments images et de scripts correspond en effet à la généralisation des connexions permanentes haut débit, qui font abandonner l'obsession des premiers temps, qui était celle de la légèreté des codes et des pages.

Du point de vue du producteur, l'activité exige l'acquisition de compétences supplémentaires, un langage de script est plus complexe à mettre en œuvre qu'un langage de description. Cela s'accompagne d'une routinisation des tâches qui est liée à un changement d'échelle et une redéfinition des rôles. Un personnage nouveau apparaît qui est le webmestre. Il n'est pas l'administrateur réseau, mais tend à défendre vivement l'exclusivité de l'accès au secteur du serveur dont il a la responsabilité et est le garant d'une charte graphique ou d'une homogénéité des codes utilisés. De plus, la production de la ressource est insérée dans la politique globale de communication de l'institution, à l'époque terriblement prudente et hésitante tant pour ce qui est de la détermination des contenus à mettre en ligne que quant aux formes de celle-ci. Du point de vue du scripteur historien, la production de ressource électronique soit disparaît de son horizon, puisque prise en charge par des intervenants dont il est le fournisseur et non l'interlocuteur, soit perd de son intérêt puisque routinière, une fois la structure mise en place, coûteuse en temps et sans vertu réflexive.

Dix ans plus tard, c'est-à-dire aujourd'hui, les formes d'écriture savante électronique des historiens sont beaucoup plus massives et diverses. Les productions électroniques des revues sont généralement prises en charge par des portails ou des prestataires extérieurs, donc des intervenants extérieurs au champ disciplinaire, inaugurant une relation symbiotique entre des entreprises privées ou publiques et les communautés savantes, qui est en cours de redéfinition aujourd'hui. L'écriture numérique des chercheurs,

dans le domaine de l'histoire comme en d'autres, tend à prendre la forme du blog, que celui-ci soit personnel ou corresponde à l'activité d'un petit groupe de projet, hébergé par une plate-forme ou autonome. Les sites de ce type ont souvent un air de famille, que l'on comprend vite lorsque l'on regarde le code de l'un de ceux-ci, par exemple < <http://cliosoc.fr/drupal/> >.

Ce code est substantiellement différent des précédents. Nous sommes dans un site dynamique, généré au moyen d'un Content Management System (CMS), Drupal en l'espèce, mais nous aurions quelque chose de très similaire si nous regardions un site fabriqué avec SPIP, ou Typo 3. L'architecture en effet est dans tous les cas similaire, reposant sur un couplage de scripts Hypertext Preprocessor (PHP) et de bases de données Structured Query Language (SQL). Le résultat le plus spectaculaire en est la prolifération du code. Il faut ici environ six cents lignes pour afficher un mot de bienvenue de dix lignes.

Là encore, nous pouvons lier cela à des évolutions techniques et industrielles. Le fonctionnement de sites de ce type suppose en effet une forte augmentation de la puissance des processeurs et des débits réseaux, l'affichage d'une page pouvant nécessiter le traitement de dizaines ou de centaines de requêtes, façon de rappeler que les pratiques scientifiques en environnement numérique ne sont pas autodéterminées. Il nous faudrait donc, si nous voulions être précis, nous interroger sur les chemine-ments des outils développés en dehors de nos communautés scientifiques et adoptés par elle, pour comprendre certaines des logiques aujourd'hui en œuvre. WordPress, Drupal ou SPIP, pour n'évoquer que certains des plus communs, n'ont pas émergé au sein des communautés de sciences humaines, pas plus qu'ils n'ont été conçus pour en favoriser les pratiques. Ainsi, les premiers sites pour lesquels SPIP a été élaboré sont des sites d'éditeurs ou d'entreprises de presse [SPIP, s.d.].

Ces outils tendent de ce fait à incorporer des modèles de distribution des tâches, distinguant des rôles, et reproduisent les formes de la division du travail caractérisant ces industries, dont l'une des manifestations est un rapport différencié au code et au serveur, codant, au sens propre, des fonctions, dont nous sommes tentés d'écrire qu'elles ont de fortes chances d'apparaître comme une quasi-nature aux utilisateurs, qui serait imposée par la technique même et non le produit de choix de programmation ou de paramétrage. De fait, nous parlons aujourd'hui de contributeurs, qui, par le biais d'une interface graphique reproduisant ou à peu près l'aspect d'un logiciel de traitement de textes, vont déposer un texte en langue naturel,

sans avoir nécessairement conscience des transformations que celui-ci va subir avant d'apparaître sur un écran.

L'administrateur du site, qui est rarement celui du serveur, peut lui gérer la structure des contenus, les modes d'affichage, les accès des utilisateurs par le moyen d'une interface graphique, sans qu'il lui soit besoin de maîtriser un langage informatique ou d'écrire une ligne de code. Il déclenche la génération d'un code mais ne l'écrit plus. La production du code informatique est redevenue, dans une large mesure, l'apanage de communautés informaticiennes qui coproduisent les machines logicielles qu'actionnent les utilisateurs/contributeurs. L'envers en somme de ce qui est souvent présenté comme une libération ou une démocratisation de l'accès à l'écriture en réseau est une très stricte taylorisation qui assigne au moyen d'écritures, non pas invisibles mais rarement vues, les utilisateurs à des rôles souvent strictement définis.

La logique de ces évolutions a là encore peu à voir avec une quelconque spécificité disciplinaire, renvoyant à l'évolution d'une organisation productive qui permet d'économiser une compétence informatique relativement rare en l'incorporant dans des machines logicielles actionnées par des opérateurs semi-qualifiés. Il est alors possible de produire à moindres frais, du moins en économisant le travail qualifié, une multitude de micro-ou de minisites standardisés. L'un des effets de cette évolution est de renforcer la séparation, un temps ébranlée dans les pratiques, du contenu informationnel ou analytique et des modalités de sa présentation et de sa transmission, alors même que le développement de l'informatique réseau s'est accompagné en d'autres champs de l'évolution inverse [Villani, 2012].

Ce ne sont là bien sûr que quelques remarques, qui ne prennent en compte ni le contenu ni les fonctions de ces objets électroniques et ne sont nourries que de l'observation de quelques exemples. Elles suffisent à montrer la possibilité d'une historicisation des écritures savantes, par le biais de leurs conditions de production et, mettant en lumière par exemple l'absence d'autonomie technique de beaucoup de sphères savantes, à en montrer l'intérêt.

Nous sommes, observant, rapidement, une certaine forme d'écriture électronique historienne, disons les sites savants mis en place par des chercheurs ou de petits collectifs d'agents patentés, confrontés à un régime complexe de temporalité. La possibilité d'une écriture savante en réseau introduit bien sûr des ruptures, ne serait-ce que par la prodigieuse diminution, pour le producteur, du coût de dissémination de l'information et l'accélération de sa transmission. De ce point de vue, d'autres aussi,

il est possible d'associer l'ouverture de la possibilité de l'écriture en réseau à de fortes discontinuités.

Fonctionnellement cependant, ces sites savants, qui ne sont nullement d'ailleurs la seule forme de présence historique en réseau, s'inscrivent aisément dans des histoires très longues qui sont celles de la communication et des échanges au sein des communautés savantes, ou bien des revues. L'enjeu en effet pour la plupart de ces sites demeure la sélection, la qualification et la dissémination de l'information au sein de groupes relativement restreints de spécialistes et d'étudiants, mais aussi la mise en relation, nous disons aujourd'hui en réseau, de microsociétés.

Faire du récit du déploiement du numérique au sein de ce secteur celui d'une rupture au sein d'une histoire longue semble alors particulièrement réducteur. La multiplicité en effet des rôles et des positions, dont tant les contenus que la distribution ont connu et connaissent des évolutions contrastées et pas toujours linéaires, conduit à penser que la perception des dynamiques à l'œuvre est fortement différenciée au sein de cet univers. Nous avons ainsi croisé des contributeurs, des administrateurs, des propriétaires aussi de serveurs, et à une autre échelle, des industriels, des marchands, des États, toute une épaisseur sociale et technique dont l'oubli est la condition même de l'ébahissement fréquent devant les objets numériques présents dans notre quotidien. Si, à l'inverse, nous ne nous intéressons qu'au producteur historique de ressources électroniques contribuant à un site savant, nous dirions que s'il a dû, durant une petite dizaine d'années, faire face à des situations exigeant une forte montée en compétences, il tend à devenir un simple fournisseur de chaînes de caractères servant des machines logicielles dont l'appropriation est de fait impossible à la quasi-totalité des opérateurs, ce qui conduit à s'interroger, du point de vue du producteur, sur sa capacité à disposer d'un accès pérenne et efficient aux documents ainsi produits, mais aussi sur sa capacité à en codéfinir les fonctions et les formes. Cette histoire d'autre part a ses sources, écrites pour certaines et disponibles, pour partie sous forme de codes, qui témoignent des dynamiques à l'œuvre.

Cette historicité du code lui-même est un point particulièrement important. Pour les historiens du contemporain, elle est à la fois l'un des principaux défis à venir et une ressource. Une ressource parce que celui-ci est signifiant (voir dans ce volume « Des archives numériques sans historiens ? »). Un défi parce que si les sources permettant l'étude des deux ou trois dernières décennies et de celles qui viennent sont pour une très large part numériques, l'évolution des formats et des codes, si même

nous ne considérons que le Web, implique que les modes de captation et de manipulation de ces ressources sont appelés à varier selon l'âge des ressources exploitées. Le Web offre un matériau changeant à la préhensibilité variable en ses formes selon les époques. Pour l'historien, cela veut dire que les conditions de son utilisation, elles-mêmes dépendantes de celles de son archivage, sont pour partie dictées par l'évolution des formats et des langages utilisés. Une collection de pages personnelles des années 1990 constitue un ensemble extrêmement hétérogène, qui exige, si l'on veut en faire une exploitation systématique, un traitement lourd, qui n'est véritablement envisageable qu'au moyen de la mise en place de traitements informatiques complexes et d'une bonne part de manipulation. Une plate-forme de blogs contemporaine, à l'inverse, est une base de données structurée de grande taille, autorisant assez facilement à extraire une information jugée pertinente dont le traitement permettra la mise en œuvre des outils de l'analyse et de la visualisation de données, mais dont la taille défie l'exploration systématique par le moyen d'une lecture cursive. Nous sommes en ce cas proche des problématiques du *big data*.

Nous sommes pris donc dans un contexte qui exige, pour une exploitation efficiente des sources primaires nécessaires à l'historien, la réunion d'une intelligence de la source et de l'historicité de l'espace étudié, d'une intelligence informatique et d'une intelligence mathématique, qui peuvent d'ailleurs être distribuées entre plusieurs acteurs et pour partie incorporées au sein de machines logicielles. La définition d'un environnement autorisant à l'historien une préhension efficace des sources primaires relatives à la période la plus contemporaine est donc aujourd'hui un enjeu dont l'acuité ne fait que croître et qui pose d'évidence la question tant de la formation à lui nécessaire, que celle de l'élaboration d'une instrumentation adaptée et des rapports que celui-ci peut ou doit entretenir avec les professionnels du maniement et du traitement de l'information engagés sur ces terrains depuis déjà longtemps [Barats, 2013].

Si nous avons pris en effet ici un exemple simple, l'incapacité de l'historien à pouvoir compter sur ses propres forces, l'impossibilité aussi pour le travailleur isolé de s'emparer de tels matériaux, apparaissent plus flagrantes encore si nous voulons qu'une analyse pertinente de tels matériaux soit possible qui prenne en compte leur caractère intrinsèquement dynamique. Le travail en cours des historiens du PIREH, Stéphane Lamassé et Benjamin Deruelle en particulier, explorant les fiches Wikipédia consacrées à des personnages historiques [Lamassé, Deruelle et Dumont, 2013] l'illustre clairement. Partant du constat que les contenus offerts par cette

encyclopédie faisaient aujourd'hui de fait autorité, et particulièrement pour leurs propres étudiants, ils ont voulu reconstituer le processus d'élaboration d'une ressource électronique de ce type. La première difficulté rencontrée est que chaque correction occasionne l'écriture d'une nouvelle fiche dans une base de données. La seule fiche « Jeanne d'Arc » est ainsi l'occasion de la création de plus de 2 000 notices au cours d'une période de dix ans. Leur extraction automatique est techniquement difficile du fait des changements de versions des langages et logiciels utilisés. L'exploitation des données finalement rassemblées suppose, elle, le passage par plusieurs dispositifs logiciels existants, qui n'acceptent pas les mêmes formats d'entrée, rendant difficile la réutilisation de scripts et exigeant donc des écritures logicielles *ad hoc*. L'exploitation rigoureuse, dans sa temporalité, d'un matériau rendant compte des modes d'écriture associés à un objet culturel massivement utilisé, qui constitue, pour une notable partie de la population, le moyen privilégié d'accès à des savoirs spécialisés, suppose non seulement l'usage, mais l'appropriation inventive d'une instrumentation adaptée tant au grain du matériau qu'au mode de questionnement propre à la discipline. Déléguer le choix, l'assemblage et l'adaptation au contexte des outils, c'est risquer l'incapacité à poser ses questions propres, parce que l'outil logiciel incorpore une modélisation des matériaux traités comme des descripteurs pertinents. C'est en un tel contexte renoncer à la maîtrise de ses concepts.

Ce sont là des questions qu'il est possible, ainsi que les autres évoquées plus haut, de traiter à d'autres niveaux, selon d'autres points de vue que celui adopté en ces pages et en les textes qui suivent. Le numérique des mondes historiens est affaire aussi d'infrastructures, de normes, pris dans les transformations des mondes universitaires et savants, elles-mêmes liées aux mutations des états et à la redéfinition des sphères publiques. Nous avons proposé en ce volume une vue autrement située, au ras du laboratoire et de la salle de cours, ancrée aussi dans une tradition disciplinaire, dont l'enjeu n'est pas de s'interroger sur l'environnement qu'il conviendrait de faire advenir, mais d'explorer l'univers numérique changeant au sein duquel se déploie l'activité du praticien, de l'historien en ce cas, et ce qu'il peut, à son échelle, y faire ou en faire. Nous espérons surtout que les plus jeunes y trouveront matière à réflexion, à invention, une incitation, aussi, au braconnage et au bricolage.

+++++

PARTIE I
ARCHÉOLOGIE
D'UNE PRATIQUE

+++++

+++++

CHAPITRE I. L'OFFRE FAITE AUX HISTORIENS. ESSAI DE DESCRIPTION DES SERVICES WEB OFFERTS AUX PROFESSIONNELS DE L'HISTOIRE CONTEMPORAINE

+++++

Une première version de ce texte est parue en juin 1998, intégrée dans un rapport rendu à la direction de l'Information scientifique et technique des bibliothèques du ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche [Guichard, 1998]. Celui-ci faisait suite à un projet de recherche, entamé au printemps 1995, qui s'était accompagné, en 1996, du lancement du serveur Barthes, qui abritait le site Clio, dédié à la recherche en histoire de l'immigration. Le petit groupe réuni autour d'Éric Guichard, qui donnerait un peu plus tard naissance à l'équipe Réseaux Savoirs & Territoires, partageait à l'époque la conviction qu'il n'était possible de développer un discours consacré aux mutations de l'informatique savante, en ce cas le développement rapide du Web, que pour autant qu'il s'accompagnait d'une pratique et de l'usage de ces nouveaux outils. En un mouvement fréquent à l'époque – trouver une information pertinente était difficile et les ressources savantes en ligne étaient assez rares –, nous avons entrepris d'inventorier et de référencer les ressources disponibles touchant à l'histoire contemporaine auxquelles nous offrions un accès par le biais d'une page de pointeurs. C'est la tenue de celle-ci, régulièrement mise à jour et enrichie durant plusieurs années, qui permit le rassemblement du matériau servant de support à ce texte. Ce texte constituant une description d'un état du réseau savant, distant de nous d'une vingtaine d'années, nous avons considéré qu'une mise à jour de celui-ci n'avait guère de sens et en avons simplement retouché l'écriture. Les ressources web décrites par ce texte ont souvent depuis lors disparu ou changé de forme, ou bien sûr d'adresse. Sauf indication contraire, nous donnons ici l'adresse valide fin 1997.

La participation à la création d'un site internet dédié à l'histoire sociale < <http://barthes.ens.fr/clio> >, la tenue d'une liste de ressources en ligne utiles aux historiens m'ont conduit tant à réfléchir aux usages possibles du réseau qu'à l'examen des services dédiés aux historiens. J'ai pu alors, de façon un peu empirique et intuitive, construire une typologie des services web s'adressant aux spécialistes de l'histoire contemporaine, qui prend en compte essentiellement les fonctions assignées aux sites par leurs concepteurs, mais fait aussi référence à la position de ceux-ci dans le champ disciplinaire et aux outils dont ils usent. Cela me permet de proposer ici une contribution à l'étude de l'appropriation de ce nouveau médium par la communauté scientifique à laquelle j'appartiens, contribution dont les limites sont évidentes. Je propose ici un tableau de l'offre faite à leurs pairs par les professionnels de l'histoire, à un instant donné, la mi-1998, alors que l'une des caractéristiques du réseau est l'évolutivité et la volatilité des ressources qu'il dispense. Ce tableau s'appuie de plus sur une expérience personnelle, une connaissance pratique du réseau, dont la précision et l'exhaustivité ne sauraient se mesurer à celle que permet d'acquérir une enquête systématique.

La première caractéristique de cette offre est la rapide évolution de ses formes. De nouveaux sites, de nouveaux usages du Web émergent en permanence. Les sites offrant un réel service et un réel contenu se font plus nombreux, alors que les vitrines destinées à prouver la virtuosité technique de celui ou de ceux qui les gèrent, toujours nombreux, ne sont plus la norme. Le constat n'est pas propre au domaine. La relative nouveauté du médium, plus encore le fait qu'il n'est connu et utilisé que depuis très peu de temps par une proportion significative de la population visée, l'explique sans doute en partie. De même, le développement d'outils (éditeurs HTML What You See Is What You Get [WYSIWYG], norme HTML 4.0) qui permettent à des non-informaticiens de se l'approprier plus aisément qu'au temps de l'écriture HTML en mode texte sous UNIX, y contribue sans doute. Il semble cependant possible, malgré ces transformations rapides et ce foisonnement, de rapprocher un grand nombre des sites visités de quelques types que nous désignerons par le biais d'une métaphore.

LE SERVEUR VITRINE

+++++

Les revues à publication papier et les centres de recherche produisent souvent des serveurs qui offrent surtout le moyen de les contacter, la liste

et le pedigree des membres de l'équipe ou du comité éditorial et quelques informations sur les publications et productions de l'institution. La logique est ici proche de celle de la plaquette, ou du journal de conseil général, il convient de donner à voir et de faire connaître l'existence du groupe. Ces sites demeurent sans doute aujourd'hui encore les plus nombreux.

LE MÉMORIAL

+++++

Souvent réalisé par des historiens amateurs, il propose un éclairage pointu sur une région, un domaine de l'histoire. Si les textes proposés sont parfois fragiles, l'intérêt provient souvent de l'iconographie dont la provenance n'est malheureusement pas toujours précisée, mais nous sommes ici dans une logique qui n'est pas celle de l'histoire érudite. On trouve par exemple en langue française le site Normandie 1900-1990 < <http://www.teaser.fr/~jpsagaire/norman/norman00.html> >.

LE CENTRE DE RESSOURCES

+++++

Un spécialiste d'un domaine, d'une sous-discipline, ou une équipe de chercheurs mettent parfois en place un serveur qui a pour mission d'offrir à une petite communauté des ressources et des informations pratiques. On y trouve fréquemment des listes de conférences à venir, des bibliographies, l'annonce de publications, des comptes rendus, parfois des manuels en ligne, des logiciels. Dans l'esprit comme dans la forme, nous sommes proches du bulletin dont la fonction est autant de mettre en évidence l'existence d'un sous-champ disciplinaire que de promouvoir ce dernier. Le support électronique n'a ici pour avantage que de permettre à des groupes ou des réseaux de taille modeste, qui n'auraient pas les moyens financiers ou humains de créer l'équivalent papier, de parvenir à ce stade. Il est de même possible de fournir pour un faible coût de revient une masse d'information supérieure à ce que permet le bulletin, de format souvent réduit. Certains commencent cependant à tirer parti de l'originalité du support en proposant des services (traitement semi-automatique de données, par exemple) aux lecteurs.

LA REVUE ÉLECTRONIQUE

+++++

Ces revues deviennent de plus en plus nombreuses dans le domaine des sciences sociales, quoique l'Histoire, et plus encore l'Histoire de langue française, soit assez peu représentée. Les formats sont extrêmement divers, allant de la mise en ligne des papiers au fur et à mesure des arrivées à l'édition à date fixe, en passant par la production irrégulière de « numéros ». Ceux-ci peuvent juxtaposer des papiers sans liens entre eux, ou être le fruit d'un regroupement thématique. Notons que, dans le monde anglo-saxon, la tendance est à rendre payante la consultation des articles qui ont longtemps été offerts gracieusement. *Annual Review Sociology Online* (< <http://www.annurev.org/soc/home.htm> >), a ainsi cessé depuis peu de permettre à tous l'accès aux textes. La plupart de ces sites ne font que mettre en ligne une production dont les règles de présentation et la conception s'écartent peu ou pas de celles de la revue papier. C'est alors la recherche d'une diminution des coûts et de la rapidité qui justifie la démarche [Harnad, 1996], quoiqu'il ne soit pas certain que les gains en ce domaine soient aussi spectaculaires que certains l'espèrent [Crosnier, 1995]. Quelques-uns cependant commencent à s'éloigner de ce modèle en proposant des services ou des fonctionnalités nouvelles. Parmi ces innovations, on note :

La mise en place de forums

Un ou quelques papiers aux thématiques proches servent de base à une discussion dont tous les éléments sont versés au dossier. L'unité de base n'est plus alors l'article lui-même, mais l'ensemble formé par celui-ci et les réactions qu'il a suscitées. Les débats organisés par *Australian Humanities Review* (< <http://www.latrobe.edu.au/library> >) fournissent un bon exemple de ces pratiques. Dans la plupart des cas, il semble cependant très probable que commentaires et réactions ont été suscités par l'équipe rédactionnelle elle-même et non produits par des lecteurs anonymes ayant soudain envie de saisir leur clavier après lecture. On peut alors se demander si l'interactivité inscrite dans les potentialités du médium ne s'actualise véritablement que lorsqu'elle est activement défendue et provoquée. La mise en ligne de prépublications est à certains égards voisine de la mise en place d'un forum. Le texte dont le caractère provisoire est affirmé peut être considéré comme un appel lancé à l'ensemble des

spécialistes du domaine à qui il est demandé de valider le texte, voire de le coécrire [Ginsparg, 1994]. La page web peut, ou non, témoigner de ce travail collectif. Dans le cas des textes offerts par l'*University of Oxford Discussion Papers in Economic and Social History* (< <http://www.nuff.ox.ac.uk/Economics/History/> >), si le caractère provisoire de certains textes est affirmé, au point qu'ils peuvent être retirés après la publication papier, nous ne trouvons pas trace des remarques ou suggestions que ces esquisses ont pu susciter. Une telle pratique n'est pas absolument neuve puisque nous y avons tous recours, sans l'usage du net, mais elle élargit le cercle des collaborateurs potentiels à l'ensemble des spécialistes. Il serait cependant intéressant de vérifier si l'ensemble formé par les auteurs des réponses et remarques diffère, ou non, de celui composé des amis, relations et connaissances de l'auteur. Si, en d'autres termes, le passage par la toile favorise la consolidation et la fluidité du fonctionnement de réseaux existants, ou si elle contribue à élargir ceux-ci et à en transformer la nature, en conduisant les membres à entrer directement en contact avec d'autres que ceux auxquels les lient déjà des relations interpersonnelles ou professionnelles.

L'accès au milieu documentaire

L'un des avantages du réseau est qu'il permet de dépasser les limitations physiques et économiques imposées par le support papier, certains en profitent pour adjoindre à un papier le dossier de recherche constitué à cette occasion, ou comme le groupe du Gresh a pu le faire, pour proposer à l'ensemble de la communauté les documents réunis lors de la préparation d'un ouvrage collectif (< <http://barthes.ens.fr/clio> >).

Le document multimédia

Certains adjoignent à leur texte des éléments de démonstration ou d'illustration que l'article papier ne permet pas d'intégrer ; cela peut aller de la bande-son au petit extrait vidéo, en passant par l'utilisation d'une profusion d'images. Cette tendance, née en dehors du champ des sciences humaines, peut déjà s'y repérer à quelques expérimentations. Elle semble amenée à s'amplifier, en particulier si le passage sur Internet de la communauté scientifique se fait dans un futur proche. Il restera cependant

à vérifier si une telle évolution mènera plus de spécialistes du domaine à prendre pour sources des documents non écrits (ce qui n'est pas une nouveauté) ou si cela conduira à la multiplication de documents joliment illustrés. La mutation en cours, qui fait de l'article un document multimédia semble cependant moins radicale que celle qui en fait l'unité de base de la recherche, consultable sans même un œil à la table des matières ou au sommaire de la revue. Cela transforme radicalement tant la notion d'auteur que celle de revue. Certains sites poussent cette logique jusqu'à ses conséquences ultimes puisqu'ils sont structurés comme des bases de données interrogeables qui renvoient simplement les textes correspondant aux requêtes [Charton, 1996]. Ce n'est cependant pas encore le cas en sciences sociales.

LE DÉPANNEUR

+++++

Les *newsgroups* et les listes de diffusion permettent depuis déjà longtemps à des chercheurs opérant dans le même champ de poser à l'ensemble de la communauté connectée des questions d'ordre divers, lorsqu'ils rencontrent un problème ou une difficulté au cours de leur recherche. Certains sites intègrent déjà cette fonction en reprenant souvent une architecture proche de celle des *newsgroups*. C'est en particulier le cas des pages du réseau H-net (< <http://h-net2.msu.edu/> >), qui proposent en ligne un accès aux archives des listes de diffusion historiques. L'avantage est sans doute qu'en dehors des cercles depuis longtemps habitués à l'usage du réseau, seuls les pages web et le courriel sont véritablement utilisés. Là encore, une évolution du procédé est probable ou possible. Certaines grandes entreprises utilisent sur leur intranet des systèmes de blasons, chacune des personnes susceptibles de répondre est définie par une liste de compétences et des indications concernant sa disponibilité. Il est alors possible, en fonction de la question posée, de sélectionner les quelques personnes (il est facile de fixer un seuil) les plus à même de répondre et de leur envoyer un courrier, puisque la plupart des navigateurs en usage incorporent aujourd'hui la fonction courrier. Cela évite aux participants d'être encombrés de courriels qui ne les concernent pas, effet secondaire courant de l'inscription à une liste de diffusion. L'efficacité du système dépend cependant du degré d'implication des connectés, de la variété des compétences qu'ils déploient et de leur nombre, ce qui explique sans

doute pourquoi ce système, dont la mise en place ne pose pas de problèmes techniques, n'a pas encore débordé l'entreprise.

L'ÉDITION NUMÉRIQUE

+++++

Des acteurs institutionnels variés entreprennent la numérisation de vastes corpus imprimés ou documentaires. La mise à disposition de l'ensemble de la communauté de matériaux difficilement accessibles est au principe de la démarche. La tentative la plus connue en ce domaine est celle initiée par la Bibliothèque nationale de France (BnF), avec le projet Gallica, le XIX^e siècle francophone, qui permet un accès direct à une très riche collection de textes francophones. De tels projets ne peuvent pour le moment prétendre à l'exhaustivité, du fait de l'ampleur de la tâche. Cela conduit ces acteurs à des choix et à proposer, soit des dossiers thématiques, soit des échantillons dont la logique de constitution n'est pas toujours claire. Quelques initiatives privées, émanant de chercheurs ou de simples passionnés, complètent ces tentatives. Une association de bénévoles s'est ainsi attelée à la numérisation des textes du domaine public francophone, et propose plus de 200 titres en ligne. Il est possible de les consulter sur le site de l'Association des bibliophiles universels (ABU, < <http://cedric.cnam.fr/ABU> >). L'échelle n'est pas la même. Cependant, dans les deux cas, deux options s'ouvrent : soit numériser le document de départ et proposer un fichier image ; soit, lorsqu'il s'agit d'un texte, proposer un texte numérique, ce qui change radicalement la nature du document puisqu'il est alors totalement indexable et manipulable au moyen d'outils d'analyse textuelle ou statistique.

LE MÉTASITE

+++++

L'extraordinaire croissance du réseau, sa taille, l'absence de hiérarchie de l'information, qui empêche les moteurs de recherche classiques de référencer plus de la moitié du Web, conduisent à une multiplication de sites spécialisés dont la seule fonction est de favoriser la navigation sur le réseau, plus exactement pour la plupart d'entre eux sur le Web. La forme la plus rudimentaire en est la liste de sites consacrés à tel ou tel thème. Elle peut être éventuellement complétée par des conseils, voire de véritables manuels destinés à favoriser les premiers pas de l'internaute

débutant. Les sites les plus élaborés structurent l'information sous forme de base de données interrogeables, mais ils sont rares encore en ce cas.

LE WEB INTERNE

+++++

Les serveurs de nombreux départements universitaires mettent en ligne les informations pratiques et administratives à l'usage de leurs étudiants et professeurs. La définition de ce qu'il est utile d'y porter varie cependant beaucoup d'un lieu à l'autre. Il arrive qu'on y trouve, outre la présentation d'un enseignement, la liste des documents qui seront étudiés dans l'année, la bibliographie d'un cours, voire le résumé de celui-ci. L'un des sites proposant la plus riche documentation concernant la Révolution française, Course Materials, French (< <http://tuna.uchicago.edu/homes/jack/course.materials.html> >), est de ce type.

Cette typologie a ses limites. Un même site peut remplir plusieurs des fonctions ici définies. Nous pensons cependant que le canevas dessiné peut avoir son utilité et servir de support à une grille de description des pages offertes à l'un ou l'autre secteur des sciences humaines, dans le cadre d'une étude plus systématique. D'autant plus que l'on peut combiner à ce premier axe décrivant les fonctions assignées au site deux autres facteurs. Il paraît en effet également possible de classer les pages web d'un domaine, selon leur degré de complexité. Nous irions alors de l'usage du Web comme nouveau moyen de diffuser un matériau préexistant (dont ni la nature ni la présentation ne changent) aux sites expérimentaux, dont les auteurs semblent parfois plus préoccupés par l'exploration des possibilités du médium (interactivité, instantanéité de la communication, possibilités d'indexation et de recherche, intégration d'éléments venant de médias divers) que par la définition des informations et services mis en place. Un troisième axe permettrait de décrire la position institutionnelle des créateurs du site (origine géographique, nature de l'institution de rattachement). Combinant ces trois axes, voire quelques autres, il serait alors possible de déterminer avec plus de précisions qui dans ce champ produit quoi et à l'aide de quels outils.

+++++

CHAPITRE II. ÉCRIRE ET PENSER UN SITE

+++++

Le site Clio, né en 1997 et régulièrement mis à jour durant une dizaine d'années, était l'un des premiers sites savants d'histoire en langue française. La nécessité d'inventer une forme électronique adaptée aux besoins d'une petite communauté de chercheurs, sans guère disposer de modèles, la difficulté aussi à s'adapter en permanence aux rapides transformations des outils et des langages, conduisait à mêler production des dispositifs et activité réflexive, en partie orientée par les besoins de la recherche elle-même. Ce texte en témoigne, qui procède du récit de pratiques. Il a été présenté pour la première fois lors du colloque *Comprendre les usages d'Internet* tenu en 1999 à l'École normale supérieure (ENS), et une version en a été publiée dans les actes de celui-ci [Guichard, 2001b, pp. 83-89].

Nous proposons ici un récit de pratiques décrivant la naissance du site Clio, avant de dresser le bilan de ses trois années d'existence et d'exposer les perspectives qui s'offrent à ses créateurs. Nous espérons que cette contribution, décrivant les réalisations, les espoirs, mais aussi les doutes des membres d'une communauté scientifique confrontée à l'internet, permettra d'entamer avec les usagers, les scientifiques, voire les politiques, un dialogue fructueux. Le site naquit en 1997, à l'initiative de chercheurs de l'ENS et de l'École des hautes études en sciences sociales (EHESS) partageant un intérêt commun pour l'histoire de l'immigration, parmi lesquels Olivier Le Guillou et Gérard Noiriel. Plusieurs logiques présidaient à la création de ce site, émanation d'un groupe qui lui préexistait. Il faut faire sa place parmi elles à la curiosité. Plusieurs d'entre nous avaient en commun une modeste culture scientifique, acquise avant de bifurquer vers les sciences sociales, qui les destinait peut-être à prêter attention à la naissance de ce nouvel outil. Encore fallait-il, pour que cette curiosité soit excitée, que nous ayons connaissance des possibilités qu'il offrait. Notre appartenance au laboratoire des sciences sociales de

l'ENS, où les échanges entre chercheurs de culture différente sont fréquents, nous le permet. Éric Guichard, en l'occurrence, joua les passeurs. Propagandiste enthousiaste et soutien généreux, il put, grâce à sa double qualité de chercheur en sciences sociales et de praticien de l'informatique, être l'indispensable médiateur qui nous permit de nous lancer dans l'aventure. Il sut de plus, parce que, l'un des nôtres, il partageait nos préoccupations, nous montrer que ce médium permettait de répondre à certaines d'entre elles. Nous savions en effet que l'apparition d'un nouvel outil ne peut suffire à faire progresser la recherche. L'histoire de l'introduction de l'informatique en histoire est à cet égard éclairante, pleine de récits de construction de bases de données géantes, dont certaines ne purent être exploitées, leurs concepteurs ayant oublié de se demander quels usages elles pouvaient susciter. Si donc nous avons eu recours au réseau, c'est parce qu'il paraissait susceptible de nous aider à trouver des réponses à quelques problèmes auxquels nous confrontait notre activité.

DES BESOINS...

+++++

Du fait de l'hypercentralisation des institutions archivistiques et scientifiques, l'accès à l'information est peu aisé pour les historiens provinciaux, comme pour ceux qui ne sont que lâchement rattachés à un centre de recherche. Il était ainsi difficile, il y a peu de temps, à moins de se rendre à Paris, de constituer rapidement une bibliographie à jour. De plus, tous les historiens éprouvent une difficulté croissante à se tenir au courant des développements de la recherche. Son internationalisation, la multiplication des travaux, la dispersion physique de ceux-ci, l'incapacité des revues, engorgées par cet afflux, à en rendre compte rapidement, condamnait l'historien à ne disposer que d'une vision partielle et différée de l'activité scientifique en cours. Nous n'étions guère plus heureux de la diffusion de nos propres recherches. L'édition d'un texte implique souvent que l'on se soumette à des normes peu compatibles avec les exigences d'un travail scientifique, tendant à exclure de la publication tant les éléments du dossier documentaire que l'appareil critique, indispensables au spécialiste, qui y trouve, outre des matériaux utilisables, les éléments lui permettant d'évaluer la pertinence des conclusions de l'auteur. Enfin, il devenait difficile de trouver, en partie du fait de l'absence de grands chantiers collectifs comparables à ceux des décennies précédentes, des endroits où l'on puisse parler techniques, tour de main, soit de ce qui fait

le quotidien de notre activité et échanger des matériaux, des références bibliographiques, matière première de notre travail, que nous nous épuisions à accumuler chacun de notre côté. L'absence de lieux de rencontres nous paraissait également conduire à la disparition des confrontations fécondes qui marquèrent l'histoire de la discipline et mener à une histoire faite de discours parallèles produits par des chercheurs ne se connaissant pas qu'ils ne se lisent.

AU PROJET

+++++

Le recours à l'internet était un moyen de travailler à résoudre ces difficultés. D'une part, le réseau offrait, pour un prix modique, une mise en communication instantanée de chercheurs distants, et semblait donc un moyen de communication adapté au fonctionnement d'un groupe de recherche constitué d'individus dispersés du Brésil à l'Ukraine. De plus, la numérisation de l'information qu'impose sa distribution électronique aboutissait à offrir au public un matériau plus riche que l'édition papier. Disposer d'une version électronique d'un tableau de recensement permet ainsi de le retravailler, que l'on s'en serve de base de calcul ou de source d'une représentation graphique, ce qui représente quelques heures gagnées pour l'utilisateur final, qui n'a plus à saisir des rangées de chiffres minuscules qui mettent en échec les logiciels de reconnaissance de caractères. Enfin, le réseau permettait la mise à disposition de masses documentaires importantes, sans que la diffusion de cette information n'entraîne de dépenses exorbitantes. En l'état actuel des choses, un centre de recherche peut mettre à disposition du public des informations volumineuses sans grever son budget propre, et la consultation de cette information coûte peu à l'utilisateur final. L'internet représentait donc pour nous un moyen bon marché, rapide et performant de diffusion de l'information au sein de notre communauté scientifique. Il était aussi un outil facilitant la mutualisation partielle de la recherche.

Partant de ce constat, nous avons défini les contours d'un serveur répondant à nos besoins. Cela nous a amenés à dégager quelques priorités. Nous voulions que les historiens puissent trouver un lieu où il soit possible de s'informer et de débattre de ce qui les intéresse : de leurs pratiques et de leurs outils, de l'usage de l'internet, par exemple. Un tel lieu ne trouve que difficilement sa place dans les circuits classiques de la communication scientifique ; il doit susciter débat, puisque tant sa fonction que

ses effets ne sont pas inscrits dans ses caractéristiques, mais dépendront des pratiques développées par les acteurs qui y sont confrontés. De plus, nous souhaitions que notre site puisse se faire l'écho des controverses scientifiques intéressant la communauté historienne. Cela supposait bien sûr que les visiteurs de notre site soient des contributeurs et pas seulement des consommateurs, hypothèse qui, à l'usage, se révèle forte. D'autre part, nous voulions offrir l'accès à des outils utiles à l'exercice de notre profession : ressources bibliographiques, mais aussi collections de sources et de documents, ou bien encore ressources facilitant l'usage des nouvelles technologies. Le site devait enfin permettre des échanges rapides entre chercheurs engagés dans les mêmes travaux en procurant à tous des informations fréquemment actualisées.

MISE EN ŒUVRE

+++++

Force nous est de constater que, si le site rêvé et le site réalisé se ressemblent, ils sont loin de se confondre. Alors que notre cadre de référence était la revue papier, la structure même du site fait que celui-ci n'est pas l'équivalent électronique d'une revue imprimée. D'une part, il a d'autres fonctions, que les contraintes de l'édition ne permettent pas aux revues traditionnelles de remplir, en particulier la publication de sources, d'autre part, il n'est pas de même structure, puisque les unités que sont le volume ou le fascicule n'ont pas d'équivalents, les éléments pertinents étant le document et la rubrique. De plus, alors que nous prévoyions de donner naissance à un site couvrant un champ très large, Clio se présente comme une revue scientifique spécialisée, dont le public attendu est constitué des spécialistes d'histoire de l'immigration. Les conditions de production du site l'expliquent. Maintenu à temps perdu par un chercheur, en plus de ses activités de recherche habituelles, il ne peut qu'être le prolongement de celles-ci, à moins qu'il ne sacrifie une partie du temps consacré à d'autres activités de recherche, seules reconnues.

Les fonctions du site ne sont pas non plus celles que nous avons anticipées. Notre site n'est pas un lieu de débats. Cela ne signifie pas que nous ayons renoncé à ce qu'il le devienne. Simplement, la lecture des forums de nombreux sites, l'habitude des listes de diffusion, conduit à la conclusion qu'un débat n'a d'intérêt que s'il permet de confronter des points de vue argumentés et des textes structurés. Il faut dès lors que le débat soit suscité, et contrôlé, ce qui implique la participation de très bons spécialistes

du domaine, qui acceptent de rédiger des textes se répondant, qui seront, la socialisation de l'internet restant très incomplète, peut-être peu lus, et dont la publication aura un statut incertain.

La pratique fit à l'inverse émerger des fonctions auxquelles nous n'avions que peu songé. D'une part, le site joua le rôle d'un aimant, et permit à plusieurs jeunes chercheurs de prendre contact avec nous, et, pour certains, de venir nous rencontrer, voire de collaborer à nos travaux. D'autre part, nous comprîmes assez vite que l'écriture pour la toile n'allait pas de soi, et parvenir à imaginer des formes d'écriture adaptées au médium devint une fin en soi. Il nous fallut ainsi décider du sort à faire aux notes qui émaillent les contributions scientifiques en notre domaine. Peu à peu, nous nous dégagâmes de la simple reproduction d'un article papier, pour explorer les possibilités qu'offraient l'hypertexte et le multifenêtrage. Le compromis auquel nous sommes parvenus fait coexister des notes en bas du texte, ce qui permet à qui l'imprime de disposer des notes et du texte sur le même document, et une fenêtre de notes, parallèle à celle du texte, qui permet au lecteur d'avoir en même temps à l'écran le texte et la note qui lui correspond. Les tâtonnements en ce domaine sont, pour partie, liés à notre ignorance des usages des lecteurs de la Toile. Nous tentons d'adapter nos présentations ou nos écritures à ce que nous supposons être les habitudes des lecteurs, sans parfois avoir à notre disposition d'autres éléments concernant celles-ci que nos propres pratiques. Situation d'écriture bien particulière, et inhabituelle pour l'universitaire, qui écrit généralement pour des pairs dont il connaît les habitudes, voire les noms. Nous avons aussi collaboré, par l'intermédiaire d'Anne-Sophie Bruno, à la genèse d'un atlas interactif permettant au visiteur d'élaborer sa propre carte, à partir des données disponibles, qui décrivaient en ce cas les populations étrangères présentes en France durant l'entre-deux-guerres. L'entreprise ne peut avoir d'équivalent papier, puisque cet atlas contient potentiellement plusieurs milliers de cartes. Enfin, je n'ai pu personnellement résister à la tentation de réaliser un vieux fantasme d'historien, et, utilisant les ressources de l'hypertexte, je publiais un article dont les notes ne renvoyaient pas aux côtés de cartons d'archives, mais aux reproductions des documents originaux, préalablement numérisées. J'ajoute que la tentative s'est soldée par un insuccès total : il est rarissime qu'un lecteur de l'article suive l'un des liens établis.

LECTORAT

De façon générale d'ailleurs, le gestionnaire d'un site, qui dispose de moyens d'épier les réactions de ses lecteurs fort différents de ceux qu'utilise un rédacteur de revue, est souvent surpris de ce que les statistiques du serveur révèlent des pratiques de ses lecteurs. Malgré de notables variations dans le temps, certaines constantes apparaissent. Nos lecteurs viennent chercher des informations bibliographiques, des documents bruts, mais les articles scientifiques, les comptes rendus de recherche, sont plus rarement consultés.

Tableau 1. Documents consultés par les lecteurs du site Clio

	Oct. 97-nov. 98	Jan. 99-oct. 99
Sommaire	22,7	20,5
Documents	29,9	23,6
Pointeurs	26,7	19,5
Comptes rendus de recherche	7	5,8
Bibliographies	6,4	11,4
Articles	2,8	4,2
Présentation de publications	2,7	1,9
Comptes rendus d'ouvrages	0,8	10,4
Revue	0,6	0,6
Pages anglaises	0,5	2
Total	100	100

Nous butons, lorsque nous voulons interpréter ces résultats, sur notre méconnaissance tant du profil que des pratiques de nos lecteurs. Nous

pouvons proposer cependant quelques hypothèses. Les textes scientifiques offerts traitent de sujets fort pointus. De plus, ces documents sont souvent longs, et d'après les statistiques de notre serveur, plus un document est lourd, moins il a de chances d'être consulté, signe de ce que nombre d'internautes reculent encore devant le chargement de documents dépassant les 50 Ko, ou de ce que les moteurs de recherche indexent mal les textes longs. Enfin, il est possible que nos lecteurs viennent chercher sur le réseau ce que les revues du champ offrent rarement, c'est-à-dire de l'information bibliographique, ou des sources éditées, ce qui ferait alors des sites scientifiques les compléments plutôt que les concurrents des revues papier. Le relatif insuccès des articles scientifiques (mais savons-nous combien de personnes consultent les articles des revues savantes ?) peut signifier que nos lecteurs n'ont pas intégré à leurs usages la consultation de papiers électroniques, à moins qu'ils ne soient encore incertains du statut à accorder aux textes que l'on trouve sur le réseau. Souvent en effet, ceux-ci n'ont pas été validés par une instance scientifique, ou se présentent comme des ébauches.

PROJETS

+++++

Le site web s'est révélé un outil précieux pour notre équipe : il permet à celle-ci de se prolonger par un réseau ; il offre un formidable outil de communication interne, permettant des gains de temps et d'argent substantiels, même si l'un des effets secondaires, qui passe souvent inaperçu, est que le coût de la communication interne est pour partie reporté sur l'utilisateur final, c'est-à-dire sur le chercheur lui-même en tant que lecteur. Si nous voulons cependant rester fidèles à l'esprit de l'entreprise, c'est-à-dire construire un site qui permette échanges et confrontations de savoirs entre les membres d'un champ disciplinaire, dont les frontières dépassent notre microspécialité, et rende possible la diffusion de leurs travaux au-delà de la communauté scientifique *stricto sensu*, il nous faut d'une part diversifier nos thématiques, d'autre part enrichir les fonctionnalités du site, en repensant par exemple les modalités de l'organisation de débats, ou en proposant des dossiers pédagogiques réalisés par des chercheurs à partir de leurs propres travaux, qui puissent, s'adressant aux enseignants du secondaire, leur fournir des supports réutilisables en classe.

Cela suppose un changement d'échelle et l'appel à des contributeurs extérieurs, ce qui n'est pas sans poser problème. Même si des

frémissements peuvent être observés au sein de certaines institutions publiques, la reconnaissance institutionnelle des tâches de maintenance et de gestion de site, sous leur double aspect de maîtrise technique et de contrôle du contenu éditorial, n'est pas encore acquise. En l'état actuel des choses, le temps que consacre un chercheur à de telles tâches est, souvent, un temps bénévole : il n'entre pas dans ses obligations et n'est donc pas rémunéré comme tel et il ne peut espérer une reconnaissance de telles activités. Or, pour passer d'un site web fonctionnant comme le bulletin d'un centre de recherche à une revue électronique ambitieuse, il faut disposer d'un budget temps que nous ne pourrions dégager s'il nous faut compter uniquement sur les moments gracieusement abandonnés par les membres de notre équipe. Il nous faut donc trouver soit le moyen d'obtenir que le travail effectué en ce domaine soit reconnu comme une activité de recherche, et évalué comme tel, soit disposer d'un budget permettant de rémunérer des personnes prenant en charge une partie des tâches qu'implique la maintenance d'un site. Bien des sites d'ailleurs sont confrontés à des problèmes similaires, ce qui explique en partie la courte durée de vie de nombre d'entre eux. D'autre part, il sera difficile d'obtenir des contributions de qualité tant que les publications électroniques n'auront pas dans notre domaine le statut de publication scientifique. Le blocage ici ne tient pas tant à l'amateurisme éventuel des revues de ce type, puisque certaines, dont nous faisons partie, se sont dotées d'un comité scientifique, mais aux habitudes des membres de la profession, qui accordent plus de crédit à une publication papier dans une revue vénérable qu'à un texte placé dans ce domaine mouvant et encore pour beaucoup largement inconnu, qu'est la Toile. Dans ce contexte, il est peu surprenant qu'un chercheur hésite à confier un texte à un éditeur dont il ne peut encore véritablement évaluer l'impact, d'autant que celui-ci ne lui assure pas la rémunération symbolique de son travail. Il faudra donc probablement, dans une période intermédiaire, rémunérer d'une façon ou d'une autre les producteurs de contenu.

La courte histoire de notre serveur résume sans doute assez bien celle des premiers sites français de sciences sociales. Au bricolage des premiers temps, mené par quelques francs-tireurs, a succédé une exploration plus méthodique des potentialités du médium, loin d'être terminée, et une réflexion sur ses usages possibles et les contraintes qu'implique le recours à celui-ci. Beaucoup, je crois, partageraient ce constat que le développement d'un réseau de sites ambitieux, vitrine et outil de diffusion des travaux des spécialistes des sciences sociales françaises, aura un coût,

et passera par la mobilisation de moyens financiers, et surtout humains, importants. L'efficacité du médium en matière scientifique, comme la nécessité (si nous voulons ne serait-ce que pouvoir continuer à travailler en français) d'un internet de langue française fort, nous semble justifier cet investissement. Un relatif optimisme est cependant permis : certains, peu nombreux encore, ont déjà rassemblé les moyens de réaliser des projets ambitieux et novateurs.

+++++

CHAPITRE III. LES SOURCES DE L'HISTORIEN À L'HEURE D'INTERNET

+++++

Ce texte est issu d'une conférence préparée à l'occasion d'une journée de l'école doctorale d'histoire de l'université Paris I tenue en avril 2003, qu'organisait Joseph Morsel et qui avait pour thème « L'historien et ses sources ». Elle s'adressait donc à un public d'étudiants avancés auxquels il s'agissait de donner quelques repères leur permettant d'orienter leurs recherches documentaires, ainsi que quelques pistes facilitant leur exploitation. Une version en a été publiée en 2003 dans la revue *Hypothèses*, attachée à cette école doctorale [Rygiel, 2003].

