

LES NOUVELLES PRATIQUES DE LA RECHERCHE

Ces dernières années, les nouvelles technologies de l'information ont bouleversé notre façon de travailler. Le développement de la micro-informatique et du numérique, la multiplication des bases de données, la puissance exponentielle en calcul et stockage des machines, et l'explosion de l'internet ont non seulement offert des possibilités de recherche et de diffusion inconnues jusqu'ici, mais ont aussi induit de nouveaux comportements scientifiques et permis l'apparition de nouvelles façons de travailler.

Très vite l'IRHT a joué un rôle essentiel dans la réflexion sur l'utilisation de ces nouvelles technologies dans le domaine des sciences humaines en général – et des sciences de l'érudition en particulier – en prônant l'emploi de « bonnes pratiques », l'utilisation de logiciels libres, l'interopérabilité, et en se préoccupant de l'archivage pérenne des données numériques et informatiques. C'est la raison pour laquelle les ministères de la Culture et de l'Enseignement supérieur et de la recherche ont eu recours à lui, en 2005, pour élaborer un cahier des charges de la reproduction numérique des documents anciens. Ensuite, la Direction de l'Information Scientifique (DIS) du CNRS puis le TGE Adonis l'ont chargé de créer en collaboration avec l'Ecole nationale des Chartes un centre de ressources numériques (TELMA) destiné à publier des corpus de sources manuscrites provenant de différentes équipes du CNRS. Ces importantes missions n'ont pu être réalisées qu'en s'appuyant sur des services photographique, informatique et plus récemment d'éditions électroniques particulièrement performants.

1. La révolution du numérique

La photographie a occupé dès l'origine une place très importante à l'IRHT, l'un des premiers objectifs de son fondateur ayant été d'y rassembler le plus grand nombre possible de témoins manuscrits afin de faciliter le travail d'édition critique des chercheurs. Plus tard, en 1979, la Direction du livre et de la lecture (DLL) du ministère de la Culture et de la Communication confiait à l'IRHT le rôle d'opérateur national pour la reproduction des manuscrits antérieurs à 1500 conservés dans les bibliothèques municipales de France. En 1992, le ministère de l'Enseignement supérieur et de la recherche (SDBD devenue par la suite SDBIS) faisait de même pour les manuscrits des bibliothèques universitaires. Aujourd'hui la photothèque-filmothèque de l'IRHT conserve 76 553 reproductions de manuscrits de toute provenance (dont 43 856 proviennent des bibliothèques françaises) et 203 866 photographies d'enluminures. Ces quatre dernières années, le fonds s'est ainsi enrichi de quelque 2753 manuscrits et de plus de 48 000 reproductions d'enluminures.

1.1— *Des reproductions de plus grande qualité*

Dès la fin des années 1990, le service « Images » de l'IRHT (sous la responsabilité de G. Kagan) a suivi les progrès des nouvelles technologies numériques afin d'assurer au mieux la délicate transition de l'argentique au numérique. Le métier de photographe s'est totalement transformé, ces dernières années, et les membres de l'équipe ont tous fourni un important effort de formation à ces nouvelles techniques, en étroite concertation avec les chercheurs, pour être en mesure de répondre à leurs demandes. Engagé depuis plus de 30 ans dans des programmes documentaires et de catalogage lourd avec des volumétries très importantes (plusieurs millions de reproductions), l'IRHT devait à la fois s'adapter à l'usage du numérique et de l'informatique et faire coexister deux technologies s'appuyant sur des supports totalement différents (argentique et numérique).

Aujourd'hui le bilan est positif : avec la mise en place d'une nouvelle politique photographique qui consiste à s'affranchir du microfilmage noir et blanc en adoptant la numérisation intégrale et en haute résolution des manuscrits médiévaux, les campagnes permettent une étude plus exhaustive des sources manuscrites. Ainsi, par exemple, les prises de vue du décor des manuscrits, qui faisaient l'objet d'un pré-inventaire puis d'une mission spécifique, sont maintenant créées directement à partir des pleines pages des manuscrits reproduits intégralement. L'arrivée sur le marché, en 2006, de systèmes d'acquisition toujours plus performants comme l'Hasselblad H2D39 IMACON (capteur matriciel supérieur à 39 mégapixels) a permis de produire des images de très grande qualité et de s'affranchir ainsi des vues de détails traditionnellement réalisées à partir de zooms puissants. Ceci représente un gain de temps considérable pour la prise de vue comme pour le traitement des fichiers numériques et l'indexation des vues. Cette nouvelle procédure a, par ailleurs, considérablement amélioré le travail des équipes scientifiques. La reproduction numérique offre aujourd'hui une précision extraordinaire pour lire le texte y compris lorsque des parties sont légèrement effacées. L'étude de la réglure pour la mise en page y est beaucoup plus aisée. De même la fidélité du rendu des couleurs pour l'étude du décor est beaucoup plus grande ; enfin la prise de vue systématique des éléments de la reliure ou des gardes concourt à l'identification d'un possesseur ou à la datation du manuscrit.

Depuis 2005, le service « Images » a redoublé d'efforts dans la mise en place de nouveaux systèmes d'acquisition numérique ; doté d'une seule station de prises de vue numérique en 2004 (Light Phase), le service dispose aujourd'hui de trois nouvelles stations de prise de vue entièrement équipées d'appareils photographiques, de presse-livres, de flashes, de pieds colonne et d'écrans. Le service dispose aussi de huit unités centrales pour le traitement informatique des images et des métadonnées.

