

La Numérisation dans les Bibliothèques de l'Enseignement Supérieur en France

par VALERIE NEOUZE

En quelques années, l'arrivée d'Internet et le développement des nouvelles technologies ont profondément transformé le paysage des bibliothèques de l'enseignement supérieur en France, comme ailleurs. Constitué de 96 bibliothèques universitaires et de six bibliothèques de grands établissements¹, ce vaste réseau documentaire s'est enrichi d'une offre numérique de qualité qu'elle met à disposition d'un public sans cesse élargi. Poursuivant leurs missions traditionnelles d'acquisition, de signalement et de diffusion de l'information, ces bibliothèques ont depuis longtemps pris conscience des potentialités offertes par le format numérique pour diffuser à distance une offre documentaire enrichie: acquisition de documents nativement numériques, numérisation de collections matérielles au sein même des établissements, voire édition électronique de documents créés dans le cadre de l'Université. La maîtrise de ces nouvelles technologies, alliée à un savoir-faire professionnel acquis au fil du temps, entraîne ainsi une évolution du rôle de médiateur de la bibliothèque qui se double à présent d'un rôle émergent de producteur de l'information scientifique et technique. Les bibliothèques de l'enseignement supérieur, en France, s'inscrivent très précisément dans ce contexte où les frontières traditionnelles entre la production, l'édition, le traitement et la diffusion de l'information sont d'ores et déjà abolies.

L'ENVIRONNEMENT NUMÉRIQUE DES BIBLIOTHÈQUES DE L'ENSEIGNEMENT SUPÉRIEUR

Un rapide état des lieux de l'environnement numérique dans les bibliothèques de l'enseignement supérieur en France s'impose en préalable pour mieux comprendre les orientations de la numérisation telles qu'elles se dessinent aujourd'hui au sein de ces établissements. Toutes les fonctions traditionnelles de bibliothèques, de l'acquisition à la communication des documents en

passant par le signalement et la conservation, sont confrontées au changement induit par l'introduction de la dimension numérique.

En quelques années, la documentation numérique² a pris une place très importante au sein des collections des bibliothèques universitaires. En quatre ans, de 1998 à 2002, les dépenses consacrées à l'acquisition de ressources électroniques, - bases de données, périodiques et autres documents numériques -, ont connu une croissance de 88 %. Elles représentent aujourd'hui environ 16 % des dépenses documentaires totales. Le coût élevé de ces services, malgré la constitution de consortia, explique en partie ce pourcentage mais il ne doit pas occulter la réalité de ce changement. A ces acquisitions proprement dites, il faudrait ajouter une masse plus difficile à évaluer, constituée par les collections numérisées. En effet, chaque service commun de la documentation (SCD), chaque bibliothèque d'établissement spécialisé souhaite légitimement mettre en valeur un fonds précieux ou compléter son offre électronique en numérisant une partie de ses collections « papier », dans le souci de construire une politique documentaire cohérente. Cette volonté affirmée se traduit par une croissance exponentielle du nombre de projets, encore souvent isolés. Enfin, par leur expertise de la gestion documentaire et leur situation au sein des Universités - éditeurs elles-mêmes d'une documentation institutionnelle et scientifique importante -, les bibliothèques ont un rôle à jouer dans le circuit éditorial qui se met en place dans ces établissements à la faveur du développement des nouvelles technologies.

Face à l'extraordinaire développement de cette documentation numérique, les universités françaises ont commencé à se doter de système d'information documentaire (SID) porté par le SCD, sous-ensemble du système d'information global de l'établissement. Ce SID a pour vocation de permettre un accès simplifié et transparent à l'ensemble des ressources numériques pourtant hétérogènes: fonds numérisés, ressources distantes négociées par l'Université pour des usagers habilités (en particulier les bases de données et les périodiques électroniques), réseaux de cédéroms pour les différents sites de l'Université, ressources et services disponibles sur internet, cours et documents scientifiques numérisés par l'Université, documents pédagogiques. Un tiers des universités françaises ont en ce sens un projet achevé ou en cours de réalisation. Par ailleurs, ces ressources sont signalées, au niveau national, dans le catalogue collectif de l'enseignement supérieur ou système universitaire de documentation (SUDOC)³ qui recense les documents conservés dans les collections des SCD, tous types et tous supports confondus. Accessible sur internet, ce catalogue, riche de plus de 5 millions de notices et de 13 millions de localisations, fait l'objet de 8000 consultations quotidiennes.