« L'enseignement de l'écriture, ô roi, dit Theuth, accroîtra la science et la mémoire des Égyptiens ; car j'ai trouvé là le remède de l'oubli et de l'ignorance. » Le roi répondit [...] toi, père de l'écriture, tu lui attribues bénévolement une efficacité contraire à celle dont elle est capable ; car elle produira l'oubli dans les âmes en leur faisant négliger la mémoire : confiants dans l'écriture, c'est du dehors, par des caractères étrangers et non plus du dedans, du fond d'eux-mêmes, qu'ils chercheront à susciter leurs souvenirs ; tu as trouvé le moyen, non pas de retenir, mais de renouveler le souvenir, et ce que tu vas procurer à tes disciples c'est la présomption qu'ils ont la science, non la science elle-même ; car, quand ils auront beaucoup lu sans apprendre, ils se croiront très savants.

Platon, *Phèdre*

Les historiens utilisent peu Internet, du moins leurs notes mentionnent-elles rarement des publications électroniques ou des ressources documentaires fournies par le réseau. Plus rares encore sont ceux qui construisent leur appareil de sources en utilisant les possibilités du réseau. Une étude américaine, examinant les références et

renvois mobilisés par les articles parus dans quelques grandes revues américaines établit qu'en 2000, moins de 3 % des articles publiés renvoyaient à des documents numérisés présents sur le réseau [Graham, 2001]. Il y a donc quelque provocation à affirmer que cet outil nouveau peut contribuer à transformer les conditions de travail des historiens et partant leur outillage intellectuel et leurs productions. Il nous faut pour le comprendre examiner les ressources auxquelles le réseau permet d'accéder, celles aussi qui seront distribuées par ce biais dans un futur proche, et examiner les conditions de leur appropriation par l'historien, soit les modes de recherche d'informations adaptés à cet univers de ressources, et tant les difficultés posées par le document numérique que les possibilités qu'ouvre son utilisation.

LES RESSOURCES DU RÉSEAU

+++++

Il nous faut, au préalable, préciser deux points de vocabulaire, qui ne sont pas sans conséquence sur le contenu de notre exposé. Internet est un ensemble de réseaux informatiques interconnectés par le biais d'un protocole, Transmission Control Protocol (TCP/IP), qui permet que transitent entre des machines très hétérogènes des données numériques, soit de l'information au sens que les physiciens donnent à ce terme. Traiter des sources et d'Internet c'est donc d'abord évoquer, du point de vue de l'historien, les caractéristiques des données numériques. D'autre part, du sens que nous donnons au mot source, découle qu'aucun historien n'a jamais trouvé de sources sur le réseau, ni n'en trouvera. Cette attente, fréquemment exprimée par des étudiants, et forcément déçue, repose sur l'oubli d'une dimension fondamentale du travail de l'historien. La collection de sources mobilisée par une étude historique est le produit d'un travail – nécessitant de solides compétences et une bonne connaissance de son matériau – de sélection, de validation et de classement de documents, et nous prenons ce dernier terme dans un sens très large qui englobe toute trace matérielle d'une activité humaine dont l'historien entreprend l'étude. La source donc doit être construite et non trouvée, et ce travail exige un temps et une culture professionnelle que jamais un réseau ne saurait abolir, même s'il transforme certaines des propriétés de cette durée et exige certaines compétences spécifiques. Notons de plus qu'une telle définition exclut de fait les bases de données documentaires complexes dont les notices incorporent des informations puisées à plusieurs

sources primaires, tels les répertoires prosopographiques qui fleurissent aujourd'hui sur Internet ["Bright Sparcs. Encyclopedia of Australian Science", s.d.], que nous assimilons, par convention, à des éléments de bibliographie et non aux matériaux bruts du travail de l'historien, même s'ils peuvent évidemment être utilisés afin de constituer une source.

LISTES, CARTES ET PLANS

+++++

Internet offre aujourd'hui à l'historien plusieurs types de documents utilisables lors de ce travail. Nous y trouvons de nombreuses listes permettant de repérer les coordonnées géographiques et institutionnelles d'objets utiles. Les musées, les bibliothèques, les centres d'archives, voire les maisons d'édition ou les revues, sont de plus en plus nombreux à offrir des répertoires ou des catalogues. Ceux-ci tendent à devenir des répertoires dynamiques. Le client, en l'occurrence l'historien, peut présenter, à l'aide d'un formulaire, une requête et recevoir en retour une réponse, qui prend généralement la forme d'une liste de fiches répondant aux critères formulés, retrouvées dans une base de données à laquelle l'historien n'a pas accès. Cette organisation, indispensable dès lors que la quantité de données détenue devient importante, permet de formuler des requêtes extrêmement précises et complexes, puisque l'on peut poser, à l'aide d'opérateurs booléens, des conditions portant sur plusieurs rubriques de la base de données. Il m'est ainsi possible, par le biais du catalogue informatisé des collections d'archéologie, d'art contemporain, d'arts décoratifs, d'histoire, de beaux-arts et d'ethnologie appartenant aux musées de France de repérer les représentations de Louis XVI visibles dans l'un des musées nationaux de Paris et d'obtenir, pour chacune, un descriptif sommaire, voire une image, et ses coordonnées [Ministère de la Culture, s.d.]. Je puis ainsi, cherchant la réponse à une question, retrouver rapidement la trace d'un document utile, ce à quoi j'aurais peut-être renoncé s'il m'avait fallu localiser, puis consulter, un ou plusieurs volumineux répertoires papier.

Revers de la médaille cependant, il faut, pour accéder à l'information désirée, emprunter des chemins tracés par d'autres, qui ont choisi des rubriques, des modes d'indexation et construit des *thesaurus* en anticipant des besoins qui ne sont pas forcément les miens. Je ne puis ainsi, utilisant la même base Joconde, point accéder de façon immédiate à une liste des estampes en taille-douce publiées à Paris à la fin du XVIII^e siècle, alors même que le procédé de fabrication et le lieu de parution des documents

archivés figurent dans la fiche qui leur est consacrée. L'auteur de la base de données n'a pas jugé utile cependant de permettre une interrogation sur ces rubriques, ce qui contraint l'utilisateur à demander la liste de toutes les estampes de la période qu'il étudie, puis à consulter une à une chaque fiche détaillée, afin de retenir celles qui sont conformes à sa recherche. La procédure n'est sans doute pas beaucoup plus longue que s'il nous fallait utiliser un document papier pour faire le même travail, mais elle n'apporte pas de gain de temps significatif. De façon plus générale, il nous faut maîtriser, si nous voulons obtenir une réponse, le lexique et la syntaxe de la base de données, ce qu'il est généralement possible de faire à partir des indications fournies par le site, mais qui peut nécessiter un certain temps. Il nous faut ainsi savoir, reprenant le même exemple, que la rubrique désignant le type de document recherché (estampe, tableau, etc.) porte le nom de « domaine ». Cette connaissance de l'organisation de l'information est sans doute moins nécessaire, ou plus immédiate parce que correspondant aux habitudes héritées de notre formation, pour qui utilise un répertoire papier, que nous avons toujours le loisir de parcourir en sa totalité afin de nous l'approprier.

LA NUMÉRISATION DE L'ARCHIVE

+++++

De plus en plus souvent, nous parvenons à accéder non pas seulement à un répertoire d'objets, mais à une collection d'images de certains de ceux-ci. Elles sont le fruit d'une multitude d'initiatives émanant d'acteurs très divers et sont souvent sans lien ni coordination entre elles. Des chercheurs, isolés ou membres d'un collectif, numérisent des collections de documents, qu'ils les détiennent ou non. L'institut de recherche et d'histoire des textes poursuit depuis plus de vingt ans le recensement des manuscrits enluminés des bibliothèques de France et permet l'accès à une banque d'images, couplée à une base de données textuelle, offrant la possibilité de rechercher et de visualiser une partie des enluminures recensées. Des institutions muséales, des bibliothèques¹, des centres d'archives², des collectivités territoriales, des sociétés privées, voire de simples particuliers contribuent

1. La bibliothèque numérique de la BnF, Gallica, qui propose un accès à 70 000 ouvrages numérisés, à plus de 80 000 images et à plusieurs dizaines d'heures d'enregistrements sonores, demeure, dans le monde francophone, l'exemple le plus spectaculaire de telles entreprises, < <http://gallica.bnf.fr> >.

2. Les archives départementales des Yvelines offrent ainsi l'accès à une collection de documents numérisés qui comprend surtout des cartes et plans anciens.

ainsi à l'enrichissement du réseau, et permettent un accès, généralement depuis le World Wide Web, aux données ainsi constituées. Ce mouvement prend aujourd'hui de l'ampleur. Il est dans de nombreux pays, dont la France, soutenu par une volonté politique, au nom d'une nécessaire numérisation du patrimoine savant et culturel de la nation, ou du besoin de permettre l'accès à des documents fragiles. Les projets de numérisation du patrimoine en cours en région Nord-Pas-de-Calais bénéficient ainsi d'importants financements publics. Près de 8 millions de francs ont été investis entre 1997 et 2000, provenant pour moitié de fonds européens, pour un quart de l'État et pour un quart des collectivités territoriales [« Évolution du plan de numérisation des fonds d'État, bilan pratique 1996-1999, nouvelles orientations pour 2000-2001 », 1999].

Il est donc probable que les collections de ce type soient, dans les années à venir, de plus en plus nombreuses et de plus en plus riches. Il n'est pas certain cependant que l'effort se poursuive indéfiniment, ni que l'accès à celles-ci demeure gratuit. De telles opérations en effet ont un coût, et celui-ci est important, ce que l'exemple d'une entreprise de qualité, mais limitée dans son ambition, illustrera. Les historiens du Reading Area Community College se sont engagés en 1997 dans un projet dont la finalité était la numérisation et la mise en ligne d'un fonds documentaire relatif à la construction d'un canal ouvert en 1825. Il a fallu pour rendre accessible ce fond, riche de 1 100 documents comprenant des plans, des diagrammes, des croquis, embaucher trois personnes qui ont, durant plusieurs mois, sous la conduite d'un historien consacrant une bonne partie de son temps à l'avancement du projet, numérisé les documents, construit la base de données les répertoriant et le site web permettant l'accès à celle-ci [Lawlor, 2002]. Le succès de l'entreprise suppose de plus que cet établissement dispose d'un serveur dont la maintenance soit efficacement assurée. Sa pérennité dépend, quant à elle, de l'entretien du site, et en particulier du fait que soit assurée, à intervalles réguliers, une migration des données vers de nouveaux formats. La durée de vie des formats informatiques est en effet généralement assez brève, ce que savent tous ceux qui disposent d'archives numérisées vieilles de plus de dix ans, périodiquement confrontés à la quasi-impossibilité d'ouvrir d'anciens fichiers sur des machines récentes. Toute entreprise de numérisation implique donc un débours initial important, plus du fait d'ailleurs de la nécessité d'un travail humain que des dépenses liées à l'achat de matériel et génère des coûts de maintenance. L'une des questions qui se pose alors, et les enjeux en sont importants, est de savoir qui paiera et dans quelles proportions.

Un désengagement des établissements publics, ou un recours accru aux redevances d'utilisation, risquant d'orienter l'effort de numérisation vers des collections pour lesquelles existe une demande solvable, ou une forte demande institutionnelle, voire dont la sélection répond à des nécessités techniques de conservation. Ce ne sont pas nécessairement les matériaux ainsi sélectionnés qui intéressent le plus l'historien.

Malgré ces incertitudes et ces limites, de telles entreprises sont utiles aux historiens. Il nous faut cependant là encore revenir à la matérialité des opérations de construction de ces documents pour mieux en apprécier la portée et les limites. Numériser un document consiste, non à en prendre un cliché ou une empreinte, même si cela peut constituer la première étape de la numérisation, mais à produire, à partir de celui-ci, un document numérique, constitué d'une séquence de nombres, générée par un opérateur humain qui a le choix entre plusieurs outils et plusieurs stratégies de numérisation, ainsi qu'entre plusieurs possibilités quant aux modes de classement et d'accès aux documents produits. Lorsque le document est disponible par le biais du réseau, cette suite est adressée à l'ordinateur de l'historien et d'autres séquences binaires, comprenant les instructions des programmes dont il dispose sur sa machine, sont mobilisées afin de traiter le fichier reçu et d'afficher sur la surface de son écran une image lisible par lui. Nous n'assistons donc pas, consultant un site, à l'affichage d'un document d'archive, mais obtenons une représentation de celui-ci, qui résulte d'un processus d'abstraction dont les caractéristiques fixent les usages possibles de l'objet, qui sont en nombre infini, mais définis par les opérations le produisant. La nécessité de préciser ce point ne vient point d'un souci pédant d'user d'un vocabulaire exact. C'est en effet la connaissance de la genèse de l'objet que nous percevons qui nous permet de mieux comprendre ce que nous pouvons – et ne pouvons pas – en faire. Rien de choquant d'ailleurs dans cette affirmation selon laquelle des propriétés matérielles des objets manipulés par l'historien et de celles des contextes d'usage dépendent les opérations qu'il peut effectuer, et donc, *in fine*, certaines des caractéristiques de l'objet de connaissance qu'il produit. Il suffit pour s'en convaincre de se référer aux travaux des historiens du livre [Cavallo et Chartier, 1997].

En l'occurrence, cette attention aux propriétés de l'objet manipulé permet d'abord de comprendre que la consultation d'un site n'est pas du même ordre que l'examen d'un document d'archive et ne peut, dans certains cas, se substituer à elle. La représentation obtenue, parce qu'elle est abstraction, ne conserve pas certaines des propriétés du document

numérisé, ni sa texture, ni son odeur, ni ses propriétés physiques, pour ne prendre que quelques exemples, ne nous sont accessibles. Cela constitue, dans le cadre de certaines recherches, une perte sensible, pensons, par exemple, à un historien qui étudierait les techniques de production des manuscrits anciens.

D'autre part, la page-écran à laquelle nous accédons incorpore des choix faits par les opérateurs chargés de la numérisation et de la mise à disposition du document. Ces choix ne sont techniques qu'en apparence, puisqu'ils enchâssent des représentations des usages légitimes ou souhaités des documents disponibles, ainsi que des compétences possédées par les utilisateurs et déterminent les usages possibles. Prenons, pour illustrer cela, l'exemple d'une circulaire d'une page adressée par un préfet aux maires de son département. Celle-ci peut être numérisée en mode image : nous aurons alors accès à l'équivalent d'un cliché du document, qui préservera sa disposition matérielle, mais ne permettra pas de manipuler le texte de la circulaire. Il sera alors impossible, par exemple, de la citer sans recopier le passage pertinent, ou de chercher un mot dans le texte sans lire celui-ci entièrement. Une numérisation aboutissant à la seule mise à disposition du contenu textuel du document le permettra, mais nous ne pourrons, nous posant des questions sur la date du texte ou son authenticité, ni examiner sa disposition matérielle ni son aspect. Si nous choisissons cette seconde option, il nous faut ensuite savoir si nous présenterons le texte en utilisant l'HTML, qui est le format le plus courant sur le Web, mais ne permet pas de fixer l'aspect de la page-écran qu'observera l'internaute (les caractères, les couleurs, la taille des colonnes varieront d'une machine à l'autre), le XML, qui permet d'incorporer au document des métadonnées décrivant avec précision la structure de celui-ci, ce qui en facilite l'incorporation dans une base de données et l'indexation, ou le Portable Document Format (PDF). Ce dernier, format propriétaire, contraint le visiteur à quelques manipulations de plus, ce dont nous pouvons, en fonction de ce que nous savons, ou devinons, du public futur, le croire ou non capable, mais a l'avantage de garantir une impression du document à l'identique, quelle que soit la machine utilisée, et donc offre la possibilité de reproduire assez fidèlement la disposition du texte original.

En somme, et nous avons là l'une des explications du relatif dédain dans lequel les historiens semblent tenir les ressources numériques, nous sommes en permanence confrontés à un écrit-écran dont les caractéristiques ont été fixées par d'autres, qui sont rarement des historiens et prennent rarement en compte les besoins des historiens, d'ailleurs très

divers d'une recherche à l'autre. En d'autres termes, il faut, le plus souvent, ou retravailler un matériau qui n'a été produit ni par ni pour les historiens, ou accepter que celui-ci ne réponde qu'à certaines de nos attentes et ne dispense pas toujours de retourner à la source de l'objet examiné.

L'ARCHIVE NUMÉRIQUE

+++++

Ces conclusions sont également valables si nous examinons les données n'existant que sous forme numérique, qu'il s'agisse des contenus du Web, des archives des listes de diffusion, des *newsgroups* ou des données numériques produites par les entreprises et les administrations, qui, de plus en plus, n'ont pas d'équivalent papier. Ce monceau de données constitue un ensemble extrêmement riche et volumineux de dispositifs discursifs qui constituent un matériau de choix, tant pour l'historien du très contemporain que pour l'historien de demain, à condition cependant qu'émergent des dispositifs d'archivage permettant de pallier leur volatilité. Ces données en effet ne sont disponibles, et ne peuvent éventuellement être accessibles par le réseau, que tant qu'elles sont physiquement inscrites dans une mémoire de masse à laquelle il est possible d'accéder et codées en un format lisible. Or cet ensemble de conditions est fort difficile à réunir. Nous avons déjà évoqué la faible durée de vie des formats informatiques. Ajoutons que nombre de supports informatiques ne sont guère plus pérennes, les disques magnétiques s'effaçant plus vite que l'encre des livres. De plus, dans le cas des sites web, les opérateurs, qu'ils s'agissent des auteurs des sites ou de ceux qui en assurent la sauvegarde ou la mise à disposition, peuvent rapidement disparaître ou abandonner un site. Enfin, les documents mis à disposition du public par le biais du réseau, et en cela la métaphore de la page, se révèlent très inexacts, ne sont pas figés, mais au contraire fréquemment modifiés, corrigés et mis à jour sans qu'il ne reste la plupart du temps de trace des états antérieurs. La tâche est donc rude et coûteuse d'autant que, pour le seul cas du Web, la croissance du volume des données portées sur le Web est exponentielle : il naîtrait aujourd'hui de 4 à 8 millions de pages par jour alors que plusieurs millions disparaissent ou se modifient chaque seconde, tandis que la technologie du stockage semble actuellement rencontrer un palier [Farison, 2001].

Ces difficultés sont anciennes et connues de la plupart des acteurs concernés, dont beaucoup ont, dans un certain désordre, entrepris d'archiver tout ou partie des données présentes sur le réseau ou des données

numériques vitales. L'Internet Archive de Scott Kirkpatrick génère de petits robots informatiques qui prennent des instantanés de la toile, ce qui lui permet d'affirmer disposer de plus de 11 milliards de pages web archivées et parvenir à en archiver plus de 120 millions par jour. Ce site permet aussi l'accès à des collections d'images, de livres, de films et de documents sonores numérisés. Destiné à servir la recherche et les historiens, il permet la consultation gratuite de ses fonds.

L'auto-archivage quelque peu décentré et anarchique du réseau – sans compter la possibilité que seuls des organismes privés puissent posséder la mémoire de parties de celui-ci – a conduit plusieurs gouvernements à mettre en place, très récemment, des dispositifs d'archivages publics des données numériques et des informations disponibles sur Internet. Au Canada, le discours du trône du 30 septembre 2002 a créé une institution du savoir qui a pour fonction d'offrir aux Canadiens un « accès facile et intégré à leur patrimoine documentaire et au savoir sur la société canadienne » [Clarkson, 2002], ce qui comprend, entre autres missions, le prélèvement périodique d'échantillons de sites canadiens et de sites traitant du Canada. En France, une loi datée du 13 juin 2001 prévoit elle aussi l'archivage automatique de sites web, mission confiée conjointement à la BnF et à l'Institut national de l'audiovisuel (INA)³. La mise en place de tels dispositifs est cependant loin de résoudre toutes les difficultés, ne serait-ce que parce que les sites dynamiques, qui se généralisent, peuvent, en l'état actuel des choses, difficilement être automatiquement aspirés, ou parce que le cadre juridique de la mise à disposition de ces données n'apparaît pas définitivement fixé.

LES DONNÉES NUMÉRIQUES NON TEXTUELLES

+++++

Nous avons jusqu'ici évoqué des données qui peuvent donner naissance à une visualisation pouvant faire sens pour tout utilisateur, pour peu qu'il sache se servir d'un navigateur et maîtrise les codes linguistiques et sémiotiques en usage dans le monde contemporain. L'activité du réseau génère de nombreux autres documents, eux aussi virtuellement source. Nous n'en donnons ici qu'un exemple, les lecteurs intéressés pouvant se reporter aux travaux d'Éric Guichard [Guichard, 2001]. Les utilisateurs des

3. Pour les aspects juridiques des entreprises de numérisation et de mise à disposition de produits numérisés sur le réseau, voir [Lambertie, 2002].

machines abritant des données accessibles depuis le réseau ont la possibilité d'enregistrer de multiples informations concernant la provenance et les pratiques de ceux qui les utilisent, voire leur identité et ce très souvent à leur insu. Les fichiers logs ainsi constitués, qui ne sont pas des fichiers en langue naturelle susceptibles d'une lecture immédiate, permettent de multiples traitements, qui ont généralement une finalité pratique : l'optimisation des sites hébergés (évaluation de la bande passante nécessaire, refonte de l'architecture afin de mieux satisfaire les besoins des visiteurs), ou commerciale (il est possible de se constituer ainsi à peu de frais un fichier d'adresses électroniques, associant à chacune de celles-ci un profil de consommateur que l'on revendra aux spécialistes du marketing direct, ou que l'on utilisera pour promouvoir auprès d'eux ses propres produits). Ces données peuvent cependant être mobilisées par le chercheur à d'autres fins, devenant des indicateurs de l'activité des pôles universitaires français [Guichard, 2002], permettant d'évaluer la provenance géographique des visiteurs d'un site, voire, lorsque nous pouvons étudier les requêtes formulées par les internautes utilisant un grand moteur de recherche francophone, nous donnant de précieuses indications sur les modes d'appropriation et de lecture du Web des internautes français et de langue française. Là encore, les historiens poussant leurs investigations jusqu'au très contemporain peuvent trouver du grain à moudre, à condition toutefois d'être capable de localiser ces sources et d'extraire et de traiter l'information issue de tels fichiers.

LES MÉANDRES DU NET

+++++

De façon plus générale l'usage, entendons l'usage professionnel, d'Internet n'est jamais ni immédiat, ni intuitif, ni transparent, il suppose l'acquisition de compétences, nous en avons déjà évoqué quelques-unes, que n'acquière pas traditionnellement les historiens [Genet, 1994], et dont l'apprentissage suppose du temps, et cela même lorsque nous envisageons des tâches en apparence simples, telles que chercher une information pertinente et utilisable sur Internet. L'un des commentaires les plus fréquents faits par les étudiants, ou les historiens, utilisant Internet, est en effet que le réseau regorge sans doute d'informations, mais qu'elles sont fort difficiles à trouver et que leur fiabilité est sujette à caution. La structure du réseau l'explique en partie. Aucune autorité centrale ne contrôle ni ne classe les documents accessibles depuis celui-ci. De ce fait, le réseau

ressemble physiquement à une gigantesque collection de greniers, parfois rarement nettoyés, en lesquels de multiples individus déposent des documents de toute nature, à ceci près que nous avons immédiatement accès à une bonne partie de ces lieux de dépôts, qu'il est souvent possible d'écrire au propriétaire des lieux, et que les documents qu'ils recèlent sont constitués de formes (en l'occurrence des séries numériques) que les utilisateurs des machines du réseau peuvent identifier et manipuler, ce qui permet l'indexage et le catalogage d'une partie des ressources disponibles. Aucune instance centrale n'assure, même à l'échelon local, ces tâches. Les bibliothécaires et les documentalistes du réseau sont pourtant fort nombreux, même s'ils sont loin de suffire à la tâche, et extraordinairement divers. Des chercheurs, de simples particuliers, des acteurs institutionnels, des compagnies privées, tissent les cartes entrecroisées de multiples territoires virtuels et nous sont des auxiliaires indispensables. Leur travail aboutit à la production de deux types d'objets favorisant la navigation, des portails et des moteurs de recherche, dont il faut généralement combiner les indications dans le cadre d'une recherche.

PORTAILS

+++++

Les sites-portails sont construits par des spécialistes d'un domaine, des institutions, des sociétés commerciales ou de simples passionnés et recensent des ressources, généralement classées, dans le meilleur des cas présentées et évaluées, utiles aux spécialistes et aux étudiants. Ils constituent de bons points de départ pour une recherche documentaire. Je ne puis, dans le cadre de cet article, citer tous les portails francophones utiles aux historiens en quête de sources, et renvoie donc au recensement des sites-portails d'histoire effectué par Christine Ducourtieux pour l'école doctorale de Paris I. Je me contente de signaler que les médiévistes disposent, avec Ménestrel, d'un très bel outil de ce genre et que tous les sites des archives départementales sont recensés par un portail destiné aux généalogistes [« Adresses et sites des archives départementales », s. d.].

MOTEURS DE RECHERCHE

+++++

Il est souvent nécessaire de croiser entre elles les informations fournies par ces sites et de les compléter par celles obtenues par le moyen d'un

moteur de recherche, de la même façon qu'un historien approchant un sujet est contraint de compulser plusieurs répertoires de sources et de nombreux ouvrages. Contrairement aux portails, la plupart des moteurs de recherche ne sont pas conçus par et pour des historiens, mais sont créés par des sociétés commerciales et destinés à un large public. Ils reposent tous sur des principes assez similaires. Un moteur indexe automatiquement les suites de caractères, en faisant abstraction de leur sens, que ses robots rencontrent sur les pages. Ce fonctionnement permet aux moteurs de recherche de renvoyer à l'internaute un grand nombre de réponses sur tous types de requêtes, à partir du contenu de la totalité des pages stockées dans leur index. Le classement des résultats se fait selon un algorithme de pertinence vis-à-vis de la requête formulée, chaque moteur ayant son propre algorithme. Il s'en déduit, et l'étude des fichiers logs des moteurs permet de le confirmer [Guichard, 2001], que la quasi-totalité des requêtes obtiennent des réponses non pertinentes et inutilisables, car, constituées de peu de mots et de mots peu discriminants, elles génèrent des listes interminables où figurent beaucoup de pages inutiles au chercheur. Celui-ci consultant rarement plus d'une ou deux des pages concernées, il est fréquent que le moteur retourne une information pertinente dont l'internaute ne prenne pas connaissance. Il faut donc, pour pouvoir utiliser avec efficacité ces outils, développer des stratégies adaptées à son objet, et en particulier formuler des requêtes longues, si possible composées de plusieurs termes les plus discriminants possibles, voire lancer plusieurs requêtes, légèrement différentes les unes des autres, adressées à plusieurs moteurs, et avant d'explorer la liste des réponses, examiner les indications de provenance et de contenu que le moteur retourne. La requête « états généraux, Necker, Louis XVI, trône, discours » me permet ainsi, grâce à un enseignant de l'université de Hessen qui a porté en ligne l'essentiel de l'un de ses cours et des documents utilisés dans le cadre de celui-ci, de récupérer de larges extraits du discours d'ouverture des états généraux, la référence à ce document arrivant en sixième position sur la liste des réponses fournies par un moteur de recherche. J'ajoute que ce document est le premier de ceux de la liste que j'ai consulté, tant l'adresse des autres sites, que les autres indications fournies par le moteur, me faisant douter de pouvoir y trouver ce que je cherchais.

VALIDATION DE L'INFORMATION

+++++

J'ajoute qu'ayant spontanément confiance en la compétence et le sérieux d'un collègue, en celle aussi de l'administrateur réseau d'une université allemande, que j'ai pu identifier grâce à son Uniform Resource Locator (URL), je suis assez raisonnablement certain du sérieux de ma source. Du moins, je serais prêt à l'utiliser telle quelle dans le cadre d'un cours, même si je vérifierais probablement la transcription et retournerais sans doute à l'original s'il s'agissait là d'une pièce importante du dossier documentaire d'une recherche en cours. Les choses cependant ne sont pas toujours aussi simples, particulièrement pour les contemporanéistes, qui ne chercheraient pas la trace de sources anciennes, mais voudraient constituer un dossier documentaire, concernant par exemple l'activité d'une organisation politique, incluant des documents internet. En effet, la fiabilité des données accessibles depuis le réseau pose véritablement problème dans un certain nombre de cas. D'une part, nous ne disposons pas des repères physiques et institutionnels (la présence dans un dépôt d'archives, par exemple) qui nous permettent d'habitude d'identifier le document que nous utilisons.

D'autre part, du fait de sa nature même, l'information numérique est assez facile à copier et à modifier de façon indécélable par l'utilisateur et s'introduire dans un serveur afin d'en modifier les données est à la portée de nombre d'informaticiens et de bricoleurs. Un groupe non identifié de libéraux furieux a ainsi attaqué le site de la revue *Le mouvement social*, il y a quelques mois. Les visiteurs accédaient à une page blanche sur laquelle de grosses lettres noires indiquaient que ce site ne serait pas accessible tant que le mouvement social prendrait en otage la France. Un historien tant soit peu averti pouvait cependant déceler la modification et douter de ce que l'équipe de cette vénérable revue ait soudain décidé de se croiser afin de défendre la société libérale menacée par des hordes de grévistes. Boutade bien sûr, mais qui voudrait rappeler que les historiens sont parmi les mieux armés pour affronter cette incertitude et ne devraient point en être effrayés. Le document obtenu depuis Internet est, pour peu que l'on ait quelque idée de sa structure et de sa matérialité, susceptible, comme tout document, d'une critique, tant interne qu'externe. Son adresse (ou URL) donne ainsi des indications sur sa localisation physique, et parfois l'inscription institutionnelle de l'auteur du document, l'affichage des sources de la page permet, dans un certain nombre de cas, de vérifier la date de modification affichée sur celle-ci. Reste enfin le loisir de croiser le document obtenu et d'autres, qu'ils soient électroniques ou

non. Ajoutons enfin que le réseau n'est pas parcouru sans relâche par de mauvais génies employant toute leur industrie à falsifier les données intéressant les historiens.

LIRE LES SOURCES NUMÉRIQUES

+++++

Le lecteur pourra se demander à ce point, après l'évocation de tant de peines et de difficultés, si le coût de constitution d'un recueil de sources numériques est justifié. La réponse varie, dans d'importantes proportions, selon le type de sujet et de questions que traite l'historien. À la date d'aujourd'hui, les données numériques fournies par le réseau sont, dans certains cas, de modestes adjuvants, dans d'autres des sources indispensables. Il y a à cette conclusion deux prémisses. D'une part, l'ubiquité du réseau permet le rassemblement de données géographiquement dispersées, à un coût et en un temps acceptable. D'autre part, les données numériques, même lorsqu'elles ont un équivalent physique, ne sont jamais des représentations appauvries d'un référent réel, mais un objet, tout aussi matériel, mais d'une autre nature, qui rend possibles des formes de lecture ou d'appropriations spécifiques. Les fichiers numériques permettent en particulier la manipulation de gigantesques quantités de données et l'automatisation d'un certain nombre de tâches. De ce fait, elles autorisent des questions qu'un historien ne pourrait poser sans elles. Je me propose d'illustrer ce point en évoquant plusieurs recherches dont les résultats sont déjà publiés, chacune illustrant, sans que l'inventaire soit exhaustif, un usage possible d'un appareil de sources constitué pour tout ou partie à l'aide du réseau, en allant du plus simple, ou du plus familier, au plus complexe.

LE PLUS COURT CHEMIN DE L'HISTORIEN A SA SOURCE

+++++

Une étude italienne, consacrée aux modes d'autoreprésentation graphiques adoptés par les partis socialistes et communistes européens, c'est-à-dire aux logos et symboles choisis par eux, me fournit mon premier exemple [Caretti, Degl'Innocenti et Silei, 2002]. L'ouvrage se clôt par un dernier chapitre qui utilise pour sources les sites de nombreux partis socialistes et étudie les logos et les dispositifs graphiques de ceux-ci. Les outils d'analyse utilisés sont fort classiques et relèvent pour l'essentiel d'une sémiotique qui pourrait être mise en œuvre à partir de sources classiques,

et l'est d'ailleurs dans la première partie de l'ouvrage. Ce qu'apporte ici Internet est la possibilité à moindres coûts, et en un temps limité, de rassembler une documentation physiquement très dispersée et de prendre comme échelle d'observation le monde. Certes, il ne serait pas impossible de se fixer comme objectif de récupérer une collection de tels objets par d'autres moyens. Il faudrait pour cela se procurer les adresses de toutes les organisations appartenant à cette famille, puis écrire à chacune d'elles en lui demandant de bien vouloir envoyer quelques échantillons de ses dispositifs graphiques. L'opération cependant a un coût, certes encore modeste en ce cas, demande quelques compétences linguistiques, implique un délai de réponse plus ou moins long, et il est fort probable que toutes ne répondent pas. Il est ici plus facile, plus rapide et moins coûteux de constituer l'appareil de sources à partir d'Internet, même s'il était possible de le faire par d'autres moyens.

DE NOUVEAUX MODES DISCURSIFS

+++++

Plusieurs historiens et spécialistes des sciences politiques s'intéressent aujourd'hui beaucoup aux formes nouvelles nées avec la diffusion du réseau – sites web, mais aussi forum de discussion et listes de diffusion – parce qu'elles permettent d'avoir accès à des formes de communication internes au groupe ou aux réseaux sociaux, auxquelles l'historien a rarement accès : la conversation, l'échange téléphonique, voire souvent la feuille ronéotée, laissant peu de traces. Fabien Granjon a ainsi étudié les répertoires d'action de plusieurs réseaux et associations politiques appartenant à la mouvance antimondialiste à l'aide, entre autres sources, d'archives numériques qu'il a lui-même constituées, stockant les messages passant par les forums et les listes de diffusion de plusieurs associations [Granjon, 2002]. En ce cas encore, les méthodes d'analyse utilisées sont les mêmes que celles des historiens travaillant des documents papier, la source cependant ne peut être constituée que grâce au réseau, et elle permet de décrire, dans la durée, beaucoup plus finement les échanges informels au sein du groupe que n'importe quelle autre source.

ANALYSE TEXTUELLE

Les données numérisées fournies par le réseau offrent de plus la possibilité de mettre en œuvre, sur de grands corpus, des modes d'exploration et d'utilisation de l'information automatisés. De tels dispositifs sont nombreux, nous ne mentionnerons que l'analyse textuelle, qui apparaît ici particulièrement adaptée, à la fois parce que le réseau offre de vastes corpus de textes déjà numérisés et parce que les outils et les méthodologies de ces formes d'analyse sont déjà rodés et utilisés depuis longtemps, tant par des historiens que par des spécialistes de l'analyse littéraire [Bernard, 1999]. Certains voient même dans ces formes d'analyse le simple prolongement de très anciennes pratiques. Michel Bernard écrit ainsi que :

« [...] on a pratiqué une forme manuelle de la statistique textuelle depuis le XIX^e siècle [...]. Il ne s'agit donc pas de l'irruption dans le champ littéraire d'un intrus exogène et imposé de l'extérieur, mais de la rencontre assez naturelle entre les techniques de la recherche littéraire et des outils qui la facilitent [...]. » [Bernard, 1999, p. 8]

Même si ces outils permettent des questionnements nouveaux, ils sont souvent un moyen d'obtenir des éléments de réponse précis à des interrogations anciennes. Il est ainsi possible, écrit-il, de vérifier la date de première attestation de tous les mots d'un poème afin de se rendre compte de la part d'archaïsmes et de néologismes qu'il utilise, préoccupation ancienne, mais « qui aurait fait une telle démarche en feuilletant les pages d'un dictionnaire étymologique » [Bernard, 1999, p. 14].

Cette même tension entre radicale nouveauté et continuité est au cœur de tout discours dédié au trinôme que constituent l'historien, ses sources et le réseau, ou l'ordinateur – ce qui revient au fond au même – qui hésite toujours entre la métaphore rassurante qui, rapportant au connu, permet d'apprivoiser l'objet nouveau – au risque de méconnaître les ruptures qu'il peut introduire – et l'annonce prophétique. Nous n'avons pas ici échappé à la règle, puisque nous sommes tentés de conclure que l'intégration du réseau dans les stratégies de construction de sources ne change pas les fondements du métier d'historien, même si elle suppose l'apprentissage d'un certain nombre de compétences nouvelles. L'historien doit toujours constituer des cartes mentales des gisements de traces disponibles, ces traces fussent-elles traces de traces, en évaluer la pertinence, sans le

secours parfois de l'archiviste ou du documentaliste – ce qui est l'occasion de mesurer ce qu'il leur doit – à partir d'une critique qui suppose la compréhension de la genèse et du mode de conservation des documents recensés. Il doit enfin rendre compte de l'économie de ces vestiges en utilisant des moyens appropriés à leurs caractéristiques, soit savoir les lire et les interpréter. Dans la mesure cependant où certaines recherches peuvent désormais mobiliser des données numériques à une échelle sans précédent, il semble probable qu'un nombre croissant d'historiens soit appelé à mobiliser des technologies intellectuelles nouvelles, dont bon nombre ne sont sans doute pas encore nées, ou sont encore inconnues d'eux, ce qui ne saurait manquer d'avoir des effets sur les questions qu'ils pourront demain poser à leurs sources.

+++++

CHAPITRE IV.

LES SITES DES ARCHIVES DÉPARTEMENTALES FRANÇAISES

+++++

Ce texte a connu deux versions successives. Il a paru une première fois dans un ouvrage publié en français [Rygiel et Noiret, 2005] élaboré à la suite d'une rencontre entre les rédactions du *Mouvement social* et de *Memoria e Ricerca*, qui s'était déroulée dans les locaux de l'Équipe Réseaux-Savoirs-Territoires de l'ENS le 26 octobre 2002 et dont le thème était Internet et les revues d'histoire. Une seconde version, peu différente, est parue en italien sous le titre « I siti degli archivi dipartimentali francesi : una ricognizione », dans la revue *Memoria e Ricerca. Rivista di storia contemporanea* (mai-août 2006, n° 22, pp. 165-172).

Nous avons choisi de conserver pour ce texte les URL de l'époque, puisque son intérêt n'est pas de favoriser le repérage des sites des archives départementales françaises, mais de témoigner d'un état, vieux d'une dizaine d'années, de ce Web archivistique. Les URL ayant souvent entre-temps changé, il a paru que le meilleur moyen de lier le propos à la matérialité qu'il évoque était, donnant l'URL telle qu'elle était formée en 2004, d'autoriser à retrouver l'état de ces sites par l'usage d'une archive du net.

Nous voudrions ici explorer, non pas tant les promesses que le réseau offre aux historiens, que le produit, à une date donnée, de l'action d'institutions liées à l'activité historique. Il ne s'agit pas là de condamner implicitement les textes programmatiques ou théoriques qui tentent de discerner les futurs possibles, mais d'observer les produits de l'activité concrète de professionnels et d'institutions confrontés à Internet. En effet, l'usage de l'informatique et la pratique des réseaux sont aujourd'hui assez fréquents pour que nous puissions examiner des usages du réseau, normaux ou routiniers, et non seulement quelques entreprises pilotes et quelques sites expérimentaux. Ce type d'enquête, à la fois étude des formes de l'appropriation d'un dispositif technique et pour les

historiens que nous sommes, confrontés à de rapides transformations des dispositifs de production et de diffusion du savoir historique, production de matériaux nourrissant une réflexion sur les changements en cours, est rarement conduit de façon systématique, ce qui suffit à justifier son utilité.

Les sites des archives départementales françaises offrent, dans cette perspective, un bon point d'observation. Ils sont en effet nombreux, ce qui permet de repérer des tendances et de ne pas surestimer peut-être la rapidité des bouleversements que promettent quelques sites exceptionnels. D'autre part, les institutions les produisant et les nourrissant ont mêmes missions, même organisation et même statut, et sont animées par des personnels ayant bénéficié de formations similaires, même si bien sûr elles ne disposent pas toutes des mêmes moyens humains et financiers, ceux-ci dépendant tant de la taille du service que de la générosité ou de l'indifférence des conseils généraux. La comparaison en devient à la fois plus facile et plus pertinente et la diversité constatée dans la mise en œuvre des outils-réseaux nous prémunira des explications paresseuses qui rapportent aux seules caractéristiques de l'outil les processus d'appropriation.

Le dispositif d'observation que nous avons adopté est simple. Nous avons repéré les sites à l'aide de moteurs de recherche et de plusieurs portails. Lorsque cela ne permettait pas de trouver le site, cette recherche était complétée par la visualisation du site du conseil général, qui, la plupart du temps, héberge les sites des archives départementales, ou à tout le moins pointe vers ceux-ci. Nous avons pu ainsi visiter le site de toutes les archives départementales, ou trouver mention de l'inexistence de celui-ci. Chaque site a été décrit à l'aide d'une fiche qui regroupait des informations concernant son contenu¹ et quelques éléments techniques (langages utilisés, utilisation ou non de balises méta, dispositifs interactifs, hébergeur). Cette description est bien sûr incomplète, puisqu'elle ne prend pas en compte les aspects formels de ces sites (nombre de pages, disposition de l'information sur la page, par exemple), mais nous ne nous intéressons ici qu'aux fonctions assignées à ces dispositifs.

Il s'ensuit que notre propos est essentiellement descriptif. Nous n'avons pas mené d'entretiens auprès du personnel des archives départementales, ni observé en situation le fonctionnement des services chargés de la mise

1. Nous avons retenu comme variable descriptive pertinente la présence des rubriques suivantes : pages du service éducatif, guide des publications, informations pratiques, présence de documents numérisés, inventaires, informations institutionnelles, guides de recherche. Une dernière rubrique rassemblait un ensemble de remarques et de commentaires inspirés par le dépouillement du site.

en place des sites. Nous n'avons pas non plus étudié les textes émanant de la direction des Archives de France définissant les usages souhaitables des réseaux. Nous pouvons donc présenter un corpus de productions électroniques, suggérer peut-être quelques hypothèses rendant compte de certaines des formes qu'elles prennent, mais nous n'avons pas ici les moyens de tester solidement ces productions à partir de la reconstitution de la genèse de celles-ci, non plus que d'une connaissance fine des contextes institutionnels et locaux. Ajoutons que ce texte ne saurait viser à établir un palmarès, d'une part parce que tel n'est pas notre propos, d'autre part parce que l'appréciation des produits proposés ne peut avoir de sens sans la connaissance des conditions concrètes de leur réalisation.

MODESTIE ET REDONDANCES

+++++

Le premier constat est celui d'une très grande diversité. Nous rencontrons, de l'inexistence du site au projet ambitieux et déjà abouti, en passant par la simple page HTML mentionnant informations institutionnelles et pratiques, toutes les situations. Les sites modestes, en termes de contenu, à l'architecture simple et similaire sont cependant la norme. D'après notre enquête, 12 des 99 archives départementales n'avaient pas, courant juillet 2004, de site accessible. Nous comptons parmi ceux-ci les trois disposant d'une page² annonçant qu'une page d'information était en construction. La plupart des 87 autres sites sont des plaquettes électroniques, offrant, avec plus ou moins de détail, une présentation des missions et des locaux du service et permettant parfois un contact électronique par le biais d'une adresse mail. Dans quelques cas, ces données de base sont complétées par une page évoquant l'actualité de l'institution, l'organisation d'expositions, par exemple. Si nous distinguons ceux-ci de sites de service ou de travail, que nous définissons par le fait qu'ils offrent des ressources permettant à un visiteur potentiel de réaliser à distance une partie des tâches habituellement accomplies *in situ* (consultation des inventaires, consultation de documents, consultation d'un guide de recherche), alors nous disposons de 48 plaquettes de présentation et de 39 sites de service. C'est là surestimer le poids de ces derniers, puisque certains ne figurent dans cette catégorie que parce qu'ils proposent l'accès

2. Nous appelons page, dans la suite de ce texte, suivant l'usage proposé par Jacques Anis, « [...] un ensemble de données non verbales et iconiques susceptibles d'être visualisé sans rupture » [Anis, 1998, pp. 195].

à quelques documents choisis pour leur valeur illustrative, ou décrivent de façon précise les sous-séries du cadre général de l'inventaire. Seuls sept sites proposent un accès à la fois à des documents numérisés, à des inventaires et à un ou plusieurs guides de recherche, une vingtaine offrent deux de ces trois types de données.

Les dispositifs d'écriture mis en œuvre sont de même souvent simples et de conception ancienne. L'utilisation du langage de description de document XML³, qui tend pourtant à devenir la norme en matière d'informatique documentaire professionnelle, est très rare. Le recours à des scripts complexes est peu fréquent, une dizaine de concepteurs utilisant Hypertext Preprocessor (PHP), autant Active Server Pages (ASP), qui sont à l'heure actuelle les deux langages les plus fréquemment utilisés par les créateurs de sites dynamiques. Dans la majorité des cas, les sites associent HTML, feuilles de style et parfois JavaScript, même s'il existe bien sûr des exceptions. Le site des archives départementales des Alpes-Maritimes⁴ propose ainsi un accès en ligne à ses bases documentaires au moyen de dispositifs logiciels efficaces, faisant partie des neuf sites qui permettent un accès à distance à une ou des bases de données. Plus surprenant peut-être, certaines des fonctions de base de l'HTML sont très rarement utilisées. L'interactivité, même sous sa forme minimale d'envoi de formulaire HTML ou d'offre de choix de navigation, est très rare. Nous l'avons rencontré une dizaine de fois, quelques archives départementales offrant la possibilité de commander en ligne certaines de leurs publications ou de remplir à distance un formulaire d'inscription. Deux exceptions à cela, que nous signalons parce qu'elles témoignent de la présence chez certains d'une séduisante volonté exploratoire, nous sont offertes par les sites du Lot-et-Garonne⁵ et de la Mayenne. Dans le premier cas, les internautes sont appelés à participer à l'identification de lieux et de personnes figurant sur des clichés détenus par les archives départementales, dans le second, ils sont conviés à participer à l'élaboration d'une base de données recensant les remarques portées en marge des registres paroissiaux ou d'état civil par les curés et officiers d'état civil. Ces remarques sur la vie des paroisses et communes sont indexées par lieu et par date. Peu attirés par l'interactivité, les concepteurs de site sont de même réticents à utiliser

3. Le consortium W3C maintient une page d'information consacrée au projet XML

4. < <http://www.cg06.fr/culture/culture-archives.html> >.

5. < <http://www.cg47.fr/archives/> >.

les ressources de l'hypertexte⁶. Les liens externes sont extrêmement rares, même ceux renvoyant à d'autres sites d'archives.

Nous sommes tentés d'expliquer la modestie de beaucoup de réalisations par la faiblesse des moyens disponibles. Les archives départementales sont, notoirement, des institutions pauvres. Leur personnel et en premier lieu les archivistes-paléographes en poste ont généralement été formés avant que l'irruption de l'informatique documentaire et des réseaux ouverts ne transforme les conditions de conservation et de structuration des systèmes d'information. L'impression demeure cependant que la situation actuelle s'explique aussi par la difficulté des acteurs à s'approprier certains des concepts qu'Internet met en œuvre. L'absence de liens externes, conjuguée à la redondance des informations fournies par de nombreux sites en est pour nous un indice. Beaucoup, ainsi, précisent en une page les missions, telles qu'elles sont établies par les textes, des archives départementales. Près de 25 proposent l'accès au cadre de classement des séries des archives départementales, tel qu'il est fixé au niveau national et les guides de recherche, parfois très bien faits, offrent cependant souvent les mêmes informations d'un site à l'autre. Ainsi, 21 sites proposent-ils un ou plusieurs guides, pour un total de 33 guides, ce qui nous permet de disposer de 16 guides de recherches généalogiques et de 4 guides permettant de faire l'histoire de sa maison. Il ne s'agit pas ici de contester la place donnée aux généalogistes et à la généalogie, ou à l'érudition locale, mais de remarquer que l'informatique documentaire est censée éviter la reproduction des mêmes tâches et le réseau favoriser la mutualisation des ressources, logiques qui sont loin ici d'être mises en œuvre.

UNE NOUVELLE GÉNÉRATION DE SITES

+++++

Le constat cependant ne vaut pas pour tous. Il pourrait de plus prochainement ne plus être d'actualité. Quelques sites, tant par les services qu'ils offrent déjà, que par les réalisations en cours qu'ils annoncent, témoignent de ce que la carte des sites archivistiques est en plein bouleversement, ce que permettent des investissements lourds, à l'échelle des ressources des centres d'archives et le recours à des professionnels de l'informatique,

6. Nous prenons le terme en son sens le plus courant, tel que défini par l'inventeur de la notion. Il s'agit alors d'une « [...] écriture non séquentielle, texte qui se ramifie [...]. On se le représente généralement comme une série d'unités textuelles connectées par des liens » [Nelson, 1993, pp. 0-2].

dont les productions ont peu à voir avec les bricolages pionniers des passionnés dont témoignent encore certains sites. Il semble qu'il nous faille examiner ceux-ci si nous voulons savoir de quoi demain sera fait, ou du moins quelles tendances sont à l'œuvre.