La papyrologie est une discipline dans laquelle les technologies modernes de la photographie donnent des résultats particulièrement spectaculaires. Dans le cadre de travaux sur les papyrus carbonisés d'Herculanum (voir axe I.2.1), L. Capron (ingénieur associé à l'IRHT) a réalisé un cédérom qui accompagne l'édition du *Philodème, commentaires sur la musique*, édité par D. Delattre (IRHT) en 2007. Il a réalisé alors la première maquette virtuelle d'un rouleau complet de papyrus d'Herculanum, combinant dessins de Naples et photos en infrarouge (« MSI »). Cette réalisation sert aujourd'hui de référence aux travaux en cours dans le domaine de la papyrologie d'Herculanum pour l'accessibilité à la documentation photographique. En outre, depuis 2007, L. Capron et D. Delattre travaillent en collaboration avec le Center for Visualization & Virtual Environments (Université du Kentucky) pour la mise au point d'une technique fondée sur la résonance magnétique nucléaire et les rayons X de basse fréquence visant à une lecture non destructrice du support qui permettrait une numérisation en 3 dimensions du papyrus et la détection des zones

d'encre, donc du texte écrit, sans ouverture des rouleaux. En mars 2009, l'expérience doit se poursuivre à Bruxelles pour les réglages de la machine qui sera amenée à Paris en juillet 2009 pour effectuer l'expérience sur un rouleau carbonisé déposé à l'Institut de France.

Par ailleurs, la numérisation de la collection de papyrus de l'Institut de papyrologie de la Sorbonne est en cours depuis plusieurs années. L'ensemble des papyrus littéraires édités a été numérisé en haute résolution (600 dpi) et mis gratuitement à la disposition du public.

1.2— Une diffusion accrue

Cette plus grande efficacité technique s'est accompagnée d'une modernisation des systèmes de gestion documentaire et de diffusion des images.

La **base de données Medium** est une base de gestion des reproductions réalisées ou acquises par l'IRHT, tous supports confondus. Elle permet la gestion de toutes les formes de reproductions (argentiques et numériques), la préparation des campagnes photographiques, l'édition des listes et des inventaires. C'est également l'outil de gestion des commandes internes et externes. Consultable sur internet, elle offre la possibilité de commander des reproductions en ligne. Les cotes des manuscrits qui figurent dans Medium permettent aussi d'établir un lien vers les notices des bases scientifiques et la consultation des images dans la bibliothèque virtuelle des manuscrits (cf. *infra*).

En mars 2009 cette base, qui fonctionnait sous Oracle 7, a migré dans une nouvelle application internet en utilisant des langages de programmation ouverts qui ont permis d'élaborer un système d'informatisation des données ouvert et performant. Cette migration en Php MySql, confiée à la société ARES, s'est faite sous le contrôle du service informatique (L. Autin puis C. Masset) et du service « Images » (V. Trémault). Un important travail de vérification, correction et validation des données de la base (notices descriptives des manuscrits, caractéristiques techniques des reproductions, année de production, type de supports, lieux de stockage etc.) a ensuite été nécessaire.

La base Medium est étroitement liée à la **bibliothèque virtuelle des manuscrits (BVM)** avec laquelle elle forme une plate-forme de consultation des manuscrits en lien avec les bases scientifiques de notre unité. Face à l'accroissement considérable de la production numérique, l'IRHT a élaboré, à partir de 2007, un système de visualisation de toutes les reproductions des manuscrits. Aucune des différentes applications, qui étaient jusqu'à présent utilisées, ne permettait, en effet, de consulter l'intégralité des ressources photographiques. Les équipes de recherche ou les lecteurs devaient naviguer d'une base à l'autre pour retrouver une référence précise. Il devenait urgent de trouver le moyen de rassembler ces ressources et de permettre leur consultation de manière rapide et efficace.

La BVM est aujourd'hui opérationnelle et le versement des reproductions numériques, qui nécessite un important travail de référencement, se fait progressivement. Elle est actuellement consultable sur l'intranet de l'IRHT. A terme, toutes les reproductions numériques pourront y être versées et consultées en interne ; leur passage en ligne sur internet sera soumis à l'autorisation des bibliothèques détentrices des manuscrits. D'ores et déjà certaines bibliothèques françaises ont donné leur accord (la BM de Chartres par exemple). Des négociations sont en cours avec les autres. Dans le choix des corpus à mettre en ligne, priorité sera donnée aux manuscrits sur lesquels se fondent les programmes scientifiques de l'IRHT.

Pour chaque manuscrit reproduit, le nombre d'éléments descripteurs est volontairement limité. Depuis la fiche d'un manuscrit, un lien permet d'accéder à la base Medium. Les critères

de recherche sont organisés en deux groupes : des champs indiquent le signalement complet de l'ouvrage pour permettre un accès immédiat à la ressource en ligne ; d'autres fournissent le contenu indicatif (notice simplifiée) des manuscrits. Une aide en ligne, avec une description de toutes les fonctionnalités, est proposée à l'intérieur du site pour aider à comprendre tous les éléments de navigation.

Dans les grandes lignes la BVM permet de consulter de façon synchrone la numérisation intégrale des ouvrages, le décor et la reliure. Des fonctionnalités spécifiques ont été développées pour visualiser simultanément différentes pages de manuscrits avec un niveau de résolution important et pour consulter le document à sa dimension réelle notamment pour l'étude de l'écriture (paléographie).