Devant l'importance des investissements consentis par les bibliothèques pour acquérir des ressources en ligne et répondre au développement des

nouveaux usages et méthodes de travail fondés sur le format électronique, il apparaît comme essentiel d'accompagner cette évolution par une offre accrue de contenus numérisés. En ce sens le transfert des collections matérielles sous une forme numérique et la production au sein des établissements de documents numériques sont un trait d'union nécessaire entre les collections matérielles et les acquisitions de ressources en ligne, entre produits et services, dans le souci de proposer une politique documentaire enrichie et cohérente en phase avec les attentes du public.

LA NUMÉRISATION AU SERVICE DES MISSIONS TRADITIONNELLES DES BIBLIOTHÈQUES

Aux problèmes insolubles des bibliothèques chargées de conserver des collections patrimoniales tout en les diffusant, la numérisation propose des éléments de réponse, tout en suscitant d'autres difficultés. Proposés en règle générale dans le cadre des contrats quadriennaux qui régissent la vie des établissements, les projets de numérisation émanant des SCD répondent à deux missions traditionnelles des bibliothèques: préserver et valoriser les collections. La richesse, la confidentialité et l'état de conservation des fonds anciens des bibliothèques de l'enseignement supérieur expliquent la multiplication de ces projets, aux objectifs identiques mais aux formes variées. Une nécessaire structuration minimale a conduit à l'élaboration de trois critères, jugés essentiels par la Sous-direction des bibliothèques et de la documentation: la mise en place d'une procédure de validation scientifique du projet menée en collaboration avec des enseignants-chercheurs, un signalement systématique des documents numérisés dans les catalogues collectifs, un accès de ces fonds sur internet dans une perspective de service public. Ces principes ont pour objectif de favoriser la cohérence scientifique des collections numérisées, d'encourager le partage d'une expertise technique, d'accroître la visibilité de ces collections en France mais aussi au niveau européen et international, enfin d'optimiser les financements de ces projets dont chaque pays sait combien ils sont onéreux.

Parmi les projets les plus aboutis figure le projet *Liber Floridus* qui illustre parfaitement cette volonté publique de valoriser des fonds rares et précieux, difficilement communicables pour des raisons évidentes de fragilité, en les mettant à disposition des chercheurs mais aussi du grand public. Ce programme national, monté en collaboration avec l'Institut de Recherche et d'Histoire des Textes (IRHT/CNRS), a pour objectif de rendre accessibles sur internet toutes les enluminures des manuscrits médiévaux conservés dans les bibliothèques de l'enseignement supérieur, en commençant par les deux fonds les plus précieux, ceux de la Bibliothèque Sainte-Geneviève et de la

Bibliothèque Mazarine. Ce projet scientifique complexe associe la numérisation systématique en haute définition des enluminures à un double travail de signalement, des images d'une part, des manuscrits d'autre part. Ce corpus représente actuellement plus de 32 000 images: chacune d'elles est associée à une notice de description iconographique - suivant une indexation fine réalisée par le personnel des bibliothèques d'après le thesaurus Garnier - ainsi qu'à la notice du manuscrit correspondant. Ces notices sont ensuite intégrées dans une base de données qui permet des recherches multicritères sur les images ou les manuscrits. L'ouverture au public est prévue au printemps 2002. *Liber Floridus* devrait s'élargir progressivement à l'ensemble des fonds de manuscrits médiévaux enluminés conservés dans les bibliothèques de l'enseignement supérieur. Le caractère précieux ou fragile des documents, qui restreint voire interdit leur communication, s'avère ainsi un argument qui incite très largement les bibliothèques à se tourner vers la numérisation. A travers l'exemple de *Liber Floridus*, transparait la nécessité d'une collaboration entre bibliothèques et chercheurs dans un souci partagé de rendre accessibles des sources inexploitées et de permettre ainsi le développement de nouveaux thèmes de recherche. On ne peut que souhaiter et encourager une participation plus active des laboratoires universitaires dans la définition des projets de numérisation portés par les bibliothèques.