Trois sont aisément repérables, dont nous retrouvons l'écho au travers de la plupart des sites complexes : la mise à disposition des inventaires – sous forme de fichier téléchargeable ou plus rarement d'accès à une base de données –, la mise en place de programmes de numérisation des archives elles-mêmes, la possibilité d'interroger une ou plusieurs bases de données. Le premier de ces choix est aussi le plus fréquent, 26 services départementaux offrent la possibilité de consulter en ligne tout ou partie de leurs inventaires. C'est assez peu surprenant dans la mesure où, pour les inventaires récents dont existe une version informatique, la mise à disposition par le biais d'Internet est une opération aisée et peu coûteuse. Elle offre de plus l'avantage de permettre à une partie des visiteurs de réduire le temps de leur présence nécessaire sur site, ce qui, pour beaucoup de services confrontés à un afflux de visiteurs dépassant leurs capacités, constitue un avantage indéniable. De fait, lorsque peu d'inventaires sont en ligne, le choix se porte sur les inventaires des séries de l'état civil, ou des archives notariées, de loin les plus consultés de tous.

La numérisation de documents d'archives et l'offre d'un accès à ceux-ci par le biais d'un site est plus rare. Elle est cependant le fait de 22 services d'archives, qui nous offrent 40 collections de documents. Les choix opérés sont assez remarquablement similaires d'un site à l'autre. Les documents images sont extrêmement nombreux : nous pouvons consulter 7 collections de documents photographiques, 2 collections de cartes et de plans, une de gravures du XIX^e siècle, 10 dossiers présentant les pièces exceptionnelles détenues par les archives, qui font une large place à ce type de document et, mais le cas est assez particulier, quelques exemples de la collection de films amateurs détenus par les archives départementales de Loire-Atlantique⁷. Ces dossiers constituent donc un peu plus de la moitié de ceux proposés par les archives départementales françaises, alors qu'ils ne représentent qu'une part infime des fonds détenus. Nous pouvons lire dans ce choix, à la fois la confirmation de l'intérêt accru des centres d'archives pour les documents iconographiques et photographiques, la volonté de proposer des pages attrayantes, ce qui, si l'on en juge par exemple par l'évolution des manuels du secondaire, est généralement associé à la multiplication des

7. < <http://culture.cg44.fr/Archives/> >.

images. Nous faisons enfin l'hypothèse que la présence parmi les concepteurs de site d'individus ayant eu une formation de maquettiste ou de graphiste contribue à cette valorisation de l'image.

Il est également fréquent que soient proposées des archives de type sériel, en particulier le cadastre ancien (4 fois) et tout ou partie de l'état civil ancien (4 fois et l'on peut associer à ces 4 sites celui qui propose une numérisation d'une partie de ses fonds notariés). Le choix, là encore, surprend assez peu. Ces documents sont parmi les plus consultés. Le souci d'éviter leur détérioration autant que la surcharge des salles de lecture incite à en proposer la consultation à distance, d'autant que ces données comportent leur propre indexation (*a minima* par lieu et par date), ce qui facilite la construction de bases de données en autorisant la consultation à distance. Il ne faut de fait pas sous-estimer l'importance de ce facteur. La mise au point de systèmes de description et d'indexation de documents archivistiques est une opération difficile et complexe [Minuti, 2002, pp. 54-70].

De façon générale, les sites observés sont conçus en fonction des besoins ou des attentes du public de l'institution et les généalogistes sont particulièrement nombreux au sein de celui-ci, ainsi que les curieux d'histoire locale. Précisons que nous n'entendons pas ici proposer un schéma qui postulerait la transparence de besoins auxquels l'institution répondrait par la mise en place d'une solution technique neutre en termes sociaux et savants [Sfez, 2002, pp. 75-80]. Nous remarquons simplement que les dispositifs mis en place ont pour premiers destinataires des membres clairement définis du public des archives départementales. Nous en retrouvons les effets dans le choix des guides, des inventaires et du type de documents d'archives proposés au public. L'examen des périodes et des thématiques associées aux collections de documents proposés ne fait que renforcer cette conclusion. La période contemporaine, le XX^e siècle en particulier, avec les deux moments que sont la Belle Époque et la seconde guerre mondiale, qui, il suffit pour s'en convaincre de parcourir les rayons d'une librairie, suscitent beaucoup d'intérêt dans le public, se taillent la part du lion. Certains signes laissent supposer que ces choix ne sont pas seulement le produit des représentations que se font les personnels des centres d'archives des besoins de leur public, mais sont guidés par le produit de la consultation de celui-ci, voire infléchis par la participation de certains de ses éléments. Les archives départementales de la Haute-Vienne ont ainsi, dès 1999, mené une enquête auprès de leurs lecteurs, demandant à ceux-ci de préciser le type de services auxquels ils désiraient accéder à distance. Nombreux à répondre, ils ont massivement écrit

qu'ils souhaitent un site de services, rêvant de fonds numérisés et impatients d'accéder de chez eux à des instruments de recherche⁸. De fait, les évolutions constatées sont proches d'innovations explicitement souhaitées par de nombreux utilisateurs réguliers.

Il nous semble cependant que la prise en compte des lecteurs que sont les généalogistes et, dans une moindre mesure, les curieux d'histoire locale, n'est pas seulement liée à leur nombre, mais au fait que, collectivement, par le biais de leurs associations, ou individuellement, ils ont parfois su se montrer prêts à s'associer aux efforts entrepris. Plusieurs sites en témoignent. Celui des archives départementales des Côtes-d'Armor⁹ propose ainsi l'accès à une base de données rassemblant des relevés d'actes d'état civil collectés par des associations de généalogistes. Si les sites des archives semblent souvent leur être destinés, plus qu'aux professionnels de l'histoire, nous faisons volontiers l'hypothèse que cela ne reflète pas seulement l'avantage du nombre, mais aussi le fait que, que ce soit à titre collectif ou à titre individuel, les historiens n'ont guère cherché à être partie prenante de la création de ces objets.

La troisième tendance repérable au travers de l'examen de ces sites est l'apparition de modules d'interrogation de bases de données. Le cas est encore rare. Cela ne concerne que neuf sites. La relative nouveauté de la chose – les sites dynamiques ne se multiplient que depuis deux ou trois ans – autant que le coût de ces outils, dont la mise au point exige plus de temps et plus de compétences que ceux de la génération précédente, l'explique en partie. Les bases de données mises en ligne sont de trois types. Certaines décrivent des inventaires, évolution qui, d'un point de vue technique, est logique. Les services d'archives disposent aujourd'hui de logiciels, en particulier d'Arkheia, destinés à la constitution des inventaires qui produisent des documents électroniques structurés, dont le portage sur le Web est assez aisé. D'autres bases décrivent les collections de documents mis à disposition, en particulier lorsque ceux-ci sont des documents de type sériel (cadastre, état civil), données qui là encore, de par leur structure, se prêtent à la construction de bases de données. Quelques-uns enfin se sont lancés dans la construction de bases de données documentaires. C'est le cas, en particulier, des archives départementales de la Martinique¹⁰. Celles-ci nous proposent une base consacrée à l'histoire de

8. Nous remercions le personnel des archives départementales de Haute-Vienne d'avoir bien voulu nous transmettre ce document.

9. < <http://archives.cg22.fr/> >.

10. < <http://www.cg972.fr/ArchivesMartinique/index.htm> >.

Saint-Pierre. Les archives des Alpes-Maritimes, elles, proposent un choix impressionnant, dont une très étonnante base de données permettant de retrouver la trace d'un immigré italien ayant séjourné dans le département entre 1840 et 1935, données d'état civil et cote des dossiers dans lesquels il est mentionné à l'appui¹¹. Là encore, le coût et la complexité de telles opérations expliquent leur rareté, même s'il faut faire la part des stratégies et des représentations des archivistes, dont certains mettent en cause et pour des raisons qui ne sont pas triviales, la pertinence de tels investissements. Alain Paul, directeur des archives départementales des Hautes-Alpes, doute ainsi de l'intérêt de la mise à disposition du public de modules d'interrogation de bases de données documentaires, dont écrit-il :

« il ne sera jamais certain qu'elle sera bien pertinente quant au résultat, la réponse valable étant conditionnée par le bon paramétrage sémantique de la question et son étroite adéquation avec le vocabulaire documentaire utilisé. »[Paul, s.d.]

Si les tendances repérées ici sont les plus lourdes et préfigurent sans doute les réalisations des années à venir, elles n'épuisent pas le champ des possibles, ni même ne suffisent à présager de ce que sera le Web archivistique d'ici quatre ou cinq ans. D'une part en effet, tant l'évolution des outils que les transformations, plus lentes, des représentations et de la formation des personnels sont susceptibles de nous valoir quelques surprises. D'autre part, certains sites témoignent d'une volonté exploratoire et expérimentent des dispositifs, parfois techniquement simples, signe d'une réflexion en cours, qui est dans le monde des archivistes, comme dans celui des documentalistes, collective. En témoigne, par exemple, la tenue de réunions annuelles consacrées à l'informatique documentaire et à ses applications. Il existe ainsi un club des utilisateurs du logiciel d'archive Arkheia, qui tient des réunions annuelles au cours desquelles sont exposés les problèmes posés par l'informatique et l'archivistique documentaire. Nous avons évoqué, dans les lignes qui précèdent, quelques exemples d'interactivité ou des choix de documents atypiques. Appartiennent aussi à cette catégorie les essais de quelques services éducatifs – chaque centre d'archives départementales est doté d'un tel service animé par un enseignant du secondaire détaché – qui proposent, outre des informations sur leurs activités, des modules d'exercices pour la classe, ou, pour les plus

11. < <http://www.cg06.fr/os-html/immi/home.html> >.

audacieux, des exercices de paléographie. C'est en ce domaine encore le site des archives départementales des Alpes-Maritimes qui apparaît pionnier, proposant de nombreux modules d'exploration de documents d'archives destinés à des classes de niveau différent et des exercices de paléographie corrigés. Quelques services enfin proposent une ou plusieurs expositions virtuelles, à la suite souvent d'expositions tenues aux archives départementales. C'est le cas par exemple de la Haute-Vienne, qui propose une exposition consacrée au Tour de France – nous sommes au pays de Raymond Poulidor et de Pascal Hervé – ou des archives du Vaucluse, qui mettent en ligne plusieurs expositions.

L'écho de la « révolution Internet » apparaît singulièrement assourdi lorsque l'on visite les sites des archives départementales françaises. La règle demeure la modestie des ressources proposées et la simplicité des dispositifs mis en œuvre, voire l'ignorance de certains des principes de base du média qu'est le WWW. La règle cependant connaît de nombreuses exceptions et pour certains, le site n'est déjà plus un outil dont on explore les possibilités, mais un instrument de travail bénéficiant d'une attention et d'investissements soutenus. D'après Alain Paul, directeur des archives départementales des Hautes-Alpes, la conception du site des Yvelines, qui est l'un des sites les plus riches, confiée à une société informatique, aurait coûté 100 000 francs [Paul, s. d.].

Toutes les situations intermédiaires étant représentées, le visiteur peut avoir l'impression de se voir offrir une coupe géologique, saisissant d'un coup plusieurs étapes du développement du Web archivistique français, du désintérêt au site professionnel en passant par le bricolage pionnier. Le site des archives départementales de la Mayenne est ainsi hébergé par Wanadoo, vraisemblablement à l'initiative d'un de ses membres du personnel¹². L'intérêt de cette diversité n'est pas seulement historique. Nous pouvons aussi la considérer comme le produit de logiques locales extrêmement différentes, dont les déterminants sont les ressources (tant des archives que des conseils généraux) et l'intérêt des uns (les conseils généraux), pour les archives et le monde numérique, ou des autres (les archivistes) pour les outils numériques, voire tiennent aux rapports qu'entretiennent localement services d'archives et conseil général. De ce fait, et la conclusion dépasse le cas étudié, il apparaît que les formes d'appropriation du même outil peuvent être extraordinairement variées, alors même que les populations et les contextes institutionnels sont ici

12. < <http://perso.wanadoo.fr/archives.53/> >.

très similaires. Cela nous permet de souligner ce que celles-ci doivent au contexte social de leur emploi.

Même si l'exploration des possibilités du médium est loin d'être achevée, quelques tendances se dégagent de l'examen des sites les plus ambitieux. La mise en ligne des inventaires est en cours. Tout laisse à penser qu'elle sera, d'ici quelques années, la règle. Il semble probable que la mise en ligne d'une partie des fonds eux-mêmes, couplée à l'usage de bases de données, est appelée à se développer, mais elle servira sans doute plus les attentes des généalogistes que celles des historiens, hormis ceux d'entre eux qui utilisent état civil et cadastre. Nous ne voyons pas là une raison de s'indigner, mais l'illustration du vieil adage qui veut que les absents aient toujours tort. Les historiens, en tant que groupe professionnel, sont remarquablement absents des débats et des chantiers que suscitent les développements de l'informatique documentaire et d'Internet. Les évolutions qui se font sans eux ne sauraient se faire pour eux.

+++++

CHAPITRE V. PORTES, PORTAILS, TÉLÉPORTATION

+++++

Ce texte a d'abord servi de support à une intervention donnée lors de la journée Portails médiévistiques qui s'est tenue à Paris le 8 octobre 2006 à l'initiative du réseau Ménestrel. Celui-ci naît en 1997, regroupe des médiévistes et des représentants d'institutions savantes dont les activités ont à voir avec l'histoire médiévale. La principale activité du groupe, dont la cheville ouvrière est Christine Ducourtieux, est la tenue d'un répertoire de ressources électroniques intéressant les spécialistes de ce domaine. L'équipe de Ménestrel, comme toute équipe depuis longtemps engagée dans la production d'un service de ce type, a été plusieurs fois confrontée à la nécessité de refondre l'architecture du site et de réfléchir à ses fonctionnalités. Cette journée, organisée dans les locaux du Laboratoire de médiévistique occidentale de Paris (LAMOP), à l'occasion de l'une de ces nécessaires refontes, avait pour objet de réfléchir aux transformations possibles des sites de ce type, une dizaine d'années après leur lancement. Les actes de la journée, du fait de sa nature même, n'ont pas été publiés, mais quelques documents postés sur le site de Ménestrel témoignent de ces débats.

Le mot portail s'est imposé, dans le domaine francophone, dès lors qu'il s'agit de désigner un site dont les auteurs recensent des ressources électroniques pouvant intéresser les praticiens d'un domaine scientifique. La fortune du mot, autant que les conséquences de son emploi, ou les contours de la notion qui émerge, suscitent cependant des interrogations. En témoignent, tant le texte que les organisateurs de la journée consacrée aux portails médiévistiques nous avaient proposé de commenter [Vivant, 2004], que la présentation de cette même journée [LAMOP, 2005]. Cela appelle un examen de ce terme et de ses usages que nous ne pouvons qu'esquisser ici.

LE PORTAIL EST UN MOT

+++++

Cela passe bien sûr, puisque le portail est d'abord un mot, par la lecture de quelques définitions, que nous proposent dictionnaires et encyclopédies. D'après le concis *Robert pour tous*, le portail est ainsi une : « Grande porte, parfois de caractère monumental ». Nous sont proposés comme exemples « Le porche et le portail d'une cathédrale. Le portail du parc d'un château » et comme synonyme « grille » [*Le Robert pour tous*, 1994].

Le terme, ancien, renvoie donc en son premier sens à une réalité physique. S'en saisir afin de désigner un site internet est en faire un usage métaphorique, puisque nous transposons à une chose un nom qui en désigne une autre, en raison d'un rapport d'analogie. Il nous faut donc, pour comprendre ce transfert, explorer les univers de signification convoqués par ce mot. À lire la notice du Robert, et quelques autres, il apparaît que le terme désigne d'abord un élément architectural qui est le seul ou le principal point d'entrée vers un espace clos. Permettant la surveillance ou le contrôle des entrées, l'avenant portail a comme synonyme la rébarbative grille. Il manifeste, par son aspect monumental, la puissance ou la richesse d'un individu ou d'une institution, tout en permettant de donner à voir à tous images et discours. Il est en somme à la fois manifestation d'un pouvoir, qu'il soit celui de l'église ou du propriétaire, et l'un des lieux et des moyens de son exercice.

QU'EST-CE QU'UN PORTAIL INTERNET ?

+++++

Si nous partons d'une telle définition, alors nous pouvons effectivement repérer des sites qui possèdent plusieurs de ces propriétés. Il faut, pour accéder à certaines des ressources du réseau savant, passer par des sites ou des pages permettant ou non l'accès à celles-ci. Un bon exemple en est fourni par les portails de revues de langue anglaise, dont les opérateurs sont seuls habilités à permettre l'accès à un ensemble fini de ressources, rassemblées par eux, et sur lesquelles ils possèdent des droits, dûment codifiés et enregistrés. Pour ce qui est des sites francophones, et quoiqu'ils ne filtrent pas l'accès aux collections auxquelles ils donnent accès, la page d'accueil de Persée [Persée, s.d.] ou celle de Gallica présentent des propriétés très similaires. Dans bien d'autres cas cependant, l'analogie entre les sites qui se dénomment ou sont nommés portails et la réalité désignée par le sens premier du mot semble extrêmement ténue. Les animateurs de

Ménestrel, ainsi, ne contrôlent ni ne surveillent leurs utilisateurs, et n'en ont d'ailleurs pas la possibilité. De même, ils ne détiennent aucun droit sur les sites qu'ils référencent. Les pages de Ménestrel, de plus, sont loin d'être les seuls points permettant l'accès aux sites références. Enfin, le domaine qu'arpentent les collaborateurs de Ménestrel n'est pas un domaine fini. En l'absence en effet d'une autorité instituante habilitée à tracer une frontière entre ce qui est scientifique et ce qui ne l'est pas, ce qui est médiéval et ce qui ne l'est pas, les contours du domaine cartographié ne peuvent être que flous. Ménestrel n'est pas un portail, mais un guide du cyberroulard médiéviste ou une collection d'atlas à l'intention du cybermédiéviste. Je serais tenté d'ajouter que c'est probablement parce que son aboutissement n'est pas un portail que l'entreprise est si durable et connaît un tel succès.

Dans le domaine du Web ouvert en effet, les sites les plus fréquentés ou les plus référencés, si nous excluons les grands moteurs de recherche, sont des sites que l'on est tenté de définir comme des ponts, ce qu'est par excellence Ménestrel, et non comme des portes, soit des sites qui sont reliés à plusieurs domaines, constitués de sites densément reliés entre eux, et non des sites qui sont liés à toutes les pages d'un ensemble dense [Latapy et Pons, 2005]. Le lecteur pourra trouver la précision oiseuse, et considérer que le terme de portail n'est qu'une étiquette commode permettant de désigner des sites dont les créateurs ont l'ambition d'offrir à une communauté savante un guide fiable et efficace, le fait que l'usage nouveau ait peu à voir avec les usages anciens étant alors de peu d'importance.

QU'IMPLIQUE L'USAGE DU MOT PORTAIL ?

+++++

Il me semble cependant que l'usage des métaphores n'est pas sans effets pratiques, particulièrement lorsque des acteurs sociaux tentent de s'appropriier des objets nouveaux et d'en inventer les usages. Ni les médiévistes ni même les historiens ne sont en effet les seuls à recourir à des métaphores géographiques ou spatiales pour dire l'usage qu'ils font des ressources du réseau. La fortune de l'expression « autoroutes de l'information » est là pour en témoigner, et force est de reconnaître que souvent celles-ci organisent les pratiques et les productions électroniques de leurs usagers. Une observation des sites des centres d'archives départementaux, menée à l'été 2004¹, permet ainsi de repérer l'usage et l'effet de

1. Voir en ce volume « Les sites des archives départementales françaises ».

deux métaphores, celle du livre et celle de l'exposition, inscrites dans les caractéristiques des sites créés et les structurant. Comme leurs référents, ces sites sont construits comme des univers clos et autonomes et il est frappant, par exemple, de constater que beaucoup proposent des guides de recherche similaires, les mêmes présentations des cadres de l'inventaire, sans qu'aucun concepteur de site n'ait semble-t-il songé à offrir un lien aux ressources proposées par l'un de ses homologues.

Je précise que mon propos n'est pas ici d'opposer la naïveté de l'historien perdu dans le cyberspace et la rationalité du réseau et de ses concepteurs. Les ingénieurs réseau rêvent eux aussi et Lucien Sfez a pu montrer à quel point la réalité physique du réseau et l'imaginaire de ses concepteurs étaient liés [Sfez, 2002]. Simplement, historiens et ingénieurs font rarement les mêmes songes, et cela n'est pas sans effets.

Dans le cas des portails, il me semble que les représentations et les métaphores qui armaient leurs concepteurs permettent de comprendre que bien des débats relatifs à leur conception aient été tout à fait impossibles à clore, tel celui opposant les partisans de l'exhaustivité aux tenants du repérage de l'ensemble des sites de référence touchant à un domaine de recherche. Dans les deux cas, cela suppose implicitement que l'on puisse tracer les frontières d'un domaine de recherche ou d'un secteur de la toile, et renvoie à un imaginaire qui est celui du documentaliste ou de l'archiviste inventoriant un fond, alors que les animateurs d'un portail, au sens large dans lequel nous avons aujourd'hui pris le terme, tracent des routes. De même, il est tentant de relier l'impensé de certains sites-portails, conçus comme si les internautes devaient forcément passer par la page d'accueil avant d'emprunter l'arborescence du site, ce que manifeste l'absence sur chaque page d'un module permettant la circulation à l'intérieur du site, au souhait de bâtir une entrée monumentale, plus qu'à une observation des pratiques. Le *deeplinking* comme l'usage des moteurs de recherche conduisent beaucoup d'internautes à aborder un site sans prendre en compte son arborescence.

Nous en concluons que la métaphore du portail obscurcit parfois les enjeux liés à la construction d'un site ressource, soit les questions posées aujourd'hui. Celles-ci touchent à la position des acteurs concernés (créateurs dont la légitimité pose nécessairement problème, mais aussi utilisateurs) et à la nature des opérations réalisées par celui qui choisit et valide une ressource. Ce n'est pas là dire que les métaphores sont trompeuses et qu'il faut éviter de recourir à celles-ci. Elles peuvent se révéler extrêmement efficaces et fécondes. Je serais plus tenté de souhaiter leur

prolifération, moyen aussi de les connaître pour ce qu'elles sont, que de prôner leur éradication. Il est en effet tout aussi possible de comparer la page d'accueil de Ménestrel au portail de la cathédrale de Chartres que de voir le site comme une station spatiale permettant aux astromédiévistes de se ravitailler, ou comme une collection d'atlas, et ce faisant d'imaginer de nouveaux usages et de nouveaux dispositifs, à l'heure où nous commençons à percevoir les limites des listes statiques de ressources et disposons des moyens de donner des formes nouvelles aux sites appartenant aux premières générations de sites ressources.

+++++

CHAPITRE VI. DIFFUSION DES PRODUITS D'UNE RECHERCHE HISTORIQUE À L'HEURE DU WEB 2.0

+++++

Ce texte provient d'une conférence donnée lors du colloque *Contemporary History in the Digital Age* (Luxembourg, 15-16 octobre 2009). Il a ensuite été publié dans les actes de celui-ci [Clavert et Noiret, 2013]. Il est, comme d'autres textes de ce recueil, assimilable à un récit de pratiques. Il est, cependant, indice des changements intervenus dans le domaine numérique durant les années 2000, écrit depuis une position qui n'est pas celle d'un producteur de ressources numériques ou de sites, mais d'une sorte de chargé de mission ou de chef de projet, tentant, sans succès en ce cas, de coordonner l'activité d'acteurs multiples afin de parvenir à produire des ressources électroniques partagées.

Was können wir wollen, ist [...] das eigentliche Centralproblem aller Ethik.

Max Wentscher, *Ethik*, 1902

Je mène depuis assez longtemps, en marge de mon parcours d'historien, pour l'essentiel consacré à l'histoire des migrations, une activité de micro-entrepreneur numérique. En marge veut dire, ici, qu'elle fait pour moi partie de mes activités d'historien et que ses formes ne se comprennent qu'en référence aux objectifs des recherches historiques auxquelles elles sont liées et aux conditions de celles-ci. Cette activité a abouti à quelques réalisations. Ce n'est pas l'une d'elles cependant que je vais évoquer ici, mais une entreprise, ou un projet, qui échoue, ou du moins ne démarre pas et ne semble pas devoir le faire, pour des raisons qui ne tiennent pas d'ailleurs à des difficultés techniques ou à la complexité des dispositifs numériques les plus récents, mais aux conditions sociales institutionnelles et politiques de la conduite de ce type de projet, occasion de réfléchir aussi au contexte de production des outils

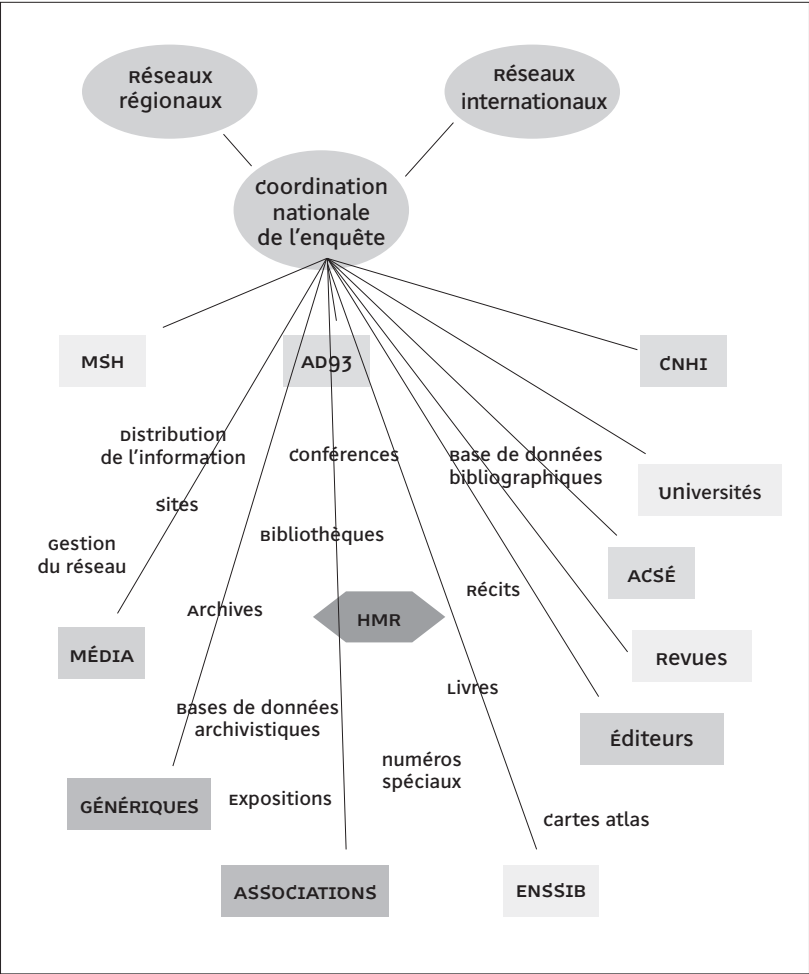
numériques, aussi déterminant quand il s'agit de les comprendre que les seules contraintes techniques.

CONDITIONS DE LA DIFFUSION DES RÉSULTATS
D'UNE ENQUÊTE

+++++

Tout part, dans le cas qui nous occupe, d'une commande passée par un organisme public français, l'Agence pour la cohésion sociale et l'égalité des chances (ACSE), sorte d'agence fédérale à la mode française, chargée de veiller à la cohésion sociale. Cela inclut aujourd'hui la lutte contre les discriminations ou l'illettrisme. Dans un contexte très incertain, puisque ses missions, son périmètre, son organisation sont alors discutés et son existence en danger, l'agence, par l'intermédiaire de sa direction recherche, lance une grande enquête dont le titre est « Histoire et mémoire des migrations en régions ». Les textes officiels présentant l'entreprise font d'elle le moyen d'accumuler un matériau et un savoir nourrissant des manifestations et des productions destinées à lutter contre les discriminations liées à l'origine en agissant sur les représentations de la population. La direction du programme, d'une durée de trois ans, est confiée à une équipe d'historiens dont je fais partie. Celle-ci a pour charge de coordonner le travail de vingt-six équipes régionales, une par région, qui doivent chacune produire un récit de l'histoire de la présence étrangère dans le cadre régional, établir un état des sources disponibles, une recension des travaux consacrés à ce thème et particulièrement de la littérature grise et mettre enfin à disposition des chercheurs futurs un certain nombre de matériaux [Bruno *et al.*, 2008].

Nous avons, au sein de la coordination scientifique de l'enquête, décidé de préparer très en amont la valorisation des résultats de l'enquête et en particulier les dispositifs numériques permettant sa diffusion. Nous avons ainsi demandé aux chercheurs en région de procéder à une saisie aussi complète que possible des résultats des recensements concernant les populations étrangères (de 1851 à aujourd'hui), prôné la constitution d'une base bibliographique indexée permettant d'intégrer les documents produits par les équipes et souhaité que les descriptifs des sources soient le plus possible normalisés, de façon à pouvoir être intégrés facilement aux bases existantes. La répartition des tâches entre les différents acteurs est très tôt décidée et acceptée et un schéma de valorisation de l'enquête émerge (cf. schéma ci-après).



ACSÉ : Agence nationale pour la cohésion sociale et l'égalité des chances

AD 93 : Archives départementales de la Seine-Saint-Denis

CNHI : Cité nationale de l'histoire de l'immigration

Enssib : École nationale supérieure des sciences de l'information et des bibliothèques

HMR : Histoire migration région (Enquête)

MSH : Maison des sciences de l'homme

Celui-ci a la particularité de prévoir une diffusion multivectorielle et multipublic. Il s'agit à la fois de permettre la diffusion de produits de l'activité scientifique réutilisables par des chercheurs, de toucher des publics non experts et d'autoriser une valorisation du travail effectué par les acteurs de l'enquête, d'où des discussions avec des acteurs du monde savant afin d'assurer la publication d'articles scientifiques, ou la réutilisation des matériaux, avec des éditeurs, afin de produire des ouvrages destinés à un public plus vaste, avec aussi un certain nombre d'acteurs associatifs, associés en régions à la diffusion des résultats de l'enquête sous forme de conférences, d'expositions ou de manifestations réutilisant ce matériau.

Une association, constituée à la suite de l'enquête, regroupant des chercheurs ayant ou non participé à l'enquête, avait pour mission de devenir l'interlocuteur des différents acteurs et de veiller particulièrement à la mise en place de ressources numériques assurant la réutilisation possible par le monde de la recherche des matériaux et résultats accumulés. Reliée à des réseaux internationaux consacrés aux mêmes thématiques, disposant de relais régionaux, elle était censée, animée par un petit groupe de jeunes chercheurs ayant accepté d'en prendre la charge, assurer aussi la diffusion des informations pertinentes au sein d'un milieu spécialisé et servir d'intermédiaire entre celui-ci et la demande d'expertise, importante dans ce champ, voire permettre, par le biais par exemple d'outils de type wiki, de produire collectivement des ressources. L'instance nouvelle était dotée dès sa naissance de l'appui des chercheurs faisant autorité dans le champ, d'appuis institutionnels lui permettant de bénéficier de locaux, d'un serveur, et de temps ingénieur (un équivalent mi-temps). Le dispositif paraissait relativement solide, d'autant que la population savante censée nourrir le dispositif était relativement homogène, au point que l'on pouvait supposer la possibilité de définir une ontologie partagée, constituait pour partie un milieu d'interconnaissance, du fait en particulier des modalités du déroulement de l'enquête et travaillait un thème faisant l'objet d'une forte demande d'intervention. Le projet consistait de fait à doter cette communauté scientifique hors-sol, au sens où elle n'est pas directement liée à une institution, d'un espace numérique ouvert et partagé permettant à la fois la circulation de l'information, la structuration du réseau et la production collective de matériaux savants.

Si cependant, deux ans après la fin officielle de l'enquête, un certain nombre de livres, d'articles, de numéros de revues sont parus, si se sont déroulés manifestations officielles, conférences, colloques et expositions,

les dispositifs numériques pensés n'ont pas vu le jour, non plus que le réseau social supposé les utiliser et les nourrir n'a émergé.

LES FACTEURS D'UN ÉCHEC

+++++

Ce qui m'a poussé, en prévision de cette conférence, à me pencher sur le pourquoi de cette non-apparition, outre un léger, quoique légitime, sentiment de frustration, est que cette entreprise avait été pensée très en amont en intégrant les contraintes matérielles techniques et politiques qui souvent condamnent à l'échec ce genre de projet. Celui-ci était doté de matériel, de compétences et d'appuis au sein d'institutions du champ garantissant *a minima* leur bienveillance. Afin d'identifier les obstacles, qui se sont pour certains révélés insurmontables, j'ai tenté d'une part de reconstituer les chaînes de traitement d'un certain nombre de matériaux et mené d'autre part une enquête assez rapide auprès des responsables des équipes régionales associés à l'enquête. Ces entretiens se sont déroulés par mail au cours de l'été 2009. Un courrier adressé à chacun des vingt-six responsables d'équipes régionales demandait quel dispositif de valorisation électronique avait été élaboré en région, et quels étaient souhaits et suggestions, dans la perspective de la mise en ligne d'un serveur destiné à assurer la diffusion de l'ensemble des résultats et matériaux accumulés au cours de l'enquête. Toutes les citations d'entretiens qui figurent dans les lignes ci-dessous proviennent de ces conversations électroniques. Nous avons simplement, du fait du faible nombre de répondants, anonymé leurs propos.

Ces échanges font d'abord apparaître la difficulté pour les équipes régionales à prendre en charge de manière autonome la diffusion électronique des résultats de l'enquête. Les initiatives locales se traduisant souvent par la mise en ligne d'un ou deux papiers au format PDF. Simultanément les attentes sont fortes, mais assez vagues. Interrogé sur ce qui lui paraît souhaitable en la matière, le responsable de l'une des régions répond : « Mise en ligne des rapports, et aussi des résultats des dépouillements. » En effet, l'accès à la source paraît fondamental pour des analyses ultérieures. La demande est de fait souvent celle de la mise à disposition de la donnée brute rassemblée par les chercheurs, plus que d'un traitement de celle-ci. Les formes cependant qui permettraient de réaliser ce vœu sont fréquemment assez mal appréhendées. Un autre responsable

d'enquête répond à la même question : « Ce qui te (vous) semblera faisable, car je ne vois pas trop ce que ça peut recouvrir. »

La modeste culture informatique d'une large portion des chercheurs en sciences humaines permet de comprendre ce type de réponse et se révèle de fait à la source de plusieurs des blocages rencontrés. Une partie des matériaux numériques accumulés se sont ainsi révélés difficiles à exploiter. Certains étaient fournis en des formats impliquant un retraitement coûteux. La généralisation du PDF en mode image, sans que soient toujours conservés les documents maîtres, considérés comme de simples brouillons, se révèle en l'occurrence avoir des effets secondaires assez désastreux. D'autres données, les tableaux de données statistiques par exemple, apparaissent au final trop peu normalisées, leur collecte ayant été confiée à des prestataires extérieurs ou à des étudiants, sans qu'un contrôle de qualité suffisant ait été effectué en cours d'enquête. L'importance de la chose n'apparaissait pas nécessairement ni aux chercheurs ni aux personnes chargées de la saisie et les indications données par la coordination de l'enquête avaient été trop imprécises. Ce n'est pas d'abord ou surtout négligence des uns et des autres, mais ignorance d'une part des contraintes que font peser en termes de format et de structuration des données des procédures automatiques d'importation et de traitement des données et oubli d'autre part de ce que les normes de fabrication d'un document numérique doivent être très précisément détaillées, plus encore quand participent à celle-ci des opérateurs qui ne font pas partie d'une microcommunauté ayant ses habitudes et ses évidences.

À ces problèmes de qualité des matériaux se sont ajoutées de vraies difficultés de communication entre ingénieurs et chercheurs. Elles dérivent là encore de l'absence d'une culture commune permettant aux uns de formuler des demandes autrement que sous forme d'attentes assez vagues et sans prise en compte de l'horizon des possibles, aux autres de proposer des dispositifs adaptés aux besoins et aux pratiques de leurs interlocuteurs. Étant donné les difficultés évoquées précédemment, surgissant lors d'interactions entre chercheurs d'un même champ participant ou non d'une culture de production numérique, cela ne doit guère surprendre. De fait, les projets de production de ressources numériques durables en sciences humaines qui parviennent à leurs fins bénéficient souvent de la présence de passeurs, participant des différentes cultures, qui assurent une fonction de truchement et/ou de médiateurs. La difficulté provient de ce que de tels profils sont encore rares, résultant souvent de hasards biographiques et, qu'en l'espèce, nous avons mis en présence des chercheurs

disposant d'une forte compétence disciplinaire et un compétent informaticien de métier, qui, au sens propre et sans mauvaise volonté des acteurs, avaient toutes les peines du monde à s'entendre et à se comprendre.

À ces facteurs, qui renvoient pour une large part au mode de participation des chercheurs en sciences humaines à la culture numérique, s'en ajoutent d'autres, liés à la structure du champ académique. Dans un contexte d'hyperconcurrence où les postes sont rares, les acteurs académiques sont tentés de maximiser les profits symboliques de leur activité de recherche. Dans le champ des sciences humaines, et particulièrement de l'histoire, demeurent particulièrement valorisées les publications papiers et les œuvres singulières, ce qui rend incertaine la rentabilité d'un investissement dans le travail collectif ou la production de formes savantes (un site, une base de données, un logiciel), non encore validées comme telles. Les plus en lisière de la sphère académique sont de plus contraints d'accepter des tâches assurant immédiatement espèces sonnantes et trébuchantes, ce qui rend difficile leur participation à des opérations de valorisation *ex-post*. Se sont ajoutées à cela, en ce cas particulier, les réticences des équipes et/ou des partenaires éditoriaux à permettre la diffusion sous forme numérique de rapports ou d'éléments réunis lors de la recherche. Il était craint que cela puisse nuire à la viabilité ou à la rentabilité des ouvrages élaborés à partir de ceux-ci.

LE POIDS DU CONTEXTE

+++++

Quelques effets de contexte ont également joué, liés aux transformations de l'agence à l'origine de l'enquête. Profondément transformée dans le cadre de la réforme générale des politiques publiques, elle a perdu une partie de ses effectifs, vu ses missions redéfinies et son organisation modifiée. Alors que le cœur de son activité était, au moment du lancement de l'enquête, l'immigration et l'intégration des populations étrangères, elle tend aujourd'hui à devenir une agence de la cohésion sociale, pour laquelle les problématiques de l'immigration ne sont plus centrales. Sa fonction d'opérateur de recherches est de plus contestée, voire abandonnée. Cela se traduit par la disparition de la direction recherche directement à l'origine de l'enquête, remplacée par une division diagnostic placée sous la responsabilité de la direction communication. Sur le terrain, ces changements se traduisent par la disparition des personnels qui avaient assuré le suivi du travail et la difficulté à en trouver de nouveaux au sein de la structure.

Le directeur de l'enquête pour la région C note ainsi que : « Pour les supports électroniques, je ne sais, mais je suis pessimiste... d'autant que les rapports sont longs. L'ACSE [...] n'arrétant pas de se restructurer, je n'ai plus de nouvelles, M. étant parti. » Ces restructurations rapides, le plan de diffusion des résultats de l'enquête n'ayant pas prévu, faute de possibilité de le faire, de solution de remplacement en cas de retrait de l'un des partenaires, ont posé parfois de délicats problèmes d'accès aux données et aux rapports d'enquête. L'institution les ayant commandées dans le cadre d'un marché public demeure en effet propriétaire de celles-ci. Elle en a de plus été logiquement le premier dépositaire et malheureusement parfois le seul. Dans le contexte de la mutation de l'agence, cela a eu pour effet la perte d'un certain nombre d'informations ou de partie des matériaux réunis, dont de fait aucune collection complète n'existe plus au niveau national. Il a résulté également de cet état de fait une certaine incertitude quant au cadre juridique d'une diffusion électronique des résultats d'une enquête demeurant propriété d'un établissement, qui n'entendait plus y participer et s'en désintéressait.

Le projet a souffert des difficultés habituellement rencontrées par toute entreprise de ce type, en particulier dans le cas français. La faiblesse de la culture numérique des chercheurs en sciences humaines, dont les effets sont multiples, la faible validation accordée par les institutions universitaires à l'élaboration de ressources numériques, la difficulté à trouver des porteurs de projet participant tant d'une culture disciplinaire que d'une culture informatique. Les effets de ces facteurs étaient ici aggravés par l'instabilité d'un contexte institutionnel soumis à des transformations brutales du fait d'évolutions macropolitiques sur lesquelles les acteurs de terrain avaient peu de prise. Il ne suffit pas en somme pour faire société d'un serveur, d'un ingénieur et d'une communauté d'expertise, tant il est vrai que le dispositif numérique exprime autant qu'il facilite des arrangements sociaux complexes et dynamiques qui lient des individus et des institutions hétérogènes, tant par leurs logiques d'action que par les ressources et compétences dont elles disposent. Une autre façon de le dire, ou une conséquence de cela, est que la complexité même de ces arrangements, autant que la diversité des compétences requises, suppose en certains cas à la fois une conduite de projet, que rien d'ailleurs n'empêche d'être démocratique, et une forte division du travail, à rebours de ce que laissent penser nombre de discours associés à Internet, qui exaltent la spontanéité et l'absolue égalité fonctionnelle des acteurs.

Il en découle, je crois, pour les acteurs de ce qui est appelé en ce moment les humanités numériques, quelques impératifs stratégiques. Le premier concerne la formation, celle des futurs chercheurs, dont il est plus que souhaitable qu'ils possèdent les rudiments d'une culture informatique, non pas pour concevoir eux-mêmes leurs outils, mais afin d'être capables de négocier leur élaboration en concertation avec des techniciens et des ingénieurs compétents. La remarque n'est pas nouvelle. Elle demeure d'actualité au regard de la modestie des dispositifs d'enseignement adaptés à ces fins dans les universités françaises. Nous aurons besoin aussi de médiateurs, indispensables à la conduite de projets de ce type et disposant aussi bien d'une réelle culture disciplinaire que de solides compétences informatiques. Des formations susceptibles de fournir de tels profils apparaissent depuis quelques années en France et il faut souhaiter que d'autres émergent. Il faudra un peu de temps cependant pour que leurs produits soient opérationnels. Surtout, l'expérience des dernières décennies le montre, la validation des compétences de ces spécialistes de l'ingénierie en sciences humaines et de leur contribution pose encore problème. Les personnels qui ont assumé ce type de fonctions dans les institutions scientifiques l'ont souvent fait en sus de leurs obligations et occupé les marges du monde savant, recrutés pour des contrats de courte durée dans le cadre de programmes de recherche, ou bien incapables de devenir ni enseignant chercheur ni ingénieur de recherche, parce que trop peu pointus pour être considérés comme des informaticiens et ne produisant pas assez de dispositifs savants respectant les formes canoniques pour obtenir une reconnaissance disciplinaire. Il en est résulté un certain nombre de départs ou, la fatigue aidant, de retraits, ce qui a parfois gêné transmission et capitalisation des expériences et des savoir-faire. La réflexion sur ce point est ouverte et dépasse largement le cadre de ce papier, reste qu'il entend souligner qu'une réflexion sur les humanités numériques et les possibles qu'elles ouvrent ne peut ignorer la question de la formation et du statut de ses agents.

+++++

PARTIE II

RÉFLEXIVITÉS

RÉTICULAIRES

+++++

+++++

CHAPITRE VII. DES ARCHIVES NUMÉRIQUES SANS HISTORIENS ?

+++++

La première version de ce texte est parue en 2005 dans le numéro 79 de la revue *Matériaux pour l'histoire de notre temps*. Celui-ci rendait compte d'un travail collectif, piloté par Françoise Blum et Rossana Vaccaro dans le cadre du Centre d'histoire sociale du XX^e siècle de l'université Paris I. Elles entamaient, avec d'autres, une réflexion sur les archives des mouvements sociaux à l'heure d'Internet. Ce travail allait connaître de nombreux et riches prolongements, dans le cadre du CHSXX, auxquels il est parfois fait allusion dans les pages qui suivent. Si j'ai suivi ceux-ci, pour avoir assuré avec Michel Pigenet la direction de ce centre durant quelques années, j'ai été, dans le contexte de cette enquête, un observateur attentif et un soutien institutionnel, et non un participant actif. L'inscription dans ce cadre a cependant contribué à définir le regard que j'ai pu porter sur les transformations de l'activité réseau des historiens et de leurs institutions.

Je n'ai, en tant qu'historien, jamais utilisé d'archives électroniques. Je suis cependant membre d'une corporation, dont je ne puis me targuer d'être le représentant, nécessairement intéressée – la nécessité étant ici hélas de droit et non de fait – aux transformations des conditions de l'archivage des matériaux qui nourriront les discours historiques futurs.

QUELLES ARCHIVES

+++++

J'ai peu à dire de ces transformations elles-mêmes, sinon que je regrette que les historiens, en tant que corps de métier, aient si peu fait entendre leur voix lors des débats portant sur l'archivage des données numériques. Ils auraient pu pourtant, forts de l'expérience du maniement et de l'interprétation de l'archive et exonérés de la nécessité de songer au coût des entreprises qu'ils proposent, relever certains des présupposés structurant les discussions en cours et les solutions adoptées par les institutions

archivistiques. Trois me retiendront ici, que je dégage du manuel pratique de conservation des archives électroniques édité par la direction des Archives de France [France, 2002]. Ce texte précise que la conservation des documents électroniques ne peut être efficace que si l'archiviste intervient très en amont du processus de conservation des données, idéalement dès leur création. Cette intervention touche à la fois la structuration même des données et la discussion de la pertinence d'un archivage à long terme. Cela s'accompagne nécessairement d'une profonde transformation du métier d'archiviste. Celui-ci est amené à devenir à la fois animateur de projet et interface entre producteurs d'archives, informaticiens et institutions intéressées à l'archivage. J'ajoute que le schéma proposé entend répondre à de très fortes contraintes, à la fois techniques et budgétaires, puisque l'archivage des données numériques est coûteux et complexe et que bien d'autres pays que la France ont adopté des perspectives voisines.

Cela pose cependant quelques problèmes. Le plus important est que les lignes de la curiosité scientifique se déplacent avec le temps, ce qui conduit à la mise en exploitation de gisements de sources dormant parfois depuis des décennies. Un exemple contemporain en est l'intérêt des historiens du social pour les dossiers individuels de toutes sortes [Spire, 2005], alors qu'il a longtemps été d'usage de ne conserver, pour trace de tels fonds, que des récapitulatifs statistiques ou des échantillons, pas toujours suffisants pour une exploitation pertinente. Si, ce qui ne serait pas illogique, une normalisation à l'échelle nationale des pratiques archivistiques se met en place et si l'évolution de la profession amène ses acteurs à multiplier les tâches et à gagner en technicité, tout en les éloignant et en les distinguant des historiens professionnels, ce qui est probable, les futurs historiens pourraient vivement déplorer la perte de certaines sources sans pouvoir sauter de dépôt départemental en dépôt départemental à la recherche de vestiges miraculeusement préservés de fonds détruits presque partout.

Nous pouvons d'ailleurs déjà prévoir la nature de certaines de ces disparitions. Les documents émanant des institutions archivistiques, qu'elles soient françaises ou étrangères, privilégient la conservation des données textuelles et iconiques. Or les machines, particulièrement celles qui sont connectées en réseau, génèrent d'énormes quantités d'informations, susceptibles d'être exploitées par les sociologues et les historiens. Les fichiers *access-log* ainsi, quoiqu'ils ne soient pas codés en langage naturel,

permettent la création d'indicateurs extrêmement puissants, qu'il s'agisse de se pencher sur l'intensité des échanges entre pôles universitaires [Barthelémy, Gondran et Guichard, 2002], ou les représentations et les pratiques cognitives des utilisateurs des ressources réseau [Guichard, 2001].

De même le choix fait, qui découle là encore de nécessités techniques et de contraintes budgétaires, de dissocier, lors de la préservation de nombreuses archives numériques, contenu de l'information et codage/présentation de celle-ci, ne peut qu'inquiéter l'historien, habitué à considérer que l'aspect et la matérialité du document font source, tout en étant souvent des éléments essentiels des procédures d'authentification.

