

La renaissance du livre ancien : bilan du projet DEBORA et perspectives d'avenir

**Nicolas Barbey, Jean Guillemain, Géraldine Pécoc'h,
Patrice Ract**

Sous la direction de François Dupuigrenet Desroussilles
Directeur de l'Ecole nationale supérieure des Sciences de l'Information et
des Bibliothèques

Sommaire

INTRODUCTION	4
PARTIE I - L'ENVIRONNEMENT DES BIBLIOTHÈQUES NUMÉRIQUES6	
1.1. Le contexte d'apparition et de développement des bibliothèques numériques	6
1.1.1. Historique des bibliothèques numériques	6
1.1.2. Les avantages des bibliothèques numériques	7
1.2. Les bibliothèques numériques : un état des lieux.....	9
1.2.1. Les bibliothèques numériques en mode texte	10
1.2.2. Les bibliothèques numériques en mode image	14
1.3. Bilan	15
PARTIE II - REPENSER L'ACCES AUX DOCUMENTS ANCIENS : LE NAVIGATEUR DEBORA	17
1.1. Le projet européen et sa réalisation (1999-2001)	17
1.1.1. Les partenaires du projet.....	17
1.1.2. Analyse des usages et sélection du corpus à numériser	18
1.1.3. Réalisation du projet : le navigateur DEBORA.....	19
1.2. DEBORA : l'informatique au service du livre ancien	21
1.2.1. La compression des images.....	22
1.2.2. Le format de structuration des données développé dans le cadre de DEBORA : AdHoc	25
1.3. Les principales fonctionnalités d'exploitation des documents	27
1.3.1. Les outils de visualisation	27
1.3.2. Les outils de recherche	28
1.3.3. Les outils d'exploitation personnelle	29
PARTIE III - LA MAGNIFICENCE ET L'HISTORIEN DU LIVRE : RECHERCHES ACTUELLES ET PERSPECTIVES.....	31
1.1. L'accès au document	31
1.1.1. Les imprimés.....	31
1.1.2. Les numérisations	34
1.2. Les usages actuels des chercheurs	35
1.2.1. Bibliographie matérielle « livre en main »	37
1.2.2. Bibliographie matérielle virtuelle.....	42
1.3. Les outils proposés par DEBORA	48
1.3.1. Les outils de visualisation	48
1.3.2. Les outils de recherche	55
1.3.3. Les outils d'exploitation personnelle	58
1.4. Les perspectives de développement : pour une nouvelle architecture de la base	61
1.4.1. Conception du catalogue	61
1.4.2. Corpus thématiques et bases associées.....	64
1.4.3. Environnement de recherche	66
CONCLUSION	69
BIBLIOGRAPHIE.....	72
TABLE DES ANNEXES	77

Remerciements

Nous remercions tous ceux qui ont porté intérêt à notre recherche, et plus particulièrement Ilaria Andreoli, Gérard Bruyère, Frank Le Bourgeois, Pierre Cubaud, Pierre Guinard, Yves Jocteur-Montrozier et Richard Bouché, sans oublier notre tuteur François Dupuigrenet Desroussilles qui nous a conseillé et encouragé tout au long de ce travail.

Introduction

En 1999, au moment où les bibliothèques numériques, à l'instar de Gallica, commencent à prendre de l'ampleur dans le domaine du livre ancien, le projet européen DEBORA (**D**igital **acc**Ess to the **BO**oks of the **RenA**issance) est lancé. Celui-ci vise non seulement l'élargissement de l'accès distant au livre du XVI^e siècle, mais également la mise à disposition d'outils de recherche inédits réalisables grâce aux progrès accomplis en matière de description et de hiérarchisation des données numériques.

Le développement du logiciel (ou navigateur) DEBORA s'est opéré dans l'optique caractéristique d'un prototype : montrer les compétences et les savoir-faire acquis. Par manque de temps, la réalisation du logiciel n'a pu être totalement achevée. Aussi, quoique l'avis de l'Union européenne ait finalement été favorable, aucune véritable exploitation de cet outil n'est encore possible à ce jour. C'est pourquoi nous avons été chargés, en nous plaçant du point de vue de l'historien du livre, d'évaluer le logiciel en testant ses fonctionnalités et en les comparant aux réalisations actuelles dans le secteur de la numérisation des livres du XVI^e siècle. La finalité d'une telle étude était de déterminer si une reprise du projet était envisageable, et si oui, de préciser les orientations souhaitables.

L'évaluation des outils proposés par le logiciel ne pouvait être faite qu'à partir de l'ouvrage numérisé dont le traitement est le plus abouti dans la base DEBORA¹ : *La magnificence de la superbe et triumpante entree de la noble et antique cité de Lyon faicte au tres chrestien roy de France Henry deuxiesme de ce nom, et a la royne Catherine son espouse, le XXIII de septembre MDXLVIII*, publié à Lyon en 1549 par le libraire Guillaume Rouillé. Cet opuscule anonyme a été attribué au poète Maurice Scève (v.1500-ap.1563), précurseur de la Pléiade et « prince de la Renaissance lyonnaise » selon l'expression de Verdun-Louis Saulnier. Dans ce livret, Scève relate point par point et avec force détails le déroulement des solennités organisées à l'occasion de l'entrée à Lyon du roi Henri II.

Du point de vue méthodologique, après avoir pris connaissance de l'ouvrage original à la bibliothèque municipale de Lyon, nous avons privilégié

¹ <http://rfv6.insa-lyon.fr/debora/client.htm>.

l'apprentissage du navigateur tout en consultant le plus grand nombre possible de sites web présentant des numérisations d'ouvrages du XVI^e siècle. Afin d'étoffer nos connaissances sur les aspects techniques de DEBORA et des bibliothèques numériques en général, nous avons ensuite contacté Frank Le Bourgeois du laboratoire Reconnaissance des Formes et Vision (INSA de Lyon), un des concepteurs du logiciel, et Pierre Cubaud, responsable de deux bibliothèques numériques : l'Association des Bibliophiles Universels (ABU) et le Conservatoire numérique (CNUM) du Conservatoire national des arts et métiers (CNAM). Nous avons également assisté à quelques conférences des « Jeudis du numérique »² sur le thème de la numérisation du patrimoine.

Toutes ces informations, complétées par une lecture attentive des différents documents issus de la mise en œuvre du projet, nous ont permis de bien comprendre et de définir notre sujet. Dès lors, se dessinaient les deux premières parties de notre mémoire, la première constituant un état des lieux en matière de numérisation et de diffusion du patrimoine livresque de la Renaissance, la seconde se focalisant sur le navigateur DEBORA et ses spécificités par rapport aux bibliothèques numériques.

Quant à la troisième partie, elle devait prendre en compte l'exemple de *La magnificence*. Nous avons donc décidé de mettre en parallèle les trois types d'outils qui étaient à notre disposition (la bibliographie, Gallica et DEBORA) en les appliquant au cas concret d'une étude de bibliographie matérielle. Cela nous permettait à la fois de situer les possibilités offertes par DEBORA par rapport aux exigences méthodologiques des historiens du livre et éventuellement de répondre à la « question subsidiaire » de notre sujet : peut-on légitimement attribuer l'impression de *La magnificence* à Jean de Tournes ? Enfin, sur la base de cette étude méthodologique, nous avons tenté de définir un certain nombre d'orientations et de changements possibles en vue d'une éventuelle poursuite du projet.

² Les « Jeudis du numérique » sont un cycle de conférences organisé par l'ISDN (Institut des Sciences du document numérique en Rhône-Alpes). Certaines communications de la journée d'étude du 7 mars 2002 portant sur la numérisation du patrimoine sont consultables en ligne (<http://isdn.enssib.fr/archives/numerisation/numerisation.htm>).

Partie I - L'ENVIRONNEMENT DES BIBLIOTHÈQUES NUMÉRIQUES

1.1. Le contexte d'apparition et de développement des bibliothèques numériques

1.1.1. Historique des bibliothèques numériques

Dès les années 1960, le monde des sciences humaines prit en compte l'informatique, lorsque les linguistes s'intéressèrent à l'impact qu'elle pourrait avoir sur le traitement des textes littéraires qu'ils étudiaient.

La possibilité de lire des textes sur un ordinateur offrit de nouvelles perspectives pour toutes les disciplines qui touchaient au texte (édition, linguistique, sociologie, traduction, etc.) et donna lieu à la constitution de bases textuelles qui se structurèrent, se normalisèrent³ et devinrent indispensables.

Par ailleurs, le domaine des sciences humaines s'intéressa aux possibilités offertes par l'informatique de stocker de grandes quantités d'informations. Cet aspect fut mis à profit pour constituer sur ce nouveau support des catalogues cumulatifs plus complets et aboutit par exemple à la base de données Medium, créée dans les années 1970 par l'IRHT (Institut de recherche et d'histoire des textes)⁴ en vue « de cumuler dans une même base toutes les informations sur les manuscrits médiévaux. »⁵

Parallèlement à ces entreprises, naquit l'idée de constituer des réservoirs de textes numérisés dans une optique plus large : les bibliothèques numériques. Le premier projet de cette nature vit le jour aux Etats-Unis en 1971 sous le nom de « Projet Gutenberg »⁶, à l'initiative de Michael Hart qui prit très tôt la mesure de l'impact potentiel de l'informatique sur la diffusion des livres, multipliables à l'infini grâce à cette « Replicator Technology ». Il s'agissait pour lui de créer une bibliothèque touchant le plus grand nombre par

³ La normalisation du codage des caractères fut initiée avec l'apparition en 1968 de la table Ascii, celle des textes en 1987 avec la création de la TEI (Text Encoding Initiative).

⁴ <http://irht.cnrs-orleans.fr> (visité le 20/05/2002)

⁵ LALOU (E.), *La numérisation des manuscrits médiévaux à l'Institut de recherche et d'histoire des textes*, Document numérique, p. 29.

⁶ <http://www.promo.net/pg/> ou <http://www.gutenberg.net> (visité le 11/06/2002).

une numérisation mise à la portée de tous sur les plans technique (adoption du code Ascii, outils simples), intellectuel (les textes choisis devant atteindre 99 % de satisfaction dans le grand public) et financier (textes dépourvus de droits, diffusion gratuite).

En France, la première bibliothèque numérique fit vraisemblablement son apparition en avril 1993 avec la création de l'ABU (Association des Bibliophiles Universels)⁷ qui, consciente des enjeux des bibliothèques numériques amplifiées par internet, s'assigna pour but « le développement et la promotion des supports numériques permettant la libre manipulation de l'information, l'application de ces techniques à la diffusion des travaux de recherche des membres et d'informations du domaine public ». Elle propose aujourd'hui sur internet, grâce au bénévolat de ses membres, un corpus de « textes numérisés en français du domaine public, représentatifs de la culture francophone. »

C'est dans le sillage de ces premières réalisations que se multiplièrent les bibliothèques numériques, initiatives d'individus ou d'associations mais aussi, de plus en plus, de bibliothèques traditionnelles, scientifiques et publiques.

1.1.2. Les avantages des bibliothèques numériques

- Aspects généraux

La numérisation des textes et leur groupement au sein d'une bibliothèque virtuelle répond non seulement aux besoins du monde de la recherche, mais aussi aux préoccupations et aux principales missions des bibliothèques traditionnelles qui sont la conservation et la communication des documents.

