



Recherches & Travaux

72 | 2008

De l'hypertexte au manuscrit

Introduction. De l'hypertexte au manuscrit : le manuscrit réapproprié

Enjeux, expérimentations, perspectives

Françoise Leriche et Cécile Meynard



Édition électronique

URL : <http://journals.openedition.org/recherchestravaux/82>

ISSN : 1969-6434

Éditeur

UGA Éditions/Université Grenoble Alpes

Édition imprimée

Date de publication : 15 juin 2008

Pagination : 9-36

ISBN : 978-2-84310-125-0

ISSN : 0151-1874

Référence électronique

Françoise Leriche et Cécile Meynard, « Introduction. De l'hypertexte au manuscrit : le manuscrit réapproprié », *Recherches & Travaux* [En ligne], 72 | 2008, mis en ligne le 15 décembre 2009, consulté le 01 mai 2019. URL : <http://journals.openedition.org/recherchestravaux/82>

Introduction

De l'hypertexte au manuscrit : le manuscrit réapproprié

Enjeux, expérimentations, perspectives

Du manuscrit à l'hypertexte – de la tablette d'argile à l'ordinateur en passant par le parchemin et le papier sous leurs formes de *volumen* et de *codex* –, l'histoire des *media* d'écriture est désormais bien explorée. Histoire des supports, des outils, et des formes de l'écriture¹, histoire du livre, de l'imprimerie, et de l'édition², analyse des impacts épistémologiques et médiologiques des techniques de communication³, les études dans ces domaines se sont multipliées depuis une trentaine

1. Pour ce qui concerne l'histoire de l'écriture manuscrite de l'Antiquité à l'époque moderne (paléographie et codicologie), voir l'énorme bibliographie (pourtant sélective) constituée par des spécialistes de l'École des chartes sur le site «Thélème», consultable à l'adresse : <http://theleme.enc.sorbonne.fr/sommaire69.html>. Parmi les travaux récents impliquant une réflexion sur l'écriture numérique : *Hypertextes et hypermédias : réalisations, outils, méthodes*, sous la direction de J.-P. Balpe *et al.*, Hermès, 1995 ; *Histoire de l'écriture, de l'idéogramme au multimédia*, sous la direction d'A.-M. Christin, Flammarion, 2001.

2. Pour l'histoire du livre et de l'édition, voir en particulier la grande *Histoire de l'édition française*, dirigée par R. Chartier et H.-J. Martin, Promodis, 1982-1986, rééd. Fayard, 1989-1991, en 4 tomes : [1] *Le Livre conquérant : du Moyen Âge au milieu du XVII^e siècle*; [2] *Le Livre triomphant : 1630-1830*; [3] *Le Temps des éditeurs : du romantisme à la Belle Époque*; [4] *Le Livre concurrencé : 1900-1950*. Voir aussi J.-F. Gilmont, *Le Livre, du manuscrit à l'ère électronique*, Liège, éditions du CLPCF, 1989; F. Baudin, *L'Effet Gutenberg*, Cercle de la Librairie, 1994, qui réfléchit à la façon dont le numérique menace une certaine esthétique de l'imprimé. – Pour l'histoire de la forme *codex*, voir notamment C. H. Roberts et T. C. Skeat, *The Birth of the Codex*, Londres, Oxford University Press, 1987; A. Blanchard éd., *Les Débuts du codex*, Turnhout, Brepols, 1989.

3. Outre les études signalées ci-dessus dans les notes 1 et 2, qui abordent bien souvent les implications épistémiques des techniques et des formes de l'écrit, voir par exemple, pour l'imprimerie,

d'années : l'apparition de l'informatique a en effet amené la communauté savante à s'interroger en retour sur le *medium* jusqu'alors majoritaire, le livre imprimé, et à s'apercevoir qu'il n'était qu'un moment dans la longue histoire des formes de l'écrit. Et à prendre conscience que l'histoire de l'écrit comme celle de la culture en général, se caractérise par la *dématérialisation* de plus en plus rapide du support et des modes d'inscription.

Devant l'économie de l'immatériel qu'inaugure l'ère numérique, les acteurs du monde du livre manifestent leur inquiétude depuis plusieurs années; de nombreux colloques ou publications tentent de déterminer si « ceci tuera cela » et de reconfigurer l'espace de l'édition et des bibliothèques entre les secteurs qui resteront du domaine de l'édition imprimée et ceux qui, inévitablement, émigreront vers l'édition électronique⁴. Il paraît toutefois acquis pour un grand nombre d'éditeurs que les publications universitaires périodiques mais aussi les éditions savantes, sans compter l'édition scolaire, se feront de plus en plus sous forme électronique⁵, la question étant alors de savoir quelles instances géreront cette économie. Or, au sein de ce système économique, s'il est un secteur de l'édition trop peu rentable pour les éditeurs commerciaux où le numérique – du fait des progrès technologiques de ces dernières années – s'avère prometteur pour les chercheurs, c'est celui de l'édition savante et en particulier, domaine plus marginal encore, celui de l'édition des manuscrits et des archives littéraires.

Pourtant, un intérêt collectif pour les *processus créateurs* ne cesse de s'affirmer. Dans le domaine de la recherche, le reflux du structuralisme a favorisé l'essor des études de genèse et l'édition de documents préparatoires d'écrivains célèbres

les travaux d'E. Eisenstein, *The Printing Press as an Agent of Change. Communications and Cultural Transformations in Early Modern Europe*, Cambridge University Press, 1979, et *The Printing Revolution in Modern Europe*, Cambridge University Press, 1983; pour le numérique, voir J. Clément, « Du texte à l'hypertexte : vers une épistémologie de la discursivité hypertextuelle », in J.-P. Balpe *et al.*, *op. cit.*, consultable à l'adresse <http://hypermedia.univ-paris8.fr/jean/articles/discursivite.htm>. – Pour l'analyse médiologique, rappelons quelques références désormais classiques : M. McLuhan, *La Galaxie Gutenberg : face à l'ère électronique, les civilisations de l'âge oral à l'imprimerie*, trad. J. Paré, Tours, Mame, 1967 (traduction intitulée à partir de 1977 *La Galaxie Gutenberg : la genèse de l'homme typographique*, Gallimard, « Idées »); R. Debray, *Cours de médiologie générale*, Gallimard, 1991 (rééd. 2001) et *Les Révolutions médiologiques dans l'histoire*, Villeurbanne, ENSSIB, 1999, qui prennent le relais de McLuhan en reconfigurant le système des « médiasphères ».

4. Ainsi, la revue des anciens élèves de l'École normale supérieure, *L'Archicube*, consacre son numéro 3 (décembre 2007) à un dossier conséquent sur cette question, « Le numérique et l'édition », auquel ont contribué de nombreux éditeurs, bibliothécaires, et responsables gouvernementaux.

5. Voir par exemple les perspectives esquissées par le directeur éditorial de Hachette Livre International J.-M. Ollier, dans « Le livre électronique : une fausse bonne idée », *L'Archicube*, *op. cit.*, p. 29-35.

(carnets, plans et scénarios⁶, esquisses et ébauches, projets abandonnés, versions inédites⁷). Cet intérêt se voit confirmé jusque dans le domaine de la consommation culturelle de masse, les DVD de films s'enrichissant désormais d'un supplément comprenant des scènes abandonnées, des documents de tournage, etc. : le «*making of*» fascine, le public demande à entrer dans les coulisses de la création... Or les manuscrits de travail⁸ des écrivains – objets mythiques acquis grâce à l'argent des contribuables ou des mécènes à des prix de plus en plus élevés –, sont soigneusement enfermés dans les réserves des bibliothèques ; quelquefois exposés, ils fascinent par leur valeur testimoniale, mais ne peuvent pas « parler », ne peuvent pas fournir aux visiteurs de réelles indications sur les processus créateurs tant qu'ils restent derrière des vitrines fermées. Il faut donc permettre au public de se réapproprier son bien, de se familiariser avec ces objets intimidants, fragiles, que seuls les « savants » ont le droit de manipuler avec des gants et d'innombrables précautions. Mais par quel(s) moyen(s) ?

Longtemps les manuscrits de travail ont été négligés comme « illisibles », les universitaires eux-mêmes ne s'intéressant qu'aux dernières étapes (« manuscrit au net », copie ou dactylographie, épreuves corrigées) pour établir le texte et fournir les « variantes » des éditions savantes. Lorsque les études génétiques se sont développées dans les années 1960-1970, spécialistes et étudiants avancés ont donc pu, dans le saint des saints des bibliothèques, analyser à loisir des archives de la création littéraire encore inexploitées et en tirer d'importantes études sur la genèse de tel ou tel texte, sur les processus d'écriture de son auteur, etc., publiées dans des ouvrages et périodiques spécialisés ; mais précisément parce que ces études s'appuyaient sur des documents connus seulement de *happy few*, leurs résultats pouvaient difficilement intéresser le plus grand nombre, le lecteur voulant en général – et à juste titre – pouvoir vérifier par lui-même et discuter les thèses et les conclusions qu'on lui propose. Après

6. Par exemple : M. Proust, *Le Carnet de 1908*, édité par Ph. Kolb, Gallimard, 1976 ; É. Zola, *Carnets d'enquête. Une ethnographie inédite de la France par Émile Zola*, textes établis et présentés par H. Mitterand, Plon, 1986 ; G. Flaubert, *Carnets de travail*, édition critique et génétique par P.-M. de Biasi, Balland, 1988 ; G. Flaubert, *Plans et scénarios de Madame Bovary*, présentation, transcription et notes par Y. Leclerc, Zulma / CNRS éditions, 1995, etc.

7. M. Proust, *Matinée chez la Princesse de Guermantes, Cahiers du Temps retrouvé*, édition critique établie par H. Bonnet avec la collaboration de B. Brun, Gallimard, 1982 ; P. Valéry, *Cahiers 1894-1914*, édition intégrale établie et annotée par N. Celeyrette-Pietri, Gallimard, depuis 1987, etc.