Si à ces remarques nous ajoutons la conscience de la fragilité des supports physiques, les dangers de toute opération de migration des données, pourtant nécessaire en l'état actuel de la technologie, et le coût élevé de l'archivage numérique et de la maintenance de l'information numérique, la seule certitude de l'historien est que le processus d'archivage électronique n'offrira pas à ses successeurs tout ce dont ils aimeraient disposer. Cela n'est scandaleux que pour celui qui oublie que l'archive ne fait toujours que conserver des traces choisies et épargnées des vicissitudes des époques anciennes. Il reste qu'il apparaît souhaitable tant de limiter l'ampleur de ces pertes, que d'en contrôler la nature, et qu'intégrer au processus d'archivage des historiens ès qualités, susceptibles d'évaluer l'intérêt pour la recherche de certaines sources, serait un moyen de le faire. De même est nécessaire la mutualisation des expériences et des ressources d'institutions et de centres de recherche engagés pour leur propre compte dans de telles opérations. J'ai conscience cependant de ce que de telles remarques, étant donné les priorités budgétaires actuelles, sont de l'ordre, émanant d'un historien, du vœu pieux.

QUELS HISTORIENS

+++++

Les responsabilités des historiens actuels sont d'un autre ordre. Elles touchent à la formation des historiens futurs, qui seront appelés à manipuler des archives numériques. Le maniement de celles-ci exigera en effet des compétences spécifiques, du fait, tant du volume d'information qu'il faudra manier, que des conditions de l'accès aux données. Même si

l'historien du futur regrettera sans doute l'imprévoyance des générations antérieures, qui n'auront pas préservé les données qui auraient pu satisfaire sa curiosité, il sera surtout confronté à la pléthore d'informations. Cela posera, et vite, tant la question du repérage de l'information pertinente, que celui de l'exploitation optimale de vastes corpus documentaires, qu'il serait dommage de traiter comme des réserves de citations destinées à illustrer un discours préconstruit. Les outils de la lexicométrie ou des statistiques en permettent, en effet, *a minima*, une description synthétique. Cela revient à dire qu'il faudra souvent, pour pouvoir utiliser ces ressources de façon efficace, tant une connaissance minimale des principes de l'informatique, qu'une compréhension des modes de structuration de l'information caractéristiques des bases de données relationnelles et une culture statistique, fût-elle modeste. Il est possible de plus que se mettent en place des formes inédites de division du travail au sein de la profession historique. Si en effet ces modifications ne déterminent pas à elles seules de transformations institutionnelles, par ce qu'elles exigent pour l'exploitation optimale des ressources l'alliance de compétences nombreuses et diverses, elles dévoilent le caractère éminemment collectif et social de la production scientifique et peuvent inciter à des formes collectives de recherche ou les nécessiter. Cela suppose qu'émergent des personnels capables de mettre en œuvre, au sein d'équipes, de telles compétences. Le paradoxe est que ces besoins apparaissent alors même que l'utilité d'une formation des historiens à ces outils est vivement contestée au sein des universités, pour des raisons qu'il n'est pas dans notre propos de développer ici. Il s'ensuit que les archives électroniques, qui font aujourd'hui l'objet de toute l'attention des documentalistes et des archivistes, pourraient attendre quelque temps leurs historiens, même s'il est fort probable que quelques pionniers autodidactes se lanceront avidement à l'assaut de ces nouvelles redoutes.

QUELLE HISTOIRE

+++++

Il n'est pas certain d'ailleurs, mais nous entrons là dans le domaine de la spéculation, qu'ils écrivent tout à fait la même histoire que leurs devanciers. Malgré la célébration depuis plusieurs décennies de l'histoire problème, les historiens sont encore souvent les experts d'un ou de quelques fonds et nos travaux sont parfois tout autant la mesure et l'exploration d'un ou

de quelques corpus, que nous connaissons bien, que celles d'un problème et d'une question. Il n'y a d'ailleurs dans mon esprit rien de répréhensible à cela, qui s'explique par de pressants impératifs pratiques. La possibilité d'accéder à de multiples banques de données structurées et d'y effectuer des recherches par mot-clé, voire en *full text*, ce qui est aujourd'hui tout à fait impossible, pourrait, au risque d'ailleurs de faire paraître moins vitale la connaissance de l'organisation et de la structure des fonds, transformer l'économie des productions historiennes, en affranchissant celles-ci, dans une certaine mesure, de l'organisation matérielle des dépôts d'archives.

Il est probable également, la tendance est déjà observable du fait de la mise en ligne de ressources bibliographiques, que les appareils documentaires et bibliographiques jugés nécessaires au sérieux d'une recherche tendent à prendre du poids, puisque le temps nécessaire à la collecte de l'information devrait être réduit, dans des proportions d'ailleurs difficiles à estimer. La maîtrise des techniques garantes de ces gains de productivité a en effet elle-même un coût en temps qui est loin d'être négligeable. Si nous ajoutons à cela le fait que la numérisation permet, mais nous quittons ici le domaine archivistique, la production de dispositifs savants complexes et distincts des formes sacro-saintes que sont l'article et le livre (dispositifs multimédias, système de gestion d'information, écriture logicielle), les produits livrés par les futurs historiens du temps présent ou de notre futur immédiat, soit par les utilisateurs de ces archives électroniques émergentes, pourraient être fort éloignés de nos normes actuelles.

C'est là conclure que les enjeux sont d'importance et que les transformations de l'archivage en cours contribuent à définir tant le profil de l'historien de demain, ou des collectifs historiens futurs, que les structures et les formes des productions historiennes. Se transforment simultanément les matériaux et les formes de l'accès aux matériaux des historiens, et les dispositifs d'écriture à leur disposition, soit les dispositifs cognitifs de ces professionnels de l'interprétation.

+++++

CHAPITRE VIII. L'ORDINATEUR, LE RÉSEAU ET L'ÉCRITURE DE L'HISTOIRE

+++++

La première version de ce texte a été présentée lors du colloque *Classification et histoire, de Gabriel Nadeau à Rameau : l'histoire face à l'ordre informatique* organisé à Nanterre par la Bibliothèque de documentation internationale contemporaine (BDIC) les 4 et 5 octobre 2005. Ce colloque rassemblait des professionnels de la documentation scientifique, bien plus tôt que les historiens confrontés aux mutations des pratiques accompagnant le déploiement de l'informatique et des spécialistes des sciences sociales, tous invités à réfléchir aux mutations méthodologiques et épistémologiques induites par l'informatisation des bibliothèques et des centres d'archives. Une première version en est parue dans *Matériaux*, la revue de la BDIC [Rygiel, 2006].

Les organisateurs de la journée *Classification et histoire* nous avaient posé deux questions : que change, pour l'écriture de l'histoire, l'entrée dans l'ordre informatique, défini, en ce contexte, par la présence conjointe de l'ordinateur et du réseau ? Quels changements sont probables dans un futur proche ?

À ces questions, les meilleures réponses que nous pouvions apporter étaient que, dans le premier cas nous n'en savions rien, d'autant que nous n'avions guère compétence que pour ce qui touchait à l'histoire contemporaine et au domaine francophone et, par la grâce d'une collaboration récente, italien. Dans le second, il était fort douteux que nous puissions le déterminer. Nous disposons en effet de très peu d'études empiriques consacrées aux usages que font les historiens des ressources électroniques et de leurs machines. Quand elles existent, il arrive que leurs auteurs concluent que les historiens s'en servent sans doute, mais tendent à gommer les traces de leur utilisation. Une étude américaine, menée de 1997 à 2000 auprès d'historiens en poste dans des universités dont les bibliothèques s'étaient dotées d'importantes ressources et collections électroniques, montrait ainsi que ceux-ci les mentionnaient ou les citaient très rarement

[Graham, 2001]. Et nous ne pouvons guère nous rabattre sur les discours des acteurs. Si les historiens multiplient les textes tenant de l'ego-histoire, ils mentionnent plus volontiers, pressés d'évoquer un parcours intellectuel, la découverte d'un fond magnifique que celle d'un outil logiciel. Les questions posées deviennent donc : que pouvons-nous savoir, et comment, de ce que les historiens font de leurs machines et que pouvons-nous deviner de ce que cela implique pour l'histoire telle qu'elle s'écrit ?

Les données disponibles se réduisent encore – mais ce constat sera vite dépassé, car de nombreux chantiers s'ouvrent [BDIC, 2005, Diminescu, 2005] – à quelques entreprises pionnières, qu'il est possible d'observer, à quelques études et discours consacrés à ces thèmes dont nous pouvons reprendre les conclusions. À ce qu'il est possible de dire à partir de cela, nous pourrions ajouter quelques hypothèses, tirées de ce que nous disent sociologues, géographes ou spécialistes des sciences de l'information, des caractéristiques du monde numérique. Ce matériau nous permettra d'abord de décrire l'histoire rêvée ou tentée par quelques historiens, ce qui se confond ici avec une réflexion sur l'élaboration du discours historique, puis celle qui effectivement se donne à lire par l'intermédiaire du réseau. Nous nous demanderons alors si les hypothèses que nous pourrions formuler suffisent à nourrir une prospective raisonnable et à définir quelques enjeux.

PRODUCTION DE L'HISTOIRE À L'ÂGE ÉLECTRONIQUE

+++++

L'observation des caractéristiques du réseau et des documents numériques a nourri, depuis déjà quelques années, des pronostics assez nombreux pour que nous puissions repérer un accord entre plusieurs auteurs sur les possibles ouverts aux historiens. Ils pourraient ainsi, par la grâce de la numérisation de la documentation et de procédures d'indexation automatisée, se libérer de la structure des fonds archivistiques ou documentaires et, confrontés à un problème historique, constituer, empruntant quelques documents à de multiples fonds, leurs propres dossiers. Le numérique viendrait ici au secours de l'histoire problème et de l'histoire comparée en les rendant enfin matériellement possibles. D'autres insistent, dans la lignée de Darnton [Darnton, 1999], sur le fait que l'article, la monographie, avec sa linéarité, n'est plus la seule forme possible d'écriture de l'histoire, qui pourrait se faire multimédiale, incorporer une multiplicité de parcours possibles, ou devenir polyphonique, sous l'effet de l'hypertexte, de

l'accroissement des capacités de stockage et de l'augmentation constante des capacités de calcul. Certains enfin ont vu, dans l'accessibilité de masses croissantes de données déjà numérisées, l'une des conditions de possibilité d'un renouveau d'une histoire sérielle empruntant, par exemple, aux techniques de la lexicographie et capable désormais de s'aventurer jusqu'aux rivages de l'histoire littéraire [Bernard, 1999].

Force est cependant de constater que si ces trois pronostics, choisis parmi d'autres, car ils touchent aux formes d'écriture et au régime épistémique de la discipline, apparaissent raisonnables au regard des possibilités ouvertes par les machines et leur mise en réseau, nous ne pouvons repérer, au mieux, que de timides prémisses, si nous voulons justifier de leur solidité.

Nous admettons cependant que notre jugement en la matière est tout à fait impressionniste, en l'absence de toute étude empirique, et fortement tributaire de notre connaissance, nécessairement parcellaire, du champ. Il nous semble, à partir de l'observation des territoires connus de nous, que nous pouvons repérer trois types d'innovations dans les pratiques de recherche et d'écriture, dont les caractéristiques ne se raccordent qu'imparfaitement aux prévisions que nous venons d'exposer. Des travaux émergent, dont les auteurs, utilisant les ressources du réseau, parviennent à constituer des dossiers documentaires dont la réunion aurait été beaucoup plus coûteuse et difficile, au point d'être parfois impossible, sans celui-ci. Une équipe d'historiens italiens a pu ainsi mener une étude sur les images et symboles utilisés aujourd'hui par les organisations socialistes en repérant les sites web de celles-ci et en observant les graphismes et symboles qu'ils affichent [Caretto, Degl'Innocenti et Silei, 2002]. Nous trouvons trace également de dispositifs savants nouveaux, sans équivalent papier, tels les atlas interactifs élaborés sous la houlette d'Éric Guichard. *L'Atlas de l'immigration en France dans l'entre-deux-guerres* ainsi, construit à partir des données de la Statistique générale de la France, permet à l'utilisateur de choisir tant les variables qu'il souhaite voir projetées sur la carte, que d'influer sur le mode de discrétisation des données [Guichard, 2005]. Il devient alors possible de générer une multitude de cartes adaptées aux besoins du lecteur ou du chercheur, ce qui revient à créer un outil dont les fonctions ne sont pas reproductibles par un atlas papier, qu'il serait de toute façon impossible de publier.

Certains auteurs enfin utilisent comme sources des corpus électroniques – textes passés sur des listes de discussion, messages déposés sur des forums – qui leur offrent des traces de pratiques et de discours infra-institutionnels, permettant un regard neuf sur le fonctionnement

des organisations militantes, voire la définition de nouveaux objets [Granjon, 2002].

De telles innovations cependant sont encore rares chez les historiens, moins enclins que les sociologues [Chateauraynaud, 2003] ou les géographes [Desbois, 2001] à explorer les possibilités de l'ordre informatique. Il ne semble pas de plus que les historiens se soient encore rués sur l'analyse textuelle, l'exploitation de grandes bases de données, ou l'analyse réseau. À l'inverse, les figures de proue de l'histoire quantitative ont multiplié, alors même que se diffusait la micro-informatique, les constats amers et déploré la désaffection dont étaient victimes les techniques quantitatives [Grenier, 1995, Lepetit, 1989]. En somme, les virtualités ne se sont pas toutes réalisées et nous ne pouvons non plus nous appuyer en toute confiance sur l'observation de ce qui se produit dans d'autres espaces pour identifier les développements futurs. Un exemple permettra de préciser ce point. Pierre-Yves Saunier est l'auteur de l'une des rares enquêtes consacrées aux pratiques informatiques des historiens dans le contexte français. Il s'est plus particulièrement intéressé à l'usage des listes électroniques et a montré que les historiens français étaient peu présents sur les listes internationales et que les listes professionnelles en langue française étaient rares. Quand celles-ci existent, elles permettent la diffusion de messages qui diffèrent, par leur teneur, de ceux que retiennent les animateurs des listes de langue anglaise. Les débats sont quasi inexistantes et le contenu des messages proposés par les listes françaises est assez strictement informatif. Pierre-Yves Saunier rapporte cela à l'organisation du champ académique français et à l'habitus national des professionnels de l'histoire [Saunier, 2005]. Ce n'est pas cependant son explication qui nous retiendra ici, mais le simple constat que les pratiques professionnelles incorporant l'usage de l'informatique diffèrent de pays à pays, et que leurs propriétés peuvent être rapportées à des facteurs sociaux et culturels. Cela nous invite à passer de la définition de virtualités à l'étude, en contexte, de pratiques, du moins à l'observation du produit de certaines d'entre elles : l'histoire, telle qu'elle est écrite pour le réseau.

HISTOIRES ÉLECTRONIQUES

+++++

Nous mènerons cette exploration à partir du cas italien. Celui-ci a en effet été très précisément décrit récemment [Cricione *et al.*, 2004] et il semble que nous puissions étendre les principales conclusions de cette

étude à la France, voire à l'Espagne [Martin, 2005]. Le premier constat des observateurs de la toile italienne est la quasi-absence de l'histoire savante sur celle-ci. Les historiens de métier laissent peu de traces sur la toile. Quand ils le font, c'est généralement parce qu'ils y déposent des inventaires de ressources hors ligne, des répertoires bibliographiques ou des informations. Les dispositifs savants, sous forme de texte ou de réalisations multimédiales ou logicielles, sont rares. Cela ne signifie pas qu'il soit impossible de trouver des récits historiques sur la toile italienne, mais quand ils existent, ils appartiennent à un tout autre registre. De multiples sites émanant de groupes identitaires, de mouvances politiques, voire de familles ou d'individus proposent sur la toile leur version de l'histoire contemporaine. Nous trouvons de ce fait plus facilement sur la toile italienne des mémoires électroniques que des histoires électroniques, ce qui fait d'ailleurs de celle-ci une source extraordinaire pour qui s'intéresse aux usages publics de l'histoire. Serge Noiret, qui l'a utilisée dans cette perspective, y a découvert de multiples romans électroniques de la nation et a pu ainsi mettre en évidence l'extraordinaire prégnance des catégories et des cadres de l'histoire nationale italienne. C'est elle en effet, pour ces groupes, qu'il s'agit de réécrire, en y incorporant le groupe que l'on défend, ou en revalorisant la participation de celui-ci [Noiret, 2005].

Pour ce que nous pouvons en voir donc, l'écho de la révolution informatique est singulièrement assourdi dans le monde de l'histoire savante. Nous ne pouvons pas en tout cas repérer de bouleversements par l'examen de la production électronique des historiens, ou les histoires électroniques disponibles. Cela ne signifie pas que les conditions concrètes de production du savoir historique n'ont pas été affectées. De rapides échanges entre collègues suffisent à se convaincre que ceux-ci connaissent et utilisent les portails donnant accès à des collections d'articles ou qu'ils utilisent abondamment le mail. Des transformations majeures sont de plus possibles ou probables. Nous disposons d'assez d'indices pour conclure que le champ ouvert aux professionnels de l'histoire est vaste. Cependant, les formes possibles de ces changements se laissent malaisément deviner, parce que les prémisses en sont timides, parce que, aussi, plusieurs facteurs rendent en ce domaine prévisions et pronostics hasardeux.

L'INNOCENCE DES HISTOIRES FUTURES

Le premier est qu'usage des ressources électroniques et productions électroniques sont fonction de très nombreux et très divers paramètres. Nous prendrons deux exemples pour illustrer ce point. Les dispositifs électroniques élaborés par les historiens sont puissamment structurés par les imaginaires de leurs créateurs et en particulier par les métaphores dont ils usent lorsqu'ils tentent de se les approprier. Les créateurs des sites des archives départementales françaises, ainsi, ont fréquemment pensé leurs sites comme des équivalents de livres ou de brochures, voire d'expositions, c'est-à-dire comme des objets clos et autonomes, et la structure même des sites en porte témoignage¹. Fréquemment redondants, ils offrent souvent aux lecteurs des guides de recherche portant sur de mêmes objets, et reproduisent à de multiples exemplaires des feuilles présentant les grands cadres de l'inventaire. Il est rare de plus que ces sites offrent des liens externes, alors même qu'il aurait été plus simple de poser un lien vers un guide de recherche élaboré par les collègues d'un autre département que d'en élaborer un nouveau. De même, les dispositifs interactifs sont extrêmement rares. Nous pourrions aussi montrer que la métaphore de la bibliothèque est très prégnante dans le Web savant, et qu'en particulier les créateurs de sites portails, tel Ménestrel, proposent des guides de ressources qui ressemblent beaucoup à des catalogues de bibliothèques.

Ce n'est pas là une spécificité des historiens. Les géographes ont noté la fréquence des métaphores géographiques et spatiales dans les représentations spontanées des activités liées à Internet, manifestation locale d'une tendance plus universelle qui conduit certains auteurs à conclure que : « la pensée contemporaine [...] semble caractérisée par la multiplication des références spatiales ou spatialisantes » [Benoist et Merlini, 2001]. Le terme portail lui-même en est un exemple, qui désigne, par analogie, un site donnant accès à un domaine thématiquement spécifié. L'usage de ces métaphores, et donc l'imaginaire des producteurs de ressources ou des utilisateurs du réseau, structure les pratiques, alors même que les analogies sur lesquelles celles-ci reposent sont parfois ténues. Il est ainsi, pour reprendre l'exemple du portail, difficile, lorsque l'on observe la topologie du réseau ouvert à partir de la structure des liens hypertextes, de repérer des sites qui seraient des portails, c'est-à-dire qui constitueraient des points d'entrée privilégiés vers des ensembles de sites densément

1. Voir le chapitre consacré aux sites des archives départementales françaises dans ce volume.

reliés entre eux. Les sites polarisant de très nombreux liens sont plutôt des sites ponts, ou passerelles, établissant des connexions entre des ensembles denses [Latapy et Pons, 2005]. L'appréhension du réseau et son usage se font donc à travers des catégories qui renvoient à des héritages culturels ou professionnels et ne peuvent donc être compris en se référant seulement à ses caractéristiques physiques. De ce fait, toute prédiction suppose que soient pris en compte non seulement les caractéristiques, d'ailleurs rapidement changeantes, des outils, mais aussi les imaginaires des acteurs et la probabilité de leur transformation.

Ce n'est là mettre en évidence qu'un déterminant parmi bien d'autres, dont l'inventaire ne saurait être exhaustif. La formation initiale et continue des opérateurs joue elle aussi bien sûr un rôle. On sait, dans le cas français, la grande et persistante faiblesse de la formation des historiens aux outils mathématiques et à la maîtrise des dispositifs techniques, déplorée depuis longtemps par les spécialistes de l'histoire quantitative [Genet, 1994, Genet, 2002]. De même, les procédures de déroulement des carrières d'historien et les cadres disciplinaires en vigueur favorisent peu, tant le travail d'équipe sans assignation possible de la production d'une œuvre ou d'un texte à un individu, que les entreprises collectives associant des chercheurs venant de champs divers. Or, nous savons d'expérience que les productions électroniques réussies supposent généralement que ces deux conditions soient réunies.

Enfin, tant les conditions macroéconomiques, qui déterminent par exemple le coût relatif de l'édition papier et de l'édition électronique, que les transformations de la demande de produits culturels ou l'évolution du cadre juridique, qui définit le régime des droits de reproduction et de diffusion des produits électroniques, sans même parler des impulsions politiques, qui peuvent se révéler décisives, sont autant de facteurs, partiellement indépendants les uns des autres, qui contribuent à délimiter le cadre des usages et des productions électroniques et ce faisant les caractères de ceux-ci. À l'échelle d'une profession, ou d'une discipline, et parce que la production scientifique est le produit d'une activité sociale, les transformations des propriétés des acteurs et celles des produits de leur activité apparaissent soumises à des chaînes de causalité multiples et en partie indépendantes et donc, sinon imprévisibles, du moins non déductibles des caractéristiques actuelles des machines, des outils logiciels ou des dispositifs existants.

Nous voudrions, en guise de conclusion, préciser ce point à l'aide d'un exemple. Pour demeurer au plus près des thématiques évoquées en

ce volume, nous évoquerons les classifications auxquelles l'historien du futur, en quête de sources ou de références, sera confronté. Nous ne partageons pas la conviction, que plusieurs ont manifestée lors des débats que nous avons tenus au cours de ces journées, que l'irruption de l'informatique et des réseaux dans les domaines de la documentation et de l'archivistique conduisait, du fait des caractéristiques de l'informatique, à l'invention d'un nouveau mode de classement dont les particularités pourraient être rapportées à celles de l'ordre numérique. En particulier, nous refusons l'idée, plusieurs fois exprimée, que la diffusion des outils informatiques conduit au triomphe de la logique binaire dans les procédures de classement et d'indexation. Nous croyons au contraire, et l'histoire de Rameau est là pour renforcer cette conviction, que l'usage de la logique binaire préexiste à l'informatique documentaire. Un enregistrement tiré d'un fichier thématique de bibliothèque est en effet déjà une suite de 0 et de 1. Porter un ou des mots-clés, un ou des mots vedettes sur une fiche revient à indiquer, pour chaque terme, l'inclusion ou la non-inclusion de l'ouvrage référencé dans la liste des documents pour lesquels le classeur établit un lien avec le terme. Plus même, les progrès des capacités de calcul, de stockage et des outils statistiques ouvrent la possibilité matérielle de s'affranchir d'une telle logique. Les outils existent qui permettent de remplacer l'inclusion ou la non-inclusion d'un ouvrage dans une rubrique par la probabilité que les praticiens du domaine jugent ce classement adéquat, obtenu en interrogeant une population témoin. Cela revient à remplacer, en termes statistiques, une variable binaire par une variable floue [Kaufmann, 1973]. Nous pouvons même envisager, parce que le temps des transactions client/serveur est aujourd'hui très petit, de faire remplir à chaque interrogateur de la base une requête précisant sa recherche, mais aussi ses propriétés institutionnelles et personnelles (âge, genre, champ disciplinaire, appartenance académique). Il est ensuite assez simple, en utilisant la technologie très rodée du *scoring*, dont le secteur bancaire ou les spécialistes du marketing font depuis longtemps abondamment usage, de créer un algorithme qui va lui retourner une liste de réponses classées en fonction de la probabilité de leur pertinence, calculée à partir des usages de la population témoin et de ses propriétés. Il est même possible d'introduire un mécanisme de *feedback* en demandant au requérant d'apprécier la pertinence des réponses proposées. Nous obtenons un classement qui est à la fois probabiliste et dynamique et qui permet de produire des réponses personnalisées. Un tel système est à la fois pensable et techniquement possible. Il n'est pas certain cependant

que nos bibliothèques universitaires ne proposent jamais de tels systèmes. Leur adoption suppose des moyens, des compétences et un accord, délicat, sur la fonction de l'institution et les rôles respectifs des acteurs engagés dans la chaîne de production du savoir.

De ce fait, si l'incidence des transformations des dispositifs techniques sur l'écriture de l'histoire est une véritable question, et un beau sujet, il n'est, pour les dernières décennies, pas traité et ne peut, dans une perspective prospective, qu'amener à la formulation de propositions indécidables. Pour pasticher Leibniz, puisque nous l'avons beaucoup évoqué au cours de ces journées, l'ère numérique ouvre aux historiens une infinité définie de possibles dont le produit est incalculable.

+++++

CHAPITRE IX. L'ENQUÊTE HISTORIQUE À L'HEURE NUMÉRIQUE

+++++

Une première version de ce texte est parue dans la revue italienne *Memoria e Ricerca* (2010, n° 35, pp. 185-197), une seconde version, plus courte, dans la *Revue d'histoire moderne et contemporaine* (2011/5, n° 58/4 bis, pp. 30-40). Nous reprenons ici l'essentiel de la version italienne, inédite en français à ce jour. Ce texte reprend la teneur de conférences données au cours des années précédentes lors de formations destinées à des étudiants en histoire ou de jeunes chercheurs n'étudiant pas forcément l'histoire contemporaine, en particulier lors d'écoles d'été thématiques du CNRS organisées avec la collaboration du LAMOP et du PIREH, dont celle intitulée « Histoire et méthodes quantitatives » tenue à Roscoff en septembre 2009.

L'HISTOIRE

+++++

Quiconque a dû un jour enseigner à des étudiants ce qu'était l'enquête historique, tenter de leur fournir quelques règles leur permettant de mener à bien leurs premiers travaux, sait que la chose est difficile et qu'aucune description simple de celle-ci n'est disponible. Souhaitant y parvenir, et c'est peut-être déjà là un réflexe d'historien, nous sommes tentés de nous tourner vers ceux qui passent pour les fondateurs de la discipline, ces Grecs, et leurs continuateurs romains qui pratiquent une *historia* dont nous savons que le nom même fait d'elle une enquête. Malheureusement, Hérodote ou Thucydide ne nous ont pas laissé de traité nous expliquant la façon dont celui qui veut mener une telle entreprise doit procéder. Il faut donc se rabattre sur un continuateur romain, Lucien de Samosate, dont de bons manuels [Cadiou *et al.*, 2005] nous disent qu'il est l'auteur du seul ouvrage

d'historiographie que l'Antiquité nous ait laissé. Son texte¹, rédigé vers 165 av. J.-C., est tout à fait adapté à notre propos. Il dresse d'abord la liste de tout ce que l'historien doit éviter de faire, avant de nous présenter l'historien idéal, de nous décrire le *modus operandi* qu'il doit suivre et de proposer quelques préceptes devant guider la rédaction de l'œuvre rendant compte de son parcours. Nous croyons d'abord, à le lire, entendre un universitaire chenu pester contre les amateurs et tous ceux qui pensent qu'ils peuvent écrire l'histoire sans se donner la peine d'apprendre les règles du métier :

« Cependant la plupart de nos gens croient n'avoir pas plus besoin de conseils pour leur entreprise qu'il ne faut d'industrie pour marcher, voir ou manger. Ils s'imaginent qu'écrire l'histoire est une chose fort aisée, à la portée de tous ceux qui peuvent exprimer clairement ce qui leur vient à l'esprit. » [Samosate, 1912, paragr. 37, p. 355]

Écrivant cela cependant, Lucien n'entend absolument pas dénoncer les amateurs et les autodidactes. Si tout le monde en effet ne peut pas écrire l'histoire, c'est parce que celle-ci demande des aptitudes particulières, nées de ce que nous nommerions aujourd'hui un parcours de vie spécifique. Il veut, pour en faire un bon historien, qu'on lui donne :

« [...] un élève tel que je le demande, prompt à concevoir et habile à s'exprimer, d'une vue pénétrante, capable de diriger les affaires, si on les lui confie, ayant l'esprit militaire, mais avec la science civile, et sachant par expérience ce que c'est que conduire une armée ; je veux, par Jupiter ! qu'il ait été dans les camps, qu'il ait vu les évolutions et les mouvements des troupes, qu'il connaisse les armes et les machines de guerre, ce que c'est qu'une aile, un front, des bataillons, des escadrons, comment ils se forment, ce qu'on entend par charge, par volte ; en un mot, je ne veux pas d'un homme qui ne soit jamais sorti de chez lui et qui s'en rapporte au témoignage des autres. » [Samosate, 1912, paragr. 37, p. 371]

1. Les citations de ce texte sont tirées de l'édition française de *Comment il faut écrire l'histoire* (165 av. J.-C. environ), de Lucien de Samosate in (*Œuvres* [texte établi et traduit en français par Jacques Bompain], Paris, Les Belles Lettres, 2003 (coll. des Universités de France), 3 t. Pour la commodité du lecteur, je renvoie cependant à l'édition en deux volumes des *Œuvres complètes de Lucien de Samosate* [traduction française nouvelle avec une introduction et des notes par Eugène Talbot], Paris, Hachette, 1912. [En ligne] : < gallica.bnf.fr >.

On ne peut donc être historien que si l'on a été d'abord soldat, général, haut fonctionnaire, voire ingénieur ou urbaniste et tout cela, même dans l'Antiquité, prend un certain temps. L'historien n'est donc pas un professionnel de l'histoire, qui en fait métier, mais, dirions-nous aujourd'hui, un homme d'État qui a une longue carrière derrière lui. Et puis, il lui faut une belle âme, réunir un ensemble rare de vertus :

« Ainsi l'historien doit être exempt de crainte, incorruptible, indépendant, ami de la franchise et de la vérité, appelant, comme dit le Comique, figue une figue, barque une barque ; ne donnant rien à la haine, ni à l'amitié, n'épargnant personne par pitié, par honte ou par respect, juge impartial, bienveillant pour tous, n'accordant à chacun que ce qui lui est dû, étranger dans ses ouvrages, sans pays, sans lois, sans prince, ne s'inquiétant pas de ce que dira tel ou tel mais racontant ce qui s'est fait. » [Samosate, 1912, paragr. 41, p. 373]

Je ne suis pas un spécialiste de l'histoire antique, et j'ai peur de commettre un anachronisme ou une bétise, mais il me semble que ce qui garantit ici la vérité de l'histoire, ce n'est pas le respect d'un ensemble de règles ou de procédures, mais l'expérience et les vertus du locuteur. Et cet homme d'élite ne peut écrire l'histoire que s'il en a été le témoin, ou du moins s'il a connu et rencontré des témoins dignes de foi, dont il examinera scrupuleusement les propos :

« Les faits ne doivent pas, non plus, être cousus au hasard, mais soumis à un examen laborieux et souvent pénible, à une critique sévère, l'auteur les aura vus, il en aura été le témoin ; sinon, il ne se fierait qu'à des gens qui racontent avec une fidélité incorruptible, et que l'on ne saurait soupçonner d'ajouter ou de retrancher rien aux événements, par faveur ou par haine. Pour cela, l'auteur doit avoir un discernement juste, et n'admettre dans son récit que les faits les plus probables. » [Samosate, 1912, paragr. 47, p. 375]

Si nous comparons aujourd'hui assez régulièrement l'historien à un juge, nous sommes ici plus proches du journaliste d'investigation, ou de l'inspecteur de police. L'historien, selon Lucien, mène une investigation relative à des faits récents dont la matière est constituée de témoignages examinés avec soin. Alors seulement il pourra en rendre compte, c'est-à-dire faire le récit de ce qui s'est passé, sans fioritures inutiles, même s'il

lui faut rédiger en style soutenu les discours de nobles personnages dont nous n'aurions pas conservé la teneur :

« Après un exorde, long ou bref, proportionné aux événements, il faut que la transition qui passe aux faits mêmes soit ménagée et conduite avec art ; et tout le reste du corps historique n'étant plus qu'un long récit. » [Samosate, 1912, paragr. 55]

« [...] Si quelquefois on est obligé de faire parler des personnages, il faut qu'ils tiennent des discours appropriés à leur caractère et aux événements, et que d'ailleurs ils s'expriment avec la plus grande clarté. Du reste, il vous est permis, en ce cas, de montrer votre talent dans l'art de bien dire et de déployer votre éloquence. » [Samosate, 1912, paragr. 58]

Il est assez facile de montrer que l'histoire ainsi conçue a peu à voir avec ce que les modernes, et plus encore nos contemporains, appellent du même nom. Il suffit pour cela de mettre ces courts extraits en regard du discours de la méthode proposé par Langlois et Seignobos², par exemple :

« Les faits passés ne nous sont connus que par les traces qui en ont été conservées. Ces traces que l'on appelle documents, l'historien les observe directement il est vrai ; mais après cela, il n'a plus rien à observer ; il procède désormais par voie de raisonnement, pour essayer de conclure aussi correctement que possible, des traces aux faits. Le document c'est le point départ ; le fait passé c'est le point d'arrivée. Entre ce point de départ et ce point d'arrivée, il faut traverser une série complexe de raisonnements enchaînés les uns aux autres, où les chances d'erreurs sont innombrables. » [Seignobos, Langlois, 1992]

Ce texte, dont l'influence fut durable, et dont nous sommes au fond encore si proches, peut être opposé, presque terme à terme, à celui de Lucien. D'un côté, nous trouvons le grand capitaine, de l'autre l'homme de cabinet, la belle âme contre le scientifique armé d'une méthode sûre, le témoin face au collectionneur de documents, et nous pourrions continuer

2. La première édition de cet ouvrage date de 1898 et servit longtemps, selon les termes de Madeleine Rébérioux dans la préface à l'édition de 1992, de bible aux étudiants d'histoire des universités françaises (p. 7).

à loisir. L'incarnation contemporaine de l'historien selon Lucien, ce sont Churchill ou de Gaulle rendant compte, au soir de leur vie, passions assoupies, à l'ombre de la mort proche, des événements dont ils furent les témoins directs ou indirects. Ce n'est pas le rat d'archives noircissant sans fin des fiches afin de rendre compte des vies infâmes des galériens ou examinant des questions triviales, touchant à la condition des femmes ou des esclaves dans la Rome antique.

DES VERSIONS PLUS MODERNES

+++++

C'est là dire que ni l'histoire ni l'enquête historique ne sont des essences qu'il s'agirait de retrouver, ou bien si ténues, si pauvres de déterminations, que leur valeur pratique serait bien faible. Pas d'histoire ou d'enquête en soi, donc. Du moins, l'histoire d'aujourd'hui n'est pas celle d'avant-hier.

En lecteurs un peu patauds de Foucault, nous sommes tentés de nous consoler en songeant que nous pouvons échanger une essence contre des systèmes de pratiques relativement stables, qu'irrigue un horizon discursif. Mais l'étrangeté même du texte de Lucien nous arme d'un doute. Lucien est l'équivalent romain d'un intellectuel brillant et connu, proche du pouvoir. Il parcourt le monde romain et y donne des conférences grassement rémunérées. Il est de plus proche, lorsqu'il écrit ce texte, des cercles impériaux. Enfin, il n'a jamais écrit de texte qui appartienne au genre de l'histoire, passant plutôt, pour la postérité, comme l'un des lointains précurseurs de la science-fiction. Imaginons, si nous voulons un exemple contemporain, un philosophe connu expliquant aux historiens comment ils devront écrire l'histoire de la guerre du Kosovo. Le texte de Lucien donc est un texte normatif, écrit justement parce que ceux qui se mêlent alors d'écrire l'histoire de la guerre contre les Parthes, auteurs dont nous ne connaissons souvent que ce que Lucien nous dit d'eux, ne sont pas ce que devraient être des historiens et ne font pas ce que l'historien est censé faire. Il n'y a, même à Rome, pas une seule façon de mener l'enquête, mais plusieurs, pensées, par Lucien du moins, comme mutuellement exclusives.

Nous pouvons alors revenir vers le contemporain armés de deux soupçons. L'un selon lequel un écrit sur l'histoire pourrait souvent être un texte prescrivant ce qu'il convient de faire et d'écrire, volontiers polémique et n'hésitant pas à disqualifier les pratiques ou les traits attachés, à tort ou à raison, aux adversaires ou aux devanciers. Il n'est pas certain alors que sa lecture nous renseigne beaucoup sur ce que fait l'historien et la façon

dont il mène l'enquête. Le second soupçon est qu'il est bien possible que coexistent des histoires tout à fait différentes non pas seulement par leurs thèmes ou leurs thèses, mais aussi par la façon dont elles sont produites et les enquêtes menées. Quelques récits de pratiques, insérés dans les textes d'historiens presque contemporains, permettent de le vérifier. Nous en confrontons ci-dessous quelques-uns.

Écouter les gens

« Que peut être une histoire faite par les gens d'en bas, en fonction de leurs besoins propres ? Une histoire qui ne donnerait aux professionnels qu'un rôle d'auxiliaire, non de dépositaires privilégiés ? Il y a déjà des tentatives et des expériences en ce sens. Aux États-Unis, on a parlé de *guerilla history*, au sens de *guerilla theater* ; une étude du passé qui soit souple, mobile, calquée sur les besoins concrets, libre des obligations rigides de la science conventionnelle et dont la rédaction autant que la lecture soit à la portée de tous. En Chine, vers 1960-1965, on appelait à pratiquer les "quatre histoires" : histoire des villages, des communes, des familles, des usines. Les expériences, les souvenirs et traditions, les matériaux écrits locaux étaient mis en commun à la base, en vue de faire le bilan du passé récent et de mesurer les enjeux du socialisme. Le passé populaire est un instrument de lutte populaire. » [Chesneaux, 1976, p. 146]

Écrire l'archive

« Lire l'archive est une chose ; trouver le moyen de la retenir en est une autre. On peut surprendre en affirmant que les heures passées en bibliothèque à consulter l'archive sont autant d'heures passées à la recopier, sans en changer un mot. » [...] « Devant l'archive manuscrite une urgence se crée, celle de se couler par le geste dans le flot saccadé des phrases, dans le débit heurté des demandes et des réponses, dans l'anarchie des mots. Se couler, mais aussi se laisser dérouter, entre familiarité et dépaysement. Le goût de l'archive passe par ce geste artisan, lent, peu rentable, où l'on recopie les textes, morceaux après morceaux, sans en transformer ni la forme,

ni l'orthographe, ni même la ponctuation. Sans trop même y penser. En y pensant continûment. Comme si la main, en reproduisant à sa façon le moulé des syllabes et des mots d'autrefois, en conservant la syntaxe du siècle passé, s'introduisait dans le temps avec plus d'audace qu'au moyen de notes réfléchies où l'intelligence aurait trié par avance ce qui lui semble indispensable et laissé de côté le surplus de l'archive [...]. L'archive recopiée à la main, sur une page blanche, est un morceau de temps apprivoisé ; plus tard, on découpera des thèmes, on formulera des interprétations. » [Farge, 1989, p. 25]

Modéliser et tester

« Le passé ne se dévoile jamais spontanément et la construction des faits historiques est inséparable des questions que les problématiques du présent suggèrent [...]. Objet construit, le modèle, par la formalisation qu'il nécessite, oblige l'historien à une plus grande rigueur dans l'expression des hypothèses interprétatives qui le sous-tendent. Objet artificiel, il est contrôlable, en sorte que l'opacité attribuable au réel en est absente et que le jeu raisonné sur les variables qu'il associe peut tenir lieu d'une expérimentation autrement improbable. Exposant des conséquences théoriques déduites aux démentis de la matière historique, il permet de rejeter les hypothèses infondées qui ont présidé à sa construction et d'enrichir les problématiques qui justifient son usage. Aboutissant à l'expression logico-mathématique d'un phénomène fonctionnel, la modélisation conduit à interroger l'usage que la discipline fait de la généralisation causale ou des catégories du vrai et du faux. » [Lepetit, 1988, pp. 3-4]

Sisyphes et les fiches

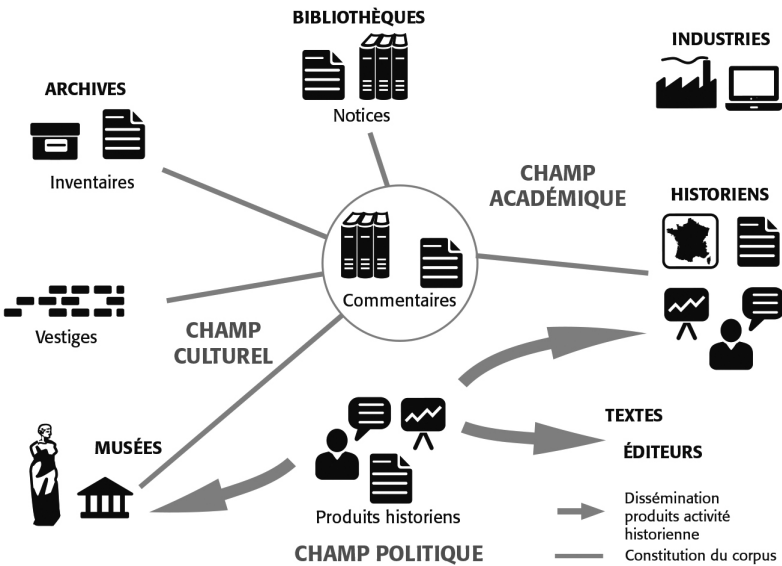
« Cela pour la rage avec laquelle nous avons opéré, copiant, lisant, aux archives, aux bibliothèques, empruntant, photocopiant – afin de prolonger l'été les lectures d'archives – par quatorze heures par jour de lectures de microfilms au péril des yeux, brûlant deux ou trois lampes par semaine sur un lecteur portatif de l'âge de pierre [...]. Il fallait essayer de comprendre et, pour cela, obtenir des

statistiques fiables. À partir de cette grille, construire une approche statistique des échanges et des mouvements [...]. Cette construction empirique que forment les sept tomes et les huit volumes de la partie statistique de Séville et l'Atlantique, trois mille huit cent quatre-vingt-dix pages grand *in-octavo*, en petits caractères très denses publiés sous la double signature d'Huguette et Pierre par l'EPHE de 1956 à 1957. Ma thèse [n'en est] que le commentaire un peu long (3 400 pages grand *in-octavo*). »[Chaunu, 1987, pp. 81-82]

Les historiens retenus ici utilisent des matériaux de nature tout à fait différente. La prise de possession de l'archive ne se fait pas de la même façon. Pierre Chaunu, d'emblée, arrache au document la donnée qui nourrira le tableau. Arlette Farge reproduit le document choisi, attentive à son unité. Les référents épistémologiques, ou du moins les critères de qualité implicitement mis en avant par les uns et les autres (l'exhaustivité du dépouillement, le lien au besoin de mémoire, la rigueur des procédures, le respect résolu du contexte), ne sont pas de même nature. Nous pouvons supposer, toute méfiance préservée envers les textes normatifs, que ces historiens ne fréquentent pas les mêmes lieux, n'ont pas les mêmes interlocuteurs et n'accomplissent pas les mêmes tâches.

Quand l'un se penche sur un manuel de statistique, l'autre écoute les paysans chinois, cependant qu'Arlette Farge se hâte vers les archives, soucieuse d'y obtenir la meilleure place. Comment alors parler d'enquête historique au singulier ? Il semble qu'il faille pour cela sortir du cadre disciplinaire, ou du moins que la chose soit plus aisée ce faisant, en empruntant par exemple librement à quelques disciplines voisines. Si nous chaussons les lunettes qu'elles nous fournissent, et nous appuyons sur les productions matérielles des historiens lorsqu'ils exercent leur activité d'historien, nous voyons des acteurs, se nommant les uns les autres historiens, qui produisent une multitude structurée de signes dont la très grande majorité n'est portée à la connaissance de personne et ces signes entretiennent des rapports complexes avec plusieurs réservoirs de signes et de symboles. Un schéma éclairera, je l'espère, ce point.

Schéma. Ceci est un historien



L'HISTORIEN CONTEMPORAIN

L'historien contemporain, si nous suivons ce schéma, apparaît d'abord comme un polygraphe hypertextuel dissimulant aux regards l'essentiel des inscriptions qu'il produit. Consultant les documents conservés par les centres d'archives, les bibliothèques, les institutions muséales, ou bien ceux qu'il a lui-même rassemblés, parce qu'il peut être lui-même créateur d'archives. Il accompagne, ou est censé le faire, ses inscriptions de l'indication du chemin d'accès au document consulté, qui peut être un vestige du passé qu'il cherche à saisir ou bien un commentaire ultérieur : extrait d'ouvrage et d'article, notice de catalogue, voire notes d'un érudit. Ce document renvoie généralement lui-même, directement ou de façon médiate, à un vestige authentifié, par l'historien ou par d'autres, d'une époque passée que l'historien tend à rapporter au référent de son discours.

Si nous voulons englober la plus grande diversité possible de pratiques, il nous faut considérer que les annotations produites par l'historien ne prennent pas forcément la forme de textes ou de fragments de texte en langage naturel. Elles peuvent aussi consister en graphiques,

enregistrements vocaux, dessins, chaînes de caractères, parfois codées et insérées dans des bases de données informatisées. Il est probable cependant, même si nous savons très peu de choses de la façon dont les historiens de fait travaillent³, que les données non textuelles qu'ils accumulent sont décrites et indexées au moyen de dispositifs textuels (listes, tableaux, arbres) plus ou moins complexes. Ainsi l'historien se trouve à la tête d'une masse d'annotations dont beaucoup sont des données textuelles ou quasi textuelles qu'il va structurer et manipuler, produisant listes et tableaux. Il use pour cela de requêtes plus ou moins complexes. Il produit également des textes nouveaux par concaténation ou extraction, voire calcul, et possiblement des dispositifs graphiques (cartes, plans, graphes, schémas), qui sont, dans le monde numérique, assimilables eux aussi aux produits de l'application d'opérateurs d'écriture à des éléments textuels [Guichard, 2010]. Ce n'est qu'ensuite qu'il va rendre compte de son activité, par un cours, une conférence, un texte, une prise de parole dans le cadre d'un débat. Il rencontre alors d'autres types d'acteurs, éditeurs, journalistes, étudiants, pour n'en citer que quelques-uns, dans le même mouvement qu'il rend publics des documents qui pourront à leur tour faire l'objet d'une appropriation et nourrir le travail d'autres historiens.

L'enquête historique devient alors la délimitation d'un corpus, par inclusion dans l'ensemble des matériaux d'inscription, référée toujours à une matérialité dont l'existence est antérieure au mouvement de l'enquête et n'est pas affecté par celle-ci, puis application d'un jeu de règles opératoires – donc production de nouveaux éléments textuels – dont il est douteux qu'elles puissent être systématiquement explicitées par les historiens ni réductibles à un ensemble d'algorithmes. Cela d'ailleurs ne distingue guère les historiens des physiciens ou d'autres spécialistes de sciences dites dures [Collins, 1995].

Il me faut, avant d'indiquer quelles conséquences je tire de cette définition, faire deux remarques. La première est que cette construction est loin d'être originale ou neuve. Michel Foucault, meilleur connaisseur qu'on ne l'a parfois écrit des travaux de ses contemporains historiens, proposait, en 1969, un schéma assez proche⁴ :

3. L'un des effets induits, et récents, de la diffusion des pratiques numériques est un intérêt nouveau pour les pratiques des historiens au travail [Dougherty et Nawrotzki, 2013].

4. Celui que nous proposons ici emprunte également à Michel de Certeau [Certeau, 2002] et Jack Goody [Goody, 1979].

« [...] par une mutation qui ne date pas d'aujourd'hui, mais qui n'est sans doute pas achevée, l'histoire a changé de position à l'égard du document : elle se donne pour tâche première non point de l'interpréter, non point de déterminer s'il dit vrai et quelle est sa valeur expressive, mais de le travailler de l'intérieur et de l'élaborer ; elle l'organise, le découpe, le distribue, l'ordonne, le répartit en niveaux, établit des séries, distingue ce qui est pertinent de ce qui ne l'est pas, repère des éléments, définit des unités, décrit des relations [...] L'histoire est le travail et la mise en œuvre d'une matérialité documentaire (livres, textes, récits, registres, actes, édifices, institutions, règlements, techniques, objets, coutumes, etc.) qui présente toujours et partout, dans toute société, des formes soit spontanées, soit organisées de rémanences. Le document n'est pas l'heureux instrument d'une histoire qui serait en elle-même et de plein droit mémoire ; l'histoire c'est une manière pour une société de donner statut et élaboration à une masse documentaire dont elle ne se sépare pas. » [Foucault, 1969, p. 14]

La seconde remarque est que cette définition n'a pas vocation à définir un être ou une essence de l'histoire telle qu'elle est pratiquée aujourd'hui. Construction, abstraction, elle laisse délibérément de côté un certain nombre de propriétés de l'acte historien et plusieurs des questions qui lui sont associées, en particulier celle du sens, mais aussi de la définition même de ce qu'est le fait historique. Nous les avons écartées parce que notre visée est d'abord stratégique. Il s'agit de réfléchir à ce que changent, ou peuvent changer, pour une profession particulière, les transformations de l'outillage des professions intellectuelles, et partant à se doter des outils permettant de définir des conduites adaptées – à titre individuel ou collectivement –, travail dont on trouvera l'écho dans d'autres textes.