Partant du postulat que seule une masse critique de documents peut amener un usager à s'approprier une bibliothèque numérique, les bibliothèques de l'enseignement supérieur se sont lancées également dans la constitution de vastes corpus de textes numérisés. Il s'agit soit d'opérations de complétude, soit d'opérations de transfert d'un support sur un autre. La Bibliothèque nationale et universitaire de Strasbourg a ainsi choisi d'acquérir la version numérisée d'œuvres littéraires françaises et allemandes du XVIIIème et du XIXème siècle, issues de la collection privée allemande Corvey et complémentaires de ses fonds dans son domaine d'excellence. Représentant 17 000 volumes et plus de 3,25 millions de pages, ce fonds sera accessible gratuitement à tous les chercheurs français par l'intermédiaire des bibliothèques universitaires et des bibliothèques de recherche.

La numérisation de collections matérielles, très consultées, en multiples exemplaires, sont également au cœur de ces projets. Les périodiques sont à ce titre particulièrement concernés par cette problématique. Parfois centenaires, ces collections s'alignent sur les étagères des bibliothèques, posant de manière cruciale la question de l'espace et du coût de stockage. Une réflexion est en cours sur l'opportunité d'une conservation partagée, à mettre en œuvre dans la perspective du plan Université 3^{ème} millénaire (U3M), corollaire de la numérisation systématique d'un exemplaire de ces collections. Ce transfert sur un support numérique permettrait ainsi de conserver un accès pour tous à des périodiques que seul un petit nombre d'établissements continuerait de conserver sous une forme matérielle. Cette numérisation s'impose d'autant

plus que les revues scientifiques, qui capitalisent et diffusent les résultats de la recherche, sont au cœur des pratiques documentaires des chercheurs. Sous leur forme papier, elles souffrent pourtant d'une absence de visibilité, encore accrue dans le cas des revues francophones. Si le domaine des sciences dures a d'ores et déjà infléchi ses méthodes de production et de diffusion de la connaissance, le domaine des sciences humaines et sociales reste à structurer. C'est pourquoi la numérisation rétrospective des périodiques en sciences humaines et sociales fait, en France, l'objet d'une attention particulière dans la perspective d'une articulation souhaitée avec des projets de passage à l'édition électronique de ces revues. L'objectif serait de proposer aux chercheurs de ces disciplines une offre numérique homogène, des premiers numéros jusqu'à la parution la plus récente. Le problème du droit d'auteur se pose dès lors dans toute son acuité: les contrats anciens entre auteurs et éditeurs - lorsqu'ils existent -, ne contiennent que rarement de formulation susceptible d'autoriser une exploitation de ce document sous une forme numérique. L'accord de chacun des auteurs ou de ses ayants droits s'avère par conséquent un préalable à leur diffusion en ligne. Entre les articles anciens tombés dans le domaine public et les articles récents pour lesquels les contrats renouvelés tendent à prévoir la cession des droits d'exploitation pour tous supports, se situe une très large partie de la production scientifique en sciences humaines et sociales. A défaut d'effectuer ce colossal travail de recherche sur les auteurs, ces articles ne sauraient être diffusés sur les réseaux, en l'état actuel de la législation et de la jurisprudence⁴. Cet obstacle juridique majeur explique la difficulté de mise en œuvre de ce projet de numérisation dont l'intérêt scientifique est pourtant avéré.

Soulignons enfin la volonté des établissements d'exploiter les potentialités offertes par la numérisation pour proposer à leur public non plus la stricte reproduction du document matériel, obtenue par le mode image, mais un document en mode texte, structuré, permettant une recherche enrichie et de nouvelles formes de travail sur le document lui-même. Les coûts induits par ce type de numérisation mis en regard des très importantes volumétries, expliquent qu'aucun projet de numérisation rétrospective à l'heure actuelle n'envisage de systématiser cette solution, pourtant en parfaite adéquation avec le cas des périodiques.

Sans rentrer dans des considérations techniques, il faudrait néanmoins insister en conclusion sur l'archivage pérenne de ces collections numérisées, qui s'impose, au même titre que la conservation des collections matérielles, comme une mission fondamentale des bibliothèques. En ce sens la numérisation, qui permet effectivement de diminuer la pression sur des documents fragiles ou précieux, pose à son tour la question de la conservation de ces documents numériques. Une question sans réponse définitive: entre émulation et migration des données, aucune règle n'est encore de rigueur. Un

certain pragmatisme préside au choix d'une solution, suivant les spécificités des documents numériques concernés.