8. On distingue les manuscrits « de travail », documents autographes où le texte s'élabore peu à peu, de ratures en reprises, des manuscrits médiévaux, qui sont des manuscrits « de diffusion » (calligraphiés par des copistes). Ils se distinguent aussi de ce qu'on pourrait appeler manuscrits « d'enregistrement », dictés à des secrétaires ou à des copistes qui, selon une pratique courante sous l'Ancien Régime, recueillent une élaboration mentale ou faite sur des documents dont on n'a pas de trace, élaboration que, par conséquent, on ne peut pas étudier.

plusieurs années d'études génétiques confidentielles, il est donc apparu nécessaire d'*éditer* ces documents inédits. C'est ainsi que, dans les années 1980, est apparue la notion d'«édition génétique⁹», le concept allant de pair avec une nouvelle ambition éditoriale : éditer non pas des «textes» mais des brouillons en expansion...

Un nouvel objet littéraire et éditorial : le texte mouvant

Almuth Grésillon, linguiste et généticienne, soulignait en 1992 toute la difficulté, voire l'impossibilité constitutive, de cette entreprise éditoriale :

[...] là où la forme du texte manifeste une structure finie et une version unique, consacrée par une édition canonique, l'avant-texte, par l'épaisseur des réécritures, se révèle radicalement incompatible avec une représentation textuelle à deux dimensions¹⁰.

Comment alors *éditer* des manuscrits de travail ? Comment rendre lisibles des ensembles complexes de brouillons inachevés, de versions fragmentaires, de notes hétéroclites, comme les manuscrits de Stendhal, de Pascal ? comment rendre lisibles des brouillons enchevêtrés (endroits, envers, marges, versions de diverses époques en regard, saut chronologique et bifurcation de projets d'une page à l'autre, comme les cahiers de Proust, entre autres exemples ?) Après une décennie de publications génétiques très diverses dans leurs protocoles de transcription et dans leurs pratiques, un colloque organisé en 1995 à Vincennes par Béatrice Didier et Jacques Neefs, «Éditer des manuscrits. Archives, complétude, lisibilité» (voir note 9), cherchait à tirer le bilan de ces travaux et posait explicitement la question de la «lisibilité» tandis que, plus radicalement, certaines contributions posaient la question de la *pertinence* de l'édition génétique dans les limites bidimensionnelles de l'édition papier.

Pierre-Marc de Biasi, distinguant deux types d'éditions génétiques, l'édition «horizontale» (édition de documents d'une même strate dans le processus d'écriture : par exemple, les notes préparatoires) et l'édition «verticale» (édition en diachronie des différentes étapes dans la rédaction du texte), concluait que

9. Pour un historique de ce concept et un classement typologique des éditions génétiques, voir P.-M. de Biasi, «Édition horizontale, édition verticale. Pour une typologie des éditions génétiques (le domaine français 1980-1995)», *Éditer des manuscrits. Archives, complétude, lisibilité*, études réunies et présentées par B. Didier et J. Neefs, Saint-Denis, Presses universitaires de Vincennes, «Manuscrits modernes», 1996, p. 159-193. Mis en ligne le 18 janvier 2007, disponible sur : <http://www.item.ens.fr/index.php?id=13346>.

10. A. Grésillon, «Ralentir : travaux», *Genesis*, n° 1, 1992, p. 9-31.

l'édition imprimée pouvait servir à l'édition horizontale d'ensembles peu volumineux et bien distincts dans le processus créateur (carnets de notes, plans et scénarios, manuscrit au net), mais que l'édition verticale, c'est-à-dire l'édition exhaustive des multiples étapes rédactionnelles, était impraticable (sauf pour de petits corpus comme des contes), d'abord pour des raisons de volume (milliers de pages) mais surtout parce que le livre imprimé ne permet pas de représenter de manière satisfaisante les versions successives d'une même unité textuelle (page, scène ou épisode), la dimension synchronique de chaque strate génétique (la continuité de chaque version du texte) entrant ainsi en conflit avec la dimension diachronique des réécritures (qui exigerait un système non linéaire pour représenter l'empilement des strates d'écriture) :

L'édition verticale est inséparable de la reconstitution raisonnée d'un processus qui met en jeu des phénomènes structuraux temporalisés dans un espace à trois dimensions. Le modèle de la pile et des strates permet d'en donner une représentation logique satisfaisante, mais difficilement transposable dans le cadre des éditions traditionnelles. *Nul doute qu'avec sa forme séquentielle et ses deux dimensions, le livre ne constitue plus aujourd'hui un support suffisant pour enregistrer et rendre intelligible ce type d'objet. L'édition électronique, enrichie de ses ressources presque illimitées de l'hypertexte et du multimédia, constitue l'horizon naturel des recherches éditoriales sur l'avant-texte.* (Art. cit., paragraphe conclusif. – Nous soulignons.)

Ce constat de 1995 laissait donc la porte ouverte à deux solutions seulement : ou bien la décision de s'en tenir à des éditions verticales « partielles », c'est-à-dire non exhaustives (édition des brouillons successifs d'un seul épisode, choisi comme « représentatif » de la méthode de travail de l'écrivain¹¹, même si la troisième dimension, celle du temps, est mal représentée par la séquentialité de l'édition papier... ; ou bien l'abandon de l'édition imprimée au profit de l'*expérimentation* de l'édition électronique hypertextuelle, dans le but de pouvoir – en mutualisant les travaux – atteindre l'exhaustivité en termes de corpus (nombre de pages, nombre de versions) et surtout de disposer d'un outil capable de représenter la *complexité des relations* entre ces diverses versions (structure rhizomatique, ou réticulaire). Comme le notait dès 1993 Jean-Louis Lebrave : « [...] l'hypertexte introduit ce qu'on serait tenté d'appeler une méta-architecture à laquelle la multiplicité des réseaux de relations transversales insuffle une vie inconnue de l'imprimé dans sa rigidité linéaire et réductrice¹² ».

11. Par exemple, *Les Comices agricoles de Gustave Flaubert*, transcription intégrale et genèse dans le manuscrit g 223 par J. Goldin, Droz, 1984.

12. J.-L. Lebrave, « L'édition génétique », *Les Manuscrits des écrivains*, sous la direction de L. Hay, Hachette/CNRS éditions, 1993, p. 222.

En effet, depuis quelques années, l'hypertexte informatique donne la possibilité technologique d'éditer des dossiers complexes, les hyperliens dynamiques permettant de lier des pages ou même des unités de manière rhizomatique là où l'édition papier obligeait à une concaténation linéaire, et de visualiser de manière synthétique toutes ces relations par des schémas hypergraphiques, etc. En outre, l'édition électronique rend caduque la distinction entre « édition horizontale » et « édition verticale », puisqu'on peut, dans une base de données, classer les textes selon leur strate respective d'appartenance, et aussi, grâce à un système d'annotation et d'indexation spécifique, permettre à l'utilisateur de les convoquer dans l'ordre diachronique des versions (voir les articles de Dirk Van Hulle, de Daniel Ferrer, Jean-Jacques Labia, Maria-Ignez Mena Barreto, ou d'Hélène Spengler). On offre ainsi des choix multiples au lecteur : celui de faire une lecture linéaire, « traditionnelle », d'une version ou d'une étape rédactionnelle donnée, ou au contraire, grâce à un moteur de recherche et/ou une navigation hypertextuelle, et à un système de superposition paradigmatique ou de multifenêtrage, d'effectuer un parcours à travers différentes versions d'un même manuscrit ou à travers l'ensemble du fonds, de visualiser et de comparer les différentes strates d'écriture.

On pourrait donc penser que l'informatique supprime les dilemmes de l'édition génétique d'antan. Cependant, du fait des possibilités offertes par la compression des données numériques et l'accroissement des capacités de stockage, une nouvelle question éditoriale est apparue, qui ne se posait pas dans le cas de l'édition imprimée : puisque c'est désormais possible, ne faut-il pas publier *aussi* les fac-similés photonumériques des originaux ? Dans ce cas, *quelle relation établir entre la transcription* (qui était dans l'édition imprimée le seul mode de représentation, *textuel*, des manuscrits) *et l'image* du document (c'est-à-dire sa représentation *visuelle*) ? Le document est-il un supplément, à valeur purement illustrative ? ou n'est-il pas le centre même du projet éditorial, la transcription – aussi fidèle que possible à la disposition spatiale de la page – ne devant apporter qu'une aide à la lecture de l'original ? Ces questions du rapport entre texte et document, posées par Hans Gabler (p. 39) et Dirk Van Hulle (p. 53), sont véritablement cruciales, car elles engagent le sens même des entreprises actuelles d'édition génétique électronique, aussi bien que la structure même des bases de données : l'édition sera-t-elle « documentaire » (donc consultable à partir de la convocation d'un document accompagné de sa transcription), ou « textuelle » (interrogeable à partir d'une chronologie des strates rédactionnelles et d'une indexation des différentes versions des unités textuelles) ? Une solution mixte est-elle possible (voir Dirk Van Hulle) ? Actuellement, devant la multiplicité des projets et des expérimentations, il est beaucoup trop tôt pour savoir si l'une de ces pratiques prévaudra sur l'autre, mais il paraît assuré que la typologie des éditions

génétiques ne se situe plus entre horizontales et verticales, mais entre documentaires et textuelles...

Le plus intéressant, c'est que ce déplacement des enjeux éditoriaux rejaillit sur l'édition imprimée : l'édition exhaustive des soixante-quinze cahiers de brouillon d'*À la recherche du temps perdu* de Proust, dont le premier tome (le Cahier 54) va paraître chez Brepols en mai 2008, est conçue selon le modèle de l'édition «documentaire», chaque folio étant reproduit en fac-similé et accompagné d'une transcription diplomatique respectant aussi fidèlement que possible la disposition des blocs textuels sur le brouillon, avec un système cartographique pour permettre au lecteur de se repérer dans le maquis des unités textuelles proustiennes (voir l'article de Nathalie Mauriac Dyer, p. 157).