Il me semble que, dans cette perspective, l'intérêt de ce détour est triple. Il permet de rappeler d'abord, après bien d'autres [Prost, 1996], que la connaissance historique est le produit d'une chaîne de production de savoirs, prise dans des institutions, des systèmes techniques et des champs sociaux, et non accumulation d'œuvres singulières. Cela nous rappelle aussi qu'il y a des médiations successives, dont nous avons parfois bien à tort postulé la transparence et qui font de la production de l'histoire une activité éminemment sociale et collective. Entre le référent auquel se rapporte le discours de l'historien et le vestige d'une activité humaine qu'il convoque à l'appui de son discours et, ensuite, entre l'incorporation de ces

vestiges dans un corpus et les descriptions multiples dont ils font l'objet, les interventions sont multiples, les chaînes de raisonnement longues et complexes et les possibilités d'erreur innombrables. De plus, nous pouvons alors penser le fait qu'une transformation des conditions de production de la connaissance historique est susceptible d'avoir de profonds effets tant sur le mode de son élaboration que sur la nature des productions émergeant dans la sphère publique. En l'occurrence, l'augmentation de la masse des données accessibles, l'existence sous des modes numériques divers d'une partie croissante d'entre elles, la diversification des modes de traitement de l'information et les transformations des modes de diffusion de la connaissance constituent tant une mutation des technologies de l'intellect que des conditions de réception du savoir et laissent augurer d'une mutation des produits de l'activité historique. Conduit à celle-ci tant une transformation de la demande, susceptible de se segmenter [Cricione *et al.*, 2004] et donc de produire une différenciation des produits et des profils historiens, que la possibilité offerte de manipulations nouvelles de corpus d'un ordre de grandeur nouveau.

Ajoutons que le schéma proposé permet de suggérer, par l'usage d'une métaphore spatiale, pistes et enjeux. S'il y a dans le monde numérique des lieux d'accumulation et de transformation de l'information, alors s'ouvre la possibilité, en même temps que d'une dissémination potentiellement universelle, d'une appropriation privée des annotations et des modes opératoires, qui, étant logiciels, sont eux-mêmes textes.

Prenons deux exemples, illustrant chacune de ces modalités et d'abord l'existence de plates-formes numériques payantes donnant accès au contenu des numéros récents des revues scientifiques. L'accès à celles-ci est affranchi des contraintes géographiques. Ces ressources peuvent être consultées de n'importe où dans le monde. La possibilité de le faire dépend cependant de l'appartenance institutionnelle des individus, de leurs capacités financières ainsi que des formes des réseaux sociaux desquels ils participent. La consultation d'un article déposé sur la plate-forme CAIRN coûte ainsi, aux dires des intéressés, cinq dollars lorsque l'on se connecte depuis le campus de l'université de Sydney. Cette transformation se produit alors que l'insertion des productions textuelles des chercheurs au sein de la bibliographie internationale est de plus en plus affirmée comme critère d'évaluation des productions scientifiques. Nous parlons ici de l'incorporation à bon escient de références à celle-ci dans les textes produits autant que de la capacité des chercheurs à obtenir que leurs propres productions soient mentionnées par les producteurs

de textes inscrits dans les corpus de référence des agences de notation. Ainsi, le passage des revues les plus prestigieuses dans le monde numérique payant tend à réserver aux plus dotés la possibilité d'accroître leur productivité et de multiplier les signes de distinction.

Il pourra alors être question d'accès, de passages, donc de territoires, de pouvoirs et de conflits. Il est alors aussi possible d'interpréter les postures et les conduites des agents du champ en référence à des positions spécifiées. L'exemple peut être ici celui des débats en cours autour de la mise à disposition des bases de données ou des dossiers documentaires élaborés par les chercheurs, lors, par exemple, de la préparation d'une thèse ou d'un ouvrage. Cela revient, si nous reprenons notre schéma d'ensemble, à dire que, puisque les annotations produites par les chercheurs sont aujourd'hui dans une large mesure nativement numériques, alors s'ouvre la possibilité de les mettre massivement à disposition de la communauté. Cela représente, par le changement d'échelle de la taille des corpus structurés disponibles, la promesse d'une augmentation de la capacité interprétative des chercheurs. Des travaux sont de fait en cours en France afin d'offrir aux acteurs de la recherche en sciences humaines et sociales l'accès à des répertoires structurés de données et de documents issus de recherches antérieures. Le principal opérateur est ici le grand équipement Adonis [Adonis, 2009].

Imaginons maintenant un chercheur travaillant sur la diffusion des problèmes arithmétiques dans l'Europe médiévale. Il peut espérer ainsi pouvoir raisonner sur trois cents formes de problèmes et non plus cinquante, couvrant une aire géographique un peu plus vaste ou une période un peu plus longue. Il peut espérer tant une plus grande fiabilité et pertinence des indicateurs statistiques qu'il construit, qu'une interprétation plus complexe, l'espace des principes interprétatifs qu'il construit se voyant doté de quelques dimensions de plus. Dans le même temps, le changement d'échelle du corpus autant que la possibilité de traitements nouveaux de ses éléments, font surgir la possibilité de questions nouvelles. Cette conjoncture apparaît à certains enthousiasmante. De fait, la revendication d'une « libération des données » est portée par nombre d'acteurs, souvent ingénieurs, ou issus des centres de recherche français les plus prestigieux, au sein de la mouvance qui tente aujourd'hui de faire émerger des humanités numériques à la française [Mounier, 2010]. Elle est relayée par un certain nombre d'acteurs au sein des institutions françaises de recherche qui prônent, pour le moins, le libre accès aux données produites par les agents du CNRS ou les thésards ayant bénéficié

de financements publics, en conformité d'ailleurs avec la déclaration de Berlin de 2003 [« Berlin Declaration on Open access to Knowledge in the Science and Humanities », 2003], signée par les représentants de plusieurs organismes de recherche européens.

Le dossier pourtant n'avance que lentement et ne semble pas générer beaucoup d'écho parmi les historiens eux-mêmes. Jouent bien sûr des difficultés techniques non triviales, touchant tant au format qu'aux structures des données. Il est permis de penser cependant que les désirs des uns et des autres sont aussi fonction de leur position au sein de la chaîne de traitement du signe que nous avons dessiné. En l'état actuel des techniques, pour réutiliser efficacement, sans que l'opération prenne trop de temps, les données structurées d'une enquête que l'on n'a pas élaborée, il faut posséder de véritables compétences informatiques (à tout le moins la maîtrise des expressions régulières, l'habitude des bases de données et une capacité minimale à programmer) en plus d'une très solide culture disciplinaire. Cela réserve l'opération aux mieux dotés en capacités cognitives ou aux membres des institutions les mieux pourvues en intelligence informatique, qui sont souvent les mêmes. Il est assez peu surprenant dès lors que de tels acteurs réclament l'ouverture des gisements de données, demande que les institutions de la recherche, tenues de promouvoir la productivité du travail scientifique, ne peuvent que prendre elles aussi à leur compte. Que beaucoup de chercheurs fassent la sourde oreille, ou ne voient pas l'intérêt de la chose, n'est guère plus surprenant. Ils participent d'une culture et d'une organisation de la recherche qui fait de l'historien un accumulateur de fiches et de données, qu'il peut espérer transmuier en position dans l'institution au moyen de la production de textes dont ce butin est le garant. Dans bien des cas, lui proposer d'accepter que l'on accède à son trésor en échange d'un droit de visite à une infinité d'autres dépôts, revient à le payer en monnaie de singe, ou, plus exactement, à le déposséder du fruit de son travail, qui ne peut plus faire office de capital.

L'accumulation et l'organisation de son dépôt n'ouvrent pas en effet, en l'état actuel des choses, droit à rémunération, ni financière ni symbolique. Les seuls produits échangeables de l'activité de l'historien – qui d'ailleurs prend généralement part à différents marchés – sont les mises en récit que, sous diverses formes, il produit à partir de son activité de rassembleur et de manipulateur de signes. Ouvrir son entrepôt à tous revient à le mettre en concurrence avec les membres les mieux équipés de sa corporation, tout en lui retirant l'exclusivité de l'accès à ses matériaux ou bien à lui proposer un libre accès à des données dont il ne

peut faire information, à moins de consentir de lourds investissements d'appropriation, qu'il ne peut pas nécessairement assumer, tout en lui ôtant l'exclusivité d'usage d'un bien rare.

De plus, sa différence, ou sa compétence spécifique, qui est souvent connaissance intime d'un fond longuement pratiqué, est brutalement dévaluée. Est affirmée – sans doute un peu vite – la possibilité que celle-ci soit codifiée, détachée de son porteur et échangeable. Ainsi s'ouvre la possibilité ou la promesse d'une division du travail, qui est aussi une hiérarchisation des producteurs et qui fait de certains ingénieurs des alliés indispensables du travail des historiens numériques. Numérisation ne rime donc pas forcément avec démocratisation et moins encore avec égalité. Ces exemples ont du moins l'avantage d'attirer l'attention sur le fait que la transformation des outils de l'historien ne conduit sans doute pas seulement à une évolution de ses productions, mais aussi à des mutations de la profession historique, voire de la figure même de l'historien.

Nous pouvons en effet, en prenant en compte l'ensemble des opérations de production du savoir historique, nous départir d'une vision qui fait de l'historien d'abord un producteur de récits. Outre que cela permet de réintroduire dans l'analyse une bonne partie de son temps de travail et un certain nombre de ses fonctions, toutes celles en particulier qui relèvent de ce que l'on appelait autrefois l'érudition (établissement d'inventaires, bibliographies, édition de documents), cela permet aussi de penser l'historien comme une fonction et un actant et non plus comme un individu. S'il accomplit des tâches logiquement décomposables en unités faisant appel à des compétences spécifiques et exigeant une technicité croissante (sélection, annotations, manipulations, mise en récit), alors s'ouvre la possibilité du collectif, mais aussi de la division du travail, à tout le moins de la coopération avec différents opérateurs porteurs de compétences propres.

Le fait n'est pas en soi nouveau. Beaucoup de grandes entreprises éditoriales et savantes, ces dernières décennies ou ces derniers siècles, étaient le produit de configurations sociales complexes associant, parfois durant un temps long, de nombreux agents dotés de compétences spécifiques et de logiques d'action propres. Pensons, pour la seule historiographie française, à l'aventure du Maitron [Maitron, s.d.], à celle des grandes histoires (de la France urbaine, de la famille, de la vie privée...) qui ont durablement marqué l'historiographie et associaient éditeurs, diffuseurs, institutions savantes, historiens de métier (et militants aussi, dans le cas du Maitron). La différence ici n'est pas de nature mais de

degré. S'ajoutent, potentiellement au moins, de nouveaux acteurs parmi lesquels différentes espèces d'ingénieurs et de producteurs de ressources électroniques (éditeurs de logiciels ou créateurs/gardiens de ressources numériques). Ceux-ci ont comme particularité, à la fois de disposer de capacités d'action sans commune mesure avec celles des autres acteurs, mais aussi de ne pas être des acteurs traditionnels de la production de savoirs dans un tel domaine. Ce sont pour certains des entreprises privées poursuivant une logique de profit et/ou cherchant à constituer des rentes dans une économie de concurrence imparfaite. La dépendance accrue à une infrastructure et un équipement complexe créent une dépendance envers ceux qui en maîtrisent l'usage et la vente (ou sans doute de plus en plus la location).

Là encore, un exemple aidera sans doute à comprendre ce point. Imaginons qu'il prenne la fantaisie à Google, ou tout autre acteur privé, de numériser l'ensemble des incunables conservés dans les bibliothèques européennes, ou toute autre ressource utile aux érudits, ce qui est loin d'être une hypothèse forte, sans en faire payer l'accès, mais, par contre, de réserver l'usage des fonctions de recherche avancée à ceux qui acceptent d'ouvrir un compte Google, permettant à la compagnie d'accumuler sur chacun de ses utilisateurs une information monnayable, nous serions placés, en tant qu'individus ou en tant que représentants d'institutions, devant d'assez redoutables dilemmes.

Nous ne sommes là pas si loin de Lucien. Définir l'activité historique est d'abord prétendre imposer une norme ou exprimer une visée stratégique. L'un des principaux apports de l'intrusion du numérique au sein de notre discipline est sans doute de nous obliger à une réflexion sur nos pratiques, de nous le rappeler et de nous inciter à nous écarter de nos cadres de pensée traditionnels, peu à même de nous permettre de penser l'irruption d'un tel changement.

La définition proposée ici, de l'historien comme agent d'une chaîne complexe de traitement de signes, n'a pas d'autre ambition. Elle n'a certainement pas celle de dévoiler une essence de l'histoire à laquelle je ne crois guère. Elle ne vise pas non plus à épuiser les déterminations permettant de penser celle-ci. Nous concevons ici l'historien comme un producteur de signes et de discours dont la validité est garantie par la référence, directe ou médiate, à des vestiges, dont la conservation et la consultation ne sont possibles qu'au prix de l'existence d'une organisation sociale complexe. Celle-ci est appuyée, depuis longtemps, sur une forte technicité dont l'existence, ou l'impact propre, n'est pas toujours interrogée. L'intérêt

de le définir ainsi réside surtout dans le fait de penser ce qui peut changer de ce métier, ou plutôt de cette activité, quand se transforment l'outillage ou les systèmes techniques dans lesquels sont pris les historiens. Ces transformations potentielles ou en cours ne sont pas minces. En font partie les configurations d'opérateurs intervenant dans la production des discours historiques, la nature des produits offerts et les conditions de leur circulation et de leur appropriation. Le chemin que nous avons pris, loin cependant de mener à une forme de déterminisme technique, met en évidence également le fait que les transformations en cours et à venir de ces pratiques sont aussi liées à des configurations économiques et sociales complexes et donc indéterminées en leurs formes, et partiellement négociables par les acteurs.

+++++

CHAPITRE X. DE QUOI LE WEB EST-IL L'ARCHIVE ?

+++++

Ce texte reprend une conférence donnée lors du colloque final de l'ANR ATHIS piloté par Jean-Philippe Genet (Lamop Paris I). Celui-ci s'est tenu à l'École française de Rome du 4 au 6 décembre 2008. Les actes en ont été publiés en 2011 [Genet et Zorzi, 2011]. L'auteur de ces lignes était le représentant du Centre d'histoire sociale du XX^e siècle dans cette ANR. La suite de rencontres internationales organisées dans ce cadre avait pour fonction l'établissement d'un état des lieux des pratiques informatiques et réseaux des historiens. Il s'agissait également de favoriser la diffusion de cette information au sein de la communauté historienne. L'article proposé ici résulte d'une commande. Le projet initial envisageait les transformations des pratiques érudites, mais ne prenait pas en compte le statut de source des documents présents sur le réseau, au moins pour de futurs historiens. La réflexion sur ce point étant dans le milieu des historiens encore embryonnaire à cette date, la réponse avait pris la forme du compte rendu d'une expérience pratique. Si le réseau d'aujourd'hui ne peut plus tout à fait s'appréhender et se travailler comme celui du milieu des années 2000, cet article conserve une certaine actualité, la réflexion sur ces questions demeurant rare au sein des milieux historiens, malgré les travaux menés à Paris I autour de Stéphane Lamassé ou d'autres chantiers en cours dont témoignent quelques publications récentes [Schafer, 2015]. Ce texte consistant en l'exploitation d'un matériau, qui, pour partie, n'est plus disponible, nous avons donné, pour les pages utilisées à l'époque, l'URL visitée alors. Il est dans certain cas possible de retrouver à partir de cela la trace du document par le moyen d'une archive du Web.

Pas de documents, pas d'histoire.
Langlois et Seignobos, *Introduction aux études historiques*, 1902

Je voudrais proposer ici quelques réflexions autour de l'usage d'Internet et plus particulièrement du World Wide Web comme source possible pour l'historien. La réflexion collective initiée par les ateliers ATHIS, qui se veut prospective, ne peut éviter de rencontrer cette question. Les pages qui suivent cependant ne peuvent proposer que quelques éléments fragiles, provisoires, témoignant d'une réflexion en cours ou plutôt d'une interrogation. Celle-ci est née d'une pratique d'enseignant, plus que de mes propres recherches. J'enseigne en effet dans un centre universitaire, le Centre d'histoire sociale de l'université Paris I, qui attire, et de plus en plus, des étudiants qui entendent travailler sur le très contemporain et parfois sur des phénomènes ou des institutions difficiles à saisir par le biais d'autres sources que les sources nativement numériques. Le Web francophone en effet, qui a aujourd'hui une quinzaine d'années d'existence, est le lieu d'expression de mouvances culturelles ou militantes très minoritaires, dont l'étude attire certains de nos étudiants. Ils se tournent donc assez logiquement vers les ressources offertes par le Web afin d'amasser l'information dont ils ont besoin. Pour ne prendre qu'un exemple, j'évoquerai une étudiante qui, il y a quelque temps, a soutenu un mémoire dont l'objet était la mouvance hip-hop en région parisienne [Fouques, 2005], petit milieu qui donne naissance à des productions culturelles qui expriment une identité définie, en partie, en référence à un substrat politique. Internet s'est révélé en ce cas réservoir de sources et d'informations utiles, mais je me suis rendu compte qu'il m'était difficile, tant d'accompagner cette recherche, que parfois d'évaluer l'usage fait des ressources électroniques.

Je pars donc d'un constat, celui d'une ignorance et d'une difficulté et je vise ici à en comprendre la teneur en précisant les termes de quelques questions simples. Puis-je trouver sur le Web des matériaux qui ont pour l'historien une utilité spécifique, c'est-à-dire auxquels je ne puis substituer d'autres matériaux ? Quelles difficultés pose leur usage et sont-elles là encore spécifiques ? Puis-je définir une posture, ou des stratégies, adaptées au matériau, qui permettent de fabriquer des sources sans renoncer aux exigences de l'historien ?

Pour éviter de donner à cette exploration un tour trop désincarné, j'ai choisi un thème de recherche, qui constitue à la fois un véritable objet d'histoire sur lequel j'ai quelques lueurs et qui offre un terrain qui se prête à une enquête partiellement menée par le Web. Je me suis donc demandé ce que la toile me permettait de dire des expulsions d'étrangers menées

par la République française au cours des deux dernières décennies et des réactions qu'elles suscitaient.

Cette politique, particulièrement brutale dans ses formes, n'est pas en soi une nouveauté. Durant les années 1930 déjà, les difficultés économiques conduisirent certains des gouvernements de l'époque à multiplier les arrêtés d'expulsion et de refoulement [Rygiel, 2009]. Cette crise présente cependant quelques traits spécifiques. L'un est qu'elle rencontre de fortes résistances au sein de la société civile qui se traduisent par la multiplication d'associations, de collectifs et de réseaux *ad hoc* dont la principale ambition est de combattre les mesures adoptées ou d'en limiter les effets. Ces structures légères sont originales à plusieurs titres. Les résistances que rencontrent les politiques des années 1930 sont le fait de partis politiques, d'organisations syndicales, qui prennent en charge la question et qui défendent tant des principes que des individus, dont certains sont membres, au moins théoriquement, de leur clientèle. D'autre part, les structures légères que nous rencontrons aujourd'hui font un usage intensif des nouvelles technologies, nous pouvons retrouver des traces de leurs activités sur la toile. Ma question devient donc, dans un premier temps, puis-je trouver, par le Web, des matériaux que ne m'offriraient pas la presse ou les archives du ministère de l'Intérieur ?

Mon premier réflexe, puisque le Web offre une multitude de discours, est d'y rechercher des tropes et des locuteurs. Je cherche à repérer ceux qui parlent des événements qui m'intéressent, à décrire leurs discours. Mes ennuis commencent cependant lorsque je veux découper un corpus, le décrire et l'analyser.

SEUL DANS UN MONDE HOSTILE

+++++

L'impossible corpus

La première difficulté est liée au nombre des éléments potentiellement pertinents. Une recherche sur Google, combinant les trois termes « étrangers », « expulsions » et « France », renvoie, en janvier 2009, à 583 000 documents, qu'il est bien sûr impossible de tous consulter et pour lesquels nous ne possédons pas non plus de description pertinente au regard de nos préoccupations. Sachant que les moteurs de recherche ne parviennent jamais qu'à indexer une partie du Web, sans d'ailleurs qu'il soit possible de déterminer ni la taille ni les propriétés des parties

manquantes, je ne peux que constater l'énormité de la tâche consistant à constituer un dossier documentaire complet et à en recenser les éléments.

Une autre difficulté pratique tient à la diversité des ressources qui me sont offertes. Je vais, fouillant les pages auxquelles je puis accéder, trouver des extraits d'émissions télévisées, des enregistrements sonores, des photographies, des textes de natures diverses (traités savants, articles de journaux, tracts ou pétitions électroniques, recueils de dispositions juridiques), soit un matériau dont l'exploitation n'obéit pas au mode de spécialisation habituel des historiens, qui sont souvent autant maîtres d'un ou de quelques fonds et types de sources qu'experts d'une question.

Le mode d'accès à l'information par le biais de requêtes sémantiques, ou pour mieux dire, par l'emploi de chaînes de caractères – qui est celui proposé par défaut par les principaux moteurs de recherche – a de plus pour effet, qu'en fonction des éléments retenus, nous obtenons des réponses différentes. Soit deux requêtes formulées le même jour sur la page de garde de Google, l'une sous la forme (expulsions + France + étrangers), l'autre sous la forme (reconduite à la frontière + France + étrangers), la première obtient 368 000 réponses, la seconde 17 600. Si nous classons au moyen d'une typologie sommaire les requêtes proposées en page 1, nous trouvons dans le premier cas 4 sites d'information – émanation internet de médias traditionnels –, 3 sites militants, un site juridique, un extrait d'encyclopédie et une page du site du Sénat. Dans le second cas, la première page nous renvoie à 8 sites juridiques, à un site d'information (marocain celui-ci) et à une encyclopédie, qui n'est d'ailleurs par là même que la précédente. Ajoutons que si nous répétons la même requête à quelques minutes d'intervalle, nous n'obtenons pas le même nombre de réponses (405 000 pour expulsions + France + étrangers) et qu'intervertir l'ordre des termes, sans modifier ni les chaînes de caractères ni les opérateurs booléens utilisés, aboutit à obtenir 627 000 réponses (expulsions + étrangers + France). Certains des éléments de ce constat ne sont pas très surprenants, étant donné la thématique choisie. Une bonne partie des luttes politiques en cours qui touchent aux questions d'immigration ont en effet pour enjeu la dénomination même des individus et des procédures et l'emploi de tel ou tel syntagme est, en soi, à la fois prise de position et intervention dans des luttes symboliques. La rafle de sans-papiers s'oppose ainsi terme à terme à l'arrestation de ressortissants étrangers en situation irrégulière. La conséquence en est cependant à la fois la démultiplication du nombre de pages potentiellement intéressantes et la

difficulté à définir un champ sémantique pertinent, et surtout à valider les choix faits, puisque cela suppose une connaissance préalable de celui-ci.

Aux médiateurs inconnus

Ce simple exemple illustre d'autres difficultés liées à l'emploi du Web et d'abord le fait que nous n'accédons jamais à celui-ci, mais obtenons d'une médiation les réponses aux requêtes que nous formulons. C'est bien sûr le cas aussi de l'historien qui travaille en archives ou en bibliothèques, dont le parcours est permis et guidé, voire infléchi, par des médiateurs. Deux différences importantes cependant subsistent. La première est que nous sommes familiers de ces médiations, en avons conscience et souvent les prenons explicitement en compte [Cerutti, Fayet et Porret, 2006], alors que l'idéologie internet, qui sacralise la transparence et l'immédiateté [Breton, 2000, p. 63 et suivantes], peut masquer à l'utilisateur le caractère toujours médié de ses vagabondages sur la toile. D'autre part, dans le cas de l'archive et de la bibliothèque, les opérations effectuées par les professionnels sont généralement documentées et explicitées dans des documents accessibles (plan d'inventaire, catalogue). Cela n'est pas vrai des procédures mises en place par les opérateurs internet. Ceux-ci en effet ne sont généralement pas des opérateurs publics, tenus de rendre compte de leurs procédures, mais des sociétés commerciales dont le succès repose en partie sur leur habilité à garder secrets les algorithmes qui font la fortune de leurs actionnaires. Je suis ainsi incapable de savoir, l'algorithme de Google n'étant pas public, pourquoi il ne me renvoie pas les mêmes réponses lorsque j'intervertis les termes d'une requête. Cette médiation a des effets sur le nombre de documents proposés : un autre moteur de recherche, Seek < www.seek.fr >, renvoie ainsi 118 réponses pour France + expulsions + étrangers, sur le choix des sites mis en valeur (parmi les dix premiers sites proposés le même jour par cet opérateur, trois sont des sites officiels français, trois des sites juridiques, deux émanent de médias, l'un est un site de particulier, un seul est un site militant), mais aussi introduit probablement en soi un biais de sélection.

Nous ne pouvons bien sûr le montrer en comparant les milliers de pages auxquelles renvoient des moteurs de recherche généraliste, mais pouvons le vérifier à partir de sites spécialisés dans l'indexation d'une partie des ressources web, en l'occurrence les blogs. Si nous utilisons Blog dimension, service francophone spécialisé dans l'exploration des blogs, nous trouvons,

en réponse à notre requête habituelle, 1 500 résultats. Parmi les premiers sites mentionnés figurent en bonne place des blogs dont les thématiques et le vocabulaire évoquent l'ultra-gauche française, ainsi que deux sites pro-palestiniens et un site personnel essentiellement consacré à la défense et à l'illustration de la révolution islamiste iranienne. À l'inverse, l'utilisation d'un moteur anglo-saxon, Blogdigger (< <http://www.blogdigger.com> >) ne me renvoie que quelques résultats, qui proviennent de pages personnelles de particuliers hébergées par le journal *Le Monde*, quand d'autres me proposent des pages m'offrant de louer une maison dans le Lot ou l'Aveyron, produites par des hôteliers amateurs. L'exemple bien sûr est malicieusement choisi. Il n'en illustre pas moins le fait que, croyant accéder au Web, je ne fais que recourir aux services d'un ou de plusieurs médiateurs. Cela a des effets tant sur le volume des ressources auxquelles je puis accéder, que sur la nature de celles qui sont mises en avant et en valeur et affecte probablement la nature du corpus auquel je puis accéder, sans qu'il me soit possible ni de définir ni d'évaluer les biais introduits.

Le même exemple peut nous permettre d'illustrer une difficulté, pour l'historien, d'une tout autre nature. Je n'ai, lorsque je consulte un document internet, pas accès à un document, mais à une vue calculée selon des procédures complexes, qui sont définies par les caractéristiques à la fois des machines utilisées et des codages informatiques employés, à partir de fichiers informatiques auxquels je ne peux souvent pas accéder. Tout un chacun a pu faire l'expérience de ce que, selon la machine ou l'écran utilisé, l'apparence de bien des pages web varie. Nous sommes peut-être moins conscients de ce que, selon le moment de la consultation, ou la position de la machine que nous utilisons dans le réseau, les éléments textuels obtenus peuvent eux aussi varier lorsque nous effectuons une même requête à partir d'un site dynamique, qui, sauf exception, ne nous permet pas d'accéder aux tables ou aux fichiers sources utilisés pour produire la réponse à une demande. Le moyen le plus simple de le vérifier est de proposer à un même moteur de recherche une requête identique à quelques heures d'intervalle, ou depuis deux postes n'appartenant pas au même sous-réseau. Il s'ensuit que, visualisant une page web, ou plutôt accédant à une URL, je ne consulte pas un document, ou fraction d'un document, au sens par exemple où je pourrai examiner une liste de brevets, ou une liste de feux, mais obtiens la réponse à une question. De plus, ce que je lis ne peut avoir le statut de document, au sens où les historiens habituellement l'entendent, que si j'enregistre, en utilisant un protocole qui en respecte

l'aspect, ce que je viens moi-même de coproduire et l'archive en en garantissant ainsi l'identité.

Les difficultés de la critique

Nous touchons là à une des caractéristiques premières des pages web, et d'autres ressources électroniques d'ailleurs, qui est la dissociation du support d'inscription de l'information et du support de visualisation de celle-ci, en l'occurrence la page web. La page n'existe que par l'instruction donnée par l'internaute, qui déclenche un processus génératif dont des pans, parfois très larges, lui sont non seulement de fait inconnus, mais aussi très souvent inaccessibles. Les conséquences pour l'historien en sont multiples. L'une d'elle en est que je ne peux mener une critique externe du document qui en établisse l'authenticité ou l'intégrité à partir de la réunion et de l'interprétation d'indices matériels. Les traces incorporées au document de sa genèse ne sont, dans le cas du document informatique, ni de même nature que celles que porte le document papier, ni surtout accessibles à celui qui visualise une page web, parce qu'éventuellement inscrites dans les entrailles de machines auxquelles il n'a généralement pas accès. La facilité avec laquelle les spécialistes du *phishing* produisent des copies quasi parfaites de grands sites commerciaux, dont le caractère frauduleux n'est détectable par l'utilisateur qu'au moyen d'interactions directes avec des représentants de l'organisation attaquée, illustre ce point.

Ajoutons que la critique de provenance n'est, dans un certain nombre de cas, guère plus facile à mener. La première raison en est que les pages web, ou les fragments qui les composent, ne sont pas nécessairement datées. Certes il est possible, par divers moyens, dont le recours aux archives existantes du Web, de reconstituer divers états d'un site, mais d'une part cela nécessite à chaque fois enquête, d'autre part cela n'est possible ni pour tous les sites ni pour toutes les parties d'un site. L'archivage d'un site stable et bien référencé se fait en effet généralement par capture, à intervalles plus ou moins réguliers, de documents générés lors d'une consultation. Cela revient, un site web, quoique n'étant pas assimilable à un média de flux en partageant certains traits, à archiver un film en prélevant une image à intervalles de quelques minutes et, du fait à la fois de l'instabilité des URL et de la multiplication au cours de ces dernières années de sites dynamiques régulièrement reconfigurés, à ne proposer de ceux-ci qu'une vue très partielle. Le mode de consultation des dépôts d'archives numériques

étant de plus généralement la définition et la documentation de sites, il est rare qu'ils permettent une recherche sémantique, ce qui interdit la constitution de dossiers documentaires rétrospectifs pertinents à partir de la définition thématique d'un champ ou par la définition d'expressions régulières *ad hoc*. En d'autres termes, je ne peux pas savoir ce que le Web permettait de lire il y a deux ans ou trois ans lorsque l'on s'intéressait aux expulsions d'étrangers par l'État français, non plus que toujours dater les éléments qu'il est possible aujourd'hui de réunir.

Pour des raisons différentes, qui tiennent autant aux usages sociaux s'accompagnant de la production de ressources destinées au Web qu'aux dispositifs techniques mis en œuvre, l'attribution à une instance ou à un auteur des documents web est parfois malaisée. Le fait est lié d'abord au développement des outils de production collaborative (qui ainsi est l'auteur d'une notice disponible sur un dispositif wiki ? et faut-il comprendre ce qu'elle nous offre comme l'émanation d'une institution dont il serait possible de définir l'identité – une forme contemporaine de *scriptorium* – ou comme la résultante d'un processus sociotechnique dont nous connaissons au fond mal les tenants). Bien d'autres productions pour le Web, particulièrement celles issues de ce qu'il est convenu de nommer le Web 2.0 [Noiret, 2011], conduisent à des questions similaires. Les blogs ainsi, particulièrement ceux qui évoquent des thèmes susceptibles d'une politisation, voient se multiplier les contributions anonymes, fragments de textes dont nous ne connaissons pas les auteurs, sauf à considérer qu'ils sont, par le simple fait qu'il ne les a pas supprimés, élément d'un discours agencé par le propriétaire du blog. Nous voilà alors confronté à une multiplication de tracts dadaïstes dont les fragments juxtaposés seraient mobiles et parfois évanescents.

Bien sûr, je force un peu le trait et certaines des difficultés rencontrées par l'historien qui explore le Web ne sont pas sans analogies avec celles qui se présentent à des historiens découvrant, dans l'arrière-salle d'un local syndicat, des piles emmêlées de tracts non signés et non datés, inventoriant patiemment les graffitis de Pompéi, ou consultant dans un fonds public une collection de lettres anonymes. Il n'en reste pas moins qu'il est des choses que le Web ne permet pas à un historien qui entend l'utiliser en respectant les principes critiques qui fondent sa discipline. Ces impossibilités, ou ces difficultés, illustrent souvent des propriétés du matériau qui s'accordent mal aux pratiques spontanées des tenants d'une discipline dont les normes furent définies en référence à des documents que caractérisaient l'identité des supports d'inscription et de visualisation,

la pérennité et la stabilité de ceux-ci – soit leur identité –, souvent garantie par une institution – archives ou bibliothèque – qui assurait leur conservation tout en facilitant leur identification, voire, souvent, en garantissait l'authenticité et l'attribution à un émetteur. Je ne peux ainsi pas constituer un corpus documentaire exploitable par l'historien à partir d'une recherche sémantique sur le Web, parce que les données qu'il me fournit sont peu structurées, ou plutôt structurées par des opérateurs qui n'appartiennent pas à la même chaîne de savoir que moi et dont je ne peux connaître le mode opératoire, mais aussi du fait de la multiplicité des ressources et de la nécessité, pour identifier et dater nombre d'entre elles, d'une enquête dont les procédures ne peuvent être automatisées. Pour le dire dans la langue native des historiens, si le Web ne me permet pas de découper un corpus thématique pertinent, c'est parce qu'il n'est ni une archive ni une bibliothèque et qu'aucune instance identifiable n'a rassemblé, structuré et décrit les documents que je peux produire le consultant, ce qui transforme profondément les conditions de leur utilisation et en particulier les conditions de la critique. Celle-ci n'est pas impossible, mais elle ne peut s'appuyer ni sur les acquis d'une érudition ni sur l'usage d'indices et de systèmes de signes familiers aux historiens. Cela ne signifie ni que l'historien ne peut pas utiliser le document électronique, ni qu'il ne le fera pas, ne serait-ce que parce que les futurs historiens de la période que nous vivons n'auront guère le choix.

Y ALLER QUAND MÊME

+++++

Des instances sans papier et des pratiques sans nom

La première raison, de fait, est que de plus en plus, l'activité d'un certain nombre d'institutions ne laisse de trace qu'électronique et, dans le cas d'un certain nombre de structures militantes, essentiellement sous la forme de ressources destinées au Web. L'historien d'ailleurs y trouvera son compte. Le Web en effet porte la trace d'échanges, de pratiques, dont nous avons pour les périodes anciennes peu de traces. La sociabilité autour d'un objet ou d'un projet, l'expression semi-publique de sentiments sur la chose publique ne sont pas ainsi choses nouvelles, mais nous n'avons pas toujours beaucoup de traces de ces pratiques. Le Web, à l'inverse, les blogs en particulier, regorgent de fragments de textes évoquant immanquablement

les conversations de comptoir ou les discussions de membres de petits groupes unis par une même passion [Rebillard, 2009].

Je ne sais pas s'il faut considérer que nous avons là des pratiques nouvelles, ou bien une visibilité nouvelle et un changement des formes de pratiques sociales depuis longtemps vivaces, mais il me semble que s'ouvre un gisement de traces nouvelles et que l'historien, dont nous savons depuis longtemps qu'il est un ogre, ne peut que le lorgner avec gourmandise, même si ces traces seront difficiles à manier. Elles supposent, par leur volume et leur nature, la définition d'une diplomatie du document électronique, d'autant plus difficile à mettre en place que ses formes évoluent rapidement dans le temps. Les experts du W3C nous annoncent (et préparent) le successeur du Web 2.0, qui prendrait la forme d'un Web sociosémantique associant les principes collaboratifs caractéristiques du Web 2.0 et la multiplication d'agents intelligents permettant l'automatisation de tâches de recherche et d'interprétation en s'appuyant sur des ontologies locales [Laublet, 2007]. De telles recherches supposent également des outils adaptés à la fois aux préoccupations des historiens et aux particularités du matériau, parmi lesquelles la taille des corpus à traiter. C'est là d'ailleurs une des particularités des documents pour le Web, et plus généralement des documents électroniques, que d'imposer et de permettre à la fois des manipulations de systèmes de signes nouveaux et des manipulations nouvelles de signes.

De nouvelles inscriptions

Ainsi, l'une des caractéristiques des pages et des sites internet est qu'ils comportent généralement des liens à des ressources externes et sont eux-mêmes la cible de liens externes. Ces liens s'apparentent aux citations dont sont pourvues les publications universitaires, mais possèdent d'autres propriétés et d'autres fonctions. Ils sont évolutifs, ce qui permet, pour peu que l'on dispose de relevés espacés dans le temps, d'en constituer un historique. Ils constituent, du point de vue de l'utilisateur, des balises permettant l'élaboration d'un cheminement et peuvent être compris, si l'on se place du point de vue du concepteur du site, à la fois comme des appendices à son propre discours, qui contribuent à définir le destinataire attendu du site, la position du locuteur (l'impartialité peut ainsi s'exprimer par le renvoi explicite à des sites illustrant des positions opposées) ou bien l'appartenance (reconnue si les liens sont symétriques, revendiquée s'il est unidirectionnel) à une

configuration sociale. Ces dimensions ne sont bien sûr pas exclusives l'une de l'autre, non plus que notre analyse exhaustive. Nous pouvons cependant à la fois reconstituer des parcours de lecture offerts aux internautes et des analyses, ou des descriptions, de réseaux sociaux, qui peuvent être conçus comme unissant des individus, des organisations ou des productions graphiques. Si nous partons d'un exemple, nous pouvons montrer assez rapidement que l'analyse peut se révéler féconde.

L'un des sites auxquels renvoie rapidement la requête (expulsions + étrangers + France) est un blog [Delaunay, s.d.] titré Immigration-Escroquerie-Sentimentale. Ouvert en novembre 2006, d'après la liste des posts précédents, cependant que le code source de la page donne comme date de création du document le 26 février 2006, il est encore actif en novembre 2008. L'en-tête de celui-ci précise que l'auteur en est une femme, Marie-Annick Delaunay, par ailleurs auteure d'un ouvrage consacré aux mariages blancs entre Françaises et étrangers [Delaunay, 2006]. Le site de l'éditeur¹, auquel renvoie le blog, nous apprend qu'elle est également la présidente d'une association, Non aux mariages et paternités de papiers, créée le 5 juillet 2004 et qui a elle-même un site internet². La consultation du code source de celle-ci, simple page HTML statique, donne comme date de création le 30 novembre 2004. Les textes lisibles sur la page de garde du blog et le premier post définissent la posture de l'auteure, mue par le souci de protéger des femmes (et exclusivement des femmes), de pratiques pouvant leur porter préjudice. Son action est selon elle une action d'information et de prévention, « [son] livre met en garde contre les escroqueries sentimentales qui se multiplient et appelle à un changement pour y mettre un terme »³.

La dimension politique de l'entreprise est cependant facilement repérable, au moyen d'une simple analyse des textes proposés à la lecture. Plusieurs de ceux présents en février 2009 sur la page de garde mettent directement en cause la politique menée par les gouvernements de gauche, coupables d'avoir facilité la tâche des escrocs. L'en-tête du site précise ainsi que « depuis la loi RESEDA de 1998 du gouvernement Jospin, on a assisté à une augmentation sans précédent des mariages mixtes contractés avec des citoyens français »⁴. Elle dénonce de plus vigoureusement

1. < www.tatamis.fr/virtuelle.php/id/111213 >, en février 2009, la page n'est plus accessible.

2. < 2jong.free.fr/associationnon >, vu le 01/04/2015.

3. < http://immigration-escroquerie-sentimentale.blogspot.fr/2006_11_01_archive.html >, en mars 2009, la page n'est plus accessible.

4. < <http://www.immigration-escroquerie-sentimentale.blogspot.fr> >.

l'indifférence des associations et institutions proches de la gauche française engagées dans la défense des droits des étrangers, écrivant ainsi que : les associations anti-racistes sont bien silencieuses chaque fois qu'un crime d'honneur, une séquestration ou des représailles à coups de couteau sont pratiqués contre une jeune fille obligée de se soumettre aux diktats de la famille, voire de la tribu. Elles sont aveugles, sourdes et muettes... la Haute autorité de lutte contre les discriminations et pour l'égalité (Halde), la Ligue des droits de l'homme (LDH), le Mouvement contre le racisme et pour l'amitié entre les peuples (MRAP), le Groupe d'information et de soutien des immigrés (GISTI), le Comité inter mouvements auprès des évacués (la Cimade)... et tous les ayatollahs de l'antiracisme⁵.

La page de garde du blog comporte, fin février 2009, six liens externes. L'un mène au site de l'éditeur de l'ouvrage mentionné plus haut, l'autre à celui de l'association dont l'auteure est la présidente. L'un conduit au site d'une association, L'enfant d'abord⁶, qui se présente, elle aussi, comme une association apolitique, soucieuse du bien des enfants et qui mène un combat contre les modalités de garde alternée prévues par la loi de mars 2002. Attentifs aux débats parlementaires, auxquels renvoie le site, les dirigeants de cette association ont obtenu en juin 2008 une audience auprès de la ministre de la Famille. Trois autres liens renvoient à des articles parus dans la presse et disponibles sur Internet, deux sont parus dans le *Figaro*, un dans *La Dépêche du Midi*. Le dernier lien est destiné à mettre en garde les Françaises se rendant en Afrique, en leur offrant de lire ce que l'on pense d'elles au Sénégal. Le texte auquel le lien renvoie⁷, d'une vulgarité crasse, évoque la « grosse conne européenne » qui :

« [...] n'est généralement pas une "fille facile" [mais] ne dédaigne pas coucher avec un "Africain" qui évidemment rêve de rester au Sénégal et n'a aucune arrière-pensée. Elle s'en sort au mieux avec des mycoses sur la chatte au pire avec un DAS qu'elle aura bien mérité. Au moins à son retour, elle coûtera cher à la sécurité sociale mais pas longtemps et ne coûtera rien à la collectivité pour sa retraite. »⁸

5. *Ibid.*

6. < www.lenfantdabord.org >.

7. < http://www.senegalaisement.com/senegal/consconnes_senegal.php >.

8. *Ibid.*

La cible du texte est politiquement définie par un petit texte en forme de question qui, placé en pied de page, précise l'acception des termes employés :

« Je suis un rebelle, j'ai campé dans le Larzac à l'appel de José Bové, suis-je un ? J'ai bien peur que oui. L'altermondialiste anarchiste est quasiment toujours un gros con persuadé de faire partie des gentils de cette planète. [il] s'est généralement fait taper sur la gueule "par des éléments incontrôlés du 9.3" lors des manifs contre le CPE et ce n'est que justice. »⁹

Le ton et les procédés linguistiques employés évoquent un anarchisme de droite à la française, qui lorgne autant vers Frédéric Dard que vers Céline. Ce texte peut être obtenu depuis un site qui se présente comme une encyclopédie en ligne du Sénégal et un site d'information utile tant aux Sénégalais de l'étranger qu'aux touristes prévoyant un séjour dans ce pays. Les descriptions des parcs nationaux y côtoient les listes d'astuces permettant d'obtenir des billets d'avion bon marché. Un article, préservé par une archive numérique sénégalaise [Presse sénégalaise, 2003], nous apprend que l'auteur, un Français propriétaire du site, a été condamné par contumace pour diffamation au Sénégal et, résidant en Espagne, est à l'abri des poursuites. Plusieurs de ses textes sont repris par les auteurs de sites personnels ou de blogs qui souvent, même lorsqu'ils sont rattachés au très sérieux journal *Le Monde*, pensent qu'il s'agit là de l'expression d'une opinion répandue parmi les Sénégalais¹⁰. Un certain nombre de sites et de forums, exprimant explicitement des thèses d'extrême droite, reprennent également ces textes¹¹.

Le dernier lien enfin renvoie au site personnel¹² d'un autre journaliste et auteur, André Bercoff, qui propose en ligne de courtes séances vidéo le mettant en scène commentant l'actualité du moment. Son blog comporte une dizaine de liens externes dont la plupart renvoient aux sites de médias généralistes et à quelques sites au positionnement idéologique brouillé, tel celui d'une Riposte laïque dont la page de garde arbore une belle cocarde tricolore et défend une république laïque et sociale dont l'un des héros est le parlementaire hollandais Geert Wilders, célèbre pour ses propos

9. *Ibid.*

10. < <http://coirault-neuburger.blog.lemonde.fr/2007/12/09> >.

11. < <http://www.france-echos.com/> > en mars 2009, le site n'est plus accessible aujourd'hui.

12. < <http://andrebercoff.com> >.

islamophobes [Hilout, 2009]. Les liens entrants¹³ sont plus difficiles et plus longs à analyser, du fait de leur nombre (90 en février 2009). La signification de ceux-ci, et donc leur interprétation et les manipulations des liens repérés, est sujette à variation dans le temps. La multiplication des blogs ainsi a pour conséquence qu'un webmestre peut obtenir un lien vers son site en allant lui-même inscrire sur le blog d'un confrère un lien vers celui-ci. Les contributions étant cependant modérées par le propriétaire du site, nous pouvons considérer que qui ne dit mot consent et que le maintien du lien vaut assentiment à l'arrimage des deux blogs.

Quelques traits peuvent cependant être dégagés. Ces liens émanent de sites plus ouvertement politiques que le site auquel ils renvoient, soit des sites souvent de militants actifs, d'élus ou d'organisations appartenant soit aux franges les plus conservatrices de la droite de gouvernement, soit aux diverses sensibilités de l'extrême droite. Radio Courtoisie, un blog de la droite nationaliste wallonne¹⁴, ou un site estampillé Mouvement national républicain (MNR)¹⁵, dirigent l'internaute vers le site de l'association, ainsi qu'un fort intéressant répertoire de sites à l'usage des droites européennes, proposant l'accès à 1 000 sites francophones de réinformation permettant de lutter contre :

« Les idéologies imposées par les médias [qui], sous différentes variantes, se résument essentiellement à deux idéologies principales : le socialisme et le mondialisme/anti-conservatisme/anti-christianisme. La première idéologie veut imposer les utopies étatiques que sont l'«anti-discrimination» et le nivellement par l'égalitarisme et le collectivisme au moyen d'un État tentaculaire et omniprésent, mais souvent irresponsable avec son hyper fiscalité et ses bureaucraties pléthoriques nationales et bientôt européennes. »¹⁶

Ces données nous permettent alors plusieurs constats. L'un tient aux stratégies d'énonciation de l'auteur du blog. Celui-ci met en circulation, ou relaie, un certain nombre de figures ou de récits qui confortent une idéologie qui est déniée en tant que telle, d'où une stratégie d'établissement de liens externes qui renvoient à des ressources documentaires garantes

13. Le module Inlinks de Yahoo, le plus efficace alors, avait été utilisé en 2009.

14. < <http://reconquista.over-blog.com> >.

15. Le Mouvement national républicain est un parti d'extrême droite créé en 1999 à la suite d'une scission du Front national.

16. < http://www.esprit-europeen.fr/agora_liens_1000_sites_francophones > en mars 2009, le site n'est plus accessible aujourd'hui, mais est référencé par Internet Archive.

du sérieux de l'information, mais à l'affichage idéologique discret ou délibérément brouillé. À l'inverse, les webmasters des sites émanant – et se présentant comme tels – de l'extrême droite francophone renvoient à ces récits, que leur positionnement définit comme objectifs et neutres, afin de conforter par des faits des postures idéologiques. L'exemple suggère une stratégie de manipulation de l'opinion et la coordination – qui ne suppose ici ni inter-connaissance ni délibération – de groupes institués et de francs-tireurs idéologiques déposant discrètement des toxines électroniques. Leur zone de diffusion correspond à un espace linguistique et non à un espace national. Le constat appelle bien sûr d'autres questions, tenant à la fréquence et à l'efficacité de ce travail idéologique, auxquelles cependant nous ne pouvons répondre sans changer d'échelle et de mode d'observation.