En effet, fournir une version numérique de ses collections est une façon de limiter la consultation de celles-ci sous leur forme physique et par conséquent leur dégradation⁸. Cela constitue surtout un moyen d'accroître la diffusion de l'information. D'une part, puisque le document numérique est

⁷ <http://abu.cnam.fr/>. Cf. CUBAUD (P.), GIRARD (D.), *ABU : une bibliothèque et son public* [en ligne]. Disponible sur le site: <http://jasmin.cnam.fr:8081/RR9813/> (visité le 11/06/2002).

⁸ Le choix de la bibliothèque municipale de Lisieux (<http://www.bmlisieux.com/>) de numériser massivement les brochures, tirés à part et opuscules en sa possession peut se justifier par la fragilité de ces supports.

multipliable à l'infini, l'utilisateur n'a plus à se soucier des problèmes de disponibilité du document et des risques qu'il soit épuisé, perdu, déjà emprunté ou consulté par quelqu'un d'autre, incommunicable pour cause de mauvais état ou de train de reliure. D'autre part, l'immatérialité du document électronique abolit les contraintes géographiques et, depuis la généralisation d'internet, permet le repérage et l'accès au document quels que soient le lieu de consultation et celui de conservation du document original.

La numérisation des collections des bibliothèques traditionnelles constitue donc une opportunité unique de tendre vers l'idéal d'une bibliothèque encyclopédique et universelle, grâce au partage d'un même langage informatique et à la mise en commun de l'information qu'il rend possible. C'est dans cette optique que l'on peut replacer Gallica, la bibliothèque numérique de la Bibliothèque nationale de France⁹, qui propose depuis 1997, en collaboration avec d'autres institutions¹⁰, 80 000 documents s'étendant du Moyen-Age au XX^e siècle, à destination du chercheur et du curieux.

L'ambition des autres bibliothèques numériques est beaucoup plus modeste et consiste avant tout à mieux faire connaître des fonds de nature locale (bibliothèque municipale de Lisieux), ou spécialisés (CNUM, Conservatoire numérique des arts et métiers)¹¹. Créé en 2000 pour illustrer à l'aide de documents numérisés les manifestations du bicentenaire de la pile Volta, le CNUM poursuit son œuvre de diffusion et, à travers ses collections, se propose de faire découvrir l'histoire des sciences et des techniques¹².

- Les documents anciens et la numérisation

Dans les programmes de numérisation, un intérêt tout particulier est porté aux ouvrages anciens, comme l'atteste l'éditorial d'un numéro spécial de *Document numérique*¹³. Les documents anciens sont souvent fragiles ou précieux et leur communication est très réglementée : il faut justifier ses demandes de consultation et manipuler les ouvrages dans certaines conditions. Leur accès est même parfois refusé sous leur forme originale quand un

⁹ <http://gallica.bnf.fr/>

¹⁰ Parmi eux : les bibliothèques du Musée de l'homme, du Muséum national d'histoire naturelle, la médiathèque du Patrimoine et de l'Architecture, les bibliothèques de l'Université d'Orsay, de l'Ecole Polytechnique, du CHU de Sainte-Anne et pour les ressources en mode texte, l'INaLF.

¹¹ <http://cnum.cnam.fr/>

¹² Outre le site, voir DEBLOCK (G.), ROZET (B.), CUBAUD (P.), *Le conservatoire numérique des arts et métiers : une création partenariale*, BBF, 2001, Paris, t. 46, n°4, p. 43-48.

¹³ *Les documents anciens*, numéro spécial de 1999 de Document numérique, p. 7.

microfilm existe¹⁴. La numérisation peut donc remplacer avantageusement les originaux puisqu'elle restitue les couleurs, offre un accès convivial et permet même dans certains cas de rendre le document plus lisible¹⁵. Par la copie ou le téléchargement, elle offre en outre à l'utilisateur la possibilité de ne plus dépendre des conditions de consultation de la bibliothèque traditionnelle et de s'approprier pleinement le document, dont il peut réaliser un tirage papier pour ses recherches personnelles.

L'accès aux documents anciens se pose de façon d'autant plus cruciale qu'ils sont bien souvent dispersés et que leur caractère unique rend souvent indispensable leur recensement et, si possible, leur consultation, dans le cadre d'une étude sur un texte, une édition, un auteur ou même un imprimeur. Sont concernés par cette situation les manuscrits mais également les livres des débuts de l'imprimerie puisque, aux variantes propres à chaque édition, s'ajoutent les caractéristiques de chaque exemplaire. La numérisation facilite leur consultation et leur comparaison.

Des initiatives de numérisation sur vidéodisques ou cédéroms avaient commencé à donner aux chercheurs un accès virtuel à certains ouvrages anciens. Internet a pris le relais à la fin des années 1990 en apportant aux bibliothèques numériques sa souplesse et son accessibilité plus grandes, devenant le réservoir par excellence de ces corpus de documents. Cela dit, ces nouvelles bibliothèques ne sont pas équivalentes, car la définition du public visé par chacune d'elles conditionne leur nature et les éventuels outils proposés.

1.2. Les bibliothèques numériques : un état des lieux¹⁶

Abstraction faite de la nature des corpus numérisés qui dépend principalement des fonds de la bibliothèque ou de la discipline du laboratoire à l'origine du projet, les bibliothèques numériques se répartissent à l'heure actuelle de façon relativement égale en deux grands ensembles. Ceux-ci se

¹⁴ LE BARS (Fabienne), MUNCK (Gilles), *La communication des documents patrimoniaux*, Protections et mise en valeur du patrimoine des bibliothèques de France : recommandations techniques, Paris, 1998, p. 45-50.

¹⁵ Des procédés techniques permettent en effet lors de la numérisation d'accentuer le contraste, de masquer les taches, etc.

¹⁶ Cette partie se fonde sur les grilles descriptives de quinze bibliothèques numériques proposant pour la plupart des livres du XVI^e siècle (Cf. tableau en annexe 1).

On trouvera une liste de sites sur la Renaissance à l'adresse suivante : <http://www.cesr.univ-tours.fr/Epistemon/Hyperb.htm> (visité le 10/05/2002).

caractérisent par le procédé qui a servi à la numérisation des documents, à savoir le mode texte ou le mode image. Cette distinction qui se fonde sur un aspect purement technique met cependant en jeu des conceptions radicalement différentes.

1.2.1. Les bibliothèques numériques en mode texte

- Les principes

En consultant les bibliothèques numériques, on constate que la très grande majorité de celles qui proposent leurs documents en mode texte sont issues du milieu de la recherche, et le plus souvent d'une discipline en relation avec la linguistique. Il peut s'agir d'une université comme celle de Poitiers pour Epistemon¹⁷, d'une institution comme le CNRS associé à l'INaLF (Institut National de la Langue Française) pour Frantext¹⁸, d'une association comme l'ABU ou même d'un chercheur comme Pierre Perroud, enseignant à Genève et concepteur du projet Athena¹⁹. Ce schéma connaît cependant des exceptions : ainsi la bibliothèque municipale de Lisieux est un cas atypique qui s'explique par sa collaboration avec Lexotor, base de données linguistique des textes lexoviens produite par l'Université de Toronto²⁰.

L'aspect nettement scientifique de ces bibliothèques tient à la nature même des préoccupations de leurs producteurs et de leurs utilisateurs, qui s'avèrent souvent être les mêmes personnes. En effet, pour mener à bien des recherches d'ordre linguistique, il est essentiel de pouvoir interroger le corps même du texte. Or, dans l'état actuel des choses, la seule façon d'y parvenir est de disposer d'un fichier de type Word.

Il existe deux procédés de numérisation en mode texte²¹. Le plus rapide consiste dans un premier temps à numériser l'original en mode image, puis à soumettre le document numérique obtenu à un OCR (Optical Character Recognition), logiciel capable de convertir en signes typographiques des configurations de points. Cependant, pour les manuscrits, incunables et livres

¹⁷ <http://www.cesr.univ-tours.fr/Epistemon> (visité le 10/05/2002).

¹⁸ <http://zeus.inalf.cnrs.fr> (visité le 10/05/2002).

¹⁹ <http://un2sg4.unige.ch/athena/html/athome.html> (visité le 10/05/2002).

²⁰ <http://www.chass.utoronto.ca/epc/langueXIX/lexotor/> (visité le 10/05/2002).

²¹ Sur ces aspects techniques, voir en particulier JACQUESSON (A.), RIVIER (A.), *Bibliothèques et documents numériques : concepts, composantes et enjeux*, Paris, Ed. du Cercle de la librairie, 1999, et notamment le chapitre I, point 4, « les techniques actuelles de numérisation ».

imprimés au XVI^e siècle, les OCR révèlent leurs limites. Sans même évoquer l'écriture manuscrite, les caractéristiques typographiques des imprimés les plus anciens ne présentent pas la régularité requise pour une identification efficace. Non seulement certains caractères n'existent plus depuis longtemps, qu'il s'agisse du s long ou même des lettres portant des abréviations, mais encore les u et les v sont confondus, les caractères sont souvent abîmés et les erreurs de composition fréquentes. Qui plus est, le recours des logiciels à des dictionnaires pour valider les formes reconnues s'avère inutile en présence de formes orthographiques fluctuantes et vieilles. Par conséquent, la saisie manuelle, beaucoup plus sûre, s'impose²².

- Les résultats proposés

La saisie manuelle, seule capable d'aboutir à une base utilisable par le chercheur en sciences du langage, trouve son expression dans des bibliothèques d'ordre exclusivement textuel où la matérialité du document apparaît finalement comme secondaire²³. L'absence d'appareil critique tient au fait que ces bibliothèques numériques sont avant tout conçues comme des outils de travail et non comme la présentation de recherches abouties. C'est ce qui ressort clairement des principes éditoriaux exposés par Epistemon : « Ces publications en ligne ont pour objet une large diffusion d'un état du texte numérisé avec la reproduction la plus exacte possible d'un exemplaire original. Il ne s'agit pas d'élaborer systématiquement des éditions hypertextuelles qui généreraient les différentes éditions d'un même texte et en fourniraient un appareil critique, bien que cela soit envisagé. »²⁴ D'ailleurs, les bibliothèques numériques en mode texte exposent généralement leurs principes de saisie, sur les pages d'accueil du site ou dans la notice de l'ouvrage. Le site Epistemon apporte toutes les précisions nécessaires à ce sujet :

- respect strict des graphies et de la ponctuation
- respect de la mise en page et du lignage
- pas d'adjonction d'alinéas, ni de guillemets, ni de tirets

²² On recourt en général à un « double keying », double saisie du texte par des personnes différentes dont les travaux sont ensuite comparés informatiquement pour relever d'éventuelles divergences.

²³ Chaque texte est précédé d'une notice bibliographique sommaire.

²⁴ Pour un exemple de génétique des textes, voir le site Galileo Galilei : http://www.mpiwg-berlin.mpg.de/Galileo_Prototype/INDEX.HTM (visité le 10/05/2002).