On peut mesurer ainsi à quel point la numérisation a modifié les problématiques de l'édition génétique depuis la fin des années 1990, et combien les choix de «lisibilité» sont au cœur des débats : une lisibilité définie soit en termes de *fidélité* au document, d'aide à la lecture du fac-similé (la transcription topographique servant d'aide à la lecture de tous les méandres de l'écriture, les unités textuelles étant transcrites à l'identique par rapport au document); soit en termes de *fluidité de lecture*, au prix d'une certaine simplification de la transcription (la présence du fac-similé permettant au lecteur de voir les simplifications du transcripteur, et de mesurer le travail réel de l'auteur, ses hésitations, l'ambiguïté dans l'insertion de certaines marginales, etc.). Or ces deux conceptions de la lisibilité ne recouvrent-elles pas deux conceptions différentes du lectorat ?

Une édition diplomatique, qui respecte totalement la présentation du feuillet manuscrit dans ses ratures, ses ajouts, ses béquets, est tellement exigeante qu'elle paraît *a priori* réservée à des spécialistes, qu'on imagine seuls capables de supporter une lecture morcelée. On a tendance à supposer que le «grand public» et les étudiants amenés à faire des recherches à travers les manuscrits d'un texte préféreront des transcriptions plus fluides, sans se formaliser de ce que l'éditeur ait occasionnellement omis des ratures ou interprété l'ordre des fragments au lieu de les restituer tels quels... Mais en ce domaine très récent, nous avons encore trop peu de recul, et il serait peut-être plus sage de penser que ce sont les éditions qui créeront et éduqueront le lecteur, plutôt que de partir *a priori* des «préférences» supposées d'un lectorat encore inexistant.

Ce qui est certain pour tous, c'est que le document, quel que soit le mode de transcription retenu, ne «parle» pas de lui-même, et que toute édition nécessite une annotation qui rende compte de la richesse, notamment intertextuelle mais pas uniquement, du travail de l'écrivain. Aussi l'intérêt d'un site dépendra-t-il en partie de la qualité du repérage et de l'information (documents d'époque, bibliothèque de l'écrivain, banque d'images, cartes, ressources bibliographiques,

etc.) qu'on aura su mettre à la disposition des lecteurs. Les outils de balisage, d'étiquetage, d'indexation, de cartographie, de navigation, permettent désormais de dépasser les maigres notes critiques ou les illisibles apparats critiques autorisés par l'imprimé, qui ne proposaient qu'une vision statique du texte mouvant. Il s'agit donc également de repenser l'annotation de l'édition traditionnelle à l'aune de ces corpus récents que sont les archives littéraires, et des possibilités technologiques d'aujourd'hui (et de demain). Une «philologie numérique» est à définir, adaptée à l'édition et à l'exploration de ces corpus.

Précisions terminologiques...

Le terme magique de «numérisation» entraîne parfois des malentendus parmi les littéraires, parfois peu au fait de la terminologie des sciences de l'information et de ses subtilités... Ainsi, une édition «numérisée» n'est pas une édition «numérique», une édition «numérique» n'est pas nécessairement une édition «en ligne», etc. Précisons... «Numériser» des données (texte ou images) consistant simplement à transférer des données sur un support numérique, il suffit, pour produire une édition «numérisée», de prendre une édition papier existante, de la scanner, et de la stocker dans la mémoire d'un ordinateur, ou sur un serveur... C'est ce que fait Gallica en produisant des fac-similés d'éditions anciennes. Il n'y a là aucun travail éditorial : la Bibliothèque nationale joue son rôle de diffusion en mettant à disposition des lecteurs sur un nouveau support des éditions existant déjà sur support imprimé. En revanche, lorsqu'on parle d'édition «numérique» ou «électronique» (les deux termes sont souvent employés de manière concurrentielle), on désigne un texte écrit et mis en forme en fonction des possibilités offertes par un logiciel d'édition. Éric-Olivier Lochard explique (ci-après, p. 83) que l'informatique amène à dissocier les opérations d'«édition» (écriture et composition des documents numériques) et de publication, la finalité de ce travail d'écriture et d'annotation des données pouvant aussi bien être la publication sous forme de livre imprimé (les données sont alors transformées avec un logiciel spécifique), la publication sur un serveur Web (grâce à un logiciel adapté), ou sous forme d'une base de données consultable sur un poste de travail. Enfin, même si le principe de l'écriture informatique consiste dans la possibilité de lier des fichiers, donc dans la liaison hypertextuelle, une édition électronique n'est pas nécessairement «hypertextuelle» au sens où la communauté littéraire l'entend (Serge Linkès, dans son article, p. 185, rappelle la polysémie de ce terme, et l'infléchissement qu'il a pris du fait de son usage dans la théorie littéraire). S'il est évident qu'une édition numérisée (Gallica, par exemple) ne saurait être

hypertextuelle, il existe des éditions électroniques qui se contentent d'un nombre de liens minimal entre les documents (textes, images, éventuellement notes critiques), et qui ne proposent aucune réelle « navigation », que ce soit sur le Web ou à travers les documents de la base elle-même. Une édition hypertextuelle, qu'elle soit en ligne ou non, est donc une édition qui organise de manière raisonnée les liens entre des fichiers, à des fins d'enrichissement critique et documentaire et/ou afin de permettre des parcours de lecture entre des documents ou au sein d'un même document – plusieurs types de liens pouvant être proposés à partir d'un même passage.

Une voie de recherche novatrice dans le champ universitaire : diversité des formules et des corpus représentés

En réunissant à Grenoble en décembre 2006 vingt-trois chercheurs, enseignants-chercheurs, ingénieurs d'étude, ingénieurs de recherche, et développeurs de logiciels impliqués dans des projets d'édition électronique *de manuscrits de travail littéraires*, notre colloque se différenciait des nombreux colloques consacrés à l'édition électronique en général (c'est-à-dire consacrés, en fait, à l'édition de textes littéraires en ligne ou sur CD-ROM), et il semble bien qu'il soit le premier colloque universitaire français à rassembler plusieurs éditeurs de dossiers génétiques « modernes » (XIX^e-XX^e siècles), leur permettant de comparer outils, méthodes, projets, à la fois entre eux et avec des projets similaires portant sur des corpus plus anciens (XVII^e-XVIII^e siècles : notes de Pascal, carnet de notes de Montesquieu).

En effet, après plusieurs programmes de recherche sur le traitement informatique des brouillons menés par des chercheurs du CNRS dans les années 1980 à 2000¹³, le séminaire général de l'ITEM (Institut des textes et manuscrits modernes) organisé à Paris par Aurèle Crasson en 2003-2004, « De l'archive manuscrite au *scriptorium* électronique », avait ouvert ces problématiques encore confidentielles de l'édition électronique des manuscrits à un cercle élargi de généticiens et de chercheurs, et toutes les questions essentielles concernant les outils (numérisation des images, possibilités de représentation et de codage du manuscrit) et les choix éditoriaux y étaient abordées¹⁴. Toutefois le principe

13. J.-L. Lebrave, qui a participé à ces programmes, en dresse un rapide historique et fait le point sur les questions techniques non encore résolues dans un récent article, « Du visible au lisible : comment représenter la genèse ? », *Genesis*, n° 27, 2006, p. 11-17.

14. Ces conférences sont consultables sous la forme d'enregistrements audio et vidéo sur le site de l'École normale supérieure, rubrique « Diffusion des savoirs », titre « Séminaire général de l'ITEM », à l'adresse suivante : <http://www.diffusion.ens.fr/index.php?res=cycles&idcycle=72>. Au moment

même du séminaire (conférences mensuelles) ne permettait pas forcément à des chercheurs géographiquement éloignés de se rencontrer pour confronter leurs expériences et leurs projets comme il est possible de le faire dans le cadre d'un colloque. De plus, beaucoup de projets ont été lancés depuis 2003 et il était essentiel de faire le point sur l'actualité de l'édition électronique des archives littéraires.

Le colloque «Édition critique/édition électronique. Bases de données textuelles. Pratique et théorie», organisé à Lyon en septembre 2005 par Antony McKenna (Université Jean-Monnet, Saint-Étienne) et Andrea Iacovella (ENS-LSH, Lyon)¹⁵, a apporté un certain nombre de réponses à cette question en rassemblant des éditeurs de fonds d'archives d'Ancien Régime, l'outil informatique étant pour eux tantôt un moyen de structuration des données en vue d'une publication imprimée, tantôt un outil d'édition et de recherche à travers des corpus très volumineux. L'édition de manuscrits d'Ancien Régime peut parfois, pour certains corpus (tel celui de Montesquieu) comportant beaucoup de documents de travail¹⁶, présenter des similarités avec l'édition génétique de fonds de manuscrits modernes ; mais étant donné les pratiques d'écriture de l'époque classique, il est rare que les étapes de brouillon et réécriture intermédiaires entre notes initiales et copies finales aient été conservées¹⁷. Cependant, outre l'intérêt intrinsèque des outils philologiques mis au point pour des corpus de correspondances, par exemple, ou de notes fragmentaires, outils applicables à des corpus modernes¹⁸, la comparaison fait mieux ressortir la spécificité des corpus, des outils existants, et la particularité de l'édition génétique dans le contexte de l'édition électronique.

Le colloque «Éditer et valoriser des fonds de manuscrits. L'apport et les limites du numérique», organisé à Grenoble en décembre 2006, est venu en ce sens apporter des réponses – et soulever des problèmes – complémentaires, les conférenciers présentant principalement des projets centrés sur des manuscrits modernes.

où nous corrigeons les épreuves du présent volume, nous apprenons la publication en volume des contributions à ce séminaire : *L'Édition du manuscrit. De l'archive de création au scriptorium électronique*, sous la direction d'A. Crasson, Bruxelles, éditions Académia Bruylant, «Au cœur des textes», 2008.