Si, en effet, les indices obtenus par les manipulations permises par l'observation des inscriptions que sont les liens externes des sites examinés sont pertinents, pour le sociologue ou l'historien et, combinés à d'autres, manifestent la possibilité d'enquêtes intégrant des sources électroniques, nous nous heurtons assez vite à trois limites. La première tient au fait que l'espace ainsi défini n'intègre pas de dimension temporelle. Les sociologues ou les spécialistes d'autres sciences sociales y pallient parfois en se faisant les archivistes de leur propre recherche, effectuant des coupes à intervalles plus ou moins réguliers, ou bien enregistrant leurs données sous la forme de flux d'informations datées [Granjon, 2002]. L'historien ne peut procéder ainsi qu'autant qu'il peut s'appuyer sur une archive électronique constituée. La seconde tient au nombre de manipulations à effectuer si nous voulons cartographier, non pas l'espace que définit un site, mais celui des prises de position sur une question ou des configurations définies par des renvois de site à site. Le relevé et le suivi manuel des liens repérés sont matériellement impossibles.

Il faut alors soit renoncer, soit se doter d'une instrumentation adaptée. Celle-ci existe [Park et Thelwall, 2005], est parfois facilement accessible et gratuite, mais elle n'est pas conçue pour les besoins des historiens. Certains de ces outils ainsi n'intègrent pas de modules permettant de prendre en compte les évolutions dans le temps des espaces décrits [Rogers, 2013]. Bien peu d'historiens de plus pourraient prendre en main ces instruments, d'abord parce qu'ils nécessitent quelques notions d'informatique, surtout parce qu'ils ne parviennent à proposer des descriptions et des représentations graphiques des configurations décrites qu'au prix d'une formalisation mathématique complexe (en l'occurrence les techniques de description des grands graphes) et constamment renouvelée.

Or, de toutes les sciences sociales, l'histoire est sans doute celle qui initie le moins ses étudiants et ses futurs maîtres à la culture mathématique et informatique. Jean Philippe Genet, constatant que « former les historiens à l'informatique n'est pas facile » [Genet, 1994], en exposait il y a près de quinze ans les raisons, qui n'ont au fond guère changé depuis.

Expressions régulières et interprétation

Ce sont là des difficultés que nous retrouvons, sous des formes plus ou moins aiguës, non seulement quand nous entendons utiliser les traces propres au Web, mais encore quand nous voulons exploiter les possibilités de recherches et de manipulation de chaînes de caractères qu'il autorise, en tant qu'il est un très grand corpus textuel ouvert. La recherche de syntagme ou de tropes, sous ses différentes formes, génétique ou bien de propagation, est une pratique érudite ancienne et efficace. Ce que permet ici le Web est de la mener à partir d'un corpus de très grande taille, ouvert, c'est-à-dire non *a priori* défini par le chercheur et sans limites physiques pesant sur le nombre de tropes identifiables et à partir d'un matériau – citation et reproduction étant des caractéristiques de l'écriture pour le Web – autorisant l'assignation d'une signification à ces pratiques en les rapportant aux pratiques sociales et aux propriétés des acteurs. Prenons un exemple simple. Nous pouvons considérer qu'une partie de l'action militante d'un certain nombre d'organisations consiste en la mise en récits de processus sociaux ou politiques, destinés, par leur dissémination, à provoquer une mobilisation ou à transformer des façons de nommer ou de classer certains types d'événements. Une analyse de ces récits peut conduire à mettre en évidence des chaînes de caractères les signalant et autoriser à évaluer leur diffusion à partir d'un point d'origine identifié. Un blog, rattaché au journal *Libération*, est ainsi utilisé par des proches ou des membres du Réseau éducation sans frontières¹⁷ qui y déposent de petits textes, souvent de courts récits, qui suscitent, ou non, des commentaires, généralement anonymes, ce qui bien sûr rend difficile leur interprétation. Il est possible par contre de suivre le devenir de certains de ces textes sur la toile au moyen de quelques mots-clés. La petite histoire narrée par Martine Vuailat en juin 2007, qui évoque sa visite dans un centre de rétention, a ainsi été abondamment reprise et commentée sur

17. < <http://www.educationsansfrontieres.org> >.

la toile. Rendant visite à une famille placée dans un centre de rétention, elle avait :

« Avant d'entrer dans la cellule de visite, [...] laissé sur le comptoir un livre d'enfant qui s'appelle *Au panier* [...] en disant avec un grand sourire "ça vous fera de la lecture". (C'est une histoire d'hommes en uniforme qui mettent au panier tout ce qui est différent). » [Arrighi, 2007]

Les fonctionnaires en poste ont semble-t-il peu apprécié l'attention, au point que « le lieutenant Sèvre [...] m'a dit qu'il me confisquait l'ouvrage pour subversion » [*ibid.*]. Retenue pour vérification d'identité, la locutrice s'est vu retirer l'autorisation de visiter le centre. Après plusieurs essais, il apparaît que la combinaison de chaînes de caractères la plus efficace pour retrouver trace de cette fable véridique est (vuailat + rétention + visite + livre). La requête effectuée sur Google permet d'atteindre plusieurs centaines de pages évoquant l'épisode. Nous ne pouvons bien sûr en mener une analyse systématique dans le cadre de cette étude. Notons que l'on retrouve trace, sans surprise, de sites émanant d'associations et de forces politiques engagées dans la défense des droits des étrangers, mais aussi de nombreux blogs et sites personnels dont les auteurs expriment souvent leur rapport au récit par un retitrage de celui-ci (nous trouvons *Le lieutenant qui censurait les livres, Attention subversion subversive ou courageuse* et bien d'autres) qui fait souvent de la censure d'une publication subversive par un officier (et non des conditions de vie et de séjour dans les centres de rétention) le pivot de l'historiette et l'objet du scandale. Nous pouvons ainsi là encore produire des sources permettant de qualifier une stratégie de communication et réfléchir à ses effets. Pour revenir à nos préoccupations premières, Internet nous offre ici les moyens d'une très classique analyse de réception, qu'il facilite et infléchit en permettant le repérage très rapide d'une masse de textes reprenant un motif, dont beaucoup sont produits par des locuteurs s'exprimant en tant qu'individus, sans pour autant que ne soient levées les interrogations tenant à la nature du corpus ainsi constitué, ni à la signification des prises de parole dont nous retrouvons la trace électronique.

De plus, là encore, si nous voulons changer d'échelle, c'est-à-dire produire une table ou une carte de distribution d'un certain nombre de syntagmes que nous pourrions rapprocher d'une table de liens, soit construire un dispositif savant inédit dont nous pouvons attendre des effets cognitifs,

il nous faudrait passer à la fois par l'explicitation de règles de manipulation (extraction, association, tri) des expressions régulières et des symboles graphiques présents dans les pages visualisées et passer, soit par l'écriture d'un algorithme, soit par l'usage d'un outil logiciel adapté. Cela suppose à la fois des compétences et des choix lourds d'enjeux, en particulier par cela qu'ils impliquent ou induisent sur les formes de division du travail et de coopération au sein de la communauté scientifique. Ils ne sont pas sans effet non plus en matière de formation ou de coopération transdisciplinaire. Le débat entre tenants d'une production locale d'outils logiciels adaptés aux besoins d'une communauté scientifique et partisans d'une appropriation d'outils génériques est ancien et récurrent et ce dans tous les champs des sciences humaines [Chateaufort, 2003].

STRATÉGIES HISTORIENNES

+++++

L'instrumentation de l'histoire

En l'état actuel des choses donc, le Web en tant que source, du fait de ses caractéristiques et de sa faible profondeur, peut rendre à l'historien du très contemporain de signalés services et promet aux historiens futurs une masse d'information dont l'exploitation pose de sérieux problèmes à qui ne veut pas renoncer aux exigences de la discipline. Nous ne pouvons, en tant qu'historien, guère que proposer quelques pistes de réflexion et quelques éléments de stratégie. Le premier constat est que les matériaux qu'offre le Web ne peuvent être maîtrisés par l'historien, c'est-à-dire manipulés selon les règles du métier et exploités dans toute leur richesse, qu'au prix du renouvellement de son instrumentation. L'un des intérêts de l'émergence du document électronique et de sa distribution en réseau est sans doute d'ailleurs de nous permettre de repenser le travail de l'historien et plus largement celui des spécialistes de sciences humaines, en le redéfinissant comme une manipulation instrumentée et réglée de symboles connectés à des traces matérielles. Les fondateurs de la discipline le savaient pertinemment qui, tels Langlois et Seignobos, ne passaient guère moins de temps à discuter de la meilleure façon d'établir et de tenir les fiches permettant l'exploitation de leurs documents-textes, qu'à réfléchir aux exigences de la critique [Seignobos, Langlois, 1992].

Adapter ces principes aux documents électroniques suppose, plus que l'achat de machines, que les historiens, une partie d'entre eux du moins,

aient quelques teintures d'une culture informatique et d'une culture mathématique qui leur permettent sinon de construire leurs propres outils, du moins de dialoguer avec des scientifiques et des ingénieurs, sous peine de devoir utiliser des dispositifs élaborés sans tenir compte de leurs besoins propres ou de ne pouvoir exploiter une bonne partie des matériaux produits par les sociétés du très contemporain. Nous retrouvons là une thématique familière aux praticiens de l'histoire quantitative et de l'informatique historique [Coll., 1995], à ceci près, et le changement est d'importance, que la question de l'instrumentation de l'historien n'est pas posée à une minorité de spécialistes traitant de questions souvent jugées absconses par une bonne partie de la profession, mais à tous ceux qui entendent ou entendront produire une connaissance du très contemporain exploitant le type d'inscriptions laissées par ses institutions, voire tous ceux, et ils seront de plus en nombreux, qui accéderont aux sources nourrissant leur recherche par le biais de documents électroniques.

La tâche est rude. Elle suppose que la formation initiale des historiens fasse une part plus grande à la culture scientifique et technique, ce qui ne pourra se faire que lentement et suppose une réflexion collective [Dallo, 2002], mais aussi qu'émergent quelques filières permettant la formation de praticiens solidement spécialisés dans l'élaboration de l'instrumentation scientifique, dont le travail soit reconnu pour ce qu'il est, un travail scientifique dont la communauté historienne a collectivement besoin, si nous voulons éviter de dépendre de l'apparition aléatoire de générations spontanées composées d'individus aux parcours atypiques, campant dans des positions marginales, ce qui est la règle, en France du moins, depuis plusieurs décennies. Elle le sera aussi parce que, contrairement au livre, dont les propriétés sont connues et stabilisées depuis longtemps, le document électronique demeure un devenir, le produit transitoire d'une évolution en cours. Les compétences qui permettent son recueil et son exploitation sont aussi coûteuses à acquérir que rapidement obsolètes, supposant, pour être constamment reconquises, plus l'accès à une culture scientifique que la maîtrise de quelques procédures rapidement abandonnées.

L'urgence archivistique

Ces mutations rapides posent un autre problème, qui est celui des conditions de l'archivage du Web et plus largement des documents sous forme électronique. La période actuelle se caractérise par la superposition

d'initiatives publiques et privées destinées à préserver ceux-ci, sans cependant que la question de la pérennité à long terme des ressources électroniques, qui pose des problèmes tant matériels que logiciels, ne soit réglée. La participation des historiens à la réflexion, déjà entamée depuis plusieurs années par les documentalistes et les archivistes apparaît nécessaire, de même que leur contribution, malheureusement trop rare, aux initiatives visant à produire, à partir de la définition d'un objet d'étude ou d'un cadre thématique, des recueils de données électroniques. Quelques exemples, mais trop peu nombreux, existent, tel le projet porté par le Centre d'histoire sociale du XX^e siècle afin de permettre l'étude future des nouveaux mouvements sociaux de la période contemporaine, dont l'action souvent s'inscrit d'abord sur des supports électroniques ne prenant pas toujours la forme de pages HTML destinées à une diffusion par le biais du Web [Blum *et al.*, 2005]. Refusant le fantasme contemporain de l'archivage universel, qui méconnaît le fait que l'archive est toujours à la fois le réceptacle de vestiges ayant échappé à un processus de destruction et le produit de choix, le projet prévoit l'identification d'acteurs clés, le recueil et l'indexation de leurs productions numériques.

Poser les questions auxquelles il est possible de répondre

La réflexion engagée, parce qu'elle est œuvre d'archivistes et de documentalistes, intègre à la fois une connaissance de l'objet appréhendé et une réflexion sur les particularités du matériau. À l'historien qui voudrait déjà, étudiant le très contemporain, s'emparer du matériau produit par l'intermédiaire d'Internet, elle offre des pistes utiles. Si l'entreprise en effet est possible, c'est parce qu'elle définit une stratégie savante qui rompt avec les pratiques spontanées de l'internaute en définissant des entités dont il est pertinent d'organiser la préservation – en collaboration avec des institutions publiques, les acteurs eux-mêmes et des historiens – et dont il serait possible de poursuivre l'étude. Le projet part en effet d'une connaissance préalable d'un champ d'intervention sociopolitique, qui permet le repérage et le choix d'acteurs que caractérisent à la fois leur usage intensif des outils électroniques et leur position au sein du champ. Cela fait de leur production électronique l'un des éléments d'un dossier portant trace de leur activité, soit un fonds d'archives, attaché à une entité dont les propriétés peuvent être définies. L'historien retrouve un cadre connu, qui est celui qui permet son activité.

Certaines des difficultés que nous avons évoquées en ouverture de cet article provenaient en effet de ce que le Web n'est pas une archive, même s'il est en partie archivable et donne accès à des archives, quoique la permanence de celles-ci ne soit pas toujours certaine. À ce titre, il est utilisable, en conjonction avec d'autres sources, par les historiens du très contemporain qui n'ont d'ailleurs guère le choix. Ce matériau impose cependant des contraintes spécifiques, qui obligent à définir des stratégies d'enquête et des tactiques manipulatoires adaptées, tant aux questionnements propres de l'historien qu'aux caractéristiques de son matériau, dont nous savons qu'elles sont parfois inédites et souvent fluctuantes. Désorientation et surcharge cognitive guettent souvent l'utilisateur de données complexes et hypertextuelles et l'historien ne fait pas exception à la règle, qui s'y aventure sans le secours d'une diplomatique exprimant une tradition érudite ni d'une sémiotique stabilisée. Il est loin d'ailleurs d'être le seul en ce cas pour ce qui est du dernier point. Les spécialistes de l'architecture des sites insistent ainsi sur la rapidité avec laquelle se transforment les modes d'organisation de présentation et de visualisation des sites web, qui sont l'expression à la fois d'une visée et d'une organisation productive [Kahn, Lenk, 2001]. Si un usage raisonné de ces ressources suppose l'appropriation, coûteuse, d'une instrumentation nouvelle, nous pouvons en attendre des bénéfices cognitifs, dont les moindres ne sont pas l'invitation à une réflexion sur nos propres pratiques et la possibilité de demeurer fidèle à la vocation de l'historien, qui est, par la manipulation inventive et réglée d'inscriptions, la production d'un savoir à la fois critique et réflexif.

+++++

CHAPITRE XI. ÉCRITURE DE L'HISTOIRE ET RÉSEAUX NUMÉRIQUES

+++++

La première version de ce texte est parue dans un ouvrage dirigé par Éric Guichard [Guichard, 2011], préparé à l'occasion du quinzième anniversaire de l'Équipe Réseaux-Savoirs-Territoires de l'ENS. Il s'inscrit dans le cadre de l'activité d'une équipe pluridisciplinaire dont l'horizon commun est constitué par les transformations des pratiques savantes à l'heure d'Internet. Contrairement à la plupart des textes réunis ici, il s'adressait à un public qui n'était pas pour l'essentiel composé d'historiens.

« La volonté de définir idéologiquement l'histoire est le propre d'une élite sociale. Elle se fonde sur une division entre les idées et le travail. Il lui était habituel de négliger également le rapport entre l'idéologie des historiens et leurs pratiques, entre les idées et leur localisation ou les conditions de leur production dans les conflits socio-économiques d'une époque, etc. »

Michel de Certeau, *L'écriture de l'histoire*, 1902

P

lusieurs personnes m'ont, au cours de ces dernières années, souvent à l'occasion d'un colloque, demandé ce que l'usage par les historiens de l'internet changeait à l'écriture de l'histoire, au prétexte que j'étais historien et que j'avais fait partie des premiers spécialistes français d'histoire contemporaine à utiliser systématiquement le réseau et à produire des contenus destinés à celui-ci. Éric Guichard, préparant ce volume, m'a posé à son tour la même question. Je me suis donc de nouveau trouvé profondément embarrassé. Je ne pense pas en effet être aujourd'hui la personne la plus qualifiée pour y répondre. De plus la question, limpide dans sa formulation, est extrêmement complexe et aussi difficile à traiter que malaisée à éluder. L'histoire en effet, telle que la pratiquent ceux qui font métier d'historien, se veut, et est, une entreprise scientifique, ou du moins, si nous voulons éviter les querelles sémantiques

liées à l'emploi de cet adjectif, entreprise de connaissance qui entend produire du savoir. Or, praticiens et épistémologues de plusieurs disciplines ont montré, au cours de la dernière décennie, que l'irruption de l'internet avait eu des effets repérables sur le volume et les caractéristiques des résultats scientifiques produits [Gauthier *et al.*, 2000]. Il est donc légitime que l'on demande à l'historien ce qu'il en est dans son champ et parce que de fait il présente au public le produit de ses travaux au moyen d'agencements structurés de signes, et le plus souvent de textes, que la question prenne, quand elle lui est adressée, la forme d'une interrogation sur les éventuelles transformations de l'écriture historique.

Les difficultés commencent alors parce que, sous un même énoncé, liant histoire, écriture et réseaux informatiques, se trouvent rassemblées des interrogations distinctes. Il s'agit pour certains de se demander si les historiens, et particulièrement les historiens de profession, écrivent pour les usagers de l'internet et sous quelle forme [Cricione *et al.*, 2004]. D'autres considèrent que l'internet offre à l'historien du temps présent un matériau nouveau, un gisement de sources dirait un historien, et s'interrogent sur ce qu'il peut en faire et dire de neuf à partir de celui-ci et à quelles conditions¹. D'autres encore, constatant que les historiens utilisent aujourd'hui des ordinateurs connectés à l'internet dans le cadre de leurs activités professionnelles posent la question de l'incidence de ces pratiques nouvelles sur les productions des historiens et leur diffusion [Doueïhi, 2008]. Quelques-uns enfin s'interrogent sur les propriétés des discours historiques, ou se présentant tels, qu'il est possible de lire sur la toile, l'interrogation porte alors sur les transformations des usages publics de l'histoire à l'ère numérique [Noiret, 2005]. Ces questions, logiquement distinctes, mais toutes valides, peuvent être liées si l'on considère qu'elles renvoient chacune à une facette d'une même activité professionnelle, d'un métier [Minuti, 2002]. L'historien collecte et traite des matériaux, tirés ou non d'archives constituées, afin de produire des discours et des textes, qui sont pour certains plus particulièrement destinés à ses pairs, pour d'autres offerts à un public plus large, entrant alors parfois en concurrence avec d'autres discours, ou du moins cohabitant avec eux. Il arrive que les produits de son activité soient disponibles sur la toile. En somme il cherche, il écrit, il prend la parole, il met en ligne (ou on le fait pour lui). Ajoutons que souvent il enseigne et que c'est fréquemment par le biais de leurs activités pédagogiques que beaucoup d'historiens, particulièrement ceux

1. Voir dans ce volume, « De quoi le Web est-il l'archive ? ».

exerçant dans les universités anglo-saxonnes, furent pour la première fois confrontés à l'internet, voire sommés de l'utiliser [Brabazon, 2002].

S'interroger sur la transformation de l'écriture de l'histoire à l'ère du réseau revient alors à étudier les mutations d'un métier contemporaines de la transformation des dispositifs techniques auxquels il est lié. Je ne suivrai pas ici ce chemin, ou pour mieux dire ce questionnaire, quoique je le juge fécond. Plusieurs raisons m'y poussent. D'une part c'est là le sujet d'une thèse, appuyée sur un solide travail de terrain, plutôt que le thème d'un article, qui ne pourrait guère faire mieux que coudre ensemble quelques impressions directement dérivées de ce que la place du locuteur lui aura permis d'observer. D'autre part, j'ai déjà évoqué certaines de ces questions dans divers textes au cours de ces dernières années, et ma réflexion n'a guère avancé depuis la parution de ceux-ci, les hasards d'une carrière m'ayant quelque peu détourné de ces interrogations. Je voudrais ici seulement réfléchir aux pratiques informatiques des historiens et à leur pratique du réseau lorsqu'ils sont les plus proches du cœur de leur métier d'historien, c'est-à-dire lorsqu'ils sont engagés dans la production d'un texte qui, appuyé sur un dépouillement d'archives qui ne sont pas nées numériques, sera soumis à l'évaluation de leurs pairs, et aux implications de ces pratiques sur le produit de cette activité. Une telle posture appelle plusieurs remarques. Parler ici du cœur d'un métier ne veut pas dire que les historiens en poste dans les facultés passent l'essentiel, ni même beaucoup de leur temps, à scruter les archives, à hanter les bibliothèques et à préparer des textes inédits, mais ce travail est à la fois ce qui les différencie de ceux de leurs collègues qui interviennent dans d'autres champs et le seul élément de leurs activités qui est soumis aux jugements de leurs pairs. Dans la plupart des cas, ce sont ces travaux qui fondent l'identité professionnelle d'un historien. D'autre part, puisque nous nous intéressons ici à des pratiques, nous regroupons sous le terme d'historien l'ensemble des individus, quel que soit leur statut au regard des institutions universitaires, qui publient des textes, ouvrages ou articles, dans des collections ou des revues d'histoire savante, c'est-à-dire ceux dont le travail est reconnu pertinent par les instances de validation de la profession.

Si délimiter ce sujet est aisé, le traiter en historien est particulièrement difficile. L'avènement de l'internet ne constitue pas une nouveauté radicale, mais le prolongement d'un processus d'informatisation des activités intellectuelles vieux déjà de plusieurs décennies et loin d'être achevé. Ce que nous tenons pour les propriétés essentielles et intangibles de l'internet est ainsi le produit d'un équilibre politique autant que technique sans

cesse renégocié entre de multiples acteurs aux intérêts souvent divergents [Mounier, 2002]. De plus, les historiens sont peu portés à nourrir de leur expérience personnelle une méditation, fut-elle heideggérienne, sur le nombre de degrés supplémentaires introduits entre eux et l'Être historique par l'écriture à la machine. Une question d'histoire se traite par l'examen de la bibliographie existante et la délimitation d'un corpus à examiner. Or la bibliographie est ici quasi inexistante, particulièrement quand on s'intéresse aux pratiques des historiens du domaine francophone. Cela exprime sans doute le fait que la généralisation de l'usage de l'internet est récente, mais aussi une certaine absence de curiosité touchant aux systèmes de pratiques mis en œuvre par les historiens, au profit souvent de la recherche d'une cohérence idéologique des œuvres [Certeau, 2002]. Les universités françaises offrent certes aujourd'hui à leurs étudiants en histoire des cours d'historiographie. Les manuels fleurissent donc, parfois fort bien faits. Cependant, s'ils proposent souvent une analyse des grandes œuvres, un panorama des grands auteurs, parfois une histoire de l'institution historienne, il est rare qu'ils accordent beaucoup d'attention aux pratiques concrètes des historiens, à leurs conditions de travail, à leurs outils ou aux transformations de leur équipement et de leur environnement. Les consultant, le lecteur peut avoir l'impression que l'historien, quoique parfois doté de vives passions, généralement idéologiques, n'a pas plus de corps que d'outils et qu'aucun quotidien n'entrave ce travailleur résolument solitaire et sans technique.

UNE APPROPRIATION TARDIVE ET LENTE

+++++

Qui veut faire un état de la question qui nous occupe ici doit donc faire un détour et examiner les quelques travaux disponibles, élaborés dans le cadre américain. Deborah Lines Andersen, une documentaliste de formation, a mené plusieurs enquêtes, à quelques années d'intervalle, sur les usages de l'informatique des historiens présents dans les universités new-yorkaises [Andersen, 2003a]. Elle nous fournit quelques indications, à partir d'enquêtes par questionnaires et de campagnes d'entretiens, sur l'évolution des usages entre 1992 et 2003. À la lire, nous constatons que l'ordinateur fut d'abord pour les historiens une machine à écrire, avant de devenir le moyen d'envoyer des courriels, puis de consulter le Web. En 1996 cependant, ces pratiques sont toujours minoritaires, et aucun des historiens consultés n'a encore produit de contenu destiné au réseau.

Il existe cependant un petit groupe d'historiens, qui souvent ont une fonction de passeur, qui utilisent très fréquemment les ressources du réseau, et pour certains un tableur, voire un logiciel de base de données (3 % des historiens interrogés ont utilisé un tableur au cours du mois précédent l'enquête et 13 % une base de données, sans que nous sachions s'ils ne font que les consulter ou s'ils les créent [Andersen, 1998]). Une seconde enquête, menée celle-ci uniquement par questionnaire, à la fin de 2003, dans les mêmes lieux, montre que l'usage du courriel est désormais systématique. Beaucoup, parmi ceux qui ont répondu au questionnaire distribué, maintiennent un site web destiné à leurs étudiants et consultent les catalogues en ligne des bibliothèques ou utilisent le Web afin de rassembler l'information bibliographique dont ils ont besoin, voire consultent en ligne des articles savants au moyen des agrégateurs de contenus permettant l'accès aux publications universitaires (J-Stor, par exemple). La publication électronique demeure par contre extrêmement rare. Un seul des quarante-deux répondants indique avoir donné des notes de lecture à une revue électronique, aucun ne mentionne une participation à la production d'un CD-ROM, les plus jeunes exprimant la crainte qu'une production savante sous forme électronique ne soit pas prise en compte par les instances d'évaluation et de recrutement [Andersen, 2003b].

Il faut, afin d'apprécier ces résultats, savoir que la réalisation des premières enquêtes était liée à la mise en place d'un programme pilote de diffusion des outils numériques au sein des institutions étudiées. De ce fait, la diffusion de ceux-ci fut sans doute ici particulièrement précoce et large. Nous pouvons conclure que la diffusion des outils et des compétences numériques apparaît rapide, mais tardive et limitée dans son extension. L'historien new-yorkais, membre d'une institution d'avant-garde en matière d'informatisation des pratiques historiennes est, en 2003, un producteur de tapuscrits et un consommateur d'information numérique armé d'un navigateur, d'un logiciel de mail et d'un traitement de textes. L'ordinateur, même connecté au réseau, est, au milieu des années 2000, pour beaucoup d'historiens français, un substitut performant au minitel et à la machine à écrire, plus qu'un dispositif ouvrant des possibilités nouvelles. La multiplication des écrans et la croissance des réseaux rendent de plus l'historien fréquemment méfiant.

L'idéologie qui sous-tend le discours de beaucoup des promoteurs de l'internet y est sans doute pour quelque chose, qui promet la disparition des médiations de tout ordre et l'abolition de l'auteur et des hiérarchies cognitives [Breton, 2000, p. 63 et suivantes]. Mais ce n'est pas là la seule raison.

Pour beaucoup, l'informatisation du travail universitaire est d'abord le moyen permettant à leur institution d'obtenir d'eux un travail supplémentaire et non payé, la tenue d'un cours s'accompagnant de la diffusion de contenus par le biais d'une page web et de fréquents et chronophages échanges de courriels avec leurs étudiants. Les réactions de certains semblent parfois très familières à un historien des mondes ouvriers. Pour certains de ces travailleurs très spécialisés et très qualifiés que sont les historiens, l'introduction des nouveaux objets techniques que sont les dispositifs électroniques apparaît comme une menace. Elle porte en elle un risque de déqualification, de dépossession et mine les conditions de leur autonomie. Ces transformations sont vécues et dites parfois comme l'annonce d'un déclassement, qui ferait de professionnels proches par certains traits des professions libérales des salariés comme les autres de grandes organisations, de plus en plus soumises à une logique de profit [Brabazon, 2002]. Ces craintes ne sont d'ailleurs pas sans fondement. L'introduction des outils-réseaux au sein d'organisations spécialisées dans le traitement de l'information a de fait parfois conduit à « [...] un retour du travail à la chaîne, sans sens, sans compréhension globale du processus » [Frey, 2007, p. 47]. Et le développement, actuellement lent et difficile, de programmes d'enseignement en ligne peut mener l'historien à devenir un simple fournisseur de contenus intégrés au sein de dispositifs dont, faute de compétences techniques, il ne maîtrise pas l'architecture et que rien n'empêche une université de maintenir en ligne après que l'enseignant a quitté l'institution. De même, une dépendance accrue aux bases de données payantes (et fort chères) d'informations bibliographiques et d'articles scientifiques est susceptible de défavoriser les chercheurs qui ne seraient pas rattachés à une grande institution capable d'en supporter les coûts. Il n'est pas certain que ces réticences, soulignées par plusieurs auteurs [Saunier, 2005], soient propres à la communauté historienne, même si elles sont particulièrement accusées en ce cas. Couplées à une formation essentiellement littéraire – malgré la mise en place, non sans difficultés ni hésitations d'enseignements dédiés à l'outil informatique [Dallo, 2002] – et à la stabilité des usages au sein d'une communauté académique très anciennement formée, dont les membres se pensent souvent plus volontiers hommes de lettres que scientifiques, elles peuvent cependant rendre compte du développement tardif et partiel des usages de l'informatique au sein de la profession.

De même pouvons-nous comprendre que le débat sur l'informatisation de la discipline ait, particulièrement en France, surtout porté sur les modalités de la mise en ligne d'une information scientifique aux formes

quasi inchangées [Dacos, 1999], le fichier PDF remplaçant, ou redoublant, l'article papier. Le constat rend aussi un peu vaines les entreprises visant, partant de l'examen des productions scientifiques actuellement disponibles, à déduire les transformations imputables à la diffusion des usages du réseau. Celle-ci est tardive et limitée et les délais de production des documents rendant compte d'une recherche sont très longs. Entre le début d'une enquête, d'une thèse par exemple, et la publication sous forme d'ouvrage ou d'article de ses résultats, il s'écoule fréquemment de cinq à dix ans. Ajoutons qu'il semble que les historiens utilisant les ressources électroniques disponibles aient eu, au moins jusqu'au début des années 2000, tendance à gommer de leurs textes les références à celles-ci, ce qui rend le repérage de leurs pratiques et de l'incidence de celle-ci plus difficile encore [Graham, 2001] lorsque l'on utilise comme matériau les textes qu'ils livrent au public.

LES CURIOSITÉS DES CYBERHISTORIENS

+++++

Plusieurs voies s'ouvrent quand nous voulons dépasser ce premier constat. L'une consiste à examiner les pratiques, ou plutôt les récits de pratiques produits par les petits milieux d'innovateurs et d'expérimentateurs dont Deborah L. Andersen constate l'existence. Ceux-ci en effet, tant en France que dans le monde anglo-saxon, disposent depuis longtemps de leurs revues ou de leurs bulletins et y évoquent souvent tant leurs expériences que leurs difficultés. Certaines tentatives s'avèrent bien sûr sans lendemain, mais nous pouvons, examinant ce corpus, repérer quelques traits des curiosités et des désirs numériques des historiens. La revue de référence pour qui s'intéresse aux croisements entre pratiques historiennes et pratiques informatiques est, dans le monde de langue anglaise, le *Journal of the Association for History and Computing*². Cette revue électronique, dotée d'un comité de lecture, est née en juin 1998. Chaque livraison propose des articles rendant compte de réalisations ou annonçant des projets, un éditorial, des comptes rendus de lecture, des descriptions de ressources internet. La revue la plus durable fut, dans le domaine français, le *Médiéviste et l'ordinateur*, créée en 1979, sous l'égide de l'Institut de recherche et d'histoire des textes (IHRT), qui publie deux volumes papier jusqu'en 2004, puis devient uniquement électronique, en prenant la forme

2. < <http://quod.lib.umich.edu/j/jahc/> >.

d'une archive ouverte, alimentée pour la dernière fois en novembre 2006. Une numérisation rétrospective permet d'accéder aux volumes parus depuis 1989³.

Les deux entreprises ne sont pas tout à fait de même nature. Le *Médiéviste et l'ordinateur* est un bulletin destiné à des professionnels de l'histoire médiévale et son comité de rédaction se préoccupe d'abord des usages de l'informatique dans le cadre de la recherche. Le *JAHC*, qui n'est pas accolé à un laboratoire, s'adresse aux professionnels de l'histoire, quel que soit leur objet, ou la période qu'ils étudient, mais fait aussi une large place aux applications pédagogiques de l'informatique dans l'enseignement supérieur et le secondaire. En dépit de ces différences, la lecture de ces deux revues fait apparaître des éléments communs. Leurs rédacteurs expriment les mêmes convictions. Ils se définissent implicitement comme des pionniers, développant des pratiques, sinon marginales, du moins minoritaires, mais qu'ils considèrent appelées à devenir communes. Jeffrey Barlow, le premier directeur de publication du *JAHC*, évoque dans le premier éditorial la méfiance des historiens confrontés à l'internet et aux journaux électroniques avant d'affirmer que « les changements sont là. Et [le lancement d'une revue électronique] était aussi nécessaire qu'inévitable » [Barlow, 1998]. Cette position les conduit à trouver des accents souvent militants, parfois missionnaires. David J. Staley, président de l'American Association for History and Computing définit ainsi celle-ci comme une organisation visionnaire guidée par une idéologie qui la conduit à jouer un rôle pionnier dans l'exploration des nouvelles technologies destinées à la recherche et à l'enseignement de l'histoire [Staley, 2005]. Le style des rédacteurs du *Médiéviste et l'ordinateur* est généralement moins fleuri et l'appréhension des outils disponibles souvent plus critique. Il n'empêche que Lucie Fossier, longtemps âme de l'entreprise, considère en 1984 que « les médiévistes [...] se trouvent à la pointe du progrès » [Fossier, 1984] ; et les appels à l'action sont nombreux dans les pages de la revue. L'informatique y est conçue comme un ensemble de techniques que, sinon l'individu historien, du moins la profession historienne, doit prendre en compte, parce qu'elle est de fait entrée dans les mœurs et parce que les apports heuristiques de son usage sont indéniables. Son articulation aux pratiques érudites suppose cependant, outre une acculturation des individus, des transformations des institutions de la recherche et la mise en place d'un système de formation adapté.

3. < <http://lemo.irht.cnrs.fr/> >.

Le volume de 1995 titré *Les médiévistes et la politique de l'informatique*, qui s'ouvre par le compte rendu de rencontres avec les responsables français des politiques scientifiques, est révélateur de cet état d'esprit. Il se clôt par une postface qui ressemble à la feuille de route d'une organisation syndicale et demande « la mise en place de formations adaptées, une meilleure reconnaissance scientifique des pratiques informatiques des chercheurs et la mise en place d'un environnement permettant, dans le cadre des institutions de recherche, une vraie communication entre chercheurs en sciences humaines et sociales et informaticiens » [Coll., 1995].

L'évolution des sommaires des deux revues présente également des traits similaires. Les pionniers de ce que les Américains nomment la *digital history* avaient incorporé l'usage de l'informatique à leurs pratiques de recherche avant même le développement de l'internet, souvent pour explorer, par l'analyse statistique, des corpus assez volumineux structurés par des bases de données. Les titres des premiers volumes du *Médiéviste et l'ordinateur* en portent la trace. Le numéro 3 et le numéro 4 (1980) sont consacrés aux sources sérielles médiévales, le numéro 5 à l'analyse factorielle (1981), le numéro 7 aux procédures de classification automatique (1982). Malgré la stabilité de l'équipe en charge de la publication, l'évocation de ces usages se fait de plus en plus rare et ils disparaissent des sommaires après un dernier numéro consacré en 1991 aux bases de données structurées. Le *JAH* connaît une évolution similaire quoique décalée dans le temps. Les premières livraisons comportent quelques articles évoquant des recherches historiques faisant un usage systématique de l'outil statistique et des bases de données, souvent dans le cadre d'une histoire sociale dont sont issus plusieurs des piliers de l'entreprise, mais le thème disparaît quasiment après 2001. Le seul texte évoquant l'analyse statistique de données historiques assistée par ordinateur, publié en 2006, adopte de plus une perspective et un vocabulaire différents de ceux offerts par les premières livraisons. Karine Pellier, son auteur, évoque non pas les traitements statistiques effectués, mais le système de gestion de l'information (*data management*) mis au point au cours de l'enquête [Pellier, 2006].

Cela traduit un déplacement, sinon des pratiques, difficiles à saisir, du moins des interrogations, que confirme l'examen des sommaires des deux revues. En effet, la découverte du réseau conduit leurs rédacteurs à se pencher sur les transformations de l'archive et de la bibliothèque. La numérisation des ouvrages anciens et des collections archivistiques, les conditions de leur indexation et de leur mise à disposition, les transformations de l'édition scientifique et de la diffusion des résultats des travaux

érudits retiennent désormais, quoique de manière non exclusive, l'attention. Le *Médiéviste* débusque les textes médiévaux sur l'internet en 1999, se penche sur la numérisation des manuscrits médiévaux à l'automne 2001 et consacre un dossier à l'édition électronique en 2004. Le *JAHC* publie entre 2000 et 2008 huit⁴ textes traitant de l'archivistique à l'ère électronique, dont une bonne part – préoccupation partagée avec le *Médiéviste* – traitent des systèmes d'indexation contextuels des données de type XML qui en permettent l'utilisation. Les transformations des bibliothèques et des métiers associés à celles-ci sont au centre, entre 2000 et 2003, de sept textes, alors que là aussi le thème était absent des premières livraisons. Dans les deux cas, nous passons rapidement de textes prospectifs, ou de présentation de projets, à l'exposé de résultats de recherche ou à la description de dispositifs déjà matures et disponibles. C'est d'ailleurs là un trait qui caractérise l'ensemble des articles traitant de l'usage des ressources en réseau dans un contexte de recherche, dont un indice efficace est la distribution du terme future, qui signale une réflexion prospective : il apparaît cinq fois au sein des titres des soixante-neuf articles de la *JAHC* traitant de problématiques de recherche, ces cinq occurrences se produisent entre 1998 et 2003.

La période la plus récente est marquée par l'intérêt pour des systèmes de gestion d'information complexes, lointains descendants des bases de données des premiers temps de l'informatique historique et qui ont pour caractéristique, à la fois de ne pas être nécessairement pensés en lien à un traitement statistique et de pouvoir incorporer des éléments non textuels. Les systèmes d'information géographique (SIG) suscitent beaucoup de curiosité et sont de fait déjà utilisés par un certain nombre d'historiens. Le *Médiéviste et l'ordinateur* leur consacre son numéro 44 (2006) et plusieurs auteurs évoquent non pas les virtualités de leur usage, mais les résultats permis par leur utilisation lors d'études s'appuyant sur les cadastres et les terriers anciens. Le *JAHC* offre de même à ses lecteurs six articles, tous publiés entre 2006 et 2008, traitant des usages possibles par l'historien de systèmes d'information géographique, avec cette particularité que plusieurs auteurs évoquent leur utilisation dans le cadre d'un enseignement. Ces systèmes ne sont pas les seuls modes

4. Les remarques qui suivent s'appuient sur le dépouillement des sommaires de la revue. Chaque article, hors éditoriaux et compte rendu, a été décrit par une fiche qui identifie sa thématique principale et son type (réflexion théorique, présentation d'un projet, exposé des résultats d'une recherche), son titre, son ou ses auteurs, sa date de publication, et quand cela est pertinent, les outils informatiques évoqués.

de structuration informatique des données qui retiennent l'attention des rédacteurs de la revue. Ils s'intéressent aussi aux simulations tridimensionnelles (deux articles en 2003). De façon générale, les systèmes de gestion de l'information historique destinés aux chercheurs et aux équipes semblent aujourd'hui les dispositifs suscitant le plus d'intérêt, quoiqu'il n'en existe pas encore d'opérationnel malgré le lancement de divers projets, dont Historic Crossroads (histcross), qui s'appuie sur les possibilités de structuration des données par des réseaux sémantiques [Kalus, 2007].

UN DIFFICILE AJUSTEMENT À LA NOUVEAUTÉ PERMANENTE

+++++

Si nous voulons synthétiser ce mouvement, nous dirons que les enjeux identifiés par les pionniers des humanités électroniques historiennes furent d'abord les possibilités de manipulation des données historiques, par le biais particulièrement du calcul, puis ceux liés aux possibilités de numérisation et de distribution des données et de l'information historique et qu'ils s'interrogent aujourd'hui sur les possibilités offertes à l'historien désireux de structurer efficacement l'abondante matière première numérique offerte par les archives et les bibliothèques de l'ère numérique. Le constat est bien sûr à la fois simplificateur et biaisé, du fait d'effets de sources. Simplificateur parce que, si les accents se déplacent d'une période à l'autre, les urgences ressenties ne sont pas nécessairement les mêmes, les questions anciennes ne disparaissent pas et celles à l'ordre du jour ne sont pas toujours radicalement nouvelles. Manfred Thaler tentait déjà, il y a plus de vingt ans, d'élaborer un système d'information historique [Thaler, 1987].

De même, la réflexion sur l'exploration et la description statistique des données historiques, si elle retient moins l'attention des rédacteurs des deux revues utilisées, s'est poursuivie tout au long de la période, quoique les historiens soient sans doute moins nombreux à la mener au cours des années 1990 qu'au cours des années 1980 ou qu'aujourd'hui. Sa disparition des pages des deux supports de notre réflexion ne renvoie pas seulement à la dynamique des recherches, mais aussi, dans le domaine francophone du moins, à l'apparition en 1986 de la revue *Histoire & Mesure*, qui s'est imposée comme le lieu où s'expose et se réfléchit la nouveauté statistique. Si nous empruntons le vocabulaire des *literacy studies*, ou pensons aux catégories de l'anthropologue Jack Goody, nous dirons que les historiens électroniques – les petits groupes pionniers lancés dans l'aventure

numérique – se sont, depuis l'avènement de l'internet, intéressés aux transformations des systèmes d'écriture dont ils se servent, à ceux qui structurent la bibliothèque et l'archive, aux conditions de la diffusion de leurs textes et plus récemment aux dispositifs permettant d'organiser leurs écritures invisibles et quotidiennes, les traces qu'ils multiplient lorsqu'ils préparent une recherche, mais ne livrent pas au public et rarement à la communauté historienne. Ils ont par contre rarement exploré les modes de manipulation et de réarrangement des traces historiques permis par les développements de l'informatique réseau ou cherché à inventer des dispositifs graphiques nouveaux permettant la diffusion des produits de leur activité ou l'exploration des données. De même, les transformations de l'économie des échanges au sein de la profession et à ses marges, induites par l'usage systématique du mail et des sites internet, sont peu interrogées.

Tout se passe comme si les historiens, au cours de la dernière décennie, avaient réagi aux transformations rapides du contexte et aux mutations, tout aussi rapides, des outils disponibles, en tentant de les comprendre et de s'y ajuster, plutôt qu'élaboré des dispositifs spécifiques permettant d'explorer les implications heuristiques des changements en cours ou expérimenté la possibilité de nouveaux modes d'exposition au public des produits de leur activité. Le constat bien sûr mérite nuance, le *Médiéviste et l'ordinateur* a ainsi consacré sa livraison de 2002 à l'apport cognitif des techniques informatiques, témoignant de ce que l'interrogation sur les implications heuristiques de l'usage de l'informatique par les historiens n'a jamais été abandonnée par ceux-ci. L'impression demeure cependant qu'elle est paradoxalement moins vive qu'au temps où l'historien confiait à l'informaticien ses fiches à traiter.

La nécessité de s'approprier d'incessantes nouveautés – il faut maintenant, nous disent certains collègues, comprendre ce que le Web 2.0 implique de nouveau – l'explique en partie. Cela d'autant plus que l'historien confronté à l'informatique s'est, au cours des dernières décennies, souvent trouvé livré à lui-même, contraint de se former sans cesse à de nouveaux outils, de nouveaux logiciels, acquérant à grand-peine des compétences sitôt obsolètes à peine qu'acquises. L'avènement de la micro-informatique, puis celui, peu après, des réseaux s'est en effet, en France du moins, produit dans un contexte de rareté des compétences informatiques dans les centres de recherche historique. Ce dont ainsi se plaignent en 1995 les rédacteurs de l'éditorial du *Médiéviste et l'ordinateur*, ce n'est pas de manquer de machines, ou d'infrastructures, mais de ne pouvoir faire entendre aux administrateurs de la recherche que leurs premiers

besoins sont l'organisation de la formation, «la mise en place de structures et de procédures d'assistance, de concertation entre utilisateurs » [Comité de rédaction, 1995], procédures d'autant plus nécessaires que la formation des historiens fait peu de place à l'acquisition de compétences informatiques et moins encore à la culture mathématique et statistique.

En France, les deux dernières décennies ont vu de plus un raccourcissement de la durée de réalisation des thèses et un engorgement du marché universitaire, phénomènes propres à décourager nombre de thésards d'acquiescer péniblement des compétences que nul ne leur demande, voire dont la maîtrise dénote le tâcheron, englué dans la technique et donc la matière. Il faudrait assurément, pour pouvoir avancer avec quelque certitude des propositions qui soient autre chose que des impressions, des enquêtes de terrain dont nous ne disposons pas, mais la fréquentation du milieu des historiens universitaires laisse supposer que les vieilles oppositions entre le travail de la main et celui de l'esprit, la science et la technique, y sont bien vivaces et l'on y rejoue volontiers la querelle de Vadius et de Trissotin. Je ne sais guère d'autre discipline qui permettrait à un brillant thésard, exposant ses travaux dans le cadre d'un séminaire prestigieux, de répondre, sans crainte, qu'il n'a pas eu recours à la lexicométrie pour explorer son vaste corpus de textes, parce qu'il n'aime pas cela, qu'il n'y comprend rien et que de toute façon il n'y croit pas. Imagine-t-on un antiquisant expliquer qu'il n'a utilisé que les sources grecques parce qu'il n'entend rien au latin, qu'il n'aime pas les sonorités de cette langue et que les historiens romains mentent toujours ? Que l'ignorance puisse ainsi encore se porter en sautoir indique à quel point la curiosité pour le renouvellement de l'outillage à la disposition de l'historien est mal partagée et a peu de chances d'être valorisée.

Ces quelques éléments ne suffisent pas, bien sûr, à expliquer les formes de la réception de l'informatique réseau par le milieu des historiens. Nous n'avons guère les moyens de le faire et ce n'est pas tout à fait notre propos. La lecture des deux revues dont nous avons fait nos points de repère nous permet à la fois de rendre compte de ce que les textes des historiens portent peu la marque de l'usage de ces outils récents et de reformuler la question tenant aux transformations de l'écriture historique. En effet, ni la manipulation des données historiques ni l'invention de formes nouvelles d'exposition du savoir historique n'ont été des priorités au cours de la période récente pour les historiens les plus engagés dans l'étude des possibilités offertes par les mutations de leur outillage, qui sont de plus peu nombreux. Si nous ajoutons à cela le temps de réalisation des travaux,

qui se compte en années, nous pouvons comprendre que l'on chercherait vainement dans les textes produits au cours de ces dernières années par les historiens la trace d'importantes mutations liées à l'avènement de l'informatique réseau. En ce sens, tout discours sur ce sujet est aujourd'hui un discours prospectif, exercice peu prisé des historiens.

Le relevé des interrogations et des expérimentations des historiens nous permet cependant de disposer de quelques éléments permettant de porter un jugement sur la probabilité de certaines évolutions, en sachant cependant la fragilité des interpolations de tendances contemporaines et que toutes les virtualités ne se réalisent pas.

UNE ÉCONOMIE NOUVELLE DE L'ÉRUDITION

+++++

Une évolution cependant est déjà certaine, qui touche aux conditions d'accès à la bibliographie disponible. Les historiens, nous l'avons vu, sont déjà nombreux à utiliser les services des grands agrégateurs de contenus donnant accès aux textes de leurs collègues. C'est là une mutation importante, car l'historien est toujours un érudit, tenu de lire les travaux de ceux qui traitent des mêmes objets que lui, et une recherche commence souvent par la constitution d'une bibliographie et la consultation au moins partielle des travaux recensés. Cela a pour effet de confronter le chercheur à des médiateurs nouveaux, dont il perçoit sans doute moins les logiques et les pratiques que celles des bibliothécaires et des documentalistes qui étaient traditionnellement ses interlocuteurs. Il est probable que l'économie de ses lectures s'en trouve changée, l'un des effets les plus évidents de cette mutation étant que l'accès aux textes en langue anglaise se trouve facilité pour les chercheurs rattachés à une institution disposant d'un accès aux grandes bases payantes. Il nous faudrait cependant de systématiques et rétrospectives analyses de citation pour décider des transformations des pratiques bibliographiques des chercheurs. Si nous sommes tentés de considérer spontanément qu'une offre plus riche peut conduire à une diversification des pratiques de lecture, ignorant désormais tant les frontières nationales que les barrières disciplinaires, l'économie des biens culturels est là pour nous rappeler que la surabondance de l'offre peut très bien s'accompagner du triomphe de quelques produits vedettes, et d'une concentration accrue se traduisant par le quasi-monopole de quelques plates-formes d'intermédiation [Rebillard, 2009, chap. 3].