- pas d'intervention sur l'usage des majuscules
- pas de séparation des mots collés (comme *lon* pour *l'on*, *la* pour *l'a...*)²⁵
- gravures : reproduites si la qualité de l'original le permet. A défaut, leur emplacement est signalé

Toutefois, il a été indispensable pour l'informatisation de procéder aux modifications suivantes, quand elles n'affectent pas l'état de la langue :

- dissimilation des *i/j* et *u/v*, sauf pour les mots latins et les *u* surmontés d'un tréma
- adjonction d'une numérotation systématique des pages de l'original pour permettre le repérage. Elle apparaît précédée d'un \$, alors que la pagination originale apparaît dans le texte
- résolution des abréviations, y compris des tildes
- correction des coquilles évidentes, signalées en fichier séparé
- transformation des lettres ornées en lettres ordinaires²⁶

Pour mener leurs recherches, les utilisateurs de ces bibliothèques numériques ont à leur disposition les outils classiques du traitement de texte. Ils peuvent notamment effectuer des copier/coller sur tout ou partie du texte et rechercher des mots ou des expressions dans le texte. Parfois, des fonctionnalités de recherche plus poussées permettent de couvrir indifféremment un texte précis ou l'ensemble du corpus, d'utiliser les opérateurs booléens, de prendre en compte le voisinage, les occurrences ou la fréquence des mots²⁷.

La numérisation en saisie manuelle pose cependant la question de la navigation dans le texte qui, une fois saisi, apparaît d'un seul tenant. L'accès se fait donc à la première page puis en descendant progressivement dans le document. Si cela n'est pas gênant dans le cas des brochures de Lisieux, cela peut le devenir lorsqu'il s'agit de travailler sur Rabelais dans le projet

²⁵ Ce ne sont pas là les méthodes enseignées aux élèves de l'Ecole des chartes, qui ont appris à éviter de confondre les pratiques d'un auteur avec celles de son imprimeur.

²⁶ Le débat sur les principes éditoriaux reste ouvert. On consultera à ce sujet la réaction d'André Tournon au « Protocole pour l'édition des textes imprimés en moyen français » d'Yves Giraud : <http://193.52.215.193/sfdes/sfdes/Bull45.html#tournon> (visité le 10/05/2002).

²⁷ On en trouve une bonne illustration à l'adresse suivante à propos de *La Venerie* de Jacques du Fouilloux : <http://www.chass.utoronto.ca/~wulfric/rentexte/fouillou/> (visité le 10/05/2002).

Gutenberg. Pour remédier à cet inconvénient, quelques bibliothèques numériques comme le Cnum, Epistemon ou Athena ont reconstitué la table des matières de l'ouvrage en créant le plus souvent pour chaque titre de chapitre un lien hypertexte vers la partie du texte correspondant.

Les bibliothèques numériques en mode texte sont donc avant tout des outils créés pour des chercheurs spécialisés en sciences du langage, et leur présentation matérielle, comme intellectuelle, n'est guère en mesure de satisfaire un autre type de public. Compte tenu du prix de revient (une numérisation en mode texte coûte dix fois plus cher que son équivalent en mode image)²⁸ et des qualifications nécessaires (maîtrise de la langue et des graphies anciennes) pour réaliser ce type de bibliothèques, elles sont en général l'œuvre des chercheurs eux-mêmes dont elles sont l'outil de travail. Ces derniers voient d'ailleurs dans la saisie du texte une bonne occasion de se familiariser avec leur objet d'étude²⁹ même si cela a pour conséquence la lenteur de l'accroissement des corpus de ces bibliothèques. Ainsi, l'ABU qui fut pourtant la première à se mettre en place ne compte « que » 288 documents et sollicite des contributions sur la base du volontariat. Parmi les inconvénients de ces bibliothèques, on peut aussi signaler qu'en l'absence de principes éditoriaux clairement définis, il est très difficile de mesurer la qualité et le sérieux du travail de saisie, ce qui a conduit l'INaLF à créer un CCRTI (Catalogue Critique des Ressources Textuelles sur Internet)³⁰ qui a « pour but d'aider les internautes en quête de textes littéraires en langue française à sélectionner, parmi les nombreux sites qui diffusent des ressources textuelles en ligne, ceux qui présentent les caractères les plus sérieux tant sur le plan du traitement éditorial que numérique des textes ». Il faut dire que cette préoccupation est d'autant plus vive dans le domaine de la numérisation en mode texte où finalement, il est relativement aisé de s'improviser « éditeur » de texte et de diffuser son travail sur internet, sans garantie aucune pour l'utilisateur.

²⁸ Cf. Gallica, menu « aide », rubrique « questions/réponses ».

²⁹ TRUDEAU (D.), *L'apport des nouvelles technologies à l'édition critique*, dans Les études françaises valorisées par les nouvelles technologies de l'information et de la communication, 12-13 mai 2000, Toronto [**en ligne**].

Disponible sur : http://www.sjsu.edu/depts/foreign_lang/Constant/apport.html.

³⁰ <http://inalf.ivry.cnrs.fr/ccrti> (visité le 20/05/2002).

1.2.2. Les bibliothèques numériques en mode image

- Les principes

Elles ont été créées le plus souvent au sein d'institutions ayant un public plus large, parmi lesquelles les bibliothèques traditionnelles. Cette situation tient tout d'abord au fait que la numérisation en mode image consiste à donner une photographie numérique du document original ou de toute prise de vue déjà effectuée (microformes, négatifs, diapositives, etc.). Elle offre par conséquent la caractéristique d'être fidèle au document original et de le restituer dans son contenu comme dans sa présentation matérielle (mais tout dépend de la qualité de la prise de vue).

Puisque les bibliothèques traditionnelles ont pour mission de communiquer les documents, le support important finalement peu, elles ont massivement adopté le mode image pour leurs bibliothèques numériques. Ce choix est d'autant plus justifié que la numérisation en mode image est beaucoup plus rapide et économique que celle en mode texte, même si le traitement des documents anciens demande de l'habileté et un matériel performant³¹.

- Les résultats proposés

Le mode image donne accès au document dans sa matérialité : il permet non seulement de lire le texte mais aussi d'en observer la mise en page, les illustrations, la typographie³². Ce type de numérisation permet ainsi de répondre à une plus grande diversité d'attentes que le mode texte : il est susceptible de s'adresser aussi bien aux spécialistes de la période et aux historiens de l'art ou du livre qu'au grand public qui peut découvrir un aspect du patrimoine écrit.

³¹ Il est nécessaire à une bonne qualité de l'image (résolution mais aussi horizontalité et éclairage uniforme du document, etc.) et au respect de l'ouvrage, en particulier de sa reliure. Pour un exemple de redressement de l'image, on se reportera à CUBAUD (P.), TOPOL (A.), *A WWW-based digital library for antiquarian collections [en ligne]*. Disponible sur le site : <http://jasmin.cnam.fr:8081/RR9909/> (visité le 22/03/2002).

³² C'est dire que l'intérêt s'est porté sur les caractéristiques d'édition, et non d'exemplaire : on souhaiterait disposer d'une prise de vue de la reliure, par exemple.

Le document original garde toutes ses caractéristiques, si l'on excepte son format initial qui se retrouve agrandi ou réduit aux dimensions de l'écran. Les seuls outils véritablement proposés portent sur la visualisation. A partir de l'image numérisée, ils permettent de masquer le navigateur, de procéder à des réductions, des agrandissements, des zooms³³ et s'avèrent très utiles pour faciliter la lecture de textes en petits caractères ou pour étudier dans le détail des éléments typographiques ou graphiques.

Cependant, aucune intervention n'est possible sur ce type de document numérisé qui forme un bloc. Les recherches plein texte ne peuvent être menées dans le corps du texte mais ne s'appliquent qu'aux notices et aux tables des matières. L'accès aux pages numérisées, correspondant aux pages physiques du document original, se fait par feuilletage. Il est donc important de souligner que si les illustrations, ornements typographiques et autres éléments graphiques sont restitués, on ne les découvre qu'au fil du document. Pour remédier à la lenteur de consultation du livre virtuel, inconvénient commun avec le mode texte, la plupart des bibliothèques numériques ont établi des tables des matières munies de liens permettant d'accéder à différents points du texte. Il arrive que certaines d'entre elles proposent également sur ce principe des tables des illustrations³⁴ ou mentionnent les figures dans les tables des matières.

Le mode image, au-delà de son aspect attrayant, est avant tout un moyen particulièrement intéressant d'obtenir un reflet fidèle du document dans son intégralité, contenu et contenant. Néanmoins, si ce mode de numérisation donne en apparence un résultat plus complet et plus immédiat que le mode texte, il ne permet aucun travail à l'intérieur du document et ne résout en rien les difficultés de compréhension que peut poser le document ancien.

1.3. Bilan

L'analyse des bibliothèques numériques conduit à distinguer deux grands ensembles régis par les modes texte et image, mais cette répartition ne

³³ Gallica fournit un bon exemple de ces outils et des formats disponibles (taille réelle, taille écran, pleine largeur, agrandissement selon un pourcentage déterminé).

³⁴ Les Ellug (éditions linguistiques et littéraires de Grenoble, <http://www-ellug.u-grenoble3.fr/ellug/>) proposent ainsi pour leur édition numérique de *La nef des folles* de Josse Bade d'accéder directement aux bois gravés par une liste.

saurait rendre compte de la très grande diversité existant dans ce domaine. En effet, face aux développements récents de ces corpus de livres virtuels, aucune norme, aucune définition de ce que doit être une bibliothèque numérique n'existe encore. Ce type d'initiatives relève donc essentiellement du bon sens et des moyens à la disposition de leurs concepteurs, ce qui explique notamment la diversité de présentation et de taille des corpus. Les différences essentielles tiennent à la démarche et aux buts poursuivis par les créateurs, selon l'usage qu'ils affectent aux textes numériques et dont découle la distinction entre mode texte et mode image. Néanmoins, ces deux modes de numérisation apparaissant finalement plus complémentaires que concurrents, des initiatives se sont développées, associant les deux formules, comme c'est le cas pour le projet Galileo Galilei³⁵ qui a pour objet l'étude d'un manuscrit scientifique, conservé à la Biblioteca Nazionale Centrale de Florence.

Quelles que soient les différentes approches des textes numériques, les pratiques des chercheurs n'ont pour l'instant guère été modifiées. Certes le mode texte a créé la linguistique informatique, et l'accès à l'information s'est trouvé amélioré (malgré les temps de téléchargements souvent très longs en mode image), mais le nombre des outils proposés et l'interactivité offerte entre le chercheur et le document demeurent limités. Ce que proposent les bibliothèques numériques correspond finalement en très grande partie aux fac-similés ou aux éditions diplomatiques. La démarche des chercheurs consiste d'ailleurs souvent à imprimer le document numérique pour l'utiliser de façon traditionnelle. On peut donc considérer que les bibliothèques numériques en *mode image*, avant la mise en route du projet DEBORA, n'apportaient aucun outil réellement novateur au monde de la recherche

³⁵ http://www.mpiwg-berlin.mpg.de/Galileo_Prototype/MAIN.HTM (visité le 28/03/2002).

Partie II - REPENSER L'ACCES AUX DOCUMENTS ANCIENS : LE NAVIGATEUR DEBORA

1.1. Le projet européen et sa réalisation (1999-2001)³⁶

Le projet européen DEBORA (**D**igital acc**Ess** to **BO**oks of the **RenA**issance) est né en janvier 1999 à l'initiative du professeur Richard Bouché de l'Enssib. Déployé sur une durée de 30 mois, il avait pour objectif de réaliser une « station de travail » inédite donnant accès à distance à des ouvrages du XVI^e siècle et permettant un travail collaboratif. L'enjeu du projet était loin d'être négligeable : élargir le public potentiel des ouvrages rares et précieux en décloisonnant les réserves des bibliothèques, souvent peu accessibles, et mettre ainsi en valeur les fonds sans nuire à leur conservation.