1.3— *Comment assurer l'archivage pérenne des données numériques?*

Sur cette question fondamentale, l'IRHT n'a cessé d'attirer l'attention de ses tutelles et de ses partenaires depuis 2004. L'enjeu est double : trouver à la fois un espace de stockage important (les images numériques de l'IRHT ont un volume de quelque 10 To en juillet 2009) et un système d'informatisation normalisé qui puisse assurer pour des décennies la lecture des images et leurs métadonnées.

Les études préalables menées par le service « Images » avec des spécialistes du Centre de recherches sur la conservation des Collections (UMR 7188), de la Mission pour la Recherche et la Technologie (MRT) et de la BnF, ont montré qu'il était nécessaire, pour répondre aux exigences d'un tel système d'archivage, d'enregistrer les images dans un format brut sans compression des données binaires, totalement ouvert (format de fichier Tiff) et pouvant englober les métadonnées internes et externes : les données techniques de la prise de vue Exifs (Exchangeable Image File Format) et les données documentaires liées au signalement des manuscrits IPTC (International Press Telecommunications Council). Plus de 550 000 fichiers ont été ainsi traités dans l'attente d'une solution d'archivage pérenne.

Face à une situation qui devenait urgente, alors que les réponses au niveau national ne sont pas encore bien définies, les solutions suivantes ont été proposées à l'IRHT par ses différents partenaires (SDBIS, BnF et Adonis) :

- En 2007 : prise en charge par la SDBIS de l'archivage pérenne de 55 000 fichiers Tiff au Centre Informatique National de l'Enseignement Supérieur (Cines) de Montpellier, représentant un volume de 6 To.
- En 2009 : ouverture d'un espace de stockage supérieur à 15 To de données à l'IN2P3 de Lyon (par l'intermédiaire du TGE Adonis) et le même volume de données à la BnF.
- En 2009 : expérimentation par Adonis d'un système d'archivage pérenne fondé essentiellement sur le protocole OAIS et qui serait confié en binôme à l'IN2P3 et au CINES. Le Centre de Ressources pour la Description de l'Oral (CRDO) est le premier centre de ressources à expérimenter cette configuration. Si cette expérimentation est positive, elle sera étendue à l'ensemble des laboratoires de l'InSHS.

2. Publier sur internet

Editer les sources anciennes et médiévales a toujours été une mission prioritaire de l'IRHT. Il y a quelques années encore, les collections sur papier étaient les seuls moyens de publication. Aujourd'hui, de nouvelles formes de publications numériques et d'édition

critique de sources anciennes sont proposées, comme moyen d'édition unique ou en parallèle avec une publication-papier (voir la liste complète en fin de chapitre). Publier sur le net des corpus de sources et des instruments de recherche nécessaires à leur exploitation, développer des sites internet pour communiquer et échanger, promouvoir les « bonnes pratiques » de l'édition électronique et former les chercheurs dans ce domaine, telles sont les principales tâches auxquelles se sont attelées différentes équipes de l'IRHT, en particulier son « Service éditorial et publications électroniques » (SEPE, créé en septembre 2008, sous la responsabilité de R. Walter).

2.1— *Des éditions de sources à géométrie variable*

L'un des avantages de l'édition électronique est la souplesse d'un système qui permet de mettre en ligne des éditions à géométrie variable tout en autorisant leur mise à jour régulière. Plusieurs plates-formes d'édition électronique ont été créées à l'IRHT et sont proposées aux collègues et plus largement à la communauté scientifique. Les éditions des différentes sources ne sont pas toutes réalisées avec le même degré de profondeur critique, sans que cela diminue leur valeur scientifique. Parfois, les éditions sont provisoires, destinées à être achevées au fur et à mesure du travail de l'éditeur. Certaines éditions proposent des reproductions commentées ou non de sources, d'éditions anciennes ; d'autres de simples transcriptions de sources sans y adjoindre la tradition textuelle et l'identification poussée des lieux et des personnes. D'autres, enfin, proposent de simples reproductions numériques de manuscrits qui vont servir de matière première à un programme scientifique.

L'intérêt de ce genre d'édition « partielle » ou de reproduction numérique est double : livrer rapidement les résultats de la recherche sans attendre l'édition critique de tout un corpus qui prendrait des décennies, et mettre immédiatement à la disposition de la communauté scientifique des transcriptions et des reproductions numériques afin de décupler l'accès aux sources anciennes. Ce travail de reproduction ou de transcription permet, par ailleurs, de favoriser la conservation du patrimoine manuscrit en dispensant parfois (mais pas toujours) les chercheurs de se rendre en bibliothèque. Dans le cas du fonds des manuscrits brûlés de Chartres, il s'agit même de mettre à la disposition des chercheurs des ensembles documentaires qui ne leur seront probablement jamais accessibles vu leur état de fragilité.