Ces transformations sont peu évoquées et peu étudiées par les historiens, sinon parfois pour dénoncer les menaces pesant sur la lecture lente et méditative, supplantée par une lecture rapide qui arrache au texte quelques fragments destinés à nourrir les articles ou les ouvrages du chercheur.

Cependant, nous ne savons guère ce qu'il en est des pratiques effectives des historiens, qui n'ont pas attendu l'internet pour consulter les productions de leurs collègues ni pour, à l'aide des index et des tables, aller chercher au détour d'une page la référence ou l'information dont ils avaient besoin. Il n'est pas certain non plus que les véritables enjeux soient de cette nature. Ce mode d'accès confronte à la pléthore, et parfois à la redondance de l'information scientifique, au bruit souvent aussi généré par toute requête qui ne soit pas très spécifique, tout en offrant, par le biais de la numérisation et de la structuration des bases bibliographiques, la possibilité de modes nouveaux de manipulation de l'information bibliographique. Un tel contexte a conduit, dans des disciplines voisines, à l'émergence d'entreprises logicistes de formalisation des productions scientifiques, appuyées sur des ontologies locales, qui ont pour fonction de sauvegarder l'exigence d'exhaustivité traditionnellement attachée aux disciplines historiques⁵. Il n'est pas certain que l'histoire, telle du moins que nous la connaissons, se prête à de semblables tentatives. Elles supposent un accord entre les praticiens du domaine sur ce qui fait objet. La discipline est cependant elle aussi confrontée à la croissance exponentielle du volume de la production scientifique et à son accessibilité croissante. Chacun en réponse picore du mieux qu'il peut, sans qu'à notre connaissance émerge une réflexion (et moins encore des pratiques réglées) qui prenne en compte cette nouvelle situation qui fait que l'érudit connaissant l'ensemble des textes et des historiens touchant à ses préoccupations est désormais une figure de légende associée à un passé révolu.

L'érudition historique est aussi affaire de connaissance des fonds et des dépôts d'archives disponibles. L'historien est souvent autant le connaisseur d'un ou de quelques fonds que l'expert d'un domaine. Cette pratique d'une archive ou de collections est, quand elle est réflexive, productrice de sens. L'archive est en effet souvent le produit d'une visée pratique et en porte la trace, tant dans le choix de ce qu'elle retient ou au contraire écarte, que par les contiguïtés établies par le classement des

5. La revue en ligne *Arkeotek* se consacre à la défense et à l'illustration de l'approche logiciste en archéologie, < <http://www.arkeotek.org/> >.

pièces réunies. L'une des principales archives permettant de mener l'histoire de la terreur nazie est ainsi le Berlin Document Center. Établi par l'occupant américain, le centre devait permettre d'identifier les responsables et les exécutants de la machine nazie par la réunion de ce qui avait pu être préservé des archives des organisations nazies. Ce souci pratique explique à la fois les caractéristiques des documents réunis et les modes d'accès à l'information suggérés par la source, qui permet par exemple d'identifier les membres d'une même organisation. Par contre, si « cette histoire renseigne [...] sur les bourreaux, les victimes sont pour ainsi dire absentes » [Dongen, 2006].

En ce sens, chaque archive définit un univers de discours possibles, que l'historien ne peut transcender que par le recours à d'autres archives ou par la transformation de traces du passé en archives dont il se fait le curateur et le premier historien. Le geste premier de l'historien est de transformer les pièces de cette ou de ces archives en sources, réunies, au moins implicitement, en un corpus aux limites plus ou moins fermement tracées. Il est des corpus définis en compréhension, qui sont ouverts et potentiellement infinis, et d'autres qui sont étroitement spécifiés. Le fait nouveau, abondamment commenté par les historiens, est que la numérisation d'une partie des matériaux leur permettant de fabriquer leurs sources est en cours, sous des formes diverses. De nombreux centres d'archives mettent au point des dispositifs matures permettant la recherche de pièces d'archives et autorisant soit à en visualiser des images, soit à accéder au texte porté par la source, lui-même parfois décrit au moyen de métadonnées structurées. Les Archives de France préparent ainsi la mise en ligne de versions étendues de certaines de leurs bases de données. Il sera bientôt possible de visualiser à distance des clichés des actes de naturalisation décrits par la base Naturalisation (NAT) des Archives de France [Archives nationales, s.d.].

Les historiens, ceux du moins qui s'expriment dans le cadre des revues dédiées à l'informatisation de la discipline, sont au fait de ces transformations et parfois même partie prenante de celles-ci. Elles rendent possibles, par le biais d'une requête utilisant des opérateurs booléens, voire d'une recherche plein texte, le repérage de pièces d'archives et de textes se rapportant à un personnage, une institution, un lieu, voire d'en visualiser une représentation, alors même qu'elles seraient conservées par de multiples institutions sans liens entre elles. Les historiens, de longue date habitués à chercher les aiguilles dissimulées dans les bottes de foin, sont particulièrement sensibles à cela. Ils semblent moins intéressés par le fait que ces

mêmes transformations autorisent à remodeler les bottes de foin et à en proposer des descriptions inédites. Il est désormais possible en effet, dans certains domaines, de composer un corpus très vaste à partir de requêtes sémantiques – si nous considérons, ce qui n'est pas une évidence, que la recherche de chaînes de caractères constitue à proprement parler une requête sémantique – ou en recourant à une ou des bases de données archivistiques. La récupération des notices descriptives, les possibilités de recherches plein texte, voire les progrès des outils de reconnaissance de formes, permettent de plus une description qui s'appuie en partie sur des mesures du corpus constitué et qui prend en compte tant les éléments contextuels attachés aux pièces du dossier que les formes et les éléments textuels offerts par celles-ci. Cela n'est pas dire que de telles pratiques sont appelées à se substituer à la recherche en archives, mais qu'elles coexisteront avec des pratiques plus anciennes, selon des modalités que nous ne connaissons pas encore. En effet, tout ne sera pas numérisé, ne serait-ce que pour des raisons de coût, et le simulacre disponible à l'écran n'est qu'une représentation en deux dimensions de la pièce d'archives dont la matérialité – les encres, le papier utilisé pour ne prendre qu'un exemple simple – peut faire sens pour l'historien. Il demeure néanmoins qu'un nouvel espace des possibles est ouvert, au prix d'un redoublement des médiations et parfois de l'absence d'un contact physique avec le support d'inscription de l'information consultée.

À la médiation de l'archiviste ou du documentaliste, qui choisit et indexe, s'ajoute en effet celle de l'informaticien, qui définit un mode d'accès aux données qui n'est pas sans effet. À l'heure actuelle, mais ce peut n'être là qu'une situation transitoire, il est rare que les dispositifs en ligne permettent l'acquisition de la base consultée et de ses descripteurs. Cela contraint l'utilisateur à formuler une requête qui souvent ne permet de visualiser qu'une collection de fiches, qu'il faut souvent examiner une à une, et non une table de données tout entière. Ces procédures supposent de plus la compréhension, parfois difficile à acquérir – parce que sa description n'est pas accessible – de l'architecture de la base et de son langage de requête, lequel peut s'avérer complexe. Nous pouvons alors supposer que cette médiation a ses effets propres, même si une telle proposition ne peut être posée que comme hypothèse, non seulement par ce qu'elle implique, mais aussi par ce qu'elle contribue à abolir. En effet, la numérisation et l'accessibilité par le réseau d'un certain nombre de fonds ont souvent une visée conservatoire. Il s'agit, pour une institution d'archives, de soustraire à la communication des documents très fragiles ou très fréquemment

utilisés, afin de garantir leur intégrité physique, ce qui est l'une de ses missions essentielles. Cela conduit dans certains cas à l'absence d'un contact direct avec les pièces d'archives comme à l'impossibilité du feuilletage d'un dossier, d'un carton, comme mode d'accès à l'information.

Or ces tâches, longues et fastidieuses, qui ont fait le quotidien de générations d'historiens, imposent à celui-ci un rapport à son matériau qui a sa temporalité propre – une lenteur qui est aussi le temps d'une imprégnation –, des postures corporelles et des mouvements, tout en l'exposant à des contiguïtés entre des pièces, des objets, produites par l'histoire de l'archive qu'il consulte. Tout historien a ainsi vu surgir des textes, des images, dont il ne savait pas qu'il les cherchait et qui parfois ont contribué à réorienter une recherche ou ont fait naître des questions nouvelles. Il est possible que nous perdions, dans le cas de certains fonds au moins, à la fois le bénéfice de ces trouvailles produites par un hasard orienté par la constitution de l'archive, mais aussi le plaisir qu'elles procurent. La fabrique de l'histoire est en effet aussi affaire de pratiques corporelles et d'affects et nous savons, d'expérience ou pour avoir lu Arlette Farge [Farge, 1989], que le travail en archive et la prise en main du document sont une expérience sensorielle, qui fait naître la rêverie, l'émotion et parfois incite au recueillement. Et si nous ne savons pas comment, ni véritablement pourquoi, nous sentons que ce que nous écrivons a partie liée aux conditions de cette expérience.

**NOUVELLES ÉCRITURES HISTORIENNES
ET LEURS CONDITIONS DE POSSIBILITÉ**

+++++

Les développements actuels suggèrent la possibilité qu'existeront, à une échéance difficile à déterminer, de vastes ensembles numérisés, disponibles sous des formats non propriétaires, permettant non seulement de construire des corpus plus vastes et autrement conçus, mais également des manipulations nouvelles, en partie automatisées, des signes et symboles qu'ils recèlent. De fait, cette possibilité existe déjà pour les historiens du temps présent, qui ont affaire à des sources nativement numériques et volumineuses fournies par le Web, les groupes de *news*, voire les collections de mails [Granjon, 2002]. Ceux-ci se prêtent tant à une description statistique précise du corpus, qu'à l'usage d'une statistique inférentielle. Les obstacles liés à la taille des corpus, qui ont longtemps interdit aux historiens l'usage pertinent d'un certain nombre d'outils

statistiques, sont ici levés. De tels corpus, même faiblement structurés et hétérogènes, peuvent également être soumis à diverses formes de fouilles de données, que celles-ci aient une visée exploratoire ou soient destinées à valider une hypothèse [Bennani et Viennet, 2007]. Ajoutons qu'un même constat peut être fait pour les écritures invisibles que multiplie l'historien au cours de son étude des sources et qui, s'il faut en croire la multiplication des ordinateurs portables dans les salles d'archives, sont de plus en plus nativement numériques.

L'accès de beaucoup à l'informatique et la diffusion d'outils logiciels simples d'accès favorisant la structuration et la fouille d'une information hétérogène permettent en effet la fabrication par les chercheurs d'archives numériques de leur cheminement propre, mêlant annotations, citations, références bibliographiques, clichés de sources, ou éléments de bases de données, voire la possibilité de partager de telles bases. Les historiens savent, puisqu'ils sont nombreux à l'écrire, que les transformations des règles de manipulation des signes, l'invention de formes nouvelles de disposition de ceux-ci, ont des effets heuristiques. Vincent Denis montre dans une thèse récente que la mise au point de techniques d'identification et de recensement des individus, qui passe par un perfectionnement des cartes et des registres employés, et les possibilités de manipulation de celles-ci au moyen d'une centralisation achevée de l'information, rendent possible une nouvelle appréhension des phénomènes sociaux, fondée sur leur quantification et permettent à l'État royal, doté de nouveaux instruments de connaissance du royaume, d'aspirer à de nouvelles formes d'intervention [Denis, 2008]. Pourtant les historiens sont peu nombreux à prêter intérêt aux outils permettant la mise en œuvre de nouvelles combinatoires à partir de vastes corpus, même parmi les plus informatisés d'entre eux. Cet intérêt, récent, n'a pas de plus débouché sur la mise au point d'outils spécifiques, alors que tel a été le cas chez les géographes, les archéologues, les documentalistes et les archivistes. Le trait est intéressant parce qu'il suggère une originalité disciplinaire qu'il n'est pas possible, lorsque la population étudiée est constituée des membres des cercles les plus engagés dans l'informatisation de la discipline, de mettre sur le compte d'une technophobie des agents. Nous ne disposons cependant, pour en rendre compte, que d'hypothèses.

Nous pouvons prendre comme point de départ l'éventuelle originalité de la discipline. Celle-ci est une connaissance par trace de mondes révolus, ce qui interdit de fait toute observation et mesure directe. Elle a de plus affaire à la temporalité, ou plutôt à des régimes de temporalité très

divers, tous les historiens ne travaillant pas à la même échelle. La discipline est de plus très peu formalisée. L'absence de procédures codifiées, d'un langage descriptif commun, de cadres théoriques partagés, ou au moins discutés, est caractéristique de celle-ci. De fait, les pratiques en son sein sont très diverses et l'on trouve parmi les historiens des auteurs de démographies rétrospectives, des spécialistes des économies anciennes ou de géographie historique dont les usages et le vocabulaire sont parfois plus proches de ceux des économistes ou des géographes du contemporain que de certains de leurs collègues. L'ambition, caractéristique des évolutions récentes de la discipline, de prendre en charge les glissements sémantiques dans le temps et les catégories indigènes ajoute encore à cette diversité. Il n'empêche que tous appartiennent à la communauté historique et sont reconnus comme tels par leurs pairs. De plus, la discipline historique progresse souvent par déplacement de ses questionnements, ce qui s'accompagne fréquemment d'un changement de vocabulaire, voire de la mobilisation de sources ou d'archives nouvelles. Or la création de systèmes experts nécessite une formalisation et une explicitation des procédures et des outils conceptuels mobilisés (et généralement leur standardisation), ce qui est de fait particulièrement difficile à accomplir dans le cadre d'une discipline que ne définit ni une *praxis*, ni un matériau, ni un vocabulaire partagé.

S'il n'est donc pas certain qu'un système d'information historique adapté à tous les usages historiens puisse voir le jour, nous pouvons supposer la possibilité qu'émergent des outillages spécifiques à la validité plus locale, partagés par des groupes de spécialistes plus étroits, aux pratiques plus homogènes. À notre connaissance, les historiens attirés par ces perspectives tentent d'adapter à leurs usages des outils conçus pour et par d'autres. Des raisons pratiques l'expliquent sans doute en partie. Parmi elles, le fait que les très grands corpus de données utilisables sont encore relativement rares.

Or la conception de tels systèmes d'information dédiés aux historiens est une opération aussi longue que coûteuse, qui exige la réunion de compétences rares et diverses. Pour y parvenir, il faudrait imaginer une enquête associant des informaticiens et des scientifiques, ou la réunion au sein d'un laboratoire des compétences nécessaires, sinon le recours à des prestataires de services extérieurs. Il est rare, quoique le cas se produise parfois [ALPAGE, s.d.], que les historiens se trouvent dans l'une ou l'autre de ces situations. Le fait est en partie lié à une organisation de

la recherche dont nous avons déjà évoqué les traits, en partie aux caractéristiques de la pratique sociale qu'est le travail historien.

Les systèmes d'information géographique – ou archéologique, pour ne rien dire des dispositifs de fouille de données –, n'ont pas émergé seulement parce que la nature des matériaux à traiter en rendait la conception aisée, mais aussi parce que l'intérêt pratique d'un système d'aide à la décision justifiait – au moins en France – des investissements coûteux ou des efforts soutenus : par exemple, afin de disposer d'éléments prédictifs adossés à une formalisation robuste en matière d'aménagement du territoire ou d'archéologie préventive. Cependant, la discipline historique n'est pas si étroitement liée à des savoirs pratiques qu'elle puisse fréquemment coproduire de tels outils. Il arrive certes que les autorités judiciaires ou le pouvoir politique demandent à l'historien d'établir des faits, en réponse à une question précise, ou que soit commandité ou approuvé un récit qui fonde une appartenance, ou justifie une politique. Ce que l'on demande cependant alors à l'historien, c'est-à-dire le travail historique qui est valorisé et financé par des opérateurs extérieurs au champ scientifique, est rarement un savoir formalisé ou une prédiction robuste, qu'il serait d'ailleurs bien en peine de fournir. La demande qui lui est parfois adressée de produire récits signifiants et mythes ne le prédispose pas non plus à s'intéresser aux formes d'écriture nouvelles permises par le développement de l'informatique et des réseaux.

Renouveler ainsi le constat que la lecture des récits historiques et la consultation de la toile ne permettent guère de repérer de changement majeur dans les orientations d'une discipline, ou les modes de son écriture, ne signifie pas que rien ne change ou ne changera. D'une part les écritures dont se servent les historiens, usagers des archives et des bibliothèques, et sans doute en partie leurs écritures quotidiennes, subissent des transformations majeures dont nous ne pouvons encore examiner l'impact. En effet, elles sont récentes et nous manquons absolument de faits construits à cette fin, même si nous pouvons définir quelques possibles conséquences en suivant les pas des premiers historiens numériques. D'autre part, les transformations d'une discipline et d'une institution ne dépendent pas uniquement de la disponibilité d'outils nouveaux. Les transformations en cours des institutions universitaires, du secteur de l'édition, peu prévisibles dans leurs formes et leurs effets, voire des usages sociaux des produits de l'activité des historiens, sont susceptibles de bouleverser les conditions d'exercice de l'activité historique et partant les modes d'écriture et de diffusion de ses produits. De plus, les durables

réticences de larges secteurs de la profession historique ne sont pas une donnée de nature, mais une disposition sociale – donc liée à un contexte – que l'on peut comprendre comme la réponse d'une institution à l'annonce de changements potentiellement porteurs d'une remise en cause de son organisation et de son fonctionnement, cependant qu'ils font planer sur les agents la menace d'un remodelage des hiérarchies fonctionnelles, voire la démonétisation de certaines compétences chèrement acquises. Les transformations de cet ordre sont lentes. De ce fait, il est probable que les expérimentations et les pratiques innovantes émergeront, pour partie, aux marges des institutions. La rapidité et le faible coût des échanges permis par le réseau, l'existence d'une population nombreuse, formée aux techniques de l'histoire mais sans espoir d'intégration dans ses institutions et sans accès aisé aux formes traditionnelles de l'édition, tout en entretenant avec elles des liens, le souci, très caractéristique du monde contemporain, de ne pas, en quelque domaine que ce soit, laisser la parole aux seuls experts patentés, l'accessibilité et la simplicité d'emploi d'outils de travail collaboratifs, laissent augurer la naissance d'historiens collectifs savants, dont l'existence ne sera pas en soi une radicale nouveauté – le monumental et collectif dictionnaire biographique du mouvement ouvrier nous le rappelle –, mais dont les productions et les écritures auront sans doute des traits spécifiques.

+++++

CHAPITRE XII. HISTOIRE@NUMÉRIQUE

+++++

La première version de ce texte est parue dans la revue *Écrire l'histoire* (numéro 11, printemps 2013, pp. 69-74). Il résulte d'une commande d'un court texte de synthèse abordant l'importance du « tournant numérique » pour l'histoire et les historiens, destiné à être publié dans une revue dédiée à l'histoire de l'histoire et à l'historiographie.

UNE ÉPOQUE NUMÉRIQUE

+++++

L'histoire que nous vivons, comme celle que nous écrivons, est aujourd'hui prise dans un environnement numérique, soit des machines interconnectées permettant le traitement automatique et quasi instantané de volumes considérables d'informations. Ce constat, ses conséquences pour la pratique historique, s'entend de plusieurs façons. La plupart des activités humaines contemporaines ne peuvent plus être comprises ni même se produire hors de ce cadre. C'est vrai de la sphère du travail [Garsten et Wulff, 2003], comme de celle de l'économie, ou de la culture [Rebillard, 2009]. Or, si ces phénomènes sont aujourd'hui inscrits d'évidence sensible dans notre quotidien, les technologies sur lesquelles ils se fondent sont au cœur du fonctionnement de beaucoup de grandes entreprises, mais aussi des grands systèmes militaro-industriels [Lecuyer, 2007], depuis plusieurs décennies. En ce sens, toute histoire du très contemporain comporte, potentiellement au moins, une dimension numérique. Par cela même, elle renvoie souvent à une histoire de très longue durée. En tant que mise en signe du monde, le langage numérique s'inscrit dans une histoire des systèmes d'écriture. Il a pour condition de possibilité aussi bien la pensée leibnizienne que l'alphabet complet élaboré pour la première fois par les anciens Grecs [Herrenschmidt, 2007]. Replacer le numérique dans le temps long, sans se cantonner à une histoire de l'informatique, de ses réseaux ou des contenus – par ailleurs nécessaire et en cours de structuration [Bourdon et Schafer, 2012] – est alors un chantier ouvert de l'histoire d'aujourd'hui. De cet usage croissant des

réseaux et des machines qui s'y raccordent, découle le fait que, pour la période récente, une bonne partie des traces des activités et des interactions humaines prennent la forme de données, numériques dès leur origine. Elles se caractérisent par leur hétérogénéité, leur surabondance et leur extrême fragilité. Leur conservation et leur exploitation sont donc de véritables défis techniques, économiques et politiques [Illien, 2011], auxquels se confrontent, depuis longtemps [Lupovici, 2005] déjà, les professionnels de la documentation et de l'archive.

Ce même environnement favorise le développement de compagnies en situation de monopole ou d'oligopole, prenant le contrôle de ressources numériques dont elles entendent ensuite monétiser l'usage. Cela fait en permanence peser de sérieuses menaces sur l'accès à la littérature scientifique, aux données patrimoniales, voire aux archives de notre temps présent. Aux États-Unis, le directeur d'une entreprise privée a ainsi pu affirmer, sans qu'il soit certain qu'il ait tort, qu'il est de fait le détenteur de notre mémoire contemporaine [Rosenzweig, 2011, p. 232].

Le plus surprenant est au fond le désintérêt, durable, de la profession historique, particulièrement en France, pour le développement des mondes numériques. Ces bouleversements affectent pourtant aussi le métier et les institutions pour lesquelles les historiens travaillent, depuis plusieurs décennies saisies par l'informatique [Gueissaz, 1999]. Les réflexions, anciennes, sur ce thème sont longtemps restées confinées en des cercles étroits. La multiplication de numéros de revues ou d'ouvrages sur les usages de l'informatique en histoire est très récente [Genet et Zorzi, 2011 ; Bourdon et Schafer, 2012]. Il n'est guère possible, dans l'espace de cet article, de revenir sur chacun des enjeux structurant ces débats, du fait de la nature même des transformations qu'ils tentent de cerner. L'informatique réseau a en effet, parmi ses propriétés, la particularité de permettre de générer automatiquement des chaînes de symboles aisément manipulables, transmissibles et reproductibles. Celles-ci peuvent être aussi bien des descriptions de n'importe quel objet du monde physique, du son aux images en passant par des propriétés physiques enregistrées par des capteurs, que la reproduction de chaînes symboliques existantes, quel que soit le support d'inscription de celles-ci. C'est l'ensemble alors de ce que l'on peut appeler, après Jack Goody, les technologies de l'intellect, qui subissent un bouleversement profond, soit les conditions de production, de manipulation, et de transmission de l'écriture et du calcul.

L'HISTORIEN INSTRUMENTÉ

+++++

Pour les historiens, qui n'ont jamais beaucoup calculé, et ont donc pu longtemps se contenter d'un outillage modeste et discret, cela signifie de fait une première confrontation à l'automatisation, à la mécanisation du travail, à la possibilité aussi d'un contrôle direct de celui-ci en son détail. Cette confrontation est synonyme de redéfinition de l'activité historique en ce qu'elle est métier, c'est-à-dire systèmes de pratiques, d'acteurs, d'institutions et organisation sociale d'une profession. Nous commençons à connaître certains aspects de ces transformations. La mutation la plus souvent mise en évidence est la possibilité, nouvelle, d'accéder immédiatement à de vastes réservoirs de données et d'articles en s'affranchissant, par le biais des recherches *full text*, en partie au moins, des logiques de classement des bibliothèques et des archives [*L'historien face à l'ordre informatique*, 2006].

La possibilité d'une communication immédiate et peu coûteuse entre spécialistes d'un même champ est souvent également soulignée [Rosental, 2000], de même que celle d'une dissémination efficace des produits de l'activité historique par le moyen de la toile [Dacos et Mounier, 2009]. Plus généralement, les transformations de la circulation des informations au sein de la communauté historique, ou les gains de productivité offerts par les dispositifs d'enregistrement, permettent de limiter la durée du passage en archives. Ces évolutions, si elles sont bien réelles, ne sont pas propres à l'histoire et sont depuis longtemps repérées et étudiées dans des champs parfois proches, comme l'économie, le droit ou la gestion [Chevalier *et al.*, 2000]. C'est également le cas des transformations des conditions d'acquisition, de structuration et de manipulation de la donnée historique par le calcul, plus explorées en d'autres domaines des sciences sociales, mais obéissant à des logiques similaires¹.

L'une des spécificités des débats entre historiens est sans doute la place tôt accordée aux implications de l'introduction de l'informatique pour l'enseignement de l'histoire. Si le débat sur l'intégration des outils numériques dans l'enseignement de l'histoire a d'abord eu pour cadre l'enseignement secondaire [Letouzey, 2003], il a touché ensuite, plus tôt aux États-Unis qu'ailleurs, les mondes universitaires [Brown, 2005]. Ces débats s'inscrivent dans une conjoncture complexe qui mêle généralisation des outils numériques, interventions de groupes industriels en quête

1. Voir dans ce volume « Les nouvelles frontières de l'historien ».

de marchés et transformations de la gouvernance d'institutions publiques dont la logique de fonctionnement tend à se rapprocher de celle de firmes privées. Certaines de ces institutions rêvent d'ailleurs à voix haute d'utiliser les possibilités technologiques pour se dispenser de l'accueil physique des étudiants et de la gestion des coûteux entrepôts que sont les bibliothèques [Tiffin et Rajasingham, 2003].

QUEL HISTORIEN NUMÉRIQUE ?

+++++

Nous nous sommes peu demandé cependant jusqu'ici quelle histoire pouvait s'enseigner à l'âge numérique et ce qu'il convenait d'enseigner aux futurs historiens. Les deux questions pourtant se posent, d'abord parce qu'il paraît aujourd'hui difficile d'enseigner, en secondaire en particulier, une histoire qui ignore les mondes techniques et scientifiques si l'on persiste à penser que cet enseignement offre des clés pour penser le contemporain. Ensuite, certaines des compétences évaluées et des exercices proposés doivent être repensés. Un simple automate connecté à Internet – ou un élève muni d'un smartphone – est aujourd'hui capable de répondre aux questions ouvertes ou aux questionnaires à choix multiples soumis aux étudiants des classes secondaires américaines avec plus de précision et de rapidité que la plupart d'entre eux [Rosenzweig, 2011]. Faire de la précision factuelle d'un élève privé de ses prothèses numériques un critère d'évaluation, ou un objectif, est alors mener une bataille déjà perdue. À l'inverse, dans un monde électronique bruissant de rumeurs, foisonnant d'erreurs, aux formes toujours changeantes, la capacité à repérer, organiser et évaluer l'information délivrée par des documents pour lesquels n'existent ni sémiologie ni diplomatique stabilisées² devient vitale. Il s'agit précisément là des compétences traditionnelles de l'historien. Nous sommes bien, et cela ne peut être sans conséquence pour l'enseignement, appelés à repenser la valeur formatrice des disciplines de culture. Cet état de fait conduit certains auteurs de langue anglaise à plaider pour un nouvel humanisme, forcément numérique [Doueih, 2011]. Cet « humanisme numérique » a pour l'instant trouvé peu d'écho en France où l'on évoque plus volontiers la naissance d'une « culture numérique », perçue comme distante des humanités traditionnelles [Berry, 2008].

2. Voir dans ce volume « De quoi le Web est-il l'archive ? ».

Les contours de ces deux concepts, au demeurant fort différents, demeurent imprécis : quelle place réserver aux sciences, à l'informatique et aux mathématiques dans la formation du nouvel humaniste ? Celui-ci doit-il être d'abord un agile utilisateur de dispositifs dont les logiques lui demeurent étrangères ? Doit-il être au contraire, dans l'idéal, un autochtone des mondes numériques en connaissant les langages, condition primordiale de son autonomie dans la civilisation moderne ? Naturellement, les réponses apportées à ces questions, que l'on voit poindre, ne seront pas sans conséquences pratiques sur la formation des futurs historiens comme des futurs enseignants.

PLACE DE L'HISTORIEN DANS LA CITÉ NUMÉRIQUE

+++++

Ces incertitudes lexicales et conceptuelles, dont il faut souhaiter qu'elles se muent en débats nécessaires, ne touchent pas seulement à la formation des futurs praticiens ou à l'enseignement qu'ils pourront dispenser un jour. Elles invitent à repenser leur fonction de médiateur dans le monde qui s'annonce. Là encore les débats sont, pour une part, communs à d'autres champs. L'idéologie internet qui accompagne le déploiement des dispositifs nouveaux est fondamentalement hostile à toute forme de médiation autorisée par un savoir incorporé. Les outils déployés, en un mouvement qui n'est pas sans rappeler les transformations du travail ouvrier, permettent d'incorporer en des objets, des machines ou des dispositifs logiciels, compétences et savoirs détachés du corps des producteurs [Metzger, 2011]. L'une des conséquences en est qu'il n'est plus besoin aujourd'hui d'une intermédiation historique pour avoir accès à de vastes collections de documents primaires, à des vues d'objets patrimoniaux [Robichaud, 2009], à une multitude de discours sur l'histoire, voire pour produire soi-même un discours historique offert à la curiosité de tous. Voisinent ainsi sur le Web des myriades d'histoires parallèles, savantes parfois, populaires souvent, quelquefois solipsistes ou reflétant les vues des tenants d'idéologies minoritaires ou de groupes vaincus [Cricione *et al.*, 2004]. Cette situation a par exemple pour conséquence l'usage intensif de Wikipédia par nos élèves, nos étudiants et le grand public. Cette encyclopédie en ligne cependant – qui ne comporte guère plus d'erreurs que la moyenne des encyclopédies papier –, accorde souvent peu de place aux objets et aux approches de l'histoire savante. Nous voyons naître une histoire en ligne, où souvent se dilue la notion d'auteur, au sein de laquelle

se forment des communautés de curiosité, dont le fonctionnement fait parfois bon marché des réputations acquises ou des positions institutionnelles [Lamassé, Deruelle et Dumont, 2013] que fréquentent peu les professionnels de l'histoire alors pourtant qu'elle devient pour beaucoup le moyen privilégié d'accès au récit historique.

Elle révèle pourtant un appétit de connaissances autant que la capacité de mobilisation de passionnés capables, par le biais du *crowdsourcing*, de coproduire des objets historiques utiles autant à l'amateur qu'au professionnel [Peccatte, 2009], même si leurs logiques d'usage diffèrent. Face à ce nouveau monde, la communauté historienne se trouve devant un choix. Elle peut assumer de se définir comme un ensemble de clercs retranchés derrière leurs murs clos, produisant une scolastique ne circulant qu'au sein de réseaux d'initiés, certes internationaux, mais cependant restreints. Elle peut aussi choisir d'inventer, sans repères, des formes d'interventions numériques permettant une présence dans une sphère publique en rapide expansion. La difficulté n'est pas tant technique que sociopolitique, puisqu'il s'agit pour l'historien de négocier une place à l'intérieur d'un système d'échanges sociaux au sein duquel son espace et son rôle, et plus généralement ceux de l'expert, ne sont pas prédéfinis [Noiret, 2011].

CHAPITRE XIII. NOUVELLES FRONTIÈRES DE L'HISTORIEN

La première version de ce texte est parue en 2014 dans le numéro 2 de la *Revue Sciences/Lettres*, revue pluridisciplinaire, dont l'objectif, selon son acte de naissance même, est d'étudier les interfaces entre les sciences et les domaines littéraires. Conçu avec Stéphane Lamassé (Paris I/LAMOP), il est pour une large part inspiré par le travail réalisé dans le cadre de l'ANR ATHIS, que portait le LAMOP, mais aussi par la longue collaboration avec les historiens de Paris I réunis au sein du Pôle informatique de recherche et d'enseignement en histoire, ainsi que par le travail mené autour de l'Equipe Matrice, piloté par Denis Peschanski et porté par le Centre d'histoire du XX^e siècle de l'université Paris I. Texte situé, dont les propositions ou les conclusions résultent plus du travail réalisé à ces frontières entre Sciences et Lettres dans un cadre précisément spécifié que d'une observation générale d'un champ auquel les locuteurs seraient extérieurs.

Ln'est plus, dans nos sociétés, d'activité savante sans usage de l'ordinateur et des réseaux électroniques. Le fait n'est pas tout à fait nouveau, même pour les historiens. La découverte des possibles ouverts par l'usage de l'informatique par de petits cercles pionniers est ancienne. Le domaine francophone voit apparaître des revues spécialisées, tels *Mémoire Vive* ou le *Médiéviste et l'ordinateur* dès les années 1980. De même, la première trace d'une activité historienne sur la toile remonterait à 1993. Cela attira peu alors l'attention de la majorité des praticiens des disciplines érudites, et particulièrement des historiens. Ils tendirent longtemps à tenir l'informatique pour une technique ancillaire, dont l'irruption ne bousculait pas les normes et les pratiques disciplinaires, et dont les spécialistes, en petit nombre, s'exprimaient en des revues spécialisées à l'audience modeste. La conjoncture s'est, au cours des deux dernières années, brutalement renversée. Plusieurs grandes revues ou collections ont consacré un numéro spécial ou un volumineux dossier au tournant numérique de l'histoire [Genet et Zorzi, 2011 ; *Le métier d'historien à l'ère numérique* :

nouveaux outils, nouvelle épistémologie ?, 2011], des séminaires d'histoire numérique naissent dans les plus prestigieuses institutions, les blogs se multiplient [Cohen, s.d.], la vénérable société des historiens modernistes et contemporanéistes a consacré sa dernière réunion annuelle à la question et plusieurs projets de recherche récemment lancés intègrent une forte dimension numérique [« Euroscientia. Localisation et circulation des savoirs d'État en Europe 1750-1850 », s.d.].

Si nous ne savons pas plus aujourd'hui qu'hier ce que les historiens font effectivement de leurs ordinateurs¹, nous sommes soudain confrontés à une avalanche de textes programmatiques ou réflexifs et claquent comme des étendards de multiples et rutilants labels. À peine averti qu'il est l'heureux possesseur d'un ordinateur connecté à un réseau, voilà l'historien sommé de décider s'il se range parmi les tenants des humanités numériques (*digital humanities*), d'une histoire numérique (*digital history*) ou incline à se définir comme un historien pris en un contexte numérique, incité aussi à se demander s'il est déjà un historien 2.0, voire un expérimentateur audacieux inventant les pratiques qui seront celles de l'érudit 3.0 [Cohen *et al.*, 2008]. Cette soudaine effervescence est liée à la conscience, nouvelle, de ce que l'informatique, sous sa forme réseau, ne permet pas seulement d'accomplir un peu mieux ou un peu plus vite les mêmes tâches qu'autrefois, mais de ce qu'elle induit une mutation des conditions de production et de diffusion des connaissances historiques, voire de la nature de celles-ci. La transformation la plus souvent mise en évidence est la possibilité nouvelle d'accéder immédiatement à de vastes réservoirs de données et d'articles. Il est probable du reste que le fait que beaucoup d'historiens, mais aussi de professeurs, peuvent aujourd'hui accéder à des textes et à des images numériques de documents d'archives utilisables dans le cadre de leurs recherches ou de leur enseignement est au principe de la prise de conscience récente de ce que l'informatique réseau n'est pas un simple perfectionnement de la machine à écrire et du téléphone.

La possibilité d'une communication immédiate et peu coûteuse entre spécialistes d'un même champ est souvent également soulignée, de même que celle d'une dissémination efficace des produits de l'activité historique par le moyen de la toile, parce que l'une comme l'autre apparaissent susceptibles de transformer tant la structuration de la profession historique que son rapport aux milieux (archivistes, industries culturelles, etc.) avec lesquels elle est en contact [Grossman, 2012]. Ces mutations, il est vrai

1. Voir dans ce volume « Écriture de l'histoire et réseaux numériques ».

spectaculaires, et d'autant plus vivement perçues qu'elles ont déjà transformé le quotidien des praticiens, sont cependant depuis longtemps identifiées, au point d'avoir donné naissance à quelques discussions canoniques, tenant aux transformations supposées des pratiques de lecture (lecture profonde ou lente qui aurait été la norme autrefois et lecture flottante ou rapide qui le serait aujourd'hui [Henriot, 2012]), à la démocratisation – notion particulièrement problématique, quoique pas toujours interrogée [Rosenzweig, 2005] – de l'écriture de l'histoire, ou bien à l'éventuel triomphe des partisans d'une pensée graphique et réticulaire sur les tenants attardés d'une réflexion séquentielle et linéaire [Liu, 2008].

Ce ne sont donc pas ces matières que nous entendons discuter, non plus que les questions, moins fréquemment abordées par les historiens, mais tout à fait essentielles de l'archivage du monde numérique [Rosenzweig, 2003], ou de l'apparition, par le biais par exemple du *crowdsourcing*, de nouveaux producteurs de récits historiques [Noiret, 2011].

CONDITIONS DE L'ACQUISITION DE LA DONNÉE HISTORIQUE

+++++

Nous nous proposons ici de revenir sur les formes contemporaines de l'accumulation, sous forme numérique, de données utilisables par l'historien et certains des possibles alors ouverts en matière de structuration de l'information et de traitement de celle-ci. Nous nous attachons particulièrement à la façon dont cela peut contribuer, pour un temps au moins, à redéfinir les échanges professionnels des praticiens de l'histoire et en particulier leur rapport aux représentants d'autres champs disciplinaires. L'aspect le plus connu de ce processus, engagé depuis déjà de nombreuses années, est la numérisation des collections des grandes bibliothèques publiques et la mise à disposition de vastes ensembles textuels, sous forme d'abord de fichiers image, puis, de plus en plus, de fichiers texte par des opérateurs tant privés que public. La croissance est en ce domaine exponentielle. La seule BnF, qui mettait à disposition en 1997 environ 3 000 volumes, offre par l'intermédiaire du portail Gallica 90 000 ouvrages numérisés et 80 000 images en 2007, et environ 400 000 documents fin 2010 [La documentation française, 2011]. Depuis peu, les institutions publiques ouvrent également l'accès à une quantité croissante de données d'enquêtes, issues souvent de la statistique publique et pour partie nativement numériques. L'équipement Données infrastructures et méthodes d'enquête en sciences humaines et sociales

(DIME-SHS) participe de ce mouvement et s'oriente, semble-t-il, vers la mise à disposition de méthodes et d'enquêtes en *open data*. De même nature est la mise à disposition de la littérature scientifique, récente ou moins récente, accessible au travers de portails souvent payants. Dans tous les cas, l'utilisateur de ces ressources constate rapidement qu'il ne peut en faire usage sans que celles-ci ne soient documentées, et leur documentation accessible et manipulable, ce qui suppose l'intervention des acteurs traditionnels que sont les archivistes, les bibliothécaires et les spécialistes de la documentation.

Plus récemment ont été lancées de grandes entreprises de numérisation de fonds d'archives [Lee, 2010 ; Potin, 2011]. En France, celles-ci touchent plus particulièrement un type de sources que l'on appelait autrefois sérielles : enregistrement, cadastres, état civil sont particulièrement susceptibles d'être disponibles sous forme numérique, et dans la plupart des cas accessibles par le biais d'Internet. La majorité des archives départementales offrent ainsi en 2012 l'accès, selon des modalités différentes, à tout ou partie de l'état civil ancien par le moyen de documents numériques [Archives de France, s.d.]. Ces plans de numérisation sont parfois plus ou moins directement liés à des entreprises d'édition scientifique. Celles-ci renouvellent profondément les méthodes de l'érudition, mais aussi parfois ses catégories : l'idée d'une possible stabilité du texte médiéval, par exemple [Bauduin et Jacquemard, 2011]. L'érudition est donc aujourd'hui déjà profondément numérique par les outils qu'elle emploie, mais aussi parce que ceux-ci peuvent être mis en lien avec une transformation des pratiques et modes de pensée constitutifs de ses disciplines.

Moins notées, parce qu'encore souvent le fait d'entreprises pilotes, sont les recherches s'accompagnant de dispositifs, au moins partiellement automatisés, permettant la production de bases de données structurées rassemblant des descripteurs des objets numérisés, ainsi parfois qu'une transcription des éléments textuels inscrits en ceux-ci, voire une explicitation de certaines des informations portées par les éléments symboliques incorporés au document, dynamiques dont témoignent plusieurs exemples. Il est aujourd'hui possible ainsi de nourrir un système de gestion d'information au moyen d'une océrisation partielle de documents portant des inscriptions manuscrites, pour peu que ceux-ci soient structurés ou semi-structurés, ce qui est le cas de nombreuses sources utilisées par les historiens du contemporain (listes, dossiers, formulaires). Des prototypes sont élaborés dès le début des années 2000 [Murshed, 2002] et des

systèmes de ce genre sont aujourd'hui utilisés de façon routinière par les entreprises et les administrations [Merlin, 2006].

Nous pouvons également caractériser des documents anciens comportant textes et figures au moyen de paramètres calculés, sans connaissance *a priori* de la structure du document ni de son contenu sémantique, c'est-à-dire produire, de manière semi-automatique, une information structurée décrivant ceux-ci et les images (lettrines, enluminures, dessins [Mickaël Coustaty *et al.*, 2009]) ou blocs de textes qui s'y trouvent, ce qui débouche sur la possibilité d'une indexation partiellement automatisée de ces matériaux et d'une analyse de leurs propriétés physiques. Cela autorise la mise au point d'algorithmes d'attribution, voire, par la détermination de classes d'objets, permet la reconnaissance de symboles et donc l'extraction d'informations, d'où la possibilité de générer des systèmes d'information associés à des documents de type carte ou plan, par exemple [Noizet et Grosso, 2011]. Ces outils sont encore des dispositifs de laboratoire, dont l'usage suppose un paramétrage fin opéré par des spécialistes au bagage mathématique conséquent (transformations de Fourier, chaînes de Markov, fractales sont parmi les outils les plus utilisés pour l'analyse semi-automatisée de documents complexes [Journet *et al.*, 2008]) : ils n'en sont pas moins appelés à se développer rapidement, du fait même de l'ampleur des programmes de numérisation qui rendent indispensables la création d'outils autorisant l'exploitation des documents disponibles dont la masse croissante rend impossible l'indexation manuelle [Mickaël Coustaty *et al.*, *ibid.*]. Sont également en cours de développement des outils permettant la description et l'indexation des flux vocaux, tels ceux capturés par les émissions de radio ou les émissions de télévision à partir de la reconstitution automatisée d'un *verbatim*, en attendant sans doute, progrès des capacités de calcul promises par les ordinateurs quantiques aidant, la naissance d'outils similaires permettant de traiter des flux vidéo importants. L'Equipex Matrice ainsi, porté par Denis Peschanski, a pour premier objectif la production d'outils efficaces permettant, à partir de sources sonores très bruitées et hétérogènes, d'améliorer la qualité des dispositifs de reconnaissance vocale, ce qui autorisera, à terme, la transcription semi-automatique de vastes réservoirs d'archives orales ou des bandes-son de flux vidéo.

La particularité de ces évolutions rapides est de n'être pas liées seulement à la purification d'algorithmes de plus en plus efficaces, non plus qu'au raffinement toujours plus grand des outils mathématiques utilisés, mais de dépendre de l'extraordinaire croissance des capacités de calcul

décrite par la loi de Moore et de l'augmentation, non moins stupéfiante, des capacités de stockage disponibles, ainsi que de celle de la rapidité du transfert des données, qu'accompagne la très rapide diminution des coûts de ces différents éléments.

DÉMULTIPLICATION DES POSSIBILITÉS DE TRAITEMENT

+++++

Ce contexte a d'autres conséquences. L'accumulation de masses d'informations utilisables par les historiens se produit alors que se démultiplient les possibilités de manipulation distribuées de vastes amas de données, et que diminuent les coûts d'appropriation d'un certain nombre d'outils logiciels, sous l'effet, en partie, du développement de solutions logicielles libres exploitant à plein la souplesse du travail en réseau.

Pour ne prendre que quelques exemples, chacun peut aujourd'hui facilement disposer sur son ordinateur personnel de systèmes de gestion d'information performants et aisément mutualisables (Evernote, par exemple ou Zotero, plus connu des historiens) permettant l'intégration de documents hétérogènes et leur indexation à la volée, alors qu'il fallait, il y a vingt ans, construire pour cela une base de données relationnelle, élaborer, souvent *a priori*, une arborescence de descripteurs et doter, si l'on voulait partager les données ainsi produites, chaque poste d'utilisateur d'une coûteuse licence. De même fallait-il, pour réaliser une simple analyse factorielle, voire obtenir rapidement une série de tableaux croisés, écrire une macro-commande, ou être capable de manipuler des logiciels professionnels, dont l'ergonomie ne semblait remarquable qu'aux adeptes, rares, du langage Fortran.

Ces mutations, qui apparaissent particulièrement rapides à l'échelle d'une carrière de chercheur, sont plus spectaculaires encore quand sont considérées des pratiques qui demeurent l'apanage de cercles plus étroits de spécialistes. L'analyse textométrique des corpus anciens a ainsi récemment subi de rapides et profondes mutations techniques et conceptuelles [Bertrand *et al.*, 2011]. L'analyse de discours se déploie aujourd'hui dans toutes les périodes [Deruelle, 2010] et des techniques comme celle de l'étiquetage morphosyntaxique, qui, il y a peu, nécessitaient un temps considérable, deviennent, en partie au moins, automatisables pour un grand nombre de langues contemporaines et anciennes². Le lien avec la discipline

2. L'étiqueteur Cordial ou le projet TreeTagger sont de bons exemples.

historique est ici particulièrement fort : les moyens de traitement de la langue ont été investis par des historiens qui infléchissent la production logicielle [Mayaffre, 2004].

Il est donc aujourd'hui possible de concevoir des chaînes d'acquisition et de traitement du signal historique, qui permettent, partant d'une source, d'un fonds, de le décrire finement au moyen de documents numériques, d'en transcrire les inscriptions et de mettre en place des outils, accessibles aux historiens de métier, permettant une navigation efficace du tout et aussi, par des manipulations automatisées de ces répertoires de signes, d'en proposer des visualisations et des descriptions synthétiques aptes à nourrir des discours historiques tout à fait conformes aux canons de la discipline. En témoignent, par exemple, les nombreuses publications récentes étudiant le Paris ancien à l'aide des outils façonnés par le projet ALPAGE mené par Hélène Noizet et le LAMOP [ALPAGE, s. d.].

L'histoire se dévoile alors comme mode de connaissance instrumenté dont les praticiens, aux prises avec une matérialité, forgent des dispositifs permettant de générer des traces graphiques interprétables. La pratique historienne, en d'autres termes, assemble des techniques. Si la conscience de cela l'est parfois, la chose n'est pas nouvelle. Les historiens du livre montrent depuis déjà longtemps que celui-ci est un objet extraordinairement complexe, dont l'évolution est liée à celle d'un secteur industriel important en même temps qu'indissociablement à des techniques de l'intellect que tout à la fois il incorpore et rend possible.

La nouveauté est que l'historien, en sa pratique quotidienne, est en contact direct avec des machines dont il ne peut saisir la complexité et que celles-ci permettent une automatisation d'une partie des tâches qu'il accomplissait sans le secours d'aucune, par exemple la production de listes, de cartes, de plans, la création d'index ou de tables. Changement d'échelle donc, de *modus operandi* et d'ordre de grandeur, qui n'abolit pas ce qui précède, mais dont les implications ne conduisent pas moins à une transformation profonde des pratiques, du métier et sans doute de la position sociale de l'historien, qui se faisait autrefois, s'il faut en croire du moins les textes normatifs d'hier, un devoir que de contrôler toutes les phases de son travail, de la collecte à la synthèse, en passant par la description et l'évaluation de ses sources. Celui-ci dépend, aujourd'hui comme hier, d'une infrastructure, c'est-à-dire d'outils, de machines, d'institutions, de procédures codifiées, du fonctionnement de puissants secteurs industriels et l'informatique réseau dévoreuse de ressources naturelles, gourmande en énergie, particulièrement polluante et important secteur d'emploi est

une industrie lourde. Elle a cependant ses propriétés propres. Elle incorpore un millefeuille de textes et de langages informatiques, qui procèdent d'outils mathématiques dont parfois seuls les plus grands spécialistes du domaine avaient, il y a quelques dizaines d'années, notion, ce qui rend l'environnement au sein duquel se déploient nos pratiques savantes infiniment plus complexe qu'autrefois, au point qu'il n'est pas certain qu'un individu puisse encore avoir une intellection globale de ce système, ni même une représentation fonctionnelle à peu près assurée de celui-ci.

Le constat n'est pas propre à la discipline historique. Claude Kergomard le faisait il y a peu pour la géographie [Kergomard, 2010], dont certains secteurs sont tôt entrés à l'ère du laboratoire latourien peuplé d'étranges boîtes noires délivrant des signes dont les usagers routiniers n'interrogent plus, et parfois ne connaissent plus, les principes de génération.