L'entreprise a, d'une certaine façon, vécu deux phases bien distinctes. Dans un premier temps, le projet à vocation industrielle présenté devant la commission européenne est rejeté. Dans un second temps, réorienté dans une optique de recherche, il est avalisé.

1.1.1. Les partenaires du projet

Trois bibliothèques-ressource ont été choisies pour la richesse de leurs fonds :

- La Biblioteca Casanatense de Rome
- La Biblioteca Geral de Coimbra
- La Bibliothèque municipale de Lyon (BML)

S'agissant de la partie technique (conception et mise en œuvre), le projet était originellement réparti entre six partenaires :

³⁶ BOUCHÉ (R.) (dir.), *DEBORA : projet européen n° LB 5608 A*, juin 2000, dactylog., 179 p.
BOUCHE (R.), MAJID (Ihadjadene), « Le Projet DEBORA », dans *Revue d'Information Scientifique et Technique*, vol. 10, n° 1-2, 2000, p. 55-61.

Développement des outils de travail collaboratif :

Département informatique de l'Université de Lancaster (Grande-Bretagne)

Analyse des besoins, usages et coûts :

Equipe de Recherche sur les Systèmes d'Information et la Communication des Organisations (ERSICO, Université Lyon 3) et Littérature, Idéologies, Représentations aux XVIII^e et XIX^e siècles (LIRE, Institut des Sciences de l'Homme, Université Lyon 2)

Travail sur l'image :

Laboratoire Reconnaissance des Formes et Vision (RFV) de l'INSA (Lyon) et l'Instituto Superior Tecnico (IST, Lisbonne)

Interfaces :

Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques (ENSSIB)

1.1.2. Analyse des usages et sélection du corpus à numériser

L'un des premiers soucis des concepteurs du projet a été de prendre en compte les usages et les besoins des chercheurs seiziémistes afin de réaliser un outil qui leur soit le mieux adapté, concernant aussi bien le corpus présent dans la base que les fonctionnalités du navigateur. Des enquêtes ont été réalisées dans les trois bibliothèques-ressource. Au total, une quarantaine d'utilisateurs des fonds du XVI^e siècle ont été retenus et ont reçu un questionnaire très précis sur leurs pratiques et leurs attentes. Leurs remarques ont surtout porté sur les outils d'une bibliothèque numérique³⁷.

La sélection du corpus à numériser a été opérée dans une optique bien particulière, conforme aux objectifs culturels de DEBORA : réunir des ouvrages

³⁷ DEBORA, Délivrable 2.1, *Etat de l'art sur les pratiques et sur les usagers des bibliothèques virtuelles*, juin 1999.

susceptibles d'intéresser aussi bien les chercheurs qu'un large public d'amateurs. Les critères de la sélection ont été regroupés en cinq catégories³⁸ :

Les critères culturels : langue (celle du pays ou le latin), contenu thématique, esthétique (lettres ornées, illustrations, typographie...)

Les critères d'usages (intérêt de l'ouvrage selon les usagers) : objet de la recherche

Les critères de technique documentaire (intérêt de l'ouvrage pour ses caractéristiques techniques) : consultation et lecture aisées, imprimeur de renom, fort tirage initial

Les critères économiques : ouvrage représentatif d'un fonds, ouvrage intéressant potentiellement un large public

Les critères techniques : adéquation de l'ouvrage avec l'opération de numérisation

Chaque bibliothèque a ainsi sélectionné une trentaine de livres au sein de ses fonds. Certes, ce corpus est peu important et sa grande diversité (de langues, de domaines et de formats) le rend peu significatif du point de vue de la recherche, mais il faut prendre en compte l'intérêt technologique et expérimental qu'il y avait à numériser des ouvrages très différents afin de prendre conscience de tous les obstacles et de chercher à les surmonter³⁹.

1.1.3. Réalisation du projet : le navigateur DEBORA

Conçu par le laboratoire RFV de l'INSA, DEBORA est un logiciel de développement⁴⁰, autrement dit, les outils qui le constituent sont le fruit d'un travail de recherche et non d'une simple intégration d'éléments existants. Le résultat est un prototype, une « vitrine technologique » tout à fait originale,

³⁸ DEBORA, Délivrable 6.1, *Critères choisis*, avril 1999.

³⁹ DEBORA, Délivrable 6.3, *Collection*, avril 1999.

⁴⁰ La première version du projet prévoyait un logiciel industriel d'intégration.

offrant des possibilités de travail nouvelles sur le texte numérisé. Son développement s'est poursuivi jusqu'en octobre 2001.

Aujourd'hui, le navigateur (ou logiciel) DEBORA demeure à l'état de « concept-car ». Il est un exemple des possibilités techniques en matière de numérisation des textes et n'est pas opérationnel du point de vue de la recherche. Sur cinq livres disponibles dans la base⁴¹, un seul, *La magnificence*, est enrichi des métadonnées⁴² et de la transcription du texte.

Concrètement, DEBORA se présente sous la forme d'une interface graphique fonctionnant sous Windows et Macintosh, comme n'importe quel logiciel. Elle se divise en quatre espaces principaux⁴³ :

- la zone de représentation arborescente de la structure du ou des livres sélectionnés (page de titre, pièces liminaires, pages de texte, images) à gauche de l'écran
- la zone d'affichage de la page consultée (plan de travail) à droite
- la zone d'affichage de la transcription
- la zone d'affichage des commentaires

Ces deux dernières zones sont facultatives. Elles peuvent être masquées lorsqu'elles ne sont pas utilisées.

Avant de présenter les principales fonctionnalités de DEBORA, il nous a semblé pertinent de montrer en quoi il diffère réellement des bibliothèques numériques « traditionnelles » ; et le fait est que cette différence se manifeste essentiellement dans la structure interne de DEBORA et dans la conception d'un format d'organisation des données tout à fait spécifique qui permet un usage nouveau et plus « dynamique » de l'image et du texte numérisés.

⁴¹ Belon (Pierre), *Les Observations...*, Paris, G. Corrozet, 1554 ; [Scève (Maurice)], *La magnificence...*, Lyon, G. Rouillé, 1549 ; Munstere (Sébastien), *La déclaration de l'instrument*, Bâle, Mareschal, 1554 ; Camões (Luis de), *Os Lusíadas*, Lisbonne, A. Gonçalves, 1572 ; Dias (Nicolau), *Vida da serenissima princessa dona Joana*, Lisbonne, A. Alvares, 1594.

⁴² Les données constituent le contenu du document. Les métadonnées constituent les règles de structuration et de présentation des données. Pour le livre ancien par exemple : incipit, colophon, privilège, lettrine, etc.

⁴³ Pour une prise de connaissance visuelle du navigateur, on se reportera au point 3.3 de ce mémoire.

1.2. DEBORA : l'informatique au service du livre ancien

Dès sa conception, le projet DEBORA voulait se placer sur un plan différent des entreprises de bibliothèques numériques existantes en proposant de nouvelles fonctionnalités définies en liaison avec les chercheurs de manière à répondre à leurs besoins. Cette orientation initiale a guidé le laboratoire RFV de l'INSA de Lyon dans son développement d'un logiciel particulier. Celui-ci a été conçu dans l'optique de développement d'un prototype et non en vue d'une commercialisation à moyen terme, ce qui explique les quelques dysfonctionnements rencontrés lors de son utilisation.

DEBORA se proposant de rendre accessibles les livres du XVI^e siècle par l'intermédiaire de leur numérisation et de leur structuration en vue d'une exploitation plus performante, il s'agissait de repenser les méthodes développées jusqu'alors par les bibliothèques numériques existantes, lesquelles ne permettent que peu d'interactivité entre le chercheur et le texte (cf. partie I). Un des buts du laboratoire RFV lorsqu'il a développé ce logiciel était en effet de *rendre vivante* l'image du livre numérisé.

Cette position résolument innovante en matière de bibliothèque numérique passait par une nouvelle conception de la chaîne de numérisation des textes puis par le développement d'outils permettant l'exploitation optimale des images binaires obtenues.

Il s'agissait donc de définir un protocole de numérisation des livres devant entrer dans la base. En effet, seule une numérisation faite selon un cahier des charges rigoureux⁴⁴, respecté par toutes les bibliothèques prenant part au projet, pouvait fournir des images dont le format normalisé pouvait être compressé selon des algorithmes déterminés.

Après cette phase essentielle de compression des images, qui seule pouvait permettre une navigation aisée à l'intérieur des ouvrages et de la base, étant donné le poids des fichiers images issus de la numérisation des livres, le second outil informatique indispensable était la mise au point d'un format spécial de description fine et pertinente des documents permettant de faciliter les recherches des spécialistes, voire de leur fournir de nouveaux moyens d'investigation dans les livres anciens.

⁴⁴ DEBORA, Délivrable 4.2.

1.2.1. La compression des images

Créer une bibliothèque numérique, c'est mettre à la disposition du lecteur un certain nombre de livres à l'écran. Le premier problème à résoudre était posé par le poids des fichiers images issus de la numérisation des ouvrages. Il paraissait indispensable de compresser ces fichiers pour permettre leur consultation aisée (éviter l'attente lors de téléchargements interminables de fichiers trop volumineux) et la possibilité de les transmettre sur les réseaux.

Cependant, la compression des images devait être abordée avec discernement. En effet, certains algorithmes de compression entraînent une perte de l'information (JPEG par exemple) alors que d'autres travaillent sans perte (RLE ou LZW).

Pour être viable dans le cas d'une numérisation de livres anciens, la compression devait être la plus grande possible sans pour autant sacrifier les renseignements contenus dans la page imprimée. Il fallait donc trouver les moyens de perdre l'information la moins utile en vue de la restitution et de l'interprétation par l'ordinateur.

Afin de résoudre ce problème, l'INSA a décidé d'appliquer des taux de compression différents selon le type d'information contenue dans la page numérisée. La méthode de séparation des différents plans de la page avait déjà été mise au point par le format de compression DjVu. Cependant, celui-ci n'utilisait pas la redondance des formes sur tout le livre mais effectuait la compression page par page de façon indépendante. C'est pourquoi le laboratoire RFV a préféré développer un type de compression inédit, s'appuyant certes sur les avancées du format DjVu, mais utilisant la forte redondance des caractères à l'intérieur d'un livre.

Le préalable à ce travail de compression modulée selon le type d'information était la segmentation des images entre, d'une part, les zones textuelles et, de l'autre, les zones graphiques, car chacune serait par la suite compressée de manière différente. Cette opération fait apparaître deux types de plans dans l'image, auxquels il faut ajouter l'arrière-plan, c'est-à-dire l'image du support papier.