De la « simple » reproduction numérique d'un manuscrit à l'édition critique la plus savante, l'édition électronique est pratiquée à l'IRHT sur ses différentes plates-formes de publication : la BVM (Bibliothèque Virtuelle des Manuscrits), d'une part, dont il a déjà été question, et **TELMA (Traitement électronique des manuscrits et des archives)** d'autre part. TELMA (<http://www.cn-telma.fr>) est une plate-forme d'éditions électroniques, créée et pilotée par l'IRHT et l'Ecole nationale des chartes. Elle peut accueillir en format XML (Extensible Markup Language) et en respectant l'initiative internationale d'encodage des textes TEI (Text Encoding Initiative), toute édition électronique de source ancienne. TELMA contient à ce jour 6 corpus propres de l'IRHT, 8 corpus de l'ENC et un corpus commun (le Dictionnaire de latin médiéval Du Cange). Mais cette plate-forme est aussi destinée à accueillir les publications électroniques d'autres laboratoires du CNRS (2 corpus en cours de publication, l'un sur les lois romaines et l'autre sur des textes en langue guèze). Actuellement, toute l'infrastructure technique de ce centre est hébergée et gérée à l'IRHT et 8 nouveaux corpus sont en cours de réalisation et seront publiés en 2009-2010. De nouvelles pratiques de lecture ou d'interrogation naissent de la constitution de ces corpus relativement normalisés, via des modules d'interrogation relativement fine, en plein texte ou sur des « champs » particuliers interrogeables spécifiquement, seuls ou les uns avec les autres, de manière croisée. On peut certes encore lire les textes sur écran les uns après les autres, de haut en bas,

écran après écran, « à l'ancienne »... mais on peut aussi accéder plus rapidement au contenu par ces modules d'interrogation.

Les corpus deviennent de plus en plus grands et ne peuvent être parcourus que par l'assistance d'aides techniques. Il devient donc impératif d'avoir des outils performants de recherche et de navigation. L'une des missions de TELMA et de l'IRHT, dans les années à venir, sera de développer un maximum d'outils d'interrogation et de parcours, afin d'affiner nos modes d'entrée dans les corpus écrits. On ajoutera des modules spécialisés par thème de recherche (en littérature hagiographique, en diplomatique...), des modules lexicographiques ou morpho-syntaxiques, des modules décuplant les capacités sémantiques des listes d'autorité (listes de noms de lieux anciens avec coordonnées GPS actuelles, bases de données prosopographiques multiples...). Ces outils sont d'ores et déjà en cours de conception en partenariat avec les différentes initiatives dans le domaine (TGE Adonis, TEI Council).

2.2— *Recenser, inventorier, cataloguer*

Sur la plate-forme TELMA ou sur d'autres espaces spécifiques (répertoires d'images ou bases de données), les équipes de l'IRHT ont publié et publient des **répertoires de données**, lointains héritiers des ensembles de fiches cartonnées qui ont fait la réputation du laboratoire depuis ses origines. Mais plusieurs caractéristiques font de ces répertoires virtuels de données des outils bien plus efficaces que leurs prédécesseurs matériels. D'abord ils sont publiés en ligne, **accessibles à tous, de manière libre et gratuite**, via l'internet. Tout chercheur en France ou à l'étranger peut y accéder, sans avoir besoin de verser de droit de consultation¹. Ainsi la base de données PINAKES est-elle largement sollicitée par les hellénistes des deux côtés de l'Atlantique et la base CARTULR, qui recense tous les cartulaires médiévaux manuscrits en France, est-elle utilisée par un grand nombre de médiévistes.

Ensuite, ce sont des répertoires publiés en l'état et en constants accroissements et améliorations. Jusqu'ici, l'usage était de publier une fois pour toutes des répertoires considérés comme définitifs et exhaustifs, après de longues années de travail. Aujourd'hui, chacun des répertoires publiés en ligne fait l'objet de **misés à jour régulières**. Ces nouvelles éditions sont signalées automatiquement aux utilisateurs de la plate-forme et aux spécialistes intéressés par courrier électronique et fil RSS. De la sorte, un répertoire même non terminé peut être publié, pourvu que les utilisateurs soient avertis de son état d'inachèvement. La base de données CARTULR (XML, sur TELMA) connaîtra au cours de 2010 une seconde édition qui doublera son volume. De même, le catalogue des manuscrits liturgiques (J.-B. Lebigue) a connu, depuis 2006, plusieurs éditions. Progressivement les chercheurs sont invités à « s'approprier » l'outil en balisant eux-mêmes leurs données ce qui leur permet de fournir un fichier prêt à être publié sur TELMA. Afin de leur faciliter le travail, le SEPE développe actuellement, dans le cadre de TELMA, des interfaces d'assistance au balisage en XML permettant aux utilisateurs de travailler directement sur leurs fichiers d'édition sans avoir besoin de connaître la syntaxe de balisage. Ce projet d'outillage générique d'édition en TEI pour la communauté scientifique, piloté par le SEPE de l'IRHT, est un des projets phares du TGE Adonis pour l'année 2010.

¹ Certaines bases restent soumises à un accès payant, comme IN PRINCIPIO : elles avaient été conçues et publiées au début de la micro-informatique, souvent en collaboration avec Brepols, l'éditeur historique de bon nombre de bases de données en sciences historique et philologique. La nouvelle génération de bases de données ne recourt plus nécessairement à ces grands éditeurs privés.