UNE APPROCHE PROSPECTIVE: LA PRODUCTION DE DOCUMENTS NUMÉRIQUES

Diffuser et promouvoir l'information scientifique et technique est donc l'une des principales missions des bibliothèques de l'enseignement supérieur. Or une partie importante de cette information, non publiée commercialement, demeure à ce jour sous-exploitée: la littérature grise, émanation directe des activités de recherche développées au sein des Universités. Les nouvelles technologies apparaissent comme une opportunité de donner à ce matériau essentiel à la recherche mais peu accessible une place nouvelle dans le paysage de la documentation scientifique. Parmi les nombreuses opérations menées par les SCD en ce sens, il faudrait citer à titre d'exemple le projet porté par la bibliothèque universitaire de Lille 1, baptisé *Grisemine*. Ce projet s'est donné pour objectif de mettre en ligne, en texte intégral, des travaux de recherche en français de nature diverse: communications pour des colloques, rapports de recherche, supports de cours ou de travaux dirigés, travaux d'étudiants, extraits de thèses... Ces documents, confiés au SCD par leurs auteurs, qui conservent le droit de retirer leur contribution, sont numérisés, stockés au format PDF et accessibles en ligne sur le site du SCD à partir de la notice catalographique saisie au format XML. Une première version de *Grisemine* est d'ores et déjà disponible en ligne⁵.

Au sein de la littérature grise, une place particulière est donnée aux thèses, expression première du dynamisme et de la qualité de la recherche dans les universités. Les projets de mise en ligne des thèses soutenues se sont multipliés ces dernières années, portant à plus d'une vingtaine le nombre de bibliothèques motivées par cette problématique. Cette réflexion relayée au niveau ministériel a donné naissance à un projet national de diffusion électronique des thèses. Jusqu'à présent, les thèses soutenues étaient déposées par leurs auteurs sous une forme papier dans la bibliothèque de l'Université de soutenance. Elles pouvaient être consultées sur place ou mises à disposition d'autres établissements grâce au prêt entre bibliothèques, dont le coût ou la lenteur décourageaient plus d'un chercheur. D'autre part, un exemplaire papier était systématiquement microfiché par l'Atelier National de Reproduction des Thèses (ANRT), qui se chargeait de diffuser ce support de substitution, peu convivial au demeurant, auprès de chaque bibliothèque universitaire française. En offrant un accès démultiplié et à distance, le réseau internet s'imposait comme une solution au caractère restreint de la diffusion de ces documents. Une chaîne de traitement de documents numériques a donc été mise en œuvre, sur un modèle canadien développé par les Presses

Universitaires de Montréal. La thèse, rédigée suivant une feuille de style normalisée, est déposée sous une forme électronique dans un format compatible RTF (le traitement du format LaTeX figure comme une priorité dans les évolutions futures). Après la soutenance, elle est convertie dans un format d'archivage en SGML/XML puis en format de diffusion HTML ou PDF, voire en SGML en attendant des développements ultérieurs. Signalée dans le SUDOC et dans le catalogue local de la bibliothèque, elle est accessible librement et gratuitement en ligne, sous réserve que l'auteur et le jury aient donné leur accord. Cette thèse numérique est archivée par la bibliothèque de soutenance tandis qu'une copie de sécurité est archivée par le Centre Informatique National de l'Enseignement Supérieur (CINES). Une quinzaine d'établissements ont répondu à l'appel à candidature lancé par le ministère de l'Éducation nationale au mois de juillet 2001, devenant ainsi pilotes d'un projet destiné à couvrir l'ensemble du territoire français. La Sous-direction des bibliothèques et de la documentation assure le suivi de ce projet par le biais d'un comité de pilotage chargé de préparer l'ouverture de ce projet à d'autres établissements et de favoriser l'évolution technique de la chaîne de traitement.