L'image du document numérisé a donc finalement été séparée en trois plans compressés différemment grâce à un modèle de compression adapté pour chacun d'eux :

- **le plan textuel** qui ne comprend que les caractères, fortement redondants à l'échelle d'un livre et auquel on peut appliquer un taux de compression élevé en utilisant l'appariement des formes. Il faut rappeler ici que le texte est compris comme un ensemble de caractères indépendants les uns des autres qu'il s'agit de coder avec le moins d'octets possible. Le poids des fichiers contenant le texte peut ainsi être fortement réduit en utilisant la redondance des caractères. Détaillons cette méthode de compression : des prototypes de caractères sont enregistrés dans un dictionnaire de formes redondantes. Chaque prototype est relié à l'adresse des formes qu'il remplace dans l'ouvrage.
- **le plan graphique** (lettrines, bandeaux, gravures etc.) dont les éléments sont faiblement redondants et pour lequel on adopte un taux de compression sans perte (codage CCITT-G4).
- **l'arrière-plan** qui est l'image du support papier auquel on a ôté les caractères et les éléments graphiques. On peut lui appliquer une forte compression avec un format de type JPEG entraînant une perte d'information.

Enfin, pour être précis, il convient de mentionner l'existence d'un quatrième plan dans l'image, le plan compensatoire. Il contient les différences entre l'image décompressée (c'est-à-dire la reconstitution du texte grâce aux prototypes des caractères d'imprimerie) et l'image d'origine (la page du livre telle qu'elle fut imprimée) et permet de reconstituer l'image originale exacte. Ce plan est compressé avec un codeur du type CCITT-G4 qui n'entraîne aucune perte d'information. Ces fichiers compensatoires permettent notamment de faire apparaître à l'écran les caractères d'imprimerie tels qu'ils se présentent sur la page du livre, avec leurs défauts (cassures, manques), ce qui était gommé dans un premier temps dans le plan textuel qui utilisait un prototype unique de chaque caractère pour l'ensemble du livre.

Le résultat de ce travail sur les formats de compression des images aboutit à des réductions importantes de la taille des fichiers : pour les quelques livres dont les fichiers images ont été compressés par cette méthode, le ratio entre la taille des images non compressées et celle des images compressées est de 50 pour 1 en moyenne.

En outre, la méthode de compression mise en œuvre pour le plan textuel et fondée sur l'appariement des formes contient une potentialité intéressante : la transcription du texte assistée par ordinateur.

En effet, aujourd'hui la transcription d'un texte peut se faire de deux façons distinctes : soit le texte imprimé est reconnu par un logiciel de Reconnaissance Optique de Caractères (OCR) qui le transforme automatiquement en un texte au format Word, soit il faut le ressaisir manuellement dans un traitement de texte. C'est cette dernière solution qui est privilégiée dans de nombreuses bibliothèques du web (cf. 1.2.1 de la partie I) en raison du manque de performance des OCR actuels pour les livres anciens. En outre, étant donné le marché économiquement limité de ce genre de numérisation, aucune firme ne semble prête à développer des logiciels d'OCR capables de reconnaître et de transcrire automatiquement des textes dont la typographie (lettres cassées, encrage plus ou moins régulier) et la langue (multiples orthographes pour un même mot) étaient peu normalisées.

La transcription des textes fournie par le laboratoire RFV de l'INSA, pour répondre aux besoins des chercheurs, a été réalisée selon le principe suivant. La compression du plan textuel étant fondée sur la redondance des caractères, on pouvait utiliser le dictionnaire des formes rencontrées dans le livre pour faire reconstituer une grande partie du texte automatiquement et intervenir manuellement pour corriger seulement les formes non reconnues à cause, notamment, des défauts de la casse. Il s'agit donc en fait d'une transcription semi-automatique du texte.

Celle-ci permet un gain de temps important par rapport à une transcription exclusivement manuelle. À titre d'exemple, la saisie de 5% des caractères d'un livre de 200 pages contenant environ 2000 caractères par page permet sa transcription en 6 heures. Après cette reconnaissance automatique des caractères fortement redondants, le texte transcrit n'est certes pas exempt d'erreurs mais la correction est beaucoup plus rapide que la saisie intégrale du texte dans le format Word. En effet, les nombreuses formes de caractères non redondants (qu'un opérateur doit saisir manuellement) ne concernent qu'une faible proportion du texte.

La transcription permet en outre la recherche plein texte, outil très prisé des linguistes notamment. Celle-ci se fait dans la transcription puis le mot est surligné dans l'image numérisée de la page originale du livre grâce à un lien entre les coordonnées du mot dans l'image et le mot présent dans la transcription.

1.2.2. Le format de structuration des données développé dans le cadre de DEBORA : AdHoc

Les bibliothèques numériques actuelles sont structurées selon divers formats, notamment PDF (Gallica) ou HTML (pour l'ABU par exemple). Ces formats ne permettent cependant pas de faire de recherches sur des éléments particuliers des textes, sauf si ceux-ci se distinguent par une mise en forme spécifique, ce que gère par exemple HTML. Il est par exemple impossible dans ces bibliothèques numériques de faire apparaître uniquement les illustrations d'un ouvrage, voire de plusieurs ouvrages présents dans la base.

Pour décrire un document complexe tel que le livre du XVI^e siècle composé de données hétérogènes (textes, images, illustrations, commentaires), et pour permettre à l'utilisateur de DEBORA de retrouver rapidement une information particulière, la nécessité de structurer les informations contenues dans le livre s'est assez vite imposée.

Dans cette optique, le laboratoire RFV de l'INSA a développé un format spécifique permettant cette structuration. Il est issu à la fois de XML pour la hiérarchisation des données et de TIFF-FX pour la gestion des images compressées.

La logique du format XML, c'est-à-dire la description des informations selon leurs fonctions dans le document et non par leur seule mise en forme, ainsi que sa syntaxe formée d'un système de balises ou *tags* signalant les composantes de la structure, ont été adoptées et sont le canevas du format développé, baptisé AdHoc (*Auto-documented⁴⁵ and hierarchically organized contents format*).

La structure hiérarchisée du format, tout comme l'individualisation des données composant le livre, la page, etc., permettent d'accéder directement à une information particulière : le format sépare en fait les divers éléments constitutifs du document et chacun d'eux est accessible indépendamment par l'utilisateur.

Pour expliciter les avantages de ce format, signalons quelques-unes des fonctionnalités développées par l'INSA dans la version 10 du logiciel de DEBORA :

⁴⁵ Un format auto-documentant assure la gestion de données hétérogènes (son, vidéo) sans aucune limitation sur leur type (cf. DEBORA, Délivrable 4.5, p. 9).

- Vignettage : possibilité de faire apparaître rapidement toutes les illustrations du livre (gravures, bandeaux, lettrines) grâce à leur identification par la balise adéquate dans l'arborescence décrivant le livre. Cette fonctionnalité avait été souhaitée par les usagers seiziémistes car elle permet une navigation plus aisée à l'intérieur même de l'ouvrage. Ces vignettes ont été précalculées ce qui permet un gain de temps lorsque l'utilisateur demande leur affichage.
- Création de documents personnels regroupant des éléments issus d'un ou de plusieurs livres.
- Enrichissement des documents par l'ajout de commentaires (annotations) liés à la partie du livre à laquelle ils se réfèrent. Possibilité de partager ces commentaires avec son groupe de recherche et gestion de la propriété des annotations⁴⁶.

Pour l'instant, aucune description normalisée de la structure des livres du XVI^e siècle, ou plus généralement des livres anciens, n'a été officiellement établie. Un tel canevas permettant de décrire exhaustivement tout livre ancien serait aisément transposable dans une arborescence de type XML : il s'agirait alors de ce qu'on appelle une DTD (Document Type Definition ou Description du Type de Document). Une tentative similaire a été menée dans le cadre du projet européen Master⁴⁷. Il s'agissait dans ce cas précis de définir la structure de description des manuscrits. Une DTD adaptée au livre du XVI^e siècle permettrait notamment d'uniformiser la description des ouvrages et de les intégrer facilement dans une bibliothèque numérique utilisant le logiciel développé dans le cadre du projet DEBORA.

⁴⁶ Cf. 1.3 dans la troisième partie.

⁴⁷ BURNARD (L.), ROBINSON (P.), *Vers un standard européen de description des manuscrits : le projet Master*, Document numérique, vol. 3, n°1-2/1999, p. 151-169.

1.3. Les principales fonctionnalités d'exploitation des documents

Le navigateur offre au chercheur seizième siècle des possibilités tout à fait nouvelles de travail sur les ouvrages imprimés. En apprécier toute la valeur suppose bien évidemment de faire abstraction de l'état actuel du logiciel et du corpus proposé et d'imaginer leur maturité idéale, à savoir un corpus important d'ouvrages numérisés et transcrits ainsi que la connexion au réseau internet (ce qui rendrait possible le travail collaboratif à distance).

Les principales fonctionnalités du navigateur peuvent se regrouper en trois grands ensembles : les outils de visualisation, les outils de recherche et les outils d'exploitation personnelle des documents. Nous décrivons ici ces instruments en expliquant ce qui fait leur intérêt tout en précisant, si besoin est, leurs faiblesses. Pour une étude appliquée de ces dispositifs à travers l'exemple de *La magnificence*, on se reportera au point 3.3 de ce mémoire.

1.3.1. Les outils de visualisation

La première étape de toute session sur DEBORA est le chargement dans le navigateur des ouvrages sur lesquels on souhaite travailler. Ceux-ci se présentent sous la forme d'une arborescence dans la partie gauche du navigateur : pièces liminaires, pages de texte, elles-mêmes se subdivisant encore en « imagerie » lorsqu'elles comportent une illustration, quelle qu'elle soit.

Avantage précieux pour toute étude comparative, le navigateur permet l'affichage de plusieurs pages à la fois, celles-ci pouvant provenir de plusieurs ouvrages différents. Ces pages peuvent être obtenues soit par l'arborescence à gauche de l'écran⁴⁸ soit par l'intermédiaire de la fonction « vignettage » qui autorise un survol de l'ouvrage en affichant toutes les pages sous forme de vignettes dans une seule fenêtre. Il est donc aisé de prendre connaissance de l'ouvrage dans son ensemble, en repérant par exemple les différents éléments d'illustration (lettrines, bandeaux, bois gravés). Pour obtenir la page désirée, il suffit de cliquer sur la vignette.

⁴⁸ Cette zone est le résultat concret de l'adoption et de l'adaptation du format XML exposé dans la partie précédente. On y distingue en outre la séparation manifeste du texte de l'image.

La structure arborescente, en revanche, ne permet pas de distinguer au préalable de quel type d'illustration est composée la page. Une simple icône (un i encadré et bleuté) signale la présence d'un élément graphique, lequel est représenté dans l'arborescence par le terme « imagerie »⁴⁹ qui n'est pas assez précis.

Grâce à la séparation numérique du texte et de l'illustration, le chercheur peut afficher toute la page, n'afficher que le texte ou n'afficher que l'illustration. Il peut ainsi isoler l'objet de sa recherche. S'il décide de travailler sur toute la page ou sur le texte, il est en mesure de conserver ou d'ôter le fond grisé afin de faciliter sa lecture. Il peut également augmenter la luminosité du document. Le menu « afficher dans une nouvelle fenêtre » permet de disposer au mieux les différentes pages ouvertes à l'écran et de procéder au travail comparatif.

Pour le travail direct sur la page, plusieurs fonctionnalités sont disponibles : une loupe et un instrument de mesure de l'image⁵⁰. La loupe classique permet un agrandissement de toute la page et une loupe « baladeuse » sert à pointer sur les détails de la typographie comme des illustrations. Le grossissement obtenu est très appréciable.