Enfin, **ces outils sont ouverts** dans tous les sens du terme. La communauté scientifique est de plus en plus associée à leur élaboration et certains chercheurs n'hésitent plus à correspondre avec les équipes pour relever des erreurs qui seront corrigées dans la future édition du répertoire. D'autres sont associés plus directement encore à la saisie des données. Ainsi un accord vient d'être conclu entre des chercheurs romanistes italiens (université La Sapienza à Rome, universités de Chieti et de Parme) et l'équipe des romanistes de l'IRHT qui publie la base de données sur les textes en langue romane JONAS. Désormais les chercheurs italiens peuvent alimenter la base, qui leur est directement accessible en ligne (en format ouvert MySQL), avec leurs propres données qu'ils pourront récupérer en format XML s'ils le souhaitent. De plus en plus, la saisie en ligne dans les bases de données sera ainsi possible par des équipes associées par convention afin qu'elles puissent contribuer aux programmes de l'IRHT mais aussi trouver un espace pour leurs propres publications. La participation de chacun sera toujours correctement identifiée, tout comme la paternité des notices publiées. D'autres projets, en cours de publication, suivront la même logique : ainsi INITIALE, base de données qui décrit les manuscrits enluminés conservés dans les bibliothèques municipales et universitaires, archives et musées en France ou encore BIBALE, base de données sur les collections anciennes de manuscrits et d'imprimés élaborée en lien étroit avec le projet ANR Biblifram. Un autre exemple notable de répertoire de données à large focale, destiné à être mis à jour de manière régulière et complètement ouvert (par la gratuité de la consultation tout comme par l'utilisation de logiciels libres) : BUDE, mise en ligne en 2007. Il s'agit, rappelons-le, d'un répertoire d'informations bibliographiques et biographiques concernant la transmission des œuvres antiques et médiévales entre le XIVe et le XVIIe s., aussi bien par les manuscrits que par les imprimés.

2.3— Les sites de l'IRHT, outils de communication et lieux d'échanges scientifiques

L'IRHT est une unité propre de recherche qui a la particularité d'associer des équipes scientifiques à des services techniques bien formés et équipés. Depuis de longues années, des chercheurs extérieurs au laboratoire sollicitent l'expertise des membres de l'IRHT autant qu'ils consultent ses ressources. Les nouvelles technologies permettent d'étendre cette expertise et facilitent le dialogue avec les autres chercheurs, au travers de divers sites web.

Dès la fin des années 1990, l'IRHT s'est doté d'un **site web complet**, régulièrement actualisé, faisant le point sur la situation de chaque équipe, sur les programmes, les publications, l'actualité scientifique, permettant l'accès aux catalogues de la bibliothèque (PMB) et de la filmothèque (Medium). Régulièrement, de nouvelles éditions de ce site web sont préparées par T. Buquet. Une nouvelle version sur le système de gestion de contenu (CMS) Drupal devrait sortir très prochainement, sous forme dynamique avec catalogue des publications, base bibliographique, pages personnelles et actualités. Le logiciel DRUPAL donnera une meilleure réactivité au site web de l'IRHT, fournissant une information plus fouillée et plus complète à la communauté scientifique.

Sur le site de l'IRHT a été également créée, dès 2004, **une collection de publications en ligne, appelée Ædilis**. Cette collection est structurée en 5 séries, correspondant à différents types d'éditions pour différents publics, chercheurs, étudiants, enseignants, « amateurs éclairés » : Publications scientifiques (monographies en ligne, éditions de sources et autres dossiers scientifiques) ; publications pédagogiques ; bases de données et logiciels (outils et répertoires pour la recherche) ; actes (publication en ligne de séminaires, tables rondes, colloques et journées d'études organisés par l'IRHT) ; sites de programmes scientifiques (site de GDR ou de projet ANR, à la fois blogs, site de publication et actualité autour des thèmes d'un groupe donné). Chaque « volume » est numéroté dans la série. Ædilis recense ainsi toutes les publications en ligne de l'IRHT, qu'elles soient mises en ligne sur les

serveurs IRHT, sur TELMA, ou sur les sites propres des projets. En juillet 2009, Ædilis comptait 37 titres.

Un nombre croissant de responsables de programmes importants de l'IRHT ont par ailleurs demandé un site web ou l'ont monté eux-mêmes, la plupart du temps sous la tutelle du service éditorial et publications électroniques (SEPE). Ainsi les **sites web** Matthias Corvin, Tradlat et Pierre Lorfèvre. D'autres sites sont conçus davantage sur le modèle du « **blog scientifique** », destiné à permettre à des collègues autorisés d'intervenir et de publier leurs informations, commentaires, articles, recensions d'ouvrages, comptes rendus de colloques auxquels ils ont assisté, annonces de colloques ou de manifestations scientifiques à venir, appels à projet ou à candidature... Ainsi Europa humanistica.org, outil de travail pour le projet éponyme et Tradlat (montés avec le logiciel SPIP), le site du GDR « Diplomatique », monté avec le logiciel Drupal et bientôt les sites LIBRARIA, attaché au programme ANR Biblifram et ILM attaché au programme européen du même nom.

Il ne s'agit donc pas seulement de diffuser un maximum d'informations mais de promouvoir un vrai dialogue au sein du laboratoire et plus largement de la communauté scientifique. Aujourd'hui les services concernés sont intégrés aux projets dès le départ car ceux-ci comptent de plus en plus souvent un volet « technologique » important. Cette situation crée de nouvelles dynamiques au sein du laboratoire et favorise les collaborations entre les chercheurs eux-mêmes, entre les sections scientifiques et les services techniques, entre la recherche et l'instrumentation.

2.4— Les « bonnes pratiques » éditoriales

Une série de « bonnes pratiques » ont été définies par les membres du laboratoire s'occupant d'édition électronique et par le SEPE. Après une tentative peu fructueuse d'utilisation d'un logiciel « propriétaire » (SIM) pour l'élaboration d'une base de donnée sur les reproductions photographiques effectuées par l'IRHT, l'utilisation de **logiciels libres et ouverts** s'est imposée. Les limites des logiciels « propriétaires » sont, en effet, connues : obsolescence, coûts importants et difficultés à les faire évoluer, danger de disparition des firmes conceptrices... Le logiciel libre, au code source disponible, est devenu une nécessité pour des raisons de coût mais surtout de sécurité et d'adaptabilité. Le choix de la mise à disposition gratuite, libre, des données est aussi un choix politique et stratégique : il place l'IRHT parmi les grands producteurs de données textuelles anciennes au niveau européen.