15. Ce colloque, intitulé «Édition critique / édition électronique. Bases de données textuelles. Pratique et théorie», a eu lieu à l'ENS-LSH de Lyon du 21 au 23 septembre 2005. Voir la page d'introduction sur le site http://pensee-classique.ens-lsh.fr/colloques/Col_ed.%20electroniq.html.

16. Voir l'article de C. Dornier dans le présent volume, p. 139.

17. Voir *Écrire aux XVII^e et XVIII^e siècles : genèses des textes littéraires et philosophiques*, sous la direction de J.-L. Lebrave et A. Grésillon, CNRS éditions, 2000.

18. Par exemple le système d'annotation mis en place pour étudier les réseaux de socialité chez les écrivains des Lumières (voir les références données dans l'article d'É.-O. Lochar d'après, p. 83).

La diversité des formules éditoriales des projets ici présentés (bases de données internes, bases de données mises en ligne, éditions collaboratives en ligne; approches documentaires, approches plus textuelles) peut sans doute créer une impression de confusion, mais dans un domaine où tout est nouveau, il est inévitable que différents outils et différentes formules coexistent. De plus, il paraît (actuellement) illusoire d'envisager une seule et même plateforme ou un seul et même logiciel d'édition pour tous les types de manuscrits et de corpus, qui ne posent pas tous les mêmes problèmes, et n'impliquent donc ni les mêmes besoins, ni les mêmes objectifs. On ne peut parler d'une «solution électronique» unique, mais de solutions électroniques diverses et adaptées aux besoins des concepteurs.

Les communications rassemblées dans ce volume¹⁹ ont été regroupées en quatre parties : aspects théoriques; présentation d'outils; projets conjoints d'éditions électronique et papier; choix de l'édition électronique. D'autres classements auraient sans doute été possibles, mais celui-ci nous a semblé rendre compte avec le plus de pertinence des effets d'écho ou d'opposition et de la complémentarité de ces textes.

Les outils se développant rapidement, il est probable que les projets présentés, déjà remaniés depuis le colloque²⁰, auront encore évolué au moment de la parution du présent volume. À défaut d'un «bilan» qui prétendrait à une vue d'ensemble, nous proposerons à partir de ces diverses contributions un état des lieux, afin de souligner les points de convergence, de divergence, de débat, et les perspectives que ce domaine semble promettre, en 2007, pour des «littéraires» engagés dans la recherche sur l'édition génétique électronique...

État des lieux 2007 et perspectives²¹

1. Quels corpus manuscrits l'édition électronique concerne-t-elle principalement?

L'édition électronique paraît s'imposer pour les dossiers génétiques volumineux, on l'a dit, et surtout pour les brouillons correspondant à des couches

19. La plupart des articles présentés dans cet ouvrage proposent une version électronique complémentaire illustrée sur le site de la MSH-Alpes : <http://www.msh-alpes.prd.fr/Actualites/Manuscrits.htm>.

20. Par exemple, la TEI (*Text Encoding Initiative*) utilisée pour l'encodage philologique a mis sa version 5 (P5) à disposition des chercheurs en novembre 2007. Certains contributeurs de notre volume ont dû actualiser leur article pour rendre compte de la mise à jour de leur système éditorial (D. Van Hulle, C. Dornier et P.-Y. Buard).

21. Nous remercions M. Bernard de nous avoir transmis la synthèse qu'il avait faite à l'issue du colloque. Certaines sections de cet «état des lieux» s'inspirent de ses notes.

d'écriture superposées appartenant à des phases rédactionnelles différentes. Mais l'intérêt d'une édition électronique, contrairement à ce qu'on pourrait imaginer, n'est pas nécessairement proportionnel au volume du dossier génétique. Il est vrai qu'il est difficile d'envisager une édition papier pour un corpus de plusieurs milliers de pages – mais récemment, Brepols, éditeur belge spécialisé dans l'édition de corpus volumineux comme la patristique, a conclu un contrat d'édition imprimée pour les soixante-quinze cahiers de brouillon de *La Recherche* (voir Nathalie Mauriac Dyer). Inversement, une édition électronique, loin d'être cantonnée uniquement à des ensembles importants, peut s'avérer précieuse (par ses possibilités d'annotations hypertextuelles) pour des corpus de quelques dizaines ou centaines de pages, comme le montrent les projets d'édition en ligne de textes à forte composante intertextuelle tels que *Vie de Napoléon* ou *Histoire d'Espagne* de Stendhal²². Il est vrai toutefois que lorsqu'on formule le souhait ambitieux de traiter l'ensemble des archives d'un écrivain, seule l'édition électronique semble possible, même si elle ne dispense pas d'éditions imprimées de parties plus ou moins importantes de ce fonds. Le projet de publier en ligne l'intégralité des manuscrits de Stendhal conservés à la Bibliothèque municipale de Grenoble (32 000 feuillets) prévoit par exemple à terme une publication papier des «Journaux et papiers» de l'écrivain.

2. À quel public l'édition génétique électronique est-elle destinée ?

Grand public, chercheurs, étudiants... ? Est-il possible de satisfaire simultanément des publics aux exigences *a priori* si différentes ? Cette question du public visé est fondamentale, comme le rappelle Caroline Angé, p. 71. Or les chercheurs qui se lancent dans un projet d'édition électronique oublient parfois de se la poser, alors qu'elle devrait pourtant orienter leurs principes et leurs choix méthodologiques et techniques. À quel(s) besoin(s), à quelle(s) demande(s) correspond une édition électronique de fonds de manuscrits ?

L'édition papier satisfait plusieurs types de demandes :

- les *beaux livres*, avec leurs reproductions en fac-similé, s'adressent aux collectionneurs fortunés, aux bibliophiles, rencontrent aussi le goût du grand public pour la rencontre émouvante avec le manuscrit, y compris dans sa dimension plastique, et permettent au chercheur spécialisé de disposer directement des images des originaux ;
- les éditions génétiques imprimées (que leurs transcriptions soient linéarisées²³ ou diplomatiques) sont destinées essentiellement aux chercheurs.

22. Projets respectifs d'H. Spengler et de C. Meynard.

23. Transcription linéarisée : «Par le jeu d'un codage typographique, les opérations d'écriture fondamentale (addition, suppression, remplacement) sont présentées comme si elles constituaient un texte et se donnent à lire comme une suite continue de signes» (J.-L. Lebrave,

Des éditions génétiques imprimées en fac-similé avec transcription diplomatique de chaque page permettraient de combiner ces deux types de documents, donc idéalement de satisfaire les demandes de tous les publics ; mais elles sont très onéreuses, et rares sont les éditeurs prêts à publier des ouvrages si luxueux.

La technologie numérique permet d'offrir gratuitement en ligne des éditions combinant ainsi ces deux types de documents. L'hypermédia permet à l'amateur de contempler le manuscrit, d'admirer les dessins marginaux qu'il peut comporter, et au chercheur et à l'étudiant de consulter le texte du manuscrit avec ses ratures, ses ajouts, etc., de bénéficier d'une transcription pour la lecture de passages difficiles à déchiffrer, de le comparer avec l'édition originale, de lire des notes savantes, de suivre les cheminements de l'écriture de versions en versions, de variante en variante, etc.

En outre, comme le montrent Carole Dornier et Pierre-Yves Buard, les logiciels d'édition philologique permettent, grâce à un codage de couleurs, de représenter les ratures et les additions de manière plus lisible que les appareils critiques de l'édition savante traditionnelle ou le codage compliqué des transcriptions linéarisées de l'édition imprimée.

Il serait ainsi envisageable d'ouvrir à de nouveaux publics cette dimension génétique des œuvres littéraires aujourd'hui cantonnée, pour des raisons d'accessibilité des documents et de coût des éditions savantes, à la sphère des chercheurs. Certes, il est utopique d'espérer toucher le «grand public», mais certaines expériences montrent que les publics scolaires et universitaires peuvent bénéficier de cette initiation et de ce regard particulier sur l'atelier d'écriture des grands écrivains²⁴.

3. Révolution ou évolution ?

Revenons sur quelques idées reçues. S'il est vrai que pour une lecture de découverte, le livre s'impose encore de nos jours sous sa forme papier traditionnelle par sa maniabilité et sa facilité de consultation (mais c'est peut-être déjà moins vrai pour les jeunes générations habituées à lire sur l'écran de leur ordinateur), en revanche dans le cadre d'une lecture critique, les oppositions entre édition imprimée et édition électronique ne sont pas si marquées qu'on le croit.

art. cit, p. 208). C'est la solution adoptée par exemple par G. Bonaccorso, qui a poussé «jusqu'aux confins du lisible» (pour reprendre l'expression de J.-L. Lebrave) cette solution éditoriale en fournissant une édition diplomatique linéarisée d'*Un cœur simple* caractérisée par la multiplication des signes diacritiques (G. Flaubert, *Corpus flaubertianum*, I. *Un cœur simple*, édition diplomatique et génétique des manuscrits par G. Bonaccorso *et al.*, Nizet, 1991).

24. L'étude des processus de la création littéraire («le travail de l'écriture», «écrire, publier, lire») fait partie du programme officiel de l'enseignement de la littérature dans les classes de seconde littéraire. Certaines universités offrent des séminaires de Master consacrés à la génétique, mais il faudrait systématiser cet enseignement, et l'envisager dès la Licence.