1.3.2. Les outils de recherche

Sur les deux types de recherche possibles, une seule, celle sur les mots du texte, met en œuvre un « moteur » de recherche. Celle portant sur les décorations de l'ouvrage demeure purement manuelle : il revient à l'utilisateur de trouver lui-même la décoration qu'il recherche par l'intermédiaire d'un « vignetage » semblable au précédent, à ceci près qu'il rassemble uniquement les ornements typographiques et les gravures et ce, sans notion d'échelle. Néanmoins, cette fonction est très utile pour isoler toutes les décorations du texte et par exemple se faire une idée du matériel ornemental employé par l'imprimeur ainsi que des gravures (leur sujet, leur manière) présentes dans l'ouvrage. Un simple « clic » sur la vignette permet ensuite d'accéder à la page où se trouve la décoration recherchée.

⁴⁹ Cf. 1.2 de la deuxième partie sur la séparation du texte et de l'image.

⁵⁰ Cet outil est un système de coordonnées de points (X, Y) ; le curseur sert de repère. Son utilisation n'est pas très aisée.

La recherche par mots du texte est réalisable grâce à la transcription de l'ouvrage, elle est donc actuellement tributaire de l'incomplétude de celle-ci. Elle peut se faire sur un ou plusieurs ouvrages et il est possible, soit de la limiter à certains champs comme le titre ou l'auteur, soit de la faire porter sur tout le contenu du ou des textes. En outre, un certain nombre de champs, inutilisés en l'état actuel, sont d'un grand intérêt pour le chercheur : lieu de publication, langue, collation etc.

Ce genre d'outil est tout à fait approprié aux études lexicographiques (recherche des néologismes, des archaïsmes ou des régionalismes), aux études historiques (dans le cas précis d'une entrée royale, recherche de noms de personnes, recherche dans le lexique des vêtements, des décors ou des armements), ou encore aux travaux sur la typographie du XVI^e siècle (recherche par caractère).

Cette recherche s'effectue par l'intermédiaire d'une « boîte de dialogue » qui permet de spécifier le champ sur lequel porte l'interrogation. Les réponses s'affichent dans cette même « boîte » sous forme de liens hypertextuels menant aux pages de l'original où se trouve le terme recherché. Celui-ci y apparaît alors encadré de rouge.

1.3.3. Les outils d'exploitation personnelle

En travaillant sur DEBORA, le chercheur peut créer son dossier personnel et y intégrer par exemple toutes les letrines qui l'intéressent, qu'elles proviennent d'un ou de plusieurs imprimés ; en outre, il a la possibilité d'apporter directement ses propres commentaires à l'ouvrage numérisé, lesquels pourront être lus par un nombre plus ou moins grand de chercheurs en fonction des spécifications de l'annotateur.

Le grand avantage du dossier personnel est qu'il permet la constitution d'une base complémentaire indépendante contenant tout ce dont le chercheur a besoin : il est donc utilisable immédiatement (en local). L'utilisateur peut y déposer, grâce au copier/coller, des pages entières ou seulement les letrines et/ou les bois gravés qu'il désire plus particulièrement observer. Il a ensuite tout loisir de consulter son dossier et de le compléter.

Le chercheur peut matérialiser sur le document numérisé ses conclusions ou les remarques qu'il souhaite faire à propos de tel terme, tel ornement ou caractère typographique, afin de les rendre accessibles à ses collègues. La procédure est très simple, elle consiste à sélectionner sur la page la zone à commenter, puis à écrire les commentaires dans la « boîte de dialogue » qui apparaît, et enfin à décider qui sera en mesure de lire ces notes : le groupe d'appartenance du chercheur ou tous les usagers du navigateur sans distinction. Le chercheur qui souhaite lire ces remarques peut visualiser tout aussi facilement les pages de l'ouvrage comportant un commentaire grâce à une icône spéciale. Ensuite, il suffit d'ouvrir la fenêtre « commentaire » et de passer le curseur sur la page choisie, la zone annotée s'anime et le commentaire lié apparaît dans la dite fenêtre, selon un procédé comparable aux notes de bas de page dans les dernières versions du logiciel Microsoft Word.

En matière d'accès aux documents anciens, le navigateur DEBORA se distingue par son approche résolument nouvelle du travail de recherche. Alors que les bibliothèques numériques disponibles sur le web n'offrent souvent que de simples « fac-similés », DEBORA envisage la mise en place d'un travail collaboratif à distance, auquel s'ajoutent des outils tendant à apporter une plus-value non négligeable aux recherches actuellement menée, comme le montre la partie suivante.

Partie III - LA MAGNIFICENCE ET L'HISTORIEN DU LIVRE : RECHERCHES ACTUELLES ET PERSPECTIVES

Afin d'évaluer correctement les outils créés dans le cadre des fonctionnalités avancées de DEBORA, il nous a paru utile de faire le point sur la nature des recherches portant sur les livres du XVI^e siècle, c'est-à-dire de comprendre les objectifs et la méthode de lecteurs savants travaillant pour la plupart avec une documentation imprimée, qu'il s'agisse d'éditions originales, d'éditions scientifiques ou de fac-similés. Pour ce faire, notre enquête a porté sur le livret des fêtes de l'entrée de Henri II à Lyon en septembre 1548, à l'heure actuelle seul ouvrage numérisé de la base DEBORA qui soit enrichi des métadonnées et de la transcription. Dans la mesure où la condition sine qua non de toute recherche est de disposer de sources, il s'agit en premier lieu de se faire une idée de l'accessibilité au livre *La magnificence de la superbe et triomphante entree*.

1.1. L'accès au document

1.1.1. Les imprimés

L'entrée de Henri II à Lyon a donné lieu à la publication de plusieurs livrets en 1548 et 1549, sans nom d'auteur ni d'imprimeur, et pour certains sans l'adresse de l'éditeur (du « libraire », comme on disait alors), ce qui était illégal⁵¹. Le 25 janvier 1549, le libraire lyonnais Guillaume Rouillé obtint un privilège des autorités municipales et publia le plus considérable de ces livrets, en français puis en italien (seule l'édition française a fait l'objet d'une numérisation dans la base de DEBORA). D'emblée, il convient de noter le faible nombre d'exemplaires qui nous sont parvenus. Il y a près d'un siècle, Baudrier en recensait sept en français et quatorze en italien, et quoique d'après Cooper

⁵¹ COOPER (R.), *Maurice Scève, the Entry of Henri II into Lyon (september 1548) : a Facsimilé with an Introduction*, Tempe, Arizona Board of Regents for Arizona State University, 1997, 326 p. (« Medieval and Renaissance Texts and Studies », 160), p. 128-132.

il s'en trouve d'autres en Italie, en Grande-Bretagne et aux États-Unis⁵², il est évident que l'accès au document original reste réservé à un petit nombre de lecteurs, qui appartiennent au public captif des chercheurs admis dans les fonds anciens des bibliothèques.

Cependant, le récit de l'événement avait connu un certain nombre de rééditions. En 1573, Guillaume Paradin avait retranscrit le texte de l'édition de Rouillé dans ses *Mémoires de l'histoire de Lyon*⁵³, volume qui n'est point rare et qui de plus a fait l'objet d'une réimpression anastatique en 1985. Un autre contemporain, Brantôme, avait copié de larges extraits du livret, puis au XVII^e siècle Godefroy l'a republié dans son *Cérémonial français*, copie qui fut à son tour résumée en 1730 par le P. de Colonia dans l'*Histoire littéraire de la ville de Lyon*, dont il existe un reprint. Un peu plus tard, en 1752, la municipalité de Lyon fit paraître un volume de relations des entrées royales dans lequel le texte de 1549 était reproduit sous une forme abrégée⁵⁴.

Le texte de la relation de l'entrée de Henri II à Lyon n'a donc jamais été, à proprement parler, inaccessible. En revanche, les éléments graphiques qui agrémentent la lecture du livret, illustrations et ornements typographiques, ont été reproduits pour la première fois en 1927, dans l'édition de l'ouvrage faite par Georges Guigue, archiviste du département du Rhône⁵⁵. Selon les usages savants, cette édition combinait la reproduction à l'identique, par un procédé photomécanique, des illustrations, lettrines et fleurons, avec la transcription du texte. Elle était accompagnée d'un appareil critique, fait de l'établissement de variantes et de la publication d'un vaste ensemble de documents d'archives contemporains de l'événement. Ce travail présentait un intérêt évident, dans la mesure où il était censé mettre à la disposition d'un large public des matériaux pour l'histoire culturelle de la Renaissance, sous le rapport non seulement de l'histoire des fêtes royales et de la littérature, mais aussi de l'histoire du livre et de la gravure. Pourtant, l'édition de Guigue, imprimée par la Société des bibliophiles lyonnais, eut un faible tirage et de plus elle ne fut pas vraiment diffusée. De nombreux exemplaires sont en effet

⁵² BAUDRIER (H.), BAUDRIER (J.), *Bibliographie lyonnaise : recherches sur les imprimeurs, libraires, relieurs et fondeurs de lettres de Lyon au XVI^e siècle*, Lyon, L. Brun / Paris, A. Picard, t. IX, 1912, p. 164-165 ; Cooper 1997, p. 142.

⁵³ PARADIN (Guillaume), *Mémoires de l'histoire de Lyon*, Lyon, A. Gryphe, 1573, [16]-444-[14] p. 320-351.

⁵⁴ Cooper 1997, p. 142.

⁵⁵ [SCÈVE (Maurice)], *La magnificence de la superbe et triomphante entree de la noble et antique cité de Lyon faicte au tres chrestien roy de France Henry deuxiesme de ce nom, et a la royne Catherine son espouse, le XXIII de septembre MDXLVIII : relations et documents contemporains*, éd. Georges Guigue, Lyon, Société des bibliophiles lyonnais, 1927, 373 p.

actuellement stockés aux Archives municipales de Lyon, dans le fonds déposé par la Société.

Malgré tout, l'édition de 1927 a contribué à faire connaître le livret de Guillaume Rouillé. Sur la base des documents d'archives publiés par Guigue, Verdun-Louis Saulnier rendit à Maurice Scève la paternité de ce texte anonyme, sans lui accorder cependant une grande valeur littéraire dans sa thèse sur le « prince de la Renaissance lyonnaise »⁵⁶. Ce n'est que vers la fin du XX^e siècle, avec le renouveau de l'histoire culturelle et l'intérêt particulier dans les années 1970 pour les entrées triomphales⁵⁷, que l'on mesura l'importance de cette œuvre du XVI^e siècle lyonnais. Il restait à en faire l'étude. Plusieurs chercheurs s'y attachèrent, parmi lesquels Anne-Marie Lecoq, dont l'intérêt pour les fêtes de la Renaissance avait déjà donné lieu à plusieurs travaux⁵⁸.

L'ouvrage monographique d'A.-M. Lecoq sur l'entrée de Henri II, que l'on attendait au début des années 1990, n'a pas encore vu le jour et c'est à Richard Cooper, professeur de littérature française à Oxford, que l'on doit la première étude circonstanciée du livret de Maurice Scève⁵⁹. L'œuvre est désormais replacée dans son contexte, les sources littéraires et la documentation matérielle de l'auteur et de l'illustrateur ont été identifiées pour une bonne part, la validité du récit en tant que source est discutée à la lumière des autres relations contemporaines de l'événement. Après un long commentaire, la seconde partie du volume de R. Cooper met en effet à la disposition du lecteur une documentation de première main, à savoir plusieurs relations de l'événement, soit sous la forme de fac-similé (trois imprimés de 1548 et 1549, dont le texte de Scève en français), soit sous la forme de transcription (textes manuscrits conservés dans des dépôts d'archives italiens et sources imprimées).