Le choix de **langages partagés et solides** est aussi une nécessité, pour les mêmes raisons : PHP et XML sont des solutions couramment utilisées à l'IRHT : PHP pour la facilité de traitement des bases de données ; XML comme mode de stockage, d'édition et de diffusion des données. On cherche en tout cas à dissocier de plus en plus les solutions logicielles des données elles-mêmes, à éviter que données et logiciel soient imbriqués au point de ne plus pouvoir extraire les données en cas de problème (comme dans FileMakerPro par exemple).

Le choix de standards du web s'impose aussi pour les « grammaires » : il s'agit d'agencer nos données exprimées en langage XML de telle façon qu'elles soient consultables, lisibles par tous, mais aussi potentiellement interopérables avec d'autres données du laboratoire ou hors du laboratoire. Pour cela, à l'IRHT, on utilise surtout le **codage issu de l'initiative TEI** (Text encoding initiative), adapté aux textes et permettant de « baliser » efficacement tout texte, ancien ou moderne, manuscrit ou imprimé. La TEI est une grammaire XML reconnue au niveau mondial du W3C ; le TEI council préside à ses destinées et à son évolution : l'IRHT fait partie de ce conseil.

La question de la **pérennité des éditions électroniques** est aussi cruciale que celle de la pérennité des fichiers numériques. Comme pour les images, il faut assurer l'archivage des bases de données, des répertoires, des éditions électroniques sur les différentes plates-formes,

pour des dizaines de gigaoctets. L'IRHT participe de très près à la réflexion menée sur le sujet par le TGE Adonis qui est lui-même en discussion avec l'IN2P3 et le CINES.

L'expertise en matière électronique de l'IRHT l'amène à organiser des journées ou sessions de formation (voir « Formation, expertise, valorisation »). Elle lui vaut aussi une reconnaissance internationale. En témoigne sa participation à la constitution de ***Medioevo Europeo***, un projet européen de méta-portail et de réseau électronique sur les sources écrites médiévales, avec des institutions comme la SISMEL (*Società Internazionale per lo Studio del Medioevo Latino*) en Italie, les MGH (*Monumenta Germaniae Historica*) en Allemagne ou le Warburg Institute en Angleterre. Deux rencontres ont déjà eu lieu à Florence, en septembre 2008 et juin 2009, pour réfléchir à la constitution de ce réseau.

PUBLICATIONS ÉLECTRONIQUES DE L'IRHT

Les publications électroniques de l'IRHT sont regroupées dans la collection *Ædilis*. Elles sont mises en ligne soit sur le site web propre d'*Ædilis*, soit sur la plate-forme du centre de ressources numériques Telma (voir plus loin), ou sous la forme de sites web autonomes. Il peut s'agir d'éditions critiques de sources, de bases de données, de répertoires de logiciels, de publications scientifiques ou pédagogiques, de revues en ligne, de sites spécifiques d'équipes ou de projets de recherche. Toutes ces publications sont consultables librement et gratuitement. La présente liste couvre la période 2005-2009.

Parmi ces publications, un effort particulier est porté sur la mise en ligne de sources manuscrites (chartes, cartulaires, manuscrits, documents d'archives, éditions de textes antiques et médiévaux, etc.) sur la plate forme Telma, réalisée en collaboration avec l'École nationale des Chartes (EnC), centre de ressources numériques soutenu par le TGE Adonis. Telma accueille des éditions et des corpus de l'IRHT et de l'EnC mais aussi des projets d'autres équipes CNRS ou universitaires. L'IRHT apporte aussi son expertise en édition électronique à ces équipes et envisage d'accueillir dans ses collections électroniques les publications extérieures, comme notre institut le fait depuis longtemps pour ses collections papier.

1. *Ædilis* : site des éditions en ligne de l'IRHT

Collection numérique structurée en cinq séries :

Série Actes (séminaires, colloques et tables rondes)

Dernières parutions (depuis 2005) :

- Actes 7 (2005) : *Paris au Moyen Âge*. Séminaire de recherche de l'IRHT (<http://aedilis.irht.cnrs.fr/paris>)
- Actes 8 (2005) : *Les matériaux du livre médiéval : supports, encres, pigments, reliures*, séminaire de recherche de l'IRHT (<http://aedilis.irht.cnrs.fr/materiaux>)
- Actes 9 (2005) : Les manuscrits liturgiques, cycle thématique 2003-2004 de l'IRHT (<http://aedilis.irht.cnrs.fr/liturgie>)
- Actes 12 (2006) : *Le manuscrit dans tous ses états, cycle thématique 2005-2006 de l'IRHT*. (<http://aedilis.irht.cnrs.fr/manuscrit>).
- Actes 13 (2007) : 1937-2007. *L'IRHT, avenir d'une tradition. Journée d'étude du 25/1/2007*. (<http://aedilis.irht.cnrs.fr/irht-avenir-tradition/>).
- Actes 14 (2007) : *Les bibliothèques médiévales au ^{xx} siècle : bases de données et inventaires en ligne. Journée d'études du 14 décembre 2006*. (<http://aedilis.irht.cnrs.fr/bibliotheques/>).
- Actes 15 (2008) : *Colloque Matthias Corvin*. Les bibliothèques princières et la genèse de l'État moderne. 15-17 novembre 2007 (<http://corvin.irht.cnrs.fr>).