L'évolution sensible de la numérisation vers la production de documents numériques dans les bibliothèques est liée également au renouvellement des méthodes pédagogiques de l'enseignement supérieur français. L'harmonisation des formations universitaires au niveau européen, articulée autour du système European Credit Transfer System (ECTS)⁶, et le développement de l'offre de formation ouverte et à distance, impliquent un accompagnement de ce mouvement de rénovation de la part des bibliothèques universitaires. Le développement des campus numériques, pour lesquels deux appels à projet ont été lancés en 2000 et en 2001, encourage une participation active des bibliothèques universitaires à la production d'une documentation numérique adaptée aux besoins des étudiants. Le projet *MANUM*, pour lequel une étude de faisabilité est en cours sous le pilotage des Instituts d'études politiques de Grenoble et de Lyon, a ainsi pour objectif de donner un accès facile et à distance aux manuels, usuels ou photocopiés dont la lecture a été prescrite par un enseignant dans le cadre d'un cours ou d'un travail dirigé. Destinés aux étudiants de premier cycle, ces documents numériques ont pour vocation de composer la première „couche” du modèle proposé par Robert Darnton⁷ sur laquelle viendrait se greffer d'autres couches de documents, comme des liens vers d'autres sites, des exercices corrigés, des tableaux et des graphiques.... Outils de formation, donc, mais aussi aide à l'auto-formation: les bibliothèques sont amenées, en collaboration avec les enseignants-chercheurs, à donner aux étudiants les moyens d'utiliser cette documentation. Elles conçoivent des documents pédagogiques multimedia, adaptés au parcours universitaire de l'étudiant: initiation à la maîtrise de l'information pour les

premiers cycles, illustrée par le projet CERISE⁸, documents spécialisés pour les étudiants avancés ou chercheurs avec l'exemple du *Jurisguide*⁹ en droit.

CONCLUSION

Au terme de ce développement, on ne peut que constater combien les nouvelles technologies ont entraîné de modifications dans le monde de la gestion de l'information. Les bibliothèques, dont le rôle de sélection, de diffusion et de conservation de l'information scientifique et technique demeure et se renforce en dépit d'une rude concurrence, se sont emparées de ces outils pour assurer l'ensemble de leurs missions traditionnelles et les élargir. Installées au cœur des universités, les bibliothèques de l'enseignement supérieur ont vocation à accompagner l'évolution de leurs activités de recherche et d'enseignement en enrichissant leur offre documentaire électronique par la numérisation des collections matérielles. Néanmoins, on constate une inflexion notable de leur activité dans le sens d'une aide à la production accrue de documents numériques au sein même des établissements, conséquences évidentes de la conjonction d'un savoir-faire technique et d'une forte demande. Dans ce nouveau circuit éditorial, les bibliothèques ont un rôle à jouer en amont et en aval des presses universitaires lorsqu'elles existent. En leur absence, peut-être pourraient-elles se proposer pour remplir cette mission d'édition? Enfin, devant la multiplication des projets de numérisation aux objectifs variés, on ne peut que souhaiter une plus grande concertation au niveau national d'une part et au niveau européen d'autre part afin de pouvoir établir une carte documentaire virtuelle, riche et cohérente.

A l'heure où le numérique abolit les frontières physiques entre établissements, entre pays, il devient essentiel de travailler en collaboration pour trouver une solution aux deux problèmes majeurs qui sont nés de ces changements: l'archivage pérenne de ces documents numériques et les problèmes juridiques qui entravent la libre diffusion de cette information. Pour que ces évolutions ne conduisent pas à une régression.

ANNOTATIONS

- 1 Bibliothèque de l'Académie nationale de médecine, Bibliothèque de l'Institut de France, Bibliothèque du Musée de l'Homme, Bibliothèque du Museum national d'histoire naturelle, Bibliothèque Mazarine, Bibliothèque Byzantine.

La Numérisation dans les Bibliothèques de l'Enseignement Supérieur

- 2 La documentation numérique comprend ici les ressources électroniques acquises auprès des éditeurs commerciaux, les collections numérisées et les documents nativement numériques produits au sein des établissements.
- 3 <<http://www.sudoc.abes.fr>>.
- 4 Les exceptions prévues dans le cadre de la directive européenne sur le droit d'auteur sont en cours d'analyse par les associations des professionnels des bibliothèques.
- 5 <<http://bibliotheques.univ-lille1.fr/grisemine>>.
- 6 Système Européen de Transfert de Crédits
- 7 DARNTON (R.), Le nouvel âge du livre, *Le Débat*, n° 105, mai-août 1999, p. 176-184.
- 8 <<http://www.ccr.jussieu.fr/urfist/cerise/index.htm>>.
- 9 <<http://jurisguide.univ-paris1.fr>>.