À commencer par le mode de lecture : on ne «lit» pas de la même manière, dit-on, un livre imprimé et un texte électronique, lequel favoriserait un mode de lecture fragmenté, désordonné, rhizomatique... Mais en réalité, le mode de lecture dépend davantage de l'objet que du support. Outre qu'on lit bien souvent un livre en désordre, en commençant par la fin, en le feuilletant au hasard, la lecture de type universitaire nécessite de s'arrêter sur tel paragraphe ou tel chapitre pour en faire une analyse approfondie, pour lire les notes critiques, etc. De fait, les annotations dans une édition critique ou savante introduisent déjà ce fonctionnement de type hypertextuel (au sens informatique du terme) dans la structuration de l'édition et engendrent cette fameuse lecture rhizomatique. Dans la lecture studieuse, on est amené à s'interrompre à tout moment pour vérifier des références, consulter un dictionnaire, un livre d'histoire. Une édition électronique comportant des annotations ne fait donc que pousser à l'extrême des pratiques de lecture déjà favorisées par l'édition savante ; et le lecteur perd moins de temps car l'éditeur, s'il a bien fait son travail, a déjà mis en place les liens vers des définitions, des précisions historiques, des fiches biographiques...

D'autre part, on a souvent tendance à opposer la facilité de lecture d'un livre et la complexité de consultation d'une édition électronique, dont le fonctionnement n'est pas toujours aisé à saisir et à maîtriser même si les concepteurs de la base ou du site en question prétendent le contraire. En caricaturant quelque peu, on serait presque tenté de dire que, dans certains cas, les seuls lecteurs compétents risquent d'être les concepteurs de ces outils éditoriaux eux-mêmes ! Une édition électronique présente souvent une dimension labyrinthique qui risque de perturber et de désorienter le lecteur. Certaines éditions électroniques, pour pallier ce défaut, comportent donc une introduction explicative, une notice à usage de l'utilisateur, voire une rubrique d'aide en ligne²⁵. Pour autant, cela ne suffit pas pour que l'ergonomie devienne satisfaisante, et il est nécessaire, indépendamment de la qualité scientifique du travail réalisé, de bien penser les interfaces, pour permettre au lecteur de toujours se repérer dans les différents parcours possibles. Mais sans doute ces problèmes de lisibilité sont-ils liés à la jeunesse de ce nouveau *medium* éditorial. N'oublions pas que le *volumen* était lui aussi peu pratique et qu'il a fallu plusieurs siècles avant qu'on invente le *codex*, qui a permis progressivement l'élaboration de techniques de repérage (pagination, index et des tables des matières, etc.²⁶).

25. Le livre imprimé est cependant confronté, certes de façon moins marquée, au même problème, d'où la floraison de métatextes et paratextes (note sur le texte, introduction, liste des abréviations, index, division en chapitres, etc...) destinés à guider le lecteur et à lui faire comprendre les principes éditoriaux adoptés.

26. Voir la conférence de M. Melot, «Le livre comme forme symbolique» (conférence prononcée dans le cadre de l'École de l'Institut d'histoire du livre, 2004, archivée sur le site de l'Institut

Plus récemment, le même problème s'est posé pour le Web : à l'origine, l'information sur un site était apportée au fil d'une interminable – donc illisible – page Web. Les concepteurs de sites n'ont pris conscience de la nécessité de sectionner et de hiérarchiser le contenu (menus, onglets, etc.) qu'avec le temps et l'expérience. Pour les éditions électroniques, la structuration s'opérera sans doute peu à peu, chacun étant conscient qu'il faut rendre la consultation plus intuitive et plus rationnelle.

En ce sens, si le numérique est bien une révolution, l'édition électronique est plutôt une simple évolution, une extension des possibilités déjà présentes dans l'édition papier.

4. Complémentarité des médiums éditoriaux ?

Autre idée reçue : le livre est un objet maniable, accessible à tous, bon marché, résistant et pérenne. En réalité les éditions savantes sont coûteuses, parfois épuisées et non rééditées, les libraires n'ont souvent que des livres à destination du grand public, ce qui pose un réel problème d'accessibilité des éditions de référence à l'étudiant ou au chercheur qui ne dispose pas (surtout en province) d'une bibliothèque de recherche suffisamment riche ou ancienne. Quant à la publication de manuscrits au format imprimé, elle cumule à peu près tous les cauchemars éditoriaux : fabrication complexe et coûteuse, public restreint et exigeant, corpus et appareils critiques volumineux. En ce sens l'édition électronique peut bel et bien répondre à une demande explicite.

Une partie des projets présentés dans cet ouvrage articulent édition imprimée et édition électronique (voir les articles de Serge Linkès, de Carole Dornier et Pierre-Yves Buard, ou de Cécile Meynard²⁷). Mais d'autres n'auraient aucune

d'histoire du livre, <http://ihl.enssib.fr/siteihl.php?page=219>). Il rappelle que le *codex* est apparu dès la fin du I^{er} siècle, mais ne s'est imposé qu'au bout de quatre siècles face au *volumen*, et qu'il a fallu encore attendre les XII^e et XIII^e siècles pour qu'interviennent réellement les notions de division en chapitres et de hiérarchisation du *codex*. «Le livre, explique-t-il, est lui-même une arborescence». Il précise aussi que les normes SGML, HTML, etc. à leurs débuts étaient «entièrement décalquées sur les structures du livre», tout en estimant que l'ordinateur, en tant que structure ouverte, dont le contenu n'est pas délimité (comme l'est celui d'un livre), se rapproche plutôt du *volumen*.

27. La solution électronique est intéressante également pour la publication d'œuvres comportant des milliers de pages. Ainsi, *Artamène*, que Cl. Bourqui et A. Gefen ont publié intégralement en ligne (www.artamene.org) et dont ils ont procuré une édition partielle chez Flammarion, illustre bien cette complémentarité éditoriale. A. Gefen présente cette réalisation, ainsi que les projets en cours concernant *L'Astrée* et «Molière 21», dans un diaporama accessible sur le site de la MSH-Alpes : <http://www.msh-alpes.prd.fr/Actualites/Manuscrits.htm>. Dans ce cas, l'édition électronique est plus scientifique, que l'édition imprimée, car exhaustive. Il en ira de même pour le projet d'édition électronique de *L'Astrée*, mené par D. Denis et son équipe, qui procurera les différentes versions de l'œuvre, tandis que l'édition imprimée ne pourra proposer, par la force des choses, qu'une des versions publiées de l'œuvre et mentionner les variantes dans l'apparat critique, ce qui peut poser un réel problème de lisibilité quand les passages variants sont longs ou très nombreux.

viabilité économique sous la forme papier et sont orientés exclusivement vers la solution électronique (voir en particulier les articles d’Yvan Leclerc et Danielle Girard, d’Hélène Spengler, de Françoise Leriche...).

Un problème de concurrence se pose dans le cas où coexistent édition imprimée et édition électronique. Il est nécessaire de disposer d’un texte numérisé pour pouvoir rechercher sans risque d’erreur ou d’oubli des citations ou des termes précis dans un ouvrage long, établir des index, faire des recherches lexicales ou syntaxiques plus avancées avec des logiciels de lexicométrie ou des analyseurs morpho-syntaxiques... En réalité, si le support numérique *ne se pose pas* en concurrent du support papier en ce qui concerne l’édition des manuscrits et de leurs transcriptions, dans les faits, il *est* concurrent. On assiste alors parfois à ce qui s’apparente à une guerre éditoriale. En général les enjeux d’une édition imprimée sont ceux de l’accessibilité et de la maniabilité, d’où un travail éditorial de sélection (dans le cas de textes trop longs, de textes inachevés présentant différentes versions ou laissés dans le désordre) et une simplification de l’annotation, une normalisation... Toutes ces contraintes disparaissent dans le cas de l’édition électronique. Cette dernière sera nécessairement plus complète, d’où la résistance des maisons d’édition à accepter de publier en volume une œuvre dont il y aura également une édition électronique²⁸.

Avantage concurrentiel certain de l’édition électronique, l’utilisateur peut déjà, pour certains corpus, créer des anthologies et sélections à sa convenance, et l’on peut envisager que l’impression à la demande se développe par des conventions entre éditeurs commerciaux et centres de recherche ou universités. Arcane, couplée avec le logiciel TeX, sert déjà pour l’édition de la correspondance de Bayle à la Voltaire Foundation. De même, CLELIA, la base documentaire des manuscrits de Stendhal conçue par Thomas Lebarbé et Cécile Meynard, permettra de produire des éditions pseudo-diplomatiques aussi bien que linéaires des pages saisies, par exemple l’édition en cours des «Journaux et papiers» de Stendhal aux ELLUG.

Quand il n’est prévu ou qu’il n’existe qu’une édition en ligne, la question de l’annotation est traitée différemment selon les cas. Yves Leclerc et Danielle Girard ont choisi la solution de deux sites complémentaires, l’un proposant seulement les images et les transcriptions (le Site Bovary), l’autre (l’Atelier Bovary)²⁹, dépourvu d’images et fournissant les outils nécessaires à la recherche et des

28. Le fait que le site roland-barthes.com, évoqué par Cl. Coste et G. Bellon, ne devait pas faire concurrence aux publications du Seuil, est peut-être ainsi une des raisons de son échec.

29. Respectivement <http://www.bovary.fr/> et <http://flaubert.univ-rouen.fr/bovary/atelier/atelier.php>.

dossiers thématiques (normandismes, voix narratives, notes de régie, index des noms propres, cartographie...). Hélène Spengler et Paolo D'Iorio, respectivement dans le prototype d'édition de *Vie de Napoléon* et dans la plateforme *HyperNietzsche*, ne fournissent pas d'annotations critiques et laissent le lecteur aller chercher lui-même les informations qui lui manquent sur les personnages, les dates, les lieux, les idiomatismes, etc. Inversement, dans le projet de Françoise Leriche, l'hypercorrespondance de Proust se substituera aux éditions imprimées et, prenant acte des possibilités du nouveau *medium*, comportera un système (hypertextuel) très développé d'annotation et d'indexation que ne permettait pas l'édition en volume.