- Actes 16 (2009) : Séminaire publication électronique (<http://aedilis.irht.cnrs.fr/publi-elec>)

À paraître dans la série :

- Actes, 10 : *Les éditions médiévales, cycle thématique 2004-2005 de l'IRHT*.
- Actes, 11 : *Les Ymagiers*, résumés des séances du cycle de conférences de l'IRHT, depuis 1998.

Série Publications scientifiques (PS)

- PS 3 (2006) : *CartluR*. Répertoire des cartulaires médiévaux et modernes (<http://www.cn-telma.fr/cartulr>).
- PS 4 (2006-2009) : *Enquêtes menées sous les derniers capétiens*, mise à jour avec ajout de 150 enquêtes en 2009 (<http://www.cn-telma.fr/enquetes/>).
- PS 5 (2006) : *Ordonnances de l'hôtel du roi* (<http://www.cn-telma.fr/ordonnances/>).
- PS 6 (2006) : *Le cartulaire de la seigneurie de Nesle* (<http://www.cn-telma.fr/nesle/>).
- PS 7 (2006-...) : *Catalogue de manuscrits liturgiques médiévaux et modernes*, mise à jour pluriannuelle du répertoire (<http://www.cn-telma.fr/liturgie>).
- PS 8 (2009) : *Actes royaux* extraits du *Corpus Philippicum* (<http://www.cn-telma.fr/actesroyaux/>)

À paraître à l'automne 2009 :

- LEPOR. *Leges Populi Romani*. Recensement des lois comitiales romaines. Projet ANR du Centre Gustave-Glotz (UMR 8585), mis en ligne sur Telma.
- Nouvelle version de la base *CartulR* (PS 3), mise en ligne sur Telma.
- Chartes originales antérieures à 1121 conservées en France. Projet porté par le laboratoire ERL 7229 de médiévistique (Nancy) et le GDR 3177 Diplomatique, corpus mis en ligne sur Telma.
- Manuscrits et textes en langue guéze. Projet ANR « Cornafrique » du CEMAF (Centre d'études des mondes africains). Corpus mis en ligne sur Telma.

Série Publications pédagogiques (PP)

- PP 2 (2006) : *Livret du stage annuel d'initiation au manuscrit médiéval*. Dossier mis à jour chaque année en septembre. (<http://aedilis.irht.cnrs.fr/stage>).
- PP 6 (2007) : *Initiation aux manuscrits liturgiques* (<http://aedilis.irht.cnrs.fr/initiation-liturgie>).
- PP 5 (2009) : *Mise en page et mise en texte des manuscrits hébreux, grecs et latins* (<http://aedilis.irht.cnrs.fr/lexicon/>)

Série Bases de données et logiciels (BDL)

- BDL 1 (2007) : *Millesimo*. Logiciel de chronologie médiévale. Version 2 mise en ligne en 2007 (<http://aedilis.irht.cnrs.fr/millesimo>).
- BDL 2 (2005) : *Calendoscope*. Logiciel d'aide à l'identification des calendriers liturgiques médiévaux (<http://calendriers.irht.cnrs.fr>).
- BDL 3 (2006) : *Onomasticon arabicum*. Prosopographie de l'islam médiéval. (<http://aedilis.irht.cnrs.fr/onomasticon>).
- BDL 4 (2006) : *Index du catalogue des manuscrits datés de France*. Révision en 2006. (<http://aedilis.irht.cnrs.fr/cmdf>).
- BDL 5 (2008) : *Pinakès*. Textes et manuscrits grecs (<http://pinakes.irht.cnrs.fr>).
- BDL 6 (2009) : *Medium*. Gestion des reproductions de manuscrits archivées à l'IRHT (<http://medium.irht.cnrs.fr/>)
- BDL 7 (2009) *Jonas*. Répertoire des manuscrits et textes français du Moyen Âge (<http://jonas.irht.cnrs.fr>).

En cours de développement :

- *Initiale*. Répertoire des manuscrits enluminés des bibliothèques de France. (mise en ligne prévue en novembre 2009).
- *ILM* : *Islamic Law materialized*.
- *Bibale* : *Collections de manuscrits et d'imprimés de l'Europe médiévale et moderne*.

Série Sites de programmes scientifiques (SPS)

- SPS 1 (2005) : *Europa humanistica* (<http://www.europahumanistica.org>)
- SPS 2 (2008) : *De rebus diplomaticis* : site du GDR 3177 « Diplomatique » (<http://www.gdr-diplomatique.net/>)
- SPS 3 (2008) : *Tradlat* : site du groupe de recherche sur les traductions latines d'œuvres vernaculaires (<http://www.tradlat.org>).

À paraître dans la série SPS à l'automne 2009 :

- *ILM* : *Islamic law materialized*. Arabic Legal Documents (8th to 15th centuries) : site du projet européen ERC
- *Libraria* : site consacré à l'histoire des bibliothèques anciennes, financé par le projet ANR Biblifram

2. Autres ressources et bases en ligne

- Bases sur les manuscrits enluminés :
 - *Enluminures* (<http://www.enluminures.culture.fr>).
 - *Liber Floridus* (<http://liberfloridus.cines.fr>).