Ainsi, au moment de la mise en route du projet DEBORA (1999), le dossier de l'entrée de Henri II à Lyon en septembre 1548 était particulièrement

⁵⁶ SAULNIER (V.-L.), *Le prince de la Renaissance lyonnaise, initiateur de la Pléiade, Maurice Scève, italianisant, humaniste et poète (ca. 1500-1560) : les milieux, la carrière, la destinée*, Paris, C. Klincksieck, 1948-1949, t.I, p. 360-361 ; Cooper 1997, p. 142. Guillaume du Choul et Barthélemy Aneau avaient été pressentis pour élaborer le programme des fêtes, mais il ressort des documents publiés par Guigue que seul Maurice Scève en fut l'organisateur : Scève 1927, p. 149 ; Cooper 1997, p. 130-131.

⁵⁷ Par exemple, V. E. Graham, « L'humanisme lyonnais à travers les entrées triomphales au XVI^e siècle », dans *L'humanisme lyonnais au XVI^e siècle*, Grenoble, Presses universitaires de Grenoble, 1974, p. 185-199.

⁵⁸ Entre autres, A.-M. Lecoq, « La "Città festeggiante" : les fêtes publiques aux XV^e et XVI^e siècle », dans *Revue de l'art*, 33, 1976, p. 83-100.

⁵⁹ Cooper 1997.

consistant. Les lecteurs disposaient d'un certain nombre d'éditions du livret de Maurice Scève consultables en bibliothèque, mais dans quelques bibliothèques seulement en raison de la rareté de ces imprimés, qui ne sont d'ailleurs pas référencés de manière satisfaisante. Le Catalogue collectif de France (CCFR), interrogé en mai 2002 sur la localisation d'exemplaires de ces différentes éditions dans les bibliothèques publiques françaises donne des réponses déconcertantes à cause du manque d'uniformisation des notices et quelques exemplaires seulement sont signalés.

1.1.2. Les numérisations

Le livret de l'entrée de Henri II est disponible dans Gallica, bibliothèque numérique de la BNF accessible depuis 1997⁶⁰. Nous le retrouvons aisément par la page « recherche » du site en formulant la requête « *Magnificence* » dans le champ « Mots du titre » : une seule réponse est donnée, qui est pertinente. La numérisation est en mode image d'après l'exemplaire de la bibliothèque municipale de Lyon (BML), coté Rés. 355882, et d'une très haute qualité : même les pages les plus mal encrées de l'original sont bien lisibles, ce qui n'est pas toujours le cas de la réimpression anastatique de R. Cooper, pour laquelle il a fallu repasser certaines lettres au stylo⁶¹. Gallica offre aussi la version italienne du livret de 1549, *La magnifica*, avec la même qualité de reproduction, d'après l'exemplaire BML Rés. 355883. Ces numérisations signalées comme « Num. BNF de l'éd. de A Lyon : chés Guillaume Rouille, 1549 » se distinguent dans Gallica de celles qui ont été faites à partir de microfilms trop contrastés, curieusement référencées « Num. BNF de l'éd. de Cambridge (Mass.) : Omnisys, [ca. 1990] ».

Pour consulter ces ouvrages numérisés, le lecteur a la possibilité de les lire dans Gallica, ou de les télécharger. La lecture dans Gallica est la moins confortable, parce que la plus lente, dans la mesure où il faut lancer l'une après l'autre l'ouverture de chaque page du document, qui correspond à une page web. De plus, la requête se fait par l'intermédiaire d'une « pagination » qui n'est qu'une liste séquentielle ne donnant aucune information sur le contenu des pages. Dans quelques cas, une « table des matières » a été

⁶⁰ <http://gallica.bnf.fr/scripts/ConsultationTout.exe?E=0&O=N076015>. Les champs permettant les requêtes dans Gallica sont « Mots du titre », « Auteur », « Sujet » et « Recherche libre ».

⁶¹ Cooper 1997, f. Kv, avant-dernière ligne.

constituée, avec des liens hypertextuels vers les têtes de chapitre. Le téléchargement de l'ouvrage, quant à lui, est parfois extrêmement long, selon l'encombrement du web et la capacité du poste de consultation, de sorte qu'il est pratiquement nécessaire de disposer d'une connexion à haut débit. C'est dire qu'il est difficilement envisageable pour l'instant de constituer chez soi sa propre bibliothèque numérique, mais ce problème tendra à disparaître avec le développement de l'ADSL. Le téléchargement se fait selon deux formats, soit en PDF, où le texte s'affiche en continu, soit en TIFF sous Imaging®, où le texte s'affiche page par page mais accompagné d'un vignettage qui permet de se déplacer commodément, grâce à un aperçu du contenu de chaque page. Enregistré sur le disque dur, le fichier de l'ouvrage se consulte rapidement. Ces différents modes permettent en outre au lecteur d'imprimer le document sur papier et donc de créer son propre fac-similé, mais seul Imaging® donne la possibilité de travailler à partir du document numérisé (par exemple en créant des dossiers d'illustrations faits par copier-coller).

Le projet DEBORA enfin, dont on regrette qu'il n'ait abouti qu'à la réalisation d'un prototype peu connu du public, donne du livret de l'entrée de Henri II une version numérisée en mode image, là aussi à partir de l'exemplaire BML Rés. 355882, consultable page par page grâce à l'arborescence qui figure à gauche de l'écran. Le document, dont la transcription – inachevée – existe, et qui est enrichi de métadonnées, se prête à des recherches que nous détaillerons plus bas.

1.2. Les usages actuels des chercheurs

Jusqu'en 1927, on disposait donc seulement des fonds patrimoniaux des bibliothèques pour se documenter sur l'entrée de Henri II ; de 1927 à 1997, on pouvait de plus accéder indirectement à la source primaire par le biais d'une transcription ; enfin depuis 1997 le livret est accessible à tous les publics par une consultation médiate, soit sous la forme d'une reproduction photomécanique (Cooper 1997) soit par l'intermédiaire d'une numérisation en mode image (Gallica). Sans prétendre dresser une liste exhaustive des usages des lecteurs, on les regroupera en deux catégories : d'une part les modes d'investigation qui nécessitent l'examen du livre original, ou du moins de documents de substitution le plus proche possible des originaux (fac-similés,

microfilms, numérisations à l'échelle et de bonne qualité)⁶², d'autre part ceux qui se fondent sur une lecture seule, par l'intermédiaire de reproductions du contenu (édition du texte, reproduction des illustrations).

La plupart des études récentes sur Lyon à la Renaissance, sur Henri II ou sur les entrées royales, qui mettent en œuvre *La magnificence* appartiennent à la seconde catégorie. A titre d'exemple, nous mentionnerons le travail de deux architectes florentins, Giuseppe Iacono et Salvatore Ennio Furone⁶³, qui nous semble représentatif d'un certain usage des sources, à savoir l'utilisation des données factuelles. Dans un chapitre portant sur l'architecture éphémère des entrées royales à Lyon pendant la première moitié du XVI^e siècle, le récit de l'entrée de Henri II est rapporté sous la forme d'un résumé, tandis que les illustrations sont reproduites pour la plupart (mais elles n'ont qu'une valeur d'illustration, n'ayant pas fait l'objet d'études particulières).

D'autres recherches, plus techniques, portant davantage sur la forme que sur le contenu peuvent être menées à partir de tout type d'édition, et de préférence avec une édition numérisée en mode texte, en raison des facilités offertes par la recherche plein texte. C'est le cas de l'étude lexicologique et métalexicographique du récit de l'entrée de Henri II à Rouen en 1550, par Terence Russon Wooldridge⁶⁴. Pour des raisons matérielles, il ne nous est guère possible d'entreprendre une recherche similaire à partir du récit de l'entrée de Henri II à Lyon en 1548. Pour les mêmes raisons, nous renonçons à développer ici une étude sur les bois gravés du livret, mais il est évident que la numérisation en mode image pourrait rendre de grands services aux historiens de l'art. Nous évoquerons ce sujet plus loin à propos du problème de l'indexation des images⁶⁵.

En revanche, afin de mettre à l'épreuve les outils proposés par DEBORA, nous choisissons dans le présent mémoire de nous intéresser tout particulièrement à la bibliographie matérielle, c'est-à-dire de considérer le livre

⁶² Certains éditeurs de fac-similés ont inventé le concept de « quasi-original » : voir par exemple le catalogue de Moleiro editor SA, de Barcelone (www.moleiro.com). On s'interrogera à ce sujet sur la nécessité d'utiliser un « papier spécial fabriqué à la main, de texture, épaisseur et odeur identiques à celles de l'original ».

⁶³ IACONO (G.), FURONE (S. E.), *Les marchands banquiers florentins et l'architecture à Lyon au XVI^e siècle*, Paris, Publisud, 1999, p. 163-195.

⁶⁴ http://www.chass.utoronto.ca/~wulfric/articles/tdh/tdh_bib.htm. L'intérêt de la numérisation en mode texte pour les chercheurs en sciences du langage a été exposé au point 1.2.1 de la partie I.

⁶⁵ Cf. 1.4. de la troisième partie. Sur la difficulté d'accéder directement aux gravures des ouvrages numérisés, voir le point 1.2.2 de la partie I, avec l'exception de la *Nef des folles* de Josse Bade.

en tant qu'objet, du point de vue de sa fabrication. Notre objectif sera donc d'analyser les moyens qui permettent de répondre à des questions telles que : qui était l'imprimeur de *La magnificence* ? De quel matériel typographique disposait-il ? Y avait-il des pratiques d'atelier ? L'édition a-t-elle connu des variantes (états, émissions) ? Pour mener une telle enquête, il est actuellement nécessaire d'examiner les originaux eux-mêmes, mais nous envisagerons la possibilité d'une bibliographie matérielle virtuelle.

1.2.1. Bibliographie matérielle « livre en main »

L'imprimeur du livret de 1549, dont Guillaume Rouillé ne fut que l'éditeur, n'est pas connu à ce jour. Baudrier, en considérant les lettrines I, N et P de l'édition italienne, avait fait un rapprochement avec celles qu'utilisait à la même époque Jean de Tournes⁶⁶. Cooper, tout en admettant ce point de vue, a rappelé que ni le livret, ni les lettrines ne figuraient dans la bibliographie des éditions de Jean de Tournes par Cartier⁶⁷. L'hypothèse de Baudrier, fondée sur une similitude de style assez courante dans la production contemporaine, reste à vérifier par des identifications.