5. Numérisation des images

Des progrès considérables ont été faits dans le domaine de la reproduction photométrique de manuscrits. Même si certaines bibliothèques donnent encore seulement accès à des microfiches et microfilms en noir et blanc d'une très médiocre qualité et d'une consultation extrêmement limitée et contraignante, un programme national de conservation et de valorisation des archives patrimoniales a permis à de nombreuses institutions de se doter d'équipements permettant la numérisation des manuscrits. Les avantages sont incontestables : transmission rapide, stockage aisé, reproduction en couleur peu coûteuse, qualité des images numériques, navigation entre les pages, accès « démocratisé », manipulation facilitée... L'existence de ces images rend superflue, voire dans certains cas inutile, la consultation manuelle des précieux et fragiles manuscrits, et permet ainsi de les préserver des détériorations liées à la manipulation.

Cependant cet accès n'est pas encore aussi démocratique qu'on pourrait le croire : il est le plus souvent nécessaire de payer pour obtenir communication des vues numérisées de pages de manuscrits (jusqu'à 8 euros pour une vue en haute définition). Certaines images ne sont pour le moment accessibles que depuis des postes internes dans les bibliothèques, ce qui, paradoxalement, oblige les chercheurs à venir les consulter sur place, alors même que la finalité des documents numériques est la facilité et la gratuité de duplication et de transmission.

Conscientes des problèmes que cela pose aux utilisateurs et de leur mission de valorisation auprès du grand public du précieux patrimoine culturel dont elles sont dépositaires, les bibliothèques se lancent dans de vastes campagnes de mise en ligne gratuite d'images de manuscrits. La mise en ligne du manuscrit de la *Vie de Henry Brulard* et, tout récemment des cinq cahiers acquis lors de la vente Bérès en 2006, sur le site de la Bibliothèque municipale de Grenoble démontre les qualités de ce mode de reproduction. Mais le projet n'en est encore qu'à ses débuts. Aujourd'hui de nombreuses entreprises spécialisées savent numériser

d'une manière industrielle et professionnelle, mais les bibliothèques n'ont pas toujours les moyens de financer des campagnes lourdes de numérisation.

Par ailleurs, les normes de cette numérisation (résolution, formats, couleurs, éclairage) seraient à débattre entre les différents partenaires impliqués : bibliothécaires, chercheurs, informaticiens, municipalités, État. En effet les bibliothèques, obéissant aux normes établies par la Bibliothèque nationale de France, se contentent souvent d'une numérisation en 300 dpi, ce qui est d'ores et déjà insuffisant (problèmes de crénelage à l'agrandissement, d'identification d'éléments significatifs du papier, tels que le filigrane, le timbre sec, la trame...). Pour déchiffrer des mots raturés, mal écrits, d'une encre pâlie ou tracés au crayon à papier, le chercheur a besoin d'images d'excellente qualité. À l'heure actuelle, les informaticiens s'accordent pour dire qu'un minimum de 600 dpi est nécessaire pour permettre un affichage et un traitement corrects des images, ces images devant rester exploitables le plus longtemps possible par les outils informatiques qui apparaîtront dans les années à venir.

Quelles que soient leur qualité et leur définition, les images numérisées présentent et présenteront cependant toujours un défaut majeur : la numérisation nous fait perdre de vue des informations essentielles sur la reliure, les papiers, leur épaisseur, leur aspect, tout ce que révèle l'analyse codicologique (voir les communications de Hans-Walter Gabler, et de Claire Bustarret et Serge Linkès). C'est par l'analyse du papier que Serge Linkès a pu rétablir l'ordre d'écriture des pages et des versions de *Lamiel*. Une collection de pages numérisées ne peut, de ce point de vue, se substituer au manuscrit dans le cadre d'une étude approfondie. D'où l'importance de la création de bases de données codicologiques (comme MUSE), qu'il faudra apprendre à interfacer avec les éditions numériques en cours de constitution.

6. Traitement de l'image numérisée

Une fois l'image numérisée, il est désormais possible d'en tirer de nombreuses informations. Des outils perfectionnés et d'un intérêt scientifique incontestable existent déjà. Pour ne donner qu'un exemple, le laboratoire LIRIS de l'INSA de Lyon (<http://liris.cnrs.fr/imagine>) a déjà créé des outils spécifiques permettant, dans les manuscrits du Moyen-Âge et de la Renaissance, l'identification automatique de différents scripteurs, l'indexation de lettrines dans les manuscrits médiévaux, et la structuration des pages de documents dont il est possible d'extraire automatiquement une mise en page, notamment pour les documents de la Renaissance. Le manuscrit moderne, en tant que document de travail, est toutefois bien plus complexe à traiter que ces manuscrits du Moyen-Âge et de la Renaissance qui étaient des manuscrits de diffusion, donc lisibles et «normalisés». Les variations nombreuses dans l'écriture des autographes, notamment,

rendent bien plus difficile l'identification de la structure du document (titres, mots...).

D'autres travaux sont actuellement réalisés par ce même laboratoire pour tenter de supprimer sur les images le «bruit» dû à la présence de taches ou à l'affichage, sur le recto d'un feuillet, du texte figurant au verso à cause d'une trop grande transparence ou porosité du papier.

Grâce aux progrès apportés par ces nouveaux outils, il devient possible, par exemple, de reconnaître des caractères partiellement rayés, ou d'identifier un mot jusque-là jugé illisible qu'un logiciel standard de reconnaissance de caractères, efficace pour des textes imprimés mais pas pour des autographes, ne pourrait pas reconnaître. L'identification partielle ou complète d'un mot (qui pourrait aussi bien être un fragment de mot qu'une entité graphique extraite de la page) se base sur une estimation de la similarité entre formes. L'appariement de ces similarités sur une (ou plusieurs) page(s) permet ensuite de mettre en concordance des formes (simples ou composées) et de proposer à l'utilisateur une liste de possibilités pouvant permettre, par exemple, de reconnaître sous une forme plus lisible un mot jugé incompréhensible ailleurs dans le texte.

L'écriture autographe de l'écrivain moderne, hâtive, irrégulière, pose encore incontestablement des problèmes pour cette reconnaissance de graphèmes par sa variabilité (inconstance de l'échelle d'écriture, formation de traits irréguliers et épaisseur variable, dégradation des formes due à la fatigue du scripteur...), mais les chercheurs travaillent activement pour débloquer ce qui est à leurs yeux un simple verrou technologique.

7. «Zonage» et représentation du classement génétique

L'une des questions récurrentes dans nos travaux est celle du classement génétique du corpus et, une fois le classement chronologique des documents réalisé (cahiers, feuilles volantes, etc., selon une datation absolue ou relative), celle de la «granularité», c'est-à-dire du degré de finesse jusqu'où il est souhaitable (et cohérent) de pousser l'analyse dans le classement des unités textuelles. Si le folio reste une référence bibliographique de base, imposée par la configuration matérielle du manuscrit et le catalogage des bibliothèques, il ne peut, dans bien des cas, suffire au classement génétique, en particulier dans le cas de textes fragmentaires (Nietzsche, Stendhal, Barthes), pour certains types de documents (carnets de notes, cahiers constitués d'unités autonomes et/ou de fragments écrits à différentes dates), ou dans le cas de manuscrits procédant au montage de petites unités provenant de brouillons antérieurs (Proust, par exemple), qu'il convient de pouvoir indexer et lier entre elles pour suivre le travail d'écriture d'une unité à l'autre et ainsi reconstituer la chaîne des réécritures à l'intérieur d'un même document ou d'un document à l'autre.

Que l'on parle de «stematologie», à l'instar des médiévistes, ou de «rhizomes», les arbres qu'il faut envisager sont souvent si complexes que seuls les moyens informatiques permettront de les organiser et de les représenter de façon exhaustive. Dans les exemples fournis par Nathalie Mauriac Dyer, on voit que l'unité textuelle à considérer chez Proust n'est pas nécessairement la phrase, mais parfois le mot, le syntagme... Ce qui oblige alors le généticien à définir une granularité très fine, qui risque de compliquer l'indexation, même informatisée... Faut-il descendre jusqu'à un niveau d'indexation génétique aussi fin? Faut-il viser une telle exhaustivité?

Il est, certes, techniquement possible de découper graphiquement la page en petites «zones», ou inversement, de regrouper plusieurs pages en une seule zone signifiante (plateforme Hyper par exemple). La question qui se pose est donc en définitive celle du temps humain nécessaire pour effectuer des zonages tenus et, par conséquent, la question du bénéfice que peut représenter une telle exhaustivité dans le repérage des liens génétiques. Ne risque-t-on pas, à saturer une page de liens vers un grand nombre d'autres documents (en multipliant les zones à l'échelle du mot), de rendre cette annotation illisible pour l'utilisateur, perdu au milieu de tous les renvois proposés? Il faudra que des tests soient effectués sur des prototypes, afin de mesurer l'impact de tant d'annotations sur les utilisateurs.

8. Convivialité, visibilité, lisibilité

Le support numérique peut, quand les éditeurs et les développeurs ont prévu cette fonctionnalité, permettre au lecteur de choisir entre plusieurs modes d'affichage :

- linéaire, avec un défilement du texte sans affichage des ratures, corrections, ajouts, marginales ;
- linéarisé, avec l'option de n'afficher que les ratures ou que les ajouts interlinéaires, etc. ;
- diplomatique (ou «pseudo-diplomatique³⁰»), voire «ultra-diplomatique³¹».

Mais l'édition diplomatique (et même ultra-diplomatique), si elle peut restituer les finesses scripturales du manuscrit (ratures de premier niveau, de

30. La transcription pseudo-diplomatique consiste à reproduire autant que possible la mise en page de l'auteur. En l'absence de logiciel spécifique, il est cependant difficile, voire impossible, de présenter certaines marginales (écrites à l'envers, en diagonales, etc.) à leur place exacte, de respecter scrupuleusement la graphie de certains mots, de reproduire les dessins...

