

ANR/ATHIS – Ateliers Histoire et informatique, 1
De l'archive à l'*open archive* : l'historien et internet
(École française de Rome, 23-25 mars 2006)
http://www.menestrel.fr/IMG/pdf/05_smith.pdf

L'aube des archives globales ?

Marc H. Smith
École nationale des chartes

Je crois avoir compris que nous sommes tous d'accord pour considérer que l'accès en ligne aux documents offre, voire impose, à chaque historien le moyen de renouveler profondément les méthodes, les questionnaires, les problématiques mêmes qu'il applique à son corpus documentaire. Je voudrais seulement poser quelques questions sur la constitution même de ce corpus, étape fondamentale de toute recherche. Une étape à laquelle Ménestrel a vocation à contribuer ; c'est même son ambition principale, dans cette phase où les ressources en ligne se multiplient mais restent souvent méconnues ou sous-exploitées.

L'accès aux documents en ligne offre des possibilités sans précédent, c'est entendu. Mais comment l'articuler avec le recours aux documents originaux ? J'entends à la fois les originaux qui n'ont pas été numérisés — ou pas encore, mais le seront-ils un jour ? —, et ceux qui sont ou seront en même temps disponibles sur écran.

Quelques réflexions seulement, et veuillez mettre leur caractère un peu abrupt sur le compte de la brièveté.

Arnold Toynbee remarquait déjà cette particularité des historiens : par contrainte ou par facilité, leur interrogation du passé se limite aux questions dont ils savent ou estiment que la réponse pourra être trouvée dans les sources accessibles. Souvent même, le choix du corpus précède et détermine la construction de la question. Or le choix du corpus et son adéquation à la question ne sont pas libres, mais soumis à un cadre économique contraignant, autrement dit à des contingences de temps, d'argent, de distance, de carrière et d'énergie individuelle.

L'Internet documentaire relâche spectaculairement ces contraintes, ou plutôt, pour le moment du moins, il lui offre des soupapes, des valves localisées, en fonction de ce qui arrive sur le marché numérique. Il permet non seulement d'exploiter au mieux mon corpus habituel, si c'est moi-même qui le numérise et le triture en XML, mais d'accéder à certaines autres sources que j'aurais autrefois négligées comme périphériques à ma recherche propre, et que désormais je choisis précisément parce qu'elles sont en ligne, quelle que soit la conception

voire la qualité de cette mise en ligne. A l'automne dernier, par exemple, pour mieux comprendre la diffusion internationale de l'écriture de chancellerie française au XIVe siècle, j'ai pu consulter de chez moi l'ensemble des chartes médiévales de la ville de Duderstadt en Allemagne. Pourquoi Duderstadt ? Parce que ses chartes sont en ligne.

L'Internet, pour l'heure, offre essentiellement des corpus expérimentaux, des sources circonscrites, qui pour la plupart des recherches ne peuvent jouer qu'un rôle complémentaire, et ne nous dispensent donc que marginalement de visiter les archives et les bibliothèques. Mais des entreprises de numérisation massive, traitées de manière plus sommaire, sont aussi en cours à l'initiative des institutions de conservation, et quelques-unes ont même abouti. Jusqu'où iront-elles dans la couverture des fonds, dans la finesse du traitement, et avec quelle vitesse de progression ? Dans quelle mesure ces sources nouvellement accessibles et très disparates modifieront-elles, volontairement ou non, individuellement ou collectivement, les stratégies de recherche ?

La pléthore documentaire qui s'annonce ouvre-t-elle une recherche sans frontières et sans limites ? Oui et non. Si j'étais auparavant pris en tenaille entre mes finances, mon temps et le rationnement des communications dans les archives ou les bibliothèques, je serai désormais confronté avant tout à mes propres limites : ma capacité à m'orienter dans les documents, à les absorber et à les interpréter.

La numérisation brouille les catégories traditionnelles de la documentation. Non seulement la notion d'édition comme on l'a dit hier soir, mais aussi bien celle, cruciale, de l'original, en créant des doubles virtuels de la source primaire ; doubles réalisés de manière souvent disparate, voire instable, mais qui pourraient bien devenir, dans la pratique, des substituts.

L'univers du papier n'était évidemment pas non plus sans pièges ni ambiguïtés ; une édition ou même une étude, quelle que soit leur exactitude ou leur complétude, vous dispensent ou plutôt vous dissuadent de l'effort qu'exigerait le retour aux sources primaires pour tout vérifier ou pour étendre votre corpus. Pourtant les préoccupations propres de chaque chercheur ne coïncident jamais exactement avec celles de l'édition ou de l'étude qu'il utilise ; et de lecture en relecture, de glose en glose, s'accumulent les choix, les grilles d'interprétation, les extrapolations ou les erreurs, que seul le recours à l'original dissiperait — et parfois, quand on en prend la peine, dissipe.