Ce n'est pas dire là que l'atelier de l'historien, créateur solitaire, est devenu ou a vocation à devenir pour tous et bientôt un laboratoire flambant neuf, où s'affairent des professionnels en blouse blanche. Nous savons, au moins depuis Bertrand Gille [Gille, 1978], que les systèmes techniques se recouvrent plus qu'ils ne se succèdent. La conjoncture nouvelle n'abolit pas l'ancien, elle s'y combine, l'agrège, en des installations parfois incongrues et toujours instables, car notre environnement informatique est mouvant, empêchant encore l'émergence de formes pérennes. Le plus bel exemple de ces dynamiques nous est fourni là encore par l'édition scientifique des textes anciens. Alors que les formes en apparaissaient définitivement stabilisées, elle est marquée depuis une quinzaine d'années par un bouillonnement créatif dont il n'est pas certain que le terme soit proche [Bauduin et Jacquemard, 2011], de même la philologie très tôt confrontée à l'informatique et souvent partie prenante des entreprises d'édition électronique a-t-elle profondément renouvelé ses pratiques [Berra, 2014].

Simplement, pour emprunter au vocabulaire latourien [Latour, 2005], d'étranges boîtes noires trouvent leur place dans cet environnement de travail, dont certaines ont l'inquiétant aspect de très lisses monolithes. Internet foisonne aujourd'hui de traceurs, sites alléchants permettant la production automatique de graphes, de figures, d'images, offrant la possibilité à tous, et donc aussi aux historiens, de générer de fascinants objets graphiques, parfois plastiquement remarquables, l'expérimentation en ce domaine empruntant fréquemment aux arts numériques [Paul, 2004] et stimulant incontestablement l'imagination. Ils permettent de visualiser des textes, des tableaux de données, des flux d'information, et d'accrocher à

ces visualisations des discours, sans qu'il soit besoin pour cela d'aucune compétence informatique, non plus que d'aucune compréhension des opérations leur donnant naissance. Il en est beaucoup, mais le plus célèbre de ces outils au sein des communautés s'occupant d'analyse des textes est sans contexte Ngram, parce que le lancement de ce site a été accompagné d'une campagne visant à promouvoir une pratique nouvelle définie par ses promoteurs comme une science nouvelle, la *culturonomics* [Michel *et al.*, 2011].

L'initiative a suscité rapidement des critiques justifiées [Morse-Gagné, 2011], le dispositif initial n'en demeure pas moins intéressant parce qu'il révèle certaines des dynamiques en cours et des futurs possibles.

Une entreprise privée, proposait, forte de sa position dominante, un dispositif défini comme savant qu'elle est seule, puisqu'elle contrôle l'accès à Google Book, à pouvoir mettre en place. Celui-ci est au sens propre un traqueur, produisant par manipulation réglée d'un vaste réservoir de données textuelles, en réponse à une requête de l'utilisateur, une trace graphique interprétable. L'instrument n'est conforme aux canons d'aucune discipline existante, ni même à certains des principes de la démarche scientifique, du fait, en particulier, des problèmes tenant à la définition des corpus de référence. Il est proposé par une firme qui n'est pas un acteur traditionnel de la sphère savante et dont nous savons qu'elle a parfois permis ou favorisé la censure étatique, voire dont certains employés ont délibérément falsifié des données afin d'éliminer des concurrents [Gévaudan, 2012]. Elle n'en a pas moins trouvé des alliés au sein des mondes savants et l'objet proposé demeure le seul mode d'interrogation possible du plus vaste corpus de données textuelles jamais rassemblé.

Nous retrouvons là une situation, par certains aspects, familière, comparable à la généralisation des tableurs qui conduisit, il y a quelques années, à la multiplication de camemberts dont les couleurs vives égayaient les pages des premiers mémoires de lettres et d'histoire entièrement composés sur ordinateur, mais qui ne respectaient que rarement les règles de la sémiologie graphique, et moins encore celles de la statistique ; à ceci près, et la différence est de taille, qu'il était toujours possible de s'emparer d'autres outils ou d'en créer.

INSTRUMENTATION DE L'HISTORIEN

+++++

Si nous nous inscrivons dans la perspective d'une histoire attentive aux conditions d'administration de la preuve, qui se conçoive comme une

entreprise de connaissance instrumentée et dont les praticiens, confrontés à un matériau opposant une résistance à leurs entreprises, mettent en œuvre des dispositifs techniques, alors la priorité du moment est celle de l'équipement de l'historien.

À cela la seule réponse possible nous paraît être l'élaboration d'une instrumentation savante par la communauté savante elle-même, qui garantisse l'inscription de ses normes au cœur des outils proposés, c'est-à-dire d'abord la possibilité de la critique et d'une appropriation créative. Pour le dire en reprenant là encore le vocabulaire latourien, il faut que les boîtes noires puissent être ouvertes et qu'il soit possible de calibrer les instruments. Cela suppose, dans le domaine qui nous occupe, l'accès aux algorithmes et au code, la possibilité aussi de paramétrer finement les traceurs et les grapheurs appelés à se multiplier [Guichard, 2005] et qui constituent de fait les outils permettant d'adapter au nouveau contexte les exigences de nos disciplines, en particulier pour l'histoire celle de la prise en compte de la totalité du matériau pertinent disponible. Il ne s'agit pas cependant ici d'adaptation contrainte et à marche forcée. Ce contexte offre, à condition de s'approprier des outils nécessaires, outre une intensification de la circulation de l'information au sein de la communauté scientifique et de sa diffusion au-delà de celle-ci, un accroissement de nos capacités manipulatoires et donc cognitives.

La tâche est rude cependant, et pour plusieurs raisons. La production, en ce contexte, d'une instrumentation adaptée aux besoins d'une communauté scientifique suppose l'alliance de compétences qui ne peuvent être réunies en une seule personne, ni même trouvées au sein des communautés scientifiques concernées. Des instruments permettant une manipulation pertinente et contrôlée de chaînes symboliques de grande taille supposent à la fois des capacités informatiques et des compétences mathématiques pointues. Aucun collectif d'historiens ne peut espérer les réunir.

Les techniques ainsi les plus prometteuses pour l'analyse exploratoire de vastes tableaux de données à forte proportion de données manquantes et comportant des données imprécises ou incertaines, ce qui est typiquement le cas pour de nombreuses tables de données produites au cours d'enquêtes historiques, sont aujourd'hui regroupées sous le nom de méthodes neuronales [Blayo et Verleysen, 1996]. Parmi celles-ci, l'algorithme de Kohonen (SOM) se révèle particulièrement adapté [Kohonen, 1995]. Il offre la possibilité de visualiser un nuage de données multidimensionnelles à l'aide d'un document unique, aisément lisible au prix d'un apprentissage. L'algorithme s'accommode de données textuelles ou

qualitatives, voire incomplètes (données manquantes ou données connues uniquement par des distances ou des indices de dissimilarité relative) [Deruelle et Lamassé, 2011].

De même, il existe aujourd'hui des techniques prometteuses en matière de traitement des données temporelles, mais dont l'appropriation là encore, et l'incorporation en un traitement logiciel, supposent la mise en œuvre de compétences rares. Les chaînes de Markov cachées [Baum et Petrie, 1966] se révèlent ainsi particulièrement adaptées à la production de systèmes de classification (donc de sériation et de périodisation) prenant en compte une succession d'événements ou d'observations. Ce contexte invite donc l'historien à repenser, pour un temps du moins, le graphe de ses relations professionnelles. Le montage de chaînes d'acquisition de l'information implique une collaboration, habituelle, avec des archivistes, des bibliothécaires, mais aussi avec des physiciens, parce qu'elles supposent des compétences en matière de reconnaissance des formes et de traitement du signal. La production de grapheurs permettant des visualisations efficaces et pertinentes de vastes amas de données implique des échanges avec des mathématiciens, des spécialistes de la visualisation de données [Fekete, 2010] et bien sûr des informaticiens. Nous écrivons ici, à dessein, collaboration et échanges et non usage d'outils ou de dispositifs produits par ces acteurs, moins encore prestation de service. Elle est en effet d'une nécessité vitale pour l'historien, plus exactement pour la communauté historienne. Si le personnage collectif qu'est celui-ci veut demeurer acteur d'une chaîne de production de savoir, il ne peut accepter que la constitution des corpus, la structuration et la manipulation des données, soient abandonnées à des prestataires extérieurs dont les logiques et les choix lui seraient totalement impénétrables. Ce serait renoncer à la possibilité de toute critique procédurale et à l'appartenance aux mondes savants, pour n'être plus qu'un augure offrant une lecture incertaine de signes mystérieux dont seul le charisme, les phrases joliment tournées ou la fonction idéologique garantiraient le pouvoir de conviction.

Ces échanges sont possibles et de fait se produisent parfois, non pas seulement parce que les érudits ont besoin des scientifiques mais parce que les compétences de l'historien, et en particulier sa maîtrise des techniques de l'érudition, en font un partenaire nécessaire. Le fait est particulièrement évident en matière d'élaboration de procédures en partie automatisées d'acquisition du signal et d'indexation des corpus, qui présupposent une validation experte qui n'est possible qu'au prix de la

maîtrise des déterminants de la matérialité de la source et de l'organisation des structures symboliques qu'elle porte.

Nous considérons qu'il en est de même en matière de structuration et d'exploration des données, fort du fonctionnement, depuis trois ans, de l'atelier Analyse et structuration des données historiques, qui réunit autour d'un séminaire et d'un dispositif logiciel des compétences mathématiques, épistémologiques, historiennes et informatiques. La première raison en est que le rapport des praticiens des SHS à la donnée est en effet souvent spécifique, marqué, dans le cas des historiens, par la mise en place de boucles de rétroaction qui conduisent en permanence des résultats de l'analyse au réexamen des sources et à l'amélioration de celles-ci (par restructuration des données, tests de discrétisations ou de classifications nouvelles, production de variables supplémentaires). Une production logicielle adaptée aux recherches empiriques effectivement menées, ce qui est de fait la problématique des mathématiques appliquées aux sciences sociales, suppose donc là aussi un dévoilement des logiques d'action des chercheurs, qui ne peuvent être réduites d'ailleurs à des algorithmes, et donc des échanges entre praticiens.

D'autre part, la numérisation de nos sociétés et de nos vies quotidiennes produit une multiplication d'énormes gisements de données portant trace de comportements humains, dont s'emparent physiciens ou mathématiciens, soit qu'ils leur fournissent des objets complexes permettant d'imaginer ou de tester de nouvelles techniques de modélisation ou de description [Dorat *et al.*, 2007], soit qu'il s'agisse, dans le but par exemple de mieux gérer un équipement, de décrire son usage, voire d'anticiper les conséquences d'une modification de certaines de ses propriétés [Borgnat *et al.*, 2011].

Même cependant quand prime le souci de parvenir à une formalisation rigoureuse, les chemins suivis ont pour point d'origine « [...] *simple social assumptions which we show can be seen as responsible for the properties of the whole network* » [Dorat *et al.*, *ibid.*]. En d'autres termes, même la modélisation la plus formaliste des traces d'un processus social suppose le recours à des hypothèses quant aux formes et aux déterminants de celui-ci, soit une théorie sociologique, de même que tout usage d'un dispositif comme Ngram Viewer suppose une théorie de l'histoire du livre et de la langue. Les praticiens des sciences sociales, et parmi eux les historiens, ont donc un rôle à jouer, en tant que tels, dans l'élaboration des nouveaux dispositifs de connaissance, d'autant que là encore la période est marquée par une activité intense portée par une forte demande d'intégration de la dimension temporelle et spatio-temporelle au sein des procédures

d'analyse de vastes amas de données. Cela revient à incorporer à celles-ci une modélisation des systèmes de raisonnement et d'inférence géographiques et historiques [Mokrane *et al.*, 2004]. La collaboration entre historiens, mathématiciens et informaticiens peut en ce cas déboucher sur l'élaboration de systèmes de traitement de données temporelles ou séquentielles adaptés tant aux caractéristiques des données effectivement rencontrées qu'aux modes pertinents d'interprétation de celles-ci. L'exploration en effet de la série temporelle est toujours encadrée par la connaissance, non nécessairement formalisée ou formalisable, d'un contexte général dont on cherche à déterminer la relation avec la chronique. L'étude ainsi de la production normative de l'État piémont-savoyard des XVI^e et XVII^e siècles menée par Julien Alerini en collaboration avec les informaticiens du CRI et les mathématiciens de l'équipe Statistique, analyse, modélisation multidisciplinaire (SAMM) a permis l'écriture d'un package R implémentant un usage nouveau des chaînes de Markov cachées, qui s'est révélé particulièrement adapté à la description des productions normatives d'un État moderne. Il a permis une classification efficace de celles-ci, autorisant à l'historien une sériation fiable de ses données [Olteanu et Ridgway, 2012].

La piste est prometteuse donc, mais le sentier escarpé. La participation à l'invention, en cours, de l'instrumentation historique de l'âge numérique, suppose des partenariats, qui exigent une *lingua franca*, qui, de fait, emprunte aux mathématiques et à l'informatique, parce que les dispositifs logiciels de traitement de l'information consistent en des manipulations réglées de chaînes de symboles dont l'écriture s'effectue en recourant à des langages informatiques. Cela exige, non pas que les historiens, en tant qu'individus, deviennent programmeurs ou mathématiciens, mais que la discipline fasse une place, qui ne soit pas un strapontin, à des praticiens dotés d'une solide culture en ce domaine. Cela passe également par une reconnaissance, en tant qu'acte de recherche, des contributions à l'élaboration de l'instrumentation scientifique. Ces évolutions supposent aussi que les utilisateurs finaux des outils qui naissent aujourd'hui, qui ne se substituent pas aux outils de l'historien d'hier, mais souvent s'y ajoutent, puissent se les approprier. Beaucoup ont la forme de traceurs et de grapheurs. Se saisir de leurs sorties et en proposer une interprétation pertinente et rigoureuse implique une familiarité avec les problématiques, travaillées depuis longtemps par les scientifiques, de la preuve graphique, mais aussi, parce que les images proposées proviennent toujours de transformations de traces au moyen d'opérateurs logiques et mathématiques,

et non d'une disposition nouvelle de celles-ci, une acculturation minimale à ces domaines, et donc un enseignement adapté. C'est à ce prix que l'historien demeurera, au cours de la décennie qui vient, un membre des mondes savants.

BIBLIOGRAPHIE

Adonis, TGE, « Lancement de la première phase de réalisation de l'infrastructure numérique d'accès unifié aux données et documents des sciences humaines et sociales », 2009, < <http://www.huma-num.fr/editorial/lancement-de-la-premiere-phase-de-realisation-de-linfrastructure-numerique-dacces-unifie> >, cette page n'est plus accessible aujourd'hui.

« Adresses et sites des archives départementales », s. d. [En ligne] < http://www.memodoc.com/liste_archives_departementales.html >.

ALPAGE, « AnaLyse diachronique de l'espace urbain Parisien : approche GEomatique », s. d. [En ligne] < <http://alpage.huma-num.fr/index.php/fr/> >.

Andersen Deborah Lines, "Academic Historians, Electronic Information Access Technologies, and the World Wide Web. A Longitudinal Study of Factors Affecting Use and Barriers to that Use", *Journal of the Association for History and Computing*, 1998, 1 (1).

Andersen Deborah Lines, "Academic Historians Revisited", *Journal of the Association for History and Computing*, 2003a, 6 (2).

Andersen Deborah Lines (dir.), *Digital Scholarship in the Tenure, Promotion, and Review Process. History, Humanities and New Technology*, Armonk, M. E. Sharpe, 2003b.

Anis Jacques, *Texte et ordinateur. L'écriture réinventée ?*, Paris – Bruxelles, De Boeck Université, 1998.

Archives de France, « Numérisation et mise en ligne des registres paroissiaux et d'état civil dans les services publics au 21 avril 2015 ». [En ligne] < <http://www.archivesdefrance.culture.gouv.fr/ressources/en-ligne/etat-civil/> >.

Archives nationales, « Consultation des bases de données des archives nationales », s. d. [En ligne] < <http://www.archivesnationales.culture.gouv.fr/arm/> >.

Arrighi Marie-Dominique, « Une lecture pour enfants jugée subversive », *À l'école des sans-papiers. Blog Libération*, 2007. [En ligne] < http://sans-papiers.blogs.liberation.fr/2007/06/10/une_lecture_pou/ >.

Barats Christine, *Manuel d'analyse du Web en sciences humaines et sociales*, Paris, Armand Colin, 2013.

Barlow Jeffrey G., "Editorial", *Journal of the Association for History and Computing*, 1998, 1 (1). [En ligne] < <http://quod.lib.umich.edu/j/jahc/3310410.0001.109> >.

Barthélemy Marc, Gondran Bernard, Guichard Éric, "Spatial Structure of the Internet Traffic", *Physica A. Statistical Mechanics and its Applications*, 2002, n° 319, pp. 633-642.

Bauduin Pierre, Jacquemard Catherine, « Les pratiques de l'édition en ligne. Expériences et questionnement », in Jean-Philippe Genet et Andrea Zorzi (dir.), *Les historiens et l'informatique. Un métier à réinventer*, Rome, École française de Rome, 2011.

Baum Leonard E., Petrie Ted, "Statistical Inference for Probabilistic Functions of Finite State Markov Chains", *The Annals of Mathematical Statistics*, 1966, vol. 37, n° 6, pp. 1554-1563.

BDIC (dir.), *Matériaux pour l'histoire de notre temps*, 2005, n° 79, *Internet et mouvement sociaux : nouvelles pratiques militantes, nouvelles sources pour l'histoire*.

Bennani Younès, Viennet Emmanuel (dir.), *Apprentissage artificiel et fouille de données*, Toulouse, Cepaduès, 2007.

Benoist Jocelyn, Merlini Fabio, « Spatialiser, historiciser », in Jocelyn Benoist et Fabio Merlini (dir.), *Historicité et spatialité. Le problème de l'espace dans la pensée contemporaine*, Paris, Vrin, 2001, pp. 7-13.

"Berlin Declaration on Open Access to Knowledge in the Science and Humanities", 2003. [En ligne] < <http://openaccess.mpg.de/Berlin-Declaration> >.

Bernard Michel, *Introduction aux études littéraires assistées par ordinateur*, Paris, PUF, 1999.

Berra Aurélien, « Philologie à venir. Textes (anciens) et humanités numériques », 2014. [En ligne] < <http://philologia.hypotheses.org/> >.

Berry Gérard, *Pourquoi et comment le monde devient numérique*, Paris, Collège de France/Fayard, 2008.

Bertrand Jean-Marie, Boileau Pierre, Genet Jean-Philippe, Schmitt Pantel Pauline (dir.), *Langue et histoire*, Paris, Publications de la Sorbonne, 2011.

Blayo François, Verleysen Michel, *Introduction aux réseaux de neurones artificiels*, Paris, PUF, 1996.

Blum Françoise, Groppo Bruno, Vaccaro Rossana, Veyron Franck, « Internet et mouvements sociaux : nouvelles pratiques militantes, nouvelles sources pour l'histoire », *Matériaux pour l'histoire de notre temps*, 2005, n° 79, pp. 3-10.

Borgnat Pierre, Robardet Céline, Rouquier Jean-Baptiste, Abry Patrice, Fleury Éric, Flandrin Patrick, "Shared Bicycles in a City: A Signal Processing and Data Analysis Perspective", *Advances in Complex Systems*, 2011, vol. 14, n° 3, pp. 1-24.

Bourdon Jérôme, Schafer Valérie (dir.), *Le temps des médias*, 2012, n° 18, *Histoire d'Internet, Internet dans l'histoire*.

Brabazon Tara, *Digital Hemlock. Internet Education and the Poisoning of Teaching*, Sydney, University of New South Wales Press, 2002.

Breton Philippe, *Le culte de l'Internet. Une menace pour le lien social ?*, Paris, La Découverte, 2000.

"Bright Sparcs", *Encyclopedia of Australian Science*, s. d. [En ligne] < <http://www.asap.unimelb.edu.au/bsparcs/bsparcshome.htm> >.

Brown Gregory S., « Enseigner l'histoire à l'âge du multimédia », *Cahiers d'histoire*, 2005, n° 96-97, pp. 121-134.

Bruno Anne Sophie et al., « Histoire et mémoires des immigrations en régions aux XIX^e et XX^e siècles », *Hommes et migrations*, 2008, n° 1273, pp. 6-17.

Cadiou François et al., *Comment se fait l'histoire. Pratiques et enjeux*, Paris, La Découverte, 2005.

Caretti Stefano, Degl'Innocenti Maurizio, Silei Gianni (dir.), *Scrivere con la sinistra : dalla carta interstata a internet*, Rome, Piero Lacaita Editore, 2002.

Cavallo Guglielmo, Chartier Roger (dir.), *Histoire de la lecture dans le monde occidental*, Paris, Seuil, 1997.

Cerutti Mauro, Fayet Jean-François, Porret Michel (dir.), *Penser l'archive. Histoires d'archives, archives d'histoire*, Lausanne, Antipodes, 2006.

Certeau Michel de, *L'écriture de l'histoire*, Paris, Gallimard, 2002.

Charton Ghislaine, « La presse périodique scientifique sur les réseaux », in Michèle Rouchet (dir.), *Les nouvelles technologies dans les bibliothèques*, Paris, Éditions du Cercle de la Librairie, 1996.

Chateauraynaud Francis, *Prospéro. Une technologie littéraire pour les sciences humaines*, Paris, CNRS Éditions, 2003.

Chaunu Pierre, « Le fils de la morte », in Maurice Aguhlon et Pierre Nora (dir.), *Essais d'ego-histoire*, Paris, Gallimard, 1987.

Chesneaux Jean, *Du passé faisons table rase ? À propos de l'histoire et des historiens*, Paris, Maspero, 1976.

Chester Jeff, *Digital Destiny. New Media and the Future of Democracy*, New York, The New Press, 2007.

Chevalier Jean-Marie, Ekeland Ivar, Frison-Roche Marie-Anne, Kalika Michel (dir.), *Internet et nos fondamentaux*, Paris, PUF, 2000.

Clarkson Adrienne, « Discours du trône », 2002, < <https://www.poltext.org/fr/discours-du-tr%C3%B4ne-canadiens> >.

Cohen Dan, blog. [En ligne] < <http://www.dancohen.org/> >.

Cohen Daniel J., Frisch Michael, Gallagher Patrick *et al.*, "Interchange: The Promise of Digital History", *The Journal of American History*, 2008, vol. 95, n° 2, pp. 442-451.

Coll., « Postface », *Médiéviste et l'ordinateur*, 1995, 31-32.

Collins Harry M., *Changing Order. Replication and Induction in Scientific Practice*, Londres, Sage, 1995.

Comité de rédaction, « Éditorial », *Médiéviste et l'ordinateur*, 1987, n° 31-32. [En ligne] < <http://lemo.irht.cnrs.fr/31-32/mo3101.htm> >.

Coustaty Mickaël, Sidère Nicolas, Ogier Jean-Marc, Héroux Pierre, Ramel Jean-Yves *et al.*, "Content-Based Old Documents Indexing", in *Eight International Workshop on Graphics Recognition*, La Rochelle, GREC, 2009.

Coustaty Mickaël, Ogier Jean-Marc, Pareti Rudolf et Vincent Nicole, « Extraction d'informations d'images de documents anciens », in *Congrès des jeunes chercheurs en vision par ordinateur*, Trégastel, s.n., 2009. [En ligne] < <https://hal.archives-ouvertes.fr/hal-00435382> >.

Cricione Antonio, Noiret Serge, Spagnolo Carlo, Vitali Stefano, *La Storia a(l) tempo di Internet : indagine sui siti italiani di storia contemporane*, Bologne, Pàtron editore, 2004.

Crosnier Hervé Le, « Les journaux scientifiques électroniques ou la communication de la science à l'heure du réseau mondial », in CEM-GRE-SIC (dir.), *La communication de l'IST dans l'enseignement supérieur et la recherche : l'effet Renater/ Internet (Actes du colloque des 16, 17 et 18 mars 1995, Bordeaux)*, Paris, ADBS Éditions, 1995. [En ligne] : < <http://gabriel.gallezot.free.fr/Solaris/d03/3lecrosnier.html> >.

Dacos Marin, « Le numérique au secours du papier. L'avenir de l'information scientifique des historiens à l'heure des réseaux », *Cahiers d'histoire*, 1999, 44 (1). [En ligne] < <http://ch.revues.org/48> >.

Dacos Marin, Mounier Pierre, « Sciences et société en interaction sur Internet. Éléments pour une histoire de l'édition électronique en sciences humaines et sociales », *Communication et langages*, 2009, n° 159, pp. 123-135.

Dallo Alain, « Quelle informatique enseigner aux historiens. L'exemple de l'Université Paris », *Médiéviste et l'ordinateur*, 2002, 41.

Darnton Robert, "The New Age of Book", *The New York Review of Books*, 1999, 46 (5).

Delaunay Marie-Annick, *L'immigration par escroquerie sentimentale*, Paris, Tatamis, 2006.

Delaunay Marie-Annick, « Immigration-Escroquerie-Sentimentale », sd. [En ligne] < <http://www.immigration-escroquerie-sentimentale.blogspot.fr/> >.

Denis Vincent, *Une histoire de l'identité. France 1715-1815*, Paris, Champ Vallon, 2008.

Deruelle Benjamin, « Enjeux politiques et pratiques de la chanson de geste à la fin du XV^e et au XVI^e siècle », *Revue historique*, 2010/3, n^o 655, pp. 551-576.

Deruelle Benjamin, Lamassé Stéphane, « Un dépouillement lexical entre analyse factorielle et cartes de Kohonen : le vocabulaire mathématique des arithmétiques en langue vulgaire du bas Moyen Âge », in *Modèles et apprentissages en sciences humaines et sociales (MASHS)*, Université Aix-Marseille, s.n., 2011.

Desbois Henri, « Les territoires de l'Internet : suggestions pour une cybergéographie », in Éric Guichard, *Comprendre les usages de l'Internet*, Paris, Éditions rue d'Ulm, 2001, pp. 253-263.

Desbois Henri, *Les mesures du territoire. Aspects techniques, politiques et culturels des mutations de la carte topographique*, Villeurbanne, Presses de l'enssib, 2015 (coll. Papiers).

Diminescu Dana (dir.), *Migrance*, 2005, n^o 23, *Les documents numériques : méthodologie d'archivage et perspectives de recherche sur l'immigration*.

Direction des Archives de France, *Les archives électroniques. Manuel pratique*, Paris, La documentation française, 2002.

Dongen Luc van, « Enjeux sociaux de deux archives de la terreur à Berlin et à Ludwigsburg », in Mauro Cerutti, Jean-François Fayet et Michel Porret (dir.), *Penser l'archive. Histoires d'archives, archives d'histoire*, Lausanne, Antipodes, 2006, pp. 266-280.

Dorat Rémi, Latapy Mathieu, Conein Bernard, Auray Nicolas, "Multi-Level Analysis of an Interaction Network Between Individuals in a Mailing-List", *Annals of Telecommunications*, 2007, vol. 62, n^o 3-4, pp. 325-349.

Doueïhi Milad, *La grande conversion numérique*, Paris, Seuil, 2008.

Doueïhi Milad, *Pour un humanisme numérique*, Paris, Seuil, 2011.

Dougherty Jack, Nawrotzki Kristen (dir.), *Writing History in the Digital Age*, Ann Arbor, University of Michigan Press, 2013.

Duclert Vincent, *L'avenir de l'histoire*, Paris, Armand Colin, 2010.

« Euroscientia. Localisation et circulation des savoirs d'État en Europe, 1750-1850 », s. d. [En ligne] < <http://euroscientia.univ-parisl.fr/> >.

« Évolution du plan de numérisation des fonds d'État, bilan pratique 1996-1999, nouvelles orientations pour 2000-2001 », < http://www.culture.fr/culture/mrt/numerisation/fr/seminaire_191199/table_rondel.htm >.

Farge Arlette, *Le goût de l'archive*, Paris, Seuil, 1989.

Farison Violaine, « La mémoire de l'internet », Séminaire internet de Sciences-Po, 2001. [En ligne] < http://barthes.ens.fr/scpo/Présentations00-01/Farison_MemoireNet.html >.

Fekete Jean-Daniel, « Visualiser l'information pour la comprendre vite et bien », in Coll. (dir.), *L'usager numérique*, s.l., ADBS éditions, 2010, pp. 161-194.

Fossier Lucie, « Éditorial », *Médiévisite et l'ordinateur*, 1984, n° 12. [En ligne] < <http://www.menestrel.fr/spip.php?rubrique667&lang=fr#2109> >.

Foucault Michel, *L'archéologie du savoir*, Paris, Gallimard, 1969.

Fouques Mélanie, « Le hip-hop face aux "zouloous". Une histoire des enfants de l'immigration », Mémoire de maîtrise d'histoire, Paris I, 2005.

Frey Sandra, « Les TIC dans l'administration », in Philippe Ricaud et Pascal Lardellier (dir.), *Le réseau pensant. Pour comprendre la société numérique*, s.l., Édition universitaires de Dijon, 2007, pp. 39-50.

Garsten Christina, Wulff Helena (dir.), *New Technologies at Work. People, Screens and Social Virtuality*, Oxford, Berg, 2003.

Gauthier Christian, Tellier Simon, Vignes-Lebbe Régine, Reichmann François, « Acquisition et gestion des connaissances dans les sciences du vivant », in *Les savoirs déroutés. Experts, documents, supports, règles, valeurs et réseaux numériques*, Lyon, Presses de l'enssib, 2000, pp. 75-79.

Genet Jean-Philippe, « Histoire, informatique, mesure », *Histoire & Mesure*, 1986, vol. 1, n° 1, pp. 7-18.

Genet Jean-Philippe, « La formation informatique des historiens en France : une urgence », *Mémoire vive*, 1994, n° 9. [En ligne] < <http://pireh.univ-paris1.fr/mv/num9.html> >.

Genet Jean-Philippe, « La formation des historiens à l'informatique en France : espoir ou désespoir ? », *Médiéviste et l'ordinateur*, 2002, n° 31-32.

Genet Jean-Philippe, Zorzi Andrea (dir.), *Les historiens et l'informatique. Un métier à réinventer*, Rome, École française de Rome, 2011.

Gille Bertrand (dir.), *Histoire des techniques*, Paris, Gallimard, 1978 (coll. Encyclopédie de la Pléiade ; 41).

Ginsparg Paul, "First Steps Towards Electronic Research Communication", *Computers in Physics*, 1994, vol. 8, n° 4.

Goody Jack, *La raison graphique. La domestication de la pensée*, Paris, Minuit, 1979.

Gévaudan Camille, « Google, moteur d'embrouilles », *Libération*, 23 janvier 2012. [En ligne] < http://www.liberation.fr/medias/2012/01/23/google-moteur-d-embrouilles_790293 >.

Graham Suzanne, "Historians and Electronic Resources: A Second Citation Analysis", *Journal of the Association for History and Computing*, 2001, vol. 4, n° 2.

Granger Christophe, *La destruction de l'université française*, Paris, La Fabrique, 2015.

Granjon Fabien, « Les répertoires d'action télématiques du néo-militantisme », *Le mouvement social*, 2002, n° 200, pp. 11-32. [En ligne] < <https://www.cairn.info/revue-le-mouvement-social-2002-3-page-11.htm> >.

Grenier Jean-Yves, « Expliquer et comprendre. La construction du temps de l'histoire économique », in Bernard Lepetit (dir.), *Les formes de l'expérience. Une autre histoire sociale*, Paris, Albin Michel, 1995, pp. 227-251.

Grossman James, "The Beckoning Horizon: New Challenges for the AHA in the Digital Age", *Perspectives on History*, 2012. [En ligne] < <https://www.historians.org/publications-and-directories/perspectives-on-history/february-2012/the-beckoning-horizon> >.

Gueissaz Albert, *Les mondes universitaires et leur informatique. Pratiques de rationalisation*, Paris, CNRS Éditions, 1999.

Guichard Éric, « Cartographie animée des échanges du réseau Renater », 2001. [En ligne] < <http://barthes.ens.fr/atelier/geo/Renater01/> >.

Guichard Éric, *L'internet. Mesures des appropriations d'une technique intellectuelle*, Paris, École des hautes études en sciences sociales, 2002.

Guichard Éric, « L'atlas Clio. Un atlas en ligne interactif de l'immigration », in Serge Noiret et Philippe Rygiel (dir.), *Les historiens, leurs revues et Internet (France, Espagne, Italie)*, Paris, Publibook, 2005, pp. 93-102.

Guichard Éric, *L'Internet et l'écriture. Du terrain à l'épistémologie*, Thèse de doctorat, Lyon, Lyon-I, 2010.

Harnad Stevan, "Electronic Scholarly Publication: Quo Vadis", *Solaris*, 1996, n° 3.

Henriot Christian, « Essai final », 2012. [En ligne] < <http://histoirenumerique.blogspot.fr/> >.

Herrenschmidt Clarisse, *Les trois écritures. Langue, nombre, code*, Paris, Gallimard, 2007.

Hilout Pascal, « Merci M. Wilders d'avoir lancé le bon débat », *Riposte laïque*, 2009, 77. [En ligne] < <http://ripostelaique.com/Merci-M-Wilders-d-avoir-lance-le.html> >.

Illien Gildas, « Une histoire politique de l'archivage du Web », *Bulletin des bibliothèques de France*, 2011, n° 2, pp. 60-68. [En ligne] < <http://bbf.enssib.fr/consulter/bbf-2011-02-0060-012> >.

Journet Nicholas, Ramel Jean-Yves, Eglin Véronique, Mullot Rémy, « Analyse d'images de documents anciens : une approche texture », *Traitement du signal*, 2008, 24 (6), pp. 461-479.

Kahn Paul, Lenk Krzysztof, *Mapping Websites*, Amsterdam, Bis Publishers, 2001.

Kalus Maximilian, "Semantic Networks and Historical Knowledge Management: Introducing New Methods of Computer-based Research", *Journal of*

the Association for History and Computing , 2007, 10 (3). [En ligne] < <http://hdl.handle.net/2027/spo.3310410.0010.301> >.

Kaufmann Arnold, *Une introduction à la théorie des sous-ensembles flous*, Paris, Masson, 1973.

Kergomard Claude, « Cartographier un objet variable dans le temps et dans l'espace : le climat », in Éric Guichard (dir.), *Cartographie et épistémologie (Colloque, 14-16 juin 2010)*, Peyresq, s.n., 2010.

Kohonen Teuvo, *Self-Organizing Maps*, Berlin, Springer, 2001.

La Documentation française, *Les bibliothèques numériques*, Paris, La Documentation française, 2011.

[En ligne] < <http://www.ladocumentationfrancaise.fr/dossiers/internet-monde/bibliotheques-numeriques.shtml> >.

Lamassé Stéphane, Deruelle Benjamin, Dumont Léo, « Histoire et écriture numérique : l'exemple de la fiche Jeanne d'Arc sur Wikipédia », École d'été, Lyon-Villeurbanne, Enssib, 16-20 septembre 2013.

Lamberterie Isabelle de (dir.), *La numérisation pour l'enseignement et la recherche. Aspects juridiques*, Paris, Éditions de la MSH, 2002, vol. 2.

LAMOP, « Journée d'étude portails médiévistiques », 2005. [En ligne] < www.menestrel.fr/IMG/rtf/programmeportailoctobre.rtf >.

Latapy Mathieu, Pons Pascal, "Computing Communities in Large Networks Using Random Walks (Long Version) », in LNCS (dir.), *Proceedings of the 20-th International Symposium on Computer and Information Sciences ISCIS'05, 2005, Istanbul, Turquie*, 2005. [En ligne] < <http://arxiv.org/abs/physics/0512106> >.

Latour Bruno, *La science en action. Introduction à la sociologie des sciences*, Paris, La Découverte, 2005.

Laublet Philippe, « Web sémantique et ontologies », in Claire Brossaud et Bernard Reber (dir.), *Humanités numériques. Nouvelles technologies cognitives et concepts des sciences sociales*, Paris, Lavoisier, 2007, vol. 1.

Lawlor John M., "The Schuylkill Navigation System Project at Reading Area Community College: Preservation and Dissemination of an Important Collection of Transportation Documents", *Journal of the Association*

for History and Computing, 2002, V (1). [En ligne] < <http://quod.lib.umich.edu/j/jahc/3310410.0005.102/--schuykill-navigation-system-project-at-reading-area?rgn=main;view=fulltext> >.

Le métier d'historien à l'ère numérique : nouveaux outils, nouvelle épistémologie ? [Titre du dossier de la] *Revue d'histoire moderne et contemporaine*, 2011/5, n° 58-4bis.

Le Robert pour tous, Paris, Dictionnaire le Robert, 1994.

Lécuyer Christophe, *Making Silicon Valley. Innovation and the Growth of High Tech, 1930-1970*, Cambridge, Mit Press, 2007.

Lee John, "Digital History and the Emergence of Digital Historical Literacies", *Technology in Retrospect*, Richard Diem et Michael J. Berson (dir.), *Social Studies in the Information Age 1984-2009*, 2010, pp. 75-90.

Lepetit Bernard, « Histoire et modélisation », *Annales. Économies, Sociétés, Civilisations*, 1988, 43^e année, n° 1, p. 3-4.

Lepetit Bernard, « L'histoire quantitative, deux ou trois choses que je sais d'elle », *Histoire & Mesure*, 1989, IV (3-4), pp. 191-199.

Letouzey Daniel, « Enseignement de l'histoire et informatique, entretien avec Daniel Letouzey », *Le cartable de Clio*, 2003, n° 3, pp. 190-201.

Liu Alan, "When Was Linearity? The Meaning of Graphics in the Digital Age", *Digital History Project*, 2008. [En ligne] < <http://digitalhistory.unl.edu/essays/liuessay.php> >.

Lupovici Catherine, « L'archivage de l'internet à la Bibliothèque nationale de France », *Matériaux pour l'histoire de notre temps*, 2005, n° 79, pp. 80-83.

« L'historien face à l'ordre informatique », *Matériaux pour l'histoire de notre temps*, 2006, n° 82.

Maitron, Association, « Le Maitron. Dictionnaire biographique du mouvement ouvrier », s. d. [En ligne] < <http://maitron-en-ligne.univ-parisl.fr/> >.

Martin Inaki Lopez, « L'histoire contemporaine en Espagne sur Internet : problèmes et perspectives », in Philippe Rygiel et Serge Noiret (dir.), *Les historiens, leurs revues et Internet (France, Espagne, Italie)*, Paris, Publibook, 2005, pp. 21-24.

Mayaffre Damon, *Paroles de président. Jacques Chirac (1995-2003) et le discours présidentiel sous la V^e République*, Paris, Champion, 2004.

McCrank Lawrence J., *Historical Information Science. An Emerging Undis-
cipline*, Medford (New Jersey), Information Today, 2001.

Merlin Céline, *Les services publics en mutation. La Poste innove*, Paris, L'Harmattan, 2006.

Metzger Jean-Louis, « Tic et travail ; de l'étude des usages à la critique de l'impératif du changement », in Julie Denouël et Fabien Granjon (dir.), *Communiquer à l'ère numérique. Regards croisés sur la sociologie des usages*, Paris, Presses des Mines, 2011, pp. 91-125.

Michel Jean-Baptiste *et al.*, "Quantitative Analysis of Culture Using Millions of Digitized Books", *Science*, 2011, vol. 331, n^o 6014, pp. 176-182.

Ministère de la Culture, Joconde. Portail des collections des musées de France. [En ligne] < <http://www.culture.gouv.fr/documentation/joconde/fr/pres.htm> >.

Minuti Rolando, *Internet et le métier d'historien*, Paris, PUF, 2002.

Mokrane Abdenour, Laouamri Oussama, Dray Gérard, Poncelet Pascal, « Modélisation spatio-temporelle des connaissances d'un Système d'information géographique », in *International Conference. Technologies of Information and Telecommunications*, Tunis, SETIT, 2004.

Morse-Gagné Elise, "Culturomics: Statistical Traps Muddy the Data", *Science*, 2011, 332 (6025), pp. 35-37.

Mounier Pierre, *Les maîtres du réseau. Les enjeux politiques d'Internet*, Paris, La Découverte, 2002.

Mounier Pierre, « Atelier : Le chercheur et la diffusion de ses sources : nécessités, risques, contraintes, reconnaissances ? (Marin Dacos, That camp Paris) », 2010. [En ligne] < <http://tcp.hypotheses.org/382> >.

Murshed Nabeel, "Automatic Reading of Traffic Tickets", in Daniel Lopresti, Jianying Hu et Ramanujan Kashi (dir.), *Document Analysis Systems V: 5th International Workshop, DAS 2002, Princeton, NJ, USA, August 2002, Proceedings*, Princeton, Springer, 2002, pp. 66-70.

Nelson Theodor H., *Literary Machines. The Report on, and of, Project Xanadu Concerning Word Processing, Electronic Publishing, Hypertext, Thinkertoys, Tomorrow's Intellectual Revolution, and Certain Other Topics Including Knowledge, Education and Freedom*, s.l., Mindful Press, 1993.

Noiret Serge, « Histoire et mémoire dans la toile d'histoire contemporaine italienne », in Philippe Rygiel et Serge Noiret (dir.), *Les historiens, leurs revues et Internet (France, Espagne, Italie)*, Paris, Publibook, 2005, pp. 81-90.

Noiret Serge, « Y a-t-il une histoire 2.0 ? », in Jean-Philippe Genet et Andrea Zorzi (dir.), *Les historiens et l'informatique. Un métier à réinventer*, Rome, École française de Rome, 2011, pp. 235-288.

Noizet Hélène, Grosso Éric, "The ALPAGE Project: Paris and its Suburban Area at the Intersection of History and Geography (9th-19th Century)", in *Digital Proceedings of the 25th International Cartographic Conference (ICC'11)*, Paris, s.n., 2011.

Olteanu Madalina, Ridgway James, "Hidden Markov Models for Time Series of Counts with Excess Zeros", in *European Symposium of Artificial Neural Networks*, Belgique, s.n., 2012, pp. 133-138.

Park Han Woo, Thelwall Mike, "The Network Approach to Web Hyperlink Research and its Utility for Science Communication", in Christine Hine (dir.), *Virtual Methods. Issues in Social Research on the Internet*, Oxford, Berg, 2005, pp. 171-181.

Paul Alain, « Bases de données contre liens documentaires », s.d. [En ligne] < <http://www.culture.gouv.fr:80/adha/bdcontrelh.html> >.

Paul Christiane, *L'art numérique*, Paris, Thames & Hudson, 2004.

Peccatte Patrick, « Contenu généré par les utilisateurs : quelle légitimité ? », *Archimag*, 2009, n° 227.

Pellier Karine, "Contributions of Date Management to Cliometrics", *Journal of the Association for History and Computing*, 2006, 9 (1). [En ligne] < <http://hdl.handle.net/2027/spo.3310410.0009.102> >.

Persée, < <http://www.persée.fr/web/guest/home> >.

Potin Yann, « Institutions et pratiques d'archives face à la numérisation », *Revue d'histoire moderne et contemporaine*, 2011, 58 (5), pp. 57-69.

Presse sénégalaise [Agence de], « Une amende de 2 milliards de FCFA de dommages et intérêts réclamés à un internaute français pourfendeur de la destination Casamance », APS, 2003.

Prost Antoine, *Douze leçons sur l'histoire*, Paris, Seuil, 2010 [1^{re} éd. 1996].

Rebillard Franck, *Le Web 2.0 en perspective. Une analyse socio-économique de l'Internet*, Paris, L'Harmattan, 2009.

Robichaud Léon, « L'histoire, le patrimoine et le public : la diffusion de l'histoire grâce aux inventaires numériques disponibles sur le Web », *Digital Studies, le champ numérique*, 2009, vol. 1, n^o 2. [En ligne] < http://www.digitalstudies.org/ojs/index.php/digital_studies/article/view/176 >.

Rogers Richard, "Mapping Public Web Space with the Issuecrawler", in Bernard Reber et Claire Brossaud (dir.), *Digital Cognitive Technologies. Epistemology and the Knowledge Economy*, Hoboken NJ, Wiley, 2013.

Rosental Claude, « Les travailleurs de la preuve sur Internet », *Actes de la recherche en sciences sociales*, 2000, n^o 134, pp. 37-44.

Rosenzweig Roy, "Abundance? Preserving the Past in a Digital Era, Scarcity or.", *American Historical Review*, 2003, 108, pp. 735-762.

Rosenzweig Roy, "Should Historical Scholarship Be Free?", *Perspectives on History*, 2005. [En ligne] < <https://www.historians.org/publications-and-directories/perspectives-on-history/april-2005/should-historical-scholarship-be-free> >.

Rosenzweig Roy, *Clio Wired. The Future of the Past in the Digital Age*, Columbia, Columbia University Press, 2011.

Roux Magali, *Biologie. L'ère numérique*, Paris, CNRS Éditions, 2009.

Rygiel Philippe, « L'ordinateur, le réseau et l'écriture de l'histoire », *Matériaux pour l'histoire de notre temps*, 2/2006, n^o 82, pp. 75-79.

Rygiel Philippe (dir.), *Le bon grain et l'ivraie. La sélection des migrants en Occident, 1880-1939*, Paris, Publibook, 2005.

Rygiel Philippe, Noiret Serge (dir.), *Les historiens, leurs revues et Internet (France, Espagne, Italie)*, Paris, Publibook, 2005.

Samosate Lucien de, *Comment il faut écrire l'histoire*, in Eugène Talbot (dir.), *Œuvres complètes de Lucien de Samosate* [traduction nouvelle avec une introduction et des notes par Eugène Talbot], Paris, Hachette, 1912, t. 1, pp. 353-379.

Saunier Pierre-Yves, « Des listes électroniques : pourquoi, pour qui ? Notes sur les historiens français et les communautés en ligne », in Philippe Rygiel et Serge Noiret (dir.), *Les historiens, leurs revues et Internet (France, Espagne, Italie)*, Paris, Publibook, 2005, pp. 115-128.

Schafer Valérie, *Une histoire de la convergence. Les technologies de l'information et de la communication depuis les années 1950*, Habilitation à diriger des recherches, Université Paris-Sorbonne, 2015.

Seignobos Charles, Langlois Charles-Victor, *Introduction aux études historiques*, Paris, Kime, 1992 [1898].

Sfez Lucien, *Technique et idéologie. Un enjeu de pouvoir*, Paris, Seuil, 2002.

SPIP, « L'histoire minuscule et anecdotique de SPIP », s.d. [En ligne] < www.spip.net >.

Spire Alexis, *Étrangers à la carte. L'administration de l'immigration en France, 1945-1975*, Paris, Grasset, 2005.

Staley David D.J., "The AAHC: a Visionary Organisation", *Journal of the Association for History and Computing*, 2005, 8 (1). [En ligne] < <http://hdl.handle.net/2027/spo.3310410.0008.105> >.

Thaler Manfred, *A Data Base System for Historical Research. Version 1.1.1, b-test Version*, Göttingen, Göttingen, Max-Planck-Institut für Geschichte, 1987.

Tiffin John, Rajasingham Lalita, *The Global Virtual University*, Londres, Routledge, 2003.

Villani Cédric, « L'écriture des mathématiques. Réflexions sur la façon dont les mathématiciens écrivent la science », in Éric Guichard (dir.), *Écritures. Sur les traces de Jack Goody*, Villeurbanne, Presses de l'enssib, 2012, pp. 199-212.

Vivant Michel, « Gare au portail », *Sciences de l'homme et de la société*, 2004, n° 69.

PAPIERS

Créée en 2008, la collection Papiers a pour ambition d'allier la recherche universitaire et professionnelle autour des sciences de l'information et des bibliothèques. Elle rassemble des points de vue variés et originaux, des spécialistes de plusieurs disciplines, des approches contemporaines ou historiques. Car rien n'est moins évident que la civilisation de l'écrit dans laquelle nous croyons baigner naturellement !

André-Pierre Syren
directeur de collection

+++++

ENSSIB – UNIVERSITÉ DE LYON
PRESSES DE L'enssib
École nationale supérieure des sciences
de l'information et des bibliothèques
17-21 boulevard du 11 novembre 1918
69623 Villeurbanne Cedex
Tél. 04 72 44 43 43 – Fax 04 72 44 43 44
< <http://www.enssib.fr/presses> >

Cet ouvrage a été réalisé avec Métopes (méthodes et outils pour l'édition structurée XML-TEI), chaîne éditoriale développée par le pôle Document numérique de la Maison de la recherche en sciences humaines de l'université de Caen.

Secrétariat d'édition :
Silvia Ceccani

Mise en page :
Cédric Vigneault

Secrétariat administratif :
Véronique Bolinde

Conception graphique :
atelier Perluette, 69001 Lyon.
< <http://www.perluette-beaufixe.fr> >

Achevé d'imprimer en novembre 2017
imprimerie Bialec (Nancy) - n° 92185



dépôt légal : 2^e semestre 2017