31. La transcription «ultra-diplomatique» consiste à respecter toutes les caractéristiques topographiques et graphiques de l'écriture (mise en page, traits de rature, place précise des additions, des ratures, etc.). Ce degré de précision graphique nécessite le recours à des logiciels spécialisés, payants et relativement complexes, tels QuarkXpress® ou Illustrator® – ou l'invention de nouveaux logiciels si les équipes sont associées à des unités de chercheurs en informatique.

deuxième niveau, de troisième niveau, additions supralinéaires ancrées au-dessus d'un mot bien précis, graphismes d'insertion d'une addition, etc.), peut entrer en conflit avec les possibilités actuelles de balisage, qui s'accommodent parfois mieux de transcriptions linéarisées, codées aujourd'hui en XML, selon les normes de la TEI³².

La navigation entre ces différents modes d'affichage donne ainsi accès à des transcriptions bien plus lisibles que les jeux de piste abscons auxquels s'astreignent encore aujourd'hui, dans les appareils critiques des éditions savantes imprimées, les éditeurs soucieux de précision.

L'interactivité rend aussi possibles différents niveaux de lecture, du texte brut aux fenêtres ou *pop-up* qui s'ouvrent d'un simple clic ou d'un passage de la souris sur un mot ou un groupe de mots, permettant d'afficher ainsi à volonté des annotations critiques. Le balisage peut autoriser des recherches complexes sur le matériau textuel, une fois tous les éléments philologiques pertinents encodés.

La dimension visuelle et interactive s'avère incontestablement très attractive, aussi bien pour le grand public que pour les chercheurs : l'auditoire est toujours impressionné lors des démonstrations spectaculaires, comme celle de la loupe interactive d'HyperNietzsche³³, ou celle des cartographies dynamiques d'Éric-Olivier Lochard³⁴. Ainsi, il est possible, par le biais d'animations, d'afficher l'écriture de l'auteur lettre par lettre, en une sorte de reproduction mimétique des mouvements de la plume sur le papier. Tous ces aspects, qui peuvent sembler aux yeux de certains relever davantage du « gadget » que d'un véritable ajout scientifique, ne sont pour autant pas à négliger. Ils peuvent notamment attirer un public jeune, habitué à trouver ce genre d'animation sur le Web, et qui peut ainsi potentiellement se réconcilier avec la consultation (ira-t-il jusqu'à la lecture?) de sites littéraires.

La question du confort de lecture est également fondamentale. Le format des écrans actuels (paysage) pose un problème non encore résolu pour l'affichage et la lecture de pages en mode traditionnel (portrait)³⁵. Se pose aussi le problème,

32. Au sujet de l'encodage et de la difficulté des outils informatiques actuels à représenter tous les aspects sémiotiques du manuscrit, voir les analyses de J.-L. Lebrave dans « Du visible au lisible : comment représenter la genèse? », art. cit.

33. Ce système de loupe interactive (permettant de superposer un élément de l'image sur la transcription correspondante et vice-versa) se révèle ainsi fort utile en cas de doute sur la graphie d'un mot.

34. Voir dans le présent volume l'exemple « DiscussionsRussie » de son article (p. 83 et fig. 1, p. 94).

35. L'affichage en double-page résout ce problème, mais en pose un autre, celui de la taille : une double-page sur écran est minuscule.

dans le cas d'écrans de petite taille (celui d'un ordinateur portable, notamment), de la présentation des images numérisées. Il n'est toujours pas possible actuellement d'avoir une image nette et lisible *en même temps* que sa transcription : si on zoome sur l'image pour analyser un détail, la transcription n'est plus visible, etc. À l'heure actuelle, il faut souvent avoir deux écrans ou un très grand écran. Peut-être arrivera-t-on un jour à concevoir des ordinateurs à double ou triple écran pliable (diptyque, triptyque) ?

9. « Interopérabilité »

L'édition génétique numérique doit réfléchir à l'interopérabilité de ses standards et de ses normes. La reproduction numérique de l'image d'un manuscrit ne suscite plus guère de débats mais sa transcription textuelle continue, de son côté, à générer des solutions diverses, et par conséquent discutables et discutées. On ne peut qu'être frappé par la multiplicité des choix, des solutions, des ergonomies de présentation : à l'heure actuelle, il n'existe malheureusement aucune réelle homogénéité, aucune passerelle, aucune plateforme commune qui permettrait de mettre en relation les résultats obtenus. Il est plus que temps de se mettre à réfléchir et à travailler autrement, de manière à mutualiser ressources, données et outils.

En effet, à la fin des années 1990, on a vu se multiplier les projets et prototypes, chacun se présentant comme un modèle. Mais leurs ambitions se traduisaient le plus souvent par la force des choses – les logiciels libres étant encore balbutiants – par l'adoption de formats propriétaires irréductibles les uns aux autres, causant sans doute une déperdition des énergies et des investissements.

Par ailleurs, la plupart des bases de données internes (par exemple le prototype de base de données sur le corpus du *Rose et Vert* de Stendhal³⁶, développée sous FileMaker®), conservées sur le disque dur d'un ordinateur, ne peuvent être diffusées en ligne en l'état actuel de la technologie, ce qui est regrettable. Cependant, même s'il existe encore aujourd'hui des interfaces et des systèmes très différents les uns des autres, leurs concepteurs semblent avoir compris l'intérêt de se référer à des formats d'échange, et de prévoir une mise en ligne de bases, plateformes, etc. conçues jusque récemment comme internes (voir par exemple l'article d'Éric-Olivier Lochard sur « Le monde selon Arcane »). Avec les principes de gratuité, de liberté, de travail collaboratif, il semble que de nouveaux modèles soient bien en train de se mettre en place. Cela nous autorise à espérer que, dans quelques années, des passerelles entre les plateformes permettent de travailler de façon transversale sur différents corpus, d'opérer des recherches systématiques, d'établir des statistiques, de dresser des concordances, de faire des comparaisons, etc.

36. Voir l'article de D. Ferrer, J.-J. Labia et M.-I. Mena Barreto (p. 249).

L'avenir de l'édition numérique des manuscrits semble donc aujourd'hui passer par l'utilisation de plates-formes dédiées, évolutives et adaptables à des projets différents sur des corpus différents, comme Hyper, Arcane, ou, dans un avenir proche, CLELIA. Pour peu qu'elles permettent l'importation comme l'exportation de leurs bases de données et qu'elles soient dotées de moteurs de recherche ouverts, accessibles de l'extérieur, elles pourront constituer les briques d'un système de stockage gigantesque, donnant accès à de vastes bibliothèques de manuscrits et de transcriptions. Mais en tout état de cause, il est trop tôt pour dire quel modèle s'imposera...

10. Travail collaboratif

L'informatique permet ainsi d'envisager des projets d'une ampleur considérable. L'inconvénient étant que cela suppose de mettre en œuvre les moyens humains correspondants : il n'est pas imaginable de traiter des milliers de pages seul, même en y passant des années.

Par ailleurs, il est incontestable que la complexité et la lourdeur du temps de traitement de chaque page risquent de décourager les transcrip-teurs, sauf en cas de mise en ligne rapide : le chercheur peut éprouver parfois le sentiment de perdre du temps par rapport au travail philologique traditionnel, quand il doit s'astreindre à saisir des renseignements dans les champs d'une base documentaire, mais cette vue, peut-être légitime à court terme, ne doit pas faire oublier qu'à long terme la solution numérique s'avèrera plus performante, et permettra donc un gain de temps considérable pour l'édition et le travail sur les manuscrits. D'où l'importance de travailler en équipe, ce qui assure la solidarité, la complémentarité des compétences, la mutualisation du travail, et donc finalement plus d'efficacité. Certains chantiers génétiques (le plus spectaculaire de ce point de vue étant sans doute celui de la transcription des manuscrits de *Madame Bovary*) sont menés par des équipes fort nombreuses, tirant parti d'une autre caractéristique du support numérique et d'Internet, qui autorisent le travail simultané et collectif sur un même objet. Il est possible de répartir les tâches de déchiffrement entre plusieurs chercheurs mais aussi d'étendre considérablement le nombre de collaborateurs et, grâce à la facilité de communication par messagerie, de fédérer des spécialistes de différents pays, en équipes «virtuelles» internationales.

D'autre part, dans le développement du Web aujourd'hui, la distance entre producteur et utilisateur d'information, entre auteur et lecteur, entre spécialiste et amateur éclairé s'amenuise peu à peu : la mise en ligne de manuscrits, de transcriptions, d'articles de génétique, peut contribuer à susciter l'intérêt d'étudiants ou de chercheurs éloignés des centres de recherche en génétique littéraire, et à leur donner les moyens d'entrer à leur tour dans le cercle des «experts».

Il serait sans doute démagogique de promettre à tout un chacun qu'il saura déchiffrer séance tenante l'écriture de Pascal ou de Stendhal mais il n'est pas inconcevable que tel amateur de bonne volonté puisse suggérer une lecture pertinente de tel passage difficile, dans lequel la fraîcheur de sa perception aura su distinguer ce que des chercheurs plus aguerris n'avaient pas perçu.

Au minimum, il n'est pas sans intérêt que le grand public se rende compte du travail opéré par les spécialistes pour lui procurer des éditions à la fois savantes et lisibles. L'autorité et la compétence des chercheurs en seront d'autant plus reconnues et elles seront de toute manière requises, par le biais de comités de validation notamment, pour accepter ou refuser les propositions qui pourraient émaner du public.

11. Questions économiques et juridiques

S'agissant d'édition, il est nécessaire d'évoquer les aspects économiques de nos entreprises de recherche et de valorisation. Du navrant échec de l'édition en ligne du *Vivre ensemble* de Barthes³⁷, jusqu'aux aventures éditoriales plus encourageantes comme les éditions complémentaires papier et électronique de *Lamiel*³⁸, on devine l'obstacle majeur expliquant la réticence des maisons d'éditions : la gratuité des éditions en ligne.