La multiplication des fac-similés permet de ce point de vue un progrès considérable. Un grand historien (Carlo Ginzburg) a donné il y a quelques années une petite étude sur Jean Fouquet, reposant en large partie sur un document inédit qui révolutionnait la biographie du peintre. Après une tournée internationale de conférences et séminaires, il a enfin publié un livre qui, malheureusement ou heureusement, incluait une photo du document. Une photo très réduite, mais où les esprits malicieux ont aussitôt relevé deux erreurs de lecture fatales. Le document ne parlait ni de Fouquet, ni même de peinture.

Le fac-similé toutefois n'est pas non plus l'original. Très récemment, un célèbre document de l'historiographie bénédictine médiévale d'Italie du Sud (deux feuillets d'une chronique de S. Vincenzo al Volturmo), perdu depuis un demi-siècle, a été retrouvé derrière un meuble dans la bibliothèque d'un éminent paléographe. Or le retour de l'original, aussitôt décortiqué par une équipe interdisciplinaire, a montré que toutes les recherches publiées jusqu'ici, basées sur l'étude et le fac-similé parus en 1925, étaient à jeter, parce que le fragment ne datait pas de 1100, mais de l'an mille.

Même sans microscope ni analyse chimique, la matérialité du document s'impose d'elle-même à la consultation attentive d'un chercheur formé et expérimenté : le rapport physique et logique de la pièce avec les documents voisins, son intégrité codicologique, jusqu'à la texture du parchemin, tout est indice de ses significations possibles. Et le dépouillement manuel, la lecture en série de documents, même si je n'y trouve rien d'utile, mon voyage même vers un dépôt d'archives en Toscane ou dans le Rouergue, les instruments de recherche locaux, la conversation du conservateur et des autres hôtes de la salle, nourrissent ma perception du contexte documentaire mais aussi ma sensibilité au temps et au lieu. De mon bureau à Paris ou à Princeton, lirai-je le document de la même manière ?

Du point de vue des principes, j'espère que nous sommes encore d'accord au moins sur ce point : ni l'édition ni le fac-similé ne *remplacent* l'original. Et une formation spécifique à la sélection et à la critique des documents numérisés (Oliver Guyotjeannin y faisait allusion hier) est un impératif qui s'ajoute à l'arsenal traditionnel des méthodes applicables aux sources primaires. L'Internet, ici, comme ailleurs, favorisera ceux qui ont une vision globale, hiérarchisée, critique de l'information disponible, et non ceux qui y jettent leur époussette au hasard. Cet apprentissage aussi fait partie de l'ambition de Ménestrel.

Mais dans la pratique, la facilité apparente, l'illusion de tout avoir à portée de main, risque bien d'entrer en conflit avec la méthode. Combien de temps, avant que les documents originaux ne soient réduits au rôle de complément ponctuel de grands corpus en ligne prenant, usurpant, le rang de documents primaires ?

La logique des numérisations, ponctuelles ou de masse, privilégie à juste titre les documents prestigieux, les documents connus, les documents les plus utilisés. Les chercheurs numérisent, éditent et xmlisent pour leur compte ce qui les intéresse aujourd'hui, et les conservateurs numérisent ce qui est le plus menacé par la consultation. Or cette réconciliation inespérée des priorités de la recherche et de la conservation ne fait qu'accentuer l'écart entre ces fonds favorisés et les autres, déjà moins fréquentés, moins connus, souvent moins bien inventoriés — réduisant d'autant l'incitation à explorer de nouveaux corpus, réserves de renouvellements historiographiques imprévisibles ; à moins que la numérisation ne veuille viser à la mise en ligne exhaustive des fonds, mais cela paraît improbable à proche et à moyen terme.

Dans une université où la fécondité d'une recherche se juge au nombre des colloques, où l'innovation se confond avec l'informatique, où la durée de la thèse dissuade de longs et lointains dépouillements, la numérisation des documents, gage de recherches approfondies pour quelques-uns, accentue par d'autres aspects la tentation éternelle du travail rapide et superficiel ; tentation à laquelle la pesanteur de l'investissement personnel dans les originaux résiste d'elle-même. Au surf et au terrassement correspondent des états d'esprit différents. La dérive est manifeste dès maintenant dans l'évolution de la bibliographie annexée à plus d'un travail universitaire : des milliers de titres déversés d'Internet, dont cinq ou dix pour cent ont été lus.