- *Catalogue de la bibliothèque de l'IRHT* (<http://bibliotheque.irht.cnrs.fr>).
- *Budé*. Base sur l'humanisme et transmission des textes (<http://bude.irht.cnrs.fr/>).
- *Bibliothèque virtuelle des manuscrits* (BVM). Interface de consultation des manuscrits numérisés de l'IRHT. Application consultable en interne uniquement.

3. Autres sites web et revues en ligne

- *École de l'érudition en réseau* (<http://www.ecole-erudition.org>).
- *Nouvelles du livre ancien* (<http://nla.revue.org>).
- *Le Médiéviste et l'ordinateur* (<http://lemo.irht.cnrs.fr>).
Dernier numéro paru : *Formes et couleurs dans les manuscrits, élaboration d'un outil* (2006) (<http://lemo.irht.cnrs.fr/45>).
- *Gazette du livre médiéval* (<http://www.gazlimed.org>).
- <http://www.papyrologie.paris4.sorbonne.fr/>

4. Cédéroms

- *In principio*. Incipitaire de textes latins. Version 15 parue en 2009.

FILMOTHÈQUE ET PHOTOTHÈQUE DE L'IRHT

Bilan de l'accroissement 2005-2008

Entre 2005 et 2008 l'Irht a procédé à la reproduction intégrale des manuscrits conservés dans le Nord Pas de Calais, en Lorraine, en Picardie, en Bretagne, en Aquitaine, en Basse-Normandie et à Paris. Cette couverture photographique représente un total de 2422 manuscrits :

- 1787 manuscrits ont été microfilmés en noir et blanc.
- 637 manuscrits ont été numérisés intégralement en couleur.

En cinq ans l'IRHT a reproduit en couleurs et en haute résolution plus de 25 452 éléments de décors provenant de manuscrits conservés dans les bibliothèques municipales de Cambrai, Boulogne-Sur-Mer, Abbeville, Caen, Coutances, Argentan, Bayeux, Lisieux, Brest, Honfleur, Cherbourg, et 5464 vues de la bibliothèque de la Sorbonne. Ces chiffres montrent une nette progression de la numérisation directe avec notamment la reproduction intégrale du fonds prestigieux de la bibliothèque du Château de Chantilly soit 306 manuscrits (52 500 folios) ou encore avec le fonds Afforty conservé à la bibliothèque municipale de Senlis qui a été intégralement numérisé en couleur avec plus de 18 448 pages reliées dans 24 volumes. En 2009, la reproduction des manuscrits et des documents d'archives est essentiellement numérique avec la reproduction de 92 manuscrits de la bibliothèque municipale d'Avranches, 51 manuscrits provenant de la cathédrale du chapitre de Bayeux, 300 manuscrits conservés à la bibliothèque municipale de Toulouse.

Reproduction intégrale des manuscrits (manuscrits, incunables, imprimés, doc. d'archives etc.)

Traitement annuel	2005	2006	2007	2008	Prévisions 2009
microfilm (noir et blanc)	310	713	423	341	20
numérisation directe (couleurs)	50	163	226	196	468

Reproduction numérique de l'iconographie en nombre de vues (manuscrits, incunables, imprimés etc.)

Traitement annuel	2005	2006	2007	2008	Prévisions 2009
Prise de vue et traitement des enluminures (couleurs)	4359	9333	7850	3910	2063
Prise de vue partielle (couleurs)	0	0	37	1350	4826

Couverture des campagnes

Année	Reproduction intégrale	Iconographie
2005	Alençon, Saint-Mihiel, Bar-le-Duc, Nancy, Pont-à-Mousson, Neufchâteau, Épinal, Rambervillers, Remiremont, Saint-Dié, Soissons, Troyes, Mehun-sur-Yèvre, Paris	Valenciennes, Douai

2006	Épinal, Neufchâteau, Remiremont, Soissons, Compiègne, Amiens, Beauvais, Laon, Clermont, Abbeville, Château Thierry, Noyon, Morienvall, Roye, Senlis, Boulogne-sur-Mer, Chantilly, Langres, Brest, Dinan, Guingamp, Lamballe, Quimper, Rennes, Saint-Brieuc, Vannes, Vitré	Bibl. de la Sorbonne, Cambrai
2007	Bordeaux, Chantilly, Chartres	Bibl. de la Sorbonne et Paris Musée de Cluny, Boulogne-sur-Mer, Noyon
2008	Bordeaux, Agen, Bayonne, Libourne, Oloron-Sainte-Marie, Pau, Périgueux, Villeneuve-sur-Lot, Chantilly, Senlis,	Abbeville, Caen, Argentan, Bayeux, Cherbourg, Coutances, Honfleur, Lisieux, Valognes, Vire
2009	Avranches, Toulouse, Bayeux	Strasbourg

Numérisation des microfilms

BM et BU	2006 - 2008
Amiens	382
Troyes	1765
Nîmes	146
Reims	971
Chalons	154
Lille (hors convention)	334
Autres fonds de Champagne-Ardennes	513
Chantilly (2008-2009 en cours)	170
Total (manuscrits)	4435

Achat et traitement des commandes internes (**Films clients archivés dans les collections de l'IRHT**)

Type de travaux	2006	2007	2008
Nb. de commandes	183	262	233
Microfilms	154	212	208
CD R et DVD	32	77	76

Ventes et traitement des commandes de reproductions

Type de travaux	2005	2006	2007	2008
Commandes de reproductions (externes)	77	220	295	350