Les internautes ont d'ores et déjà pris l'habitude de bénéficier sans bourse délier – de produits éditoriaux remarquables, mis à la disposition de tous sur Internet. Les chercheurs eux-mêmes, habitués pourvoyeurs de marchandises éditoriales, explorent parfois ces circuits de distribution nouveaux, se dispensant d'intermédiaire entre producteur et consommateur de savoir savant. On assiste ainsi avec les éditions électroniques à un phénomène inédit dans l'histoire de l'édition : la mise à disposition gratuite et souvent sans aucune restriction d'accès de la connaissance³⁹. Mais le phénomène est si récent et si contraire aux habitudes éditoriales que l'on peut concevoir la réticence de certains chercheurs à s'engager dans cette voie non lucrative et qui n'est pas toujours source de reconnaissance pour l'ampleur et la qualité du travail effectué, même si le public est déjà là.

L'édition électronique est aussi confrontée aux problèmes traditionnels de l'édition, liés à la propriété des manuscrits et à la gestion pointilleuse qu'en font les ayants-droit. Quant aux bibliothèques publiques et aux collectionneurs, s'ils ne rechignent pas toujours à livrer au public à titre gratuit, dans le

37. Voir les articles de Cl. Coste et G. Bellon au sujet du site roland.barthes.com, qui obéissait au principe d'un abonnement annuel, et s'est avéré au bout de quelques années trop peu rentable pour être maintenu.

38. Voir ci-après l'article de S. Linkès, p. 185.

39. Voir pour une analyse détaillée de cette situation inédite, l'article de P. D'Iorio, p. 279.

cadre d'expositions temporaires, les trésors génétiques dont ils ont la garde, il n'en est pas toujours de même dès lors qu'un peu de commerce, aussi peu lucratif soit-il, se met de la partie.

Les chercheurs rencontrent de grandes difficultés à publier les manuscrits des auteurs contemporains, objets d'un marché certes flatteur pour les écrivains mais fort incommode pour la recherche scientifique. Le statut de ces manuscrits, qui n'ont pas toujours été explicitement légués aux générations futures par les écrivains, qui contiennent parfois des mentions injurieuses ou diffamatoires, ou simplement indiscrettes, dont la propriété est parfois contestée, dont les conditions de conservation ne permettent pas toujours la mise à disposition, génère parfois des conflits juridiques sérieux, qui prennent le chercheur au dépourvu et le dépossèdent trop souvent de son objet d'étude.

Il serait utile de dresser un état des bonnes pratiques dans ce domaine, de diffuser des modèles de conventions propres à protéger à la fois chercheurs et ayants-droit, et les grands organismes de recherche devraient procurer aux chercheurs des ressources à cet effet.

Outre les images numérisées des manuscrits pour lesquelles il existe d'ores et déjà des systèmes très efficaces de marquage, il faudrait aussi pouvoir protéger, sans les rendre inutilisables, les transcriptions, les annotations, tout le travail éditorial nécessaire à la présentation des matériaux génétiques, pour préserver ce qui relève du droit d'auteur. Mais dans ce domaine, tout reste à inventer.

12. Problèmes de pérennité

Le dernier problème majeur de l'édition numérique est lié aux légitimes interrogations sur sa pérennité. Passons rapidement sur les problèmes liés à la dépendance inévitable des éditions électroniques par rapport à l'informatique et aux informaticiens : un serveur en panne, un ordinateur qui « plante », un simple bug informatique, l'absence ou le départ de l'informaticien qui conçoit et fait évoluer le site ou la base de données, et c'est toute l'accessibilité à l'édition électronique qui est remise en cause. Certes la gêne est le plus souvent temporaire, mais il n'y a pas de solution de rechange contrairement à ce qui se passe dans le cas des éditions papier, pour lesquelles il existe toujours une bibliothèque ou un collègue qui peut vous dépanner.

La numérisation de nos fragiles archives de papier est aussi paradoxale pour ce qui est de leur pérennisation. D'un côté, elle apparaît comme une sauvegarde de documents périssables qui souffrent de toute communication au public, fût-il celui de chercheurs triés sur le volet et particulièrement prudents. La possibilité de séparer les missions de conservation et de communication des bibliothèques préservera les précieux documents en les mettant autant que possible à l'abri des atteintes atmosphériques, des accidents et des

manipulations. Mais, d'un autre côté, le document numérique n'est en rien un gage de pérennité. Les supports les plus courants ont une durée de vie mal connue, mais certainement inférieure à celle du papier. La durée de vie d'un CD-ROM ordinaire est ainsi de deux à quatre ans, ce qui suppose de penser régulièrement à en faire des copies pour sauvegarder ses données.

Et comment concilier l'exigence de durabilité du contenu avec l'évolution des nouvelles technologies qui rendent très vite désuets les *media* de lecture ? La rapide évolution des logiciels et des matériels informatiques rend difficile l'exploitation de supports anciens. Comment les lecteurs de demain pourront-ils lire nos données si nos bases ne peuvent plus être lues que par des outils obsolètes et introuvables ? Et même pour un seul et même outil, l'évolution est telle que des versions datant de quelques années sont parfois difficilement récupérables. Il suffit de prendre pour exemple le prototype de base de données Stendhal sous FileMaker® 3, qu'il a fallu transférer non sans mal sous FileMaker® Pro 8. Il faudra donc veiller tout particulièrement à ce que les archives soient régulièrement recopiées, transcrites sur de nouveaux supports. Un des meilleurs gages de longévité, aujourd'hui comme hier, reste la grande diffusion de ces documents. Les textes de l'Antiquité ne se sont pas conservés autrement.

La solution de la mise en ligne semble s'imposer aujourd'hui, maintenant que la bande passante disponible permet de distribuer des documents volumineux et de bonne qualité. Les avantages sont nombreux : faible coût de la publication, distribution mondiale, mise à jour facile, interactivité. En revanche, la pérennité d'un CD-ROM ou d'un DVD-ROM est plus sécurisante pour l'utilisateur. Il est vrai que de nombreux sites bien faits et fort utiles ont ainsi disparu du jour au lendemain, laissant leurs utilisateurs démunis et rendant obsolètes toutes les références qui y renvoyaient.

Les bibliothèques, en particulier, répugnent souvent à augmenter leurs fonds d'un document aussi immatériel et volatile qu'une base de données en ligne, *a fortiori* si elle est payante. L'archivage pérenne du Web devra être une préoccupation constante pour les chercheurs qui choisiront ce vecteur de diffusion de leurs travaux.

Conclusion

Le développement de l'édition numérique des manuscrits suppose l'intervention de plusieurs métiers, dont il convient de reconnaître et de respecter les spécificités. Le bibliothécaire, dans le prolongement de ses missions traditionnelles de conservation et de communication, a la responsabilité de la

numérisation des manuscrits et de la mise à disposition du public de ces matériaux. Il doit veiller, dans ce cadre, à la protection de l'original, à la bonne qualité du produit numérique, à son accessibilité, à son catalogage. Le chercheur intervient ensuite, pour accompagner cette reproduction numérique d'un appareil éditorial qui en permette la compréhension et l'utilisation : transcription, historique, annotation, organisation du dossier génétique. Il est ainsi nécessaire, mais complexe, d'établir des conventions entre bibliothèques depositaires des manuscrits, et chercheurs universitaires qui se lancent dans la constitution de bases documentaires ou de sites, et qui ont impérativement besoin des images de ces manuscrits. L'informaticien, dialoguant avec les autres intervenants, apporte son savoir-faire pour créer les outils logiciels employés dans cette chaîne qui va de la numérisation à l'édition et à la publication sur le Web, et pour adapter ceux qui existent déjà à des situations particulières. La plupart des articles de cet ouvrage démontrent l'efficacité d'équipes au sein desquelles spécialistes du texte et informaticiens travaillent ensemble, dans le respect de leurs compétences respectives.

S'il importe que chaque professionnel ait une bonne connaissance des contraintes, du langage, des procédures de ses interlocuteurs, il faut absolument éviter que chacun empiète sur les compétences de l'autre. Cela impose à ces différents partenaires une nécessaire mais parfois difficile prise de conscience de la complémentarité de leurs approches.

Ces nouveaux outils de la recherche nécessitent une formation que les pionniers ont dû acquérir par eux-mêmes, par tâtonnements. Il conviendrait que les futurs chercheurs, étudiants de Master et Doctorat, soient initiés de manière plus systématique et plus précoce. Le chercheur littéraire du XXI^e siècle ne devrait pas ignorer les problématiques de l'édition numérique et le fonctionnement de ses outils ; il devrait être capable, dans un dialogue avec bibliothécaires et informaticiens, non seulement d'utiliser les éditions numériques mais également d'en concevoir.

Le livre reste pour l'instant un élément incontournable de légitimation académique aux yeux des littéraires, même si les mœurs évoluent dans ce domaine⁴⁰. Tout travail de chercheur diffusé uniquement sur Internet fait l'objet d'une suspicion, pas toujours avouée mais bien réelle (« s'il publie sur le

40. Pour l'instant, on trouve en ligne essentiellement des articles déjà publiés dans des revues, par exemple sur [revue.org](http://www.revue.org). Quelques revues électroniques existent (entre autres exemples, *Fabula* *LHT* (*Littérature, histoire, théorie*), <http://www.fabula.org/lht> ou *Recto/Verso*, <http://www.revue-rectoverso.com>) mais elles restent encore minoritaires en lettres et sciences humaines et sociales.

net, c'est qu'il n'a pas trouvé d'éditeur, que son travail n'est pas sérieux», etc.). De plus, mettre une édition en ligne fait courir le risque de la noyer au milieu de publications sans caractère scientifique attesté, alors que le lecteur non averti n'a pas toujours les moyens de faire la différence⁴¹. Il existe donc un réel enjeu de communication au sujet de ces publications en ligne de grande valeur scientifique, auprès des chercheurs, des étudiants et même du grand public.

41. Dans la rubrique Google Annuaire, il est certes possible de sélectionner uniquement les sites institutionnels mais cette fonctionnalité reste méconnue.