En rapprochant plusieurs reproductions photographiques de lettrines lyonnaises des années 1540, nous sommes en mesure d'en dire plus sur l'un des jeux d'alphabets utilisés par l'imprimeur de *La magnificence*, celui dont le décor consiste en rinceaux asymétriques au feuillage stylisé, sur fond criblé. La lettrine L, imprimée aux feuillets (f.) F3v, H2v et Iv de l'édition française, a été obtenue avec un bois utilisé en 1543 par Etienne Dolet (1509-1546) pour les *Œuvres* de Clément Marot, dont on trouvera une reproduction photographique dans une étude récente de Henri-Jean Martin⁶⁸. Comme on en reconnaît bien les légères imperfections de gravure, l'identification du bois ne fait pas de doute. Voilà donc une première orientation pour notre recherche : établir quel imprimeur lyonnais avait acquis le matériel typographique de Dolet, emprisonné à Lyon le 6 janvier 1544, fugitif le surlendemain, repris et finalement exécuté place Maubert à Paris le 3 août 1546⁶⁹. La *Bibliographie des œuvres d'Etienne Dolet* montre que notre première piste est une impasse, dans

⁶⁶ Baudrier, t. IX, p. 163.

⁶⁷ Cooper 1997, p. 135.

⁶⁸ MARTIN (Henri-Jean), *Mise en page et mise en texte du livre français : la naissance du livre moderne (XIV^e-XVII^e siècle)*, [Paris], Ed. du Cercle de la librairie, 2000, fig. 367, p. 228.

⁶⁹ Dolet disposait de son propre matériel depuis 1539 : *ibid.*, p. 227.

l'état actuel de la recherche. En effet, si le matériel typographique de Dolet est étudié dans l'ouvrage de Claude Longeon, il n'y est pas fait mention de remplois par un autre imprimeur après 1543⁷⁰.

Cette identification, il nous a été possible de l'établir parce que nous avons pu ouvrir sur la même table deux volumes, celui de Henri-Jean Martin et le fac-similé de Richard Cooper. La rencontre n'est pas tout à fait fortuite : à l'origine du processus d'identification, il y avait notre mémoire visuelle, avec tout ce qu'elle comporte de faillibilité. La vérification elle-même était laborieuse, car il est apparu que les reproductions du fac-similé de Cooper avaient été « nettoyées » avec du correcteur, afin de rendre plus lisibles les lettrines, qui ont de fait été défigurées⁷¹. L'édition de Guigue est venue à notre secours avec ses reproductions photographiques des lettrines, de meilleure qualité. Encore faut-il préciser que Guigue a donné un seul cliché par lettrine, même lorsqu'elles étaient répétées, de sorte que l'on ne sait pas à quelle page appartient celle qui est photographiée⁷².

Cet exemple servira à montrer les avantages et les limites des réimpressions anastatiques dans le cadre d'une étude des formes. Le rapprochement, qui a priori ne s'imposait pas (on peut difficilement songer à comparer le livret de l'entrée de Henri II à toute la production imprimée lyonnaise du milieu du XVI^e siècle), dépendait de l'accessibilité aux documents. Les fac-similés ont rendu possible ce rapprochement. Cependant, la qualité de reproduction est dans un cas tellement mauvaise que l'on ne pouvait rien conclure. Le fac-similé de *La magnificence* est issu en effet d'un microfilm dont il faut bien relever la mauvaise qualité, même en admettant que l'exemplaire du livret utilisé par R. Cooper ait été particulièrement mal encré⁷³. D'autres exemplaires montrent bien qu'il y a eu un problème d'encrage pour l'édition de 1549, mais l'étude matérielle du document pourra être conduite plus aisément à partir de la numérisation dans Gallica, dont nous avons déjà souligné la bonne lisibilité. En raison de la perte de qualité de l'image due à la reproduction photomécanique, le fac-similé de livre ancien est donc un ersatz qui ne dispense pas de la consultation immédiate de l'original, pour vérification et dans la perspective de faire une étude plus fine des formes. Néanmoins,

⁷⁰ LONGEON (Claude), *Bibliographie des œuvres d'Etienne Dolet écrivain, éditeur et imprimeur*, Genève, Droz, 1980, p. LV-LXII, pour les lettres ornées.

⁷¹ Voir par exemple le grossier P du f. G3v où disparaissent les graisses. Le S du f. L3 est absolument illisible.

⁷² Ainsi, le L identifié ci-dessus est aux p. 40, 54 et 60 de son édition.

⁷³ Oxford, Bodleian Library, Douce L 216.

nous avons imprimé une copie de *La magnificence* d'après Gallica, pour avoir sous la main une bonne reproduction du livret.

Pour tenter de répondre à la question qui nous occupe, il convenait d'entreprendre une étude systématique sur le matériel typographique mis en œuvre, à partir des différents supports à notre disposition. Par commodité, nous avons privilégié l'examen des lettrines, a priori plus facile à caractériser que des fleurons ou des caractères. Grâce à la numérisation dans Gallica, nous avons dans un premier temps identifié neuf alphabets dans les éditions française et italienne, sept sur fonds criblés et deux sur fonds rayés. Mais ce résultat n'était pas suffisant puisque les mesures des lettrines (information essentielle pour la suite de la recherche) n'étaient pas connues. Il a donc fallu consulter les exemplaires originaux à la bibliothèque municipale de Lyon pour prendre ces mesures. Ensuite, c'est la comparaison avec les illustrations du livre de Longeon, déjà cité, qui nous a permis d'établir que les alphabets numérotés 1 et 6 dans le tableau ci-dessous avaient été en usage dans l'imprimerie d'Etienne Dolet entre 1539 et 1543⁷⁴.

Tableau 1. — Lettrines des livrets de 1549

N°	mm	Fonds criblés	<i>La magnificence</i>	<i>La magnifica</i>
1	54 x 54	rincaux et figures	L	
2	47 x 47	rincaux	P	I, N, P
3	42 x 42	rincaux	M	
4	42 x 42	rincaux et figures	A	A
5	40 x 40	rincaux et figures	S	S
6	33 x 33	rincaux	A, C, L, S	A, C, D, I, P, S, V
7		rincaux	S	
N°	mm	Fonds rayés	<i>La magnificence</i>	<i>La magnifica</i>
8	24 x 24	rincaux et putti		F, H
9	16 x 16	rincaux et putti	P	

Cet inventaire étant établi, la plus grande partie de l'étude devait être comparative. L'objectif était en effet de tenter d'identifier les lettrines de *La magnificence* avec celles des livres de Guillaume Rouillé que les imprimeurs avaient signés par la formule *excudebat N.* ou *excudebant N. et N.* La *Bibliographie* de Baudrier fait connaître plusieurs imprimeurs lyonnais

⁷⁴ Longeon 1980, respectivement n° A-1, illustré p. LIX, et n° A-2, illustré p. LX.

travaillant pour le compte de Rouillé de 1547 à 1551, période à laquelle nous limitons notre investigation en raison de l'abondante production de cet éditeur : Etienne Roussin, Jean Ausoult, Philibert Rollet, Barthélemy Frein, Macé Bonhomme, Jean Pidié⁷⁵. Nous avons, dans un premier temps, travaillé au fonds ancien de la bibliothèque de Lyon, où le fichier « Imprimeurs lyonnais », qui comprend les libraires (au sens ancien d'éditeur), interrogé sur Rouillé, indique les combinaisons suivantes :

Tableau 2. — Rouillé et ses imprimeurs et associé de 1547 à 1551

	Rouillé	Rouillé & Constantin	Rouillé & Payen (1548)	Payen
Roussin & Ausoult	1547			
Rollet & Frein	1547?-1550	1548	?	1548
Rollet	1550-1551			
Bonhomme	1548-1549			
Pidié	1548			

D'après ce tableau, trois imprimeurs auraient travaillé en 1549 pour le compte de Rouillé : Macé Bonhomme et les associés Philibert Rollet et Barthélemy Frein. Nous avons examiné en priorité les productions de ces deux derniers, qui étaient les principaux imprimeurs de Rouillé à cette époque. La comparaison entre leurs livres et les éditions de l'entrée de Henri II permet d'établir qu'ils avaient utilisé deux alphabets de lettrines décrits ci-dessus (tableau 1). L'un, de 24 mm de côté, sur fond rayé (notre n° 8), est illustré par la lettre H au f. Hv de l'édition italienne du livret et à la p. 274 d'un Aristote imprimé en 1548 par Rollet et Frein pour Guillaume Rouillé et Antoine Constantin⁷⁶. Une deuxième occurrence vient confirmer cette identification, c'est un F (*La magnifica*, f. Dv) qui se retrouve dans un livre de droit de Tiraqueau (1549), non signé par les imprimeurs mais dont plusieurs lettres ornées paraissent dans les volumes de Rollet et Frein (tableau A en annexe). Un deuxième alphabet connu par *La magnificence*, sur fond criblé et de 33 mm de côté (notre n° 6) a servi pour des éditions de Rouillé de 1549, parues sans

⁷⁵ Baudrier, t. IX, p. 124 sqq.

⁷⁶ Aristote 1548, p. [300] : « Lugduni, excudebant Philibertus Rolletus et Bartholomaeus Fraenus, impensis honestissimorum virorum Gulielmi Rouillii et Antonii Constantini ».

nom d'imprimeur. Il s'agit du Tiraqueau et d'un livre de médecine de Tagaut, qui présentent des lettrines communes (24 x 24 mm) avec celles de l'Aristote et d'un Galien de 1548, tous deux signés par Rollet et Frein (tableau B en annexe). Ce deuxième alphabet est l'un de ceux qui proviennent de l'imprimerie de Dolet⁷⁷.

Par ces identifications, nous pouvons établir que Rollet et Frein ont disposé d'une partie au moins du matériel typographique mis en œuvre pour l'impression du livret de l'entrée de Henri II, y compris les bois des lettrines provenant de Dolet. Il convient alors de s'assurer que ce matériel leur appartenait, qu'il ne leur avait pas été prêté par Rouillé. Nous déduisons de l'examen des *Tabulae* de 1548 imprimées par les mêmes pour un autre éditeur, Thibaud Payen, qu'ils possédaient bien ce matériel (tableau A). De plus, nous croyons reconnaître dans ce dernier livre les caractères romains qui ont servi un an plus tard pour les légendes des illustrations de *La magnificence*, en particulier à travers les a et les e minuscules du corps supérieur. Enfin, nous apprenons par le livre de Bernard de Gordon, de 1550 (tableau A), qu'une partie au moins de ces bois est restée la propriété du seul Philibert Rollet, après la fin de son association avec Barthélemy Frein. Il restera à comprendre comment il était venu en possession des lettrines de Dolet, qui réapparaissent pendant l'année 1549, mais une chose est certaine, *La magnificence* et Tiraqueau sont les premiers livres in-4 imprimés par Rollet et Frein, et les premiers publiés par Rouillé dans ce format, ces trois hommes n'ayant jusque-là produit que des ouvrages de petit format, in-8 et in-16. Voilà, peut-être, ce qui explique les problèmes d'encrage déjà relevés, mais surtout la nécessité pour les imprimeurs de se procurer des lettrines de plus grande dimension que celles qu'ils utilisaient jusque-là.

Nous voilà donc fondés à conclure que non seulement *La magnificence*, mais encore les livres de Tagaut et de Tiraqueau sortaient des presses de Rollet et Frein, qui avaient récupéré une partie au moins des lettrines de Dolet. Ces résultats ont été obtenus par la consultation en bibliothèque de livres du XVI^e siècle, suivant une méthode « traditionnelle », agréable et sûre parce qu'elle permet le contact direct avec les documents originaux, mais qui a ses limites, bien soulignées par les tenants des bibliothèques numériques : le lecteur ne peut pas consulter autant de livres qu'il le souhaiterait et en les manipulant, il risque de les dégrader. Nous allons montrer, en abordant