L'indexation des images numériques, plus ou moins poussée, et l'accès direct à chaque feuillet, instruments de travail extraordinaires, n'en sont pas moins à double tranchant. Ils nous dispensent déjà largement de chercher à comprendre la structure organique des fonds, donc de dialoguer avec les archives selon leur propre logique, et ne laissent guère plus de temps à cette *rumination* dont parlait encore O. Guyotjeannin. Puisque nous sommes au pays qui a réinventé le *slow food*, j'ose plaider à mon tour pour les vertus digestives de la *slow research*.

Quant aux originaux désormais disponibles en reproduction, les exemples actuels de la Bibliothèque nationale de France, des Archives vaticanes et de bien d'autres montrent assez que ce sera la justification de leur enfermement définitif, opposable même aux chercheurs qui auraient su exploiter la différence entre l'original et son double, comme à ceux à qui il importerait de l'apprendre. En effet je m'inquiète particulièrement pour les jeunes chercheurs formés dans ce contexte. Alors que le monde virtuel et le monde réel s'interpénètrent déjà pour eux avec des conséquences impensables voici dix ans, cette génération aura-t-elle encore cette dernière chance de toucher de la main l'épaisseur des temps ?

Permettez-moi, pour rendre mon propos tout à fait déplacé, de poser la question en termes franchement sentimentaux, parce que, non moins que les théoriciens des passions au Seicento, que les ayatollahs ou que les publicitaires, je crois à l'affect comme moteur des comportements et de l'économie qui en découle. Puisque nous ne faisons pas notre métier par appât du lucre, il faut donc que ce soit par plaisir. Les plaisirs sont variés selon les tempéraments, depuis le goût du pouvoir jusqu'à l'odeur des vrais livres en papier. Qui osera dire combien de nos collègues, qui coulaient des jours heureux rue de Richelieu, à tort ou à raison, ont déserté la BnF après son transfert parce qu'on les avait dépossédés de leur plaisir ?

Une question par exemple où le principe, l'usage scientifique et l'affect sont étroitement mêlés : si le document numérisé n'est ni édition, ni original, quel statut donner à la découverte et à la publication d'un document inédit... quand on a trouvé ce document sur Internet ? La question n'est pas non plus entièrement nouvelle : ainsi il va paraître prochainement un important document inédit sur Léonard de Vinci,

inédit pour les spécialistes de Léonard mais résumé depuis longtemps dans le *Calendar of State Papers*. Elle risque cependant de devenir autrement plus fréquente et plus aiguë.

Beaucoup des plus importants et plus fructueux dépouillements d'archives, individuels ou collectifs, n'auraient pas été entrepris et menés à bien sans ce qu'Arlette Farge a appelé le goût de l'archive. Les archivistes, un peu plus blasés qu'elle, ont pu sourire de ses déclarations d'amour à la poussière. Mais ils n'oublient pas que le premier à avoir évoqué le même vertige du temps sédimenté en millions de parchemins et de papiers est un des leurs, Jules Michelet. C'est le célèbre passage où il convoque, dans cette pénombre des Archives nationales que les parchemins remplissent d'odeurs charnelles les soirs d'été, la « danse galvanique des morts ». J'espère que la danse numérique nous donnera encore autant de plaisir, de curiosité fébrile et d'imagination créatrice que la danse galvanique.

Cela dit, je ne voudrais pas jouer seulement la fée Carabosse de la fête, ou l'acariâtre renifleur de vieux papiers. Si je suis inquiet des dérives possibles, c'est au contraire que je vois combien de temps je suis capable de passer sur un écran, qui aurait peut-être été employé plus fructueusement aux archives ou en bibliothèque. Terminons donc plutôt avec un espoir, dans ce jeu entre la réalité et virtualité : celui que le contact avec le double virtuel des documents originaux soit utilisé à bon escient pour susciter des curiosités et des compétences nouvelles, aussi bien chez les jeunes historiens que chez ceux de leurs aînés qui précédemment ne seraient jamais sortis du document *imprimé*. C'est l'ambition de la rubrique « manuscrits et paléographie », entre autres, dans *Ménešrel*, aussi bien que celle des entreprises d'édition et de pédagogie numériques de l'École des chartes. La combinaison étroite de l'image avec un traitement critique approfondi, appuyé sur la restitution virtuelle aussi complète que possible des originaux eux-mêmes, combinaison commune à bien d'autres projets que nous avons vu présenter ici, reste sans doute la meilleure garantie contre la dissolution du texte médiéval, objet toujours unique, dans le flot indifférencié de l'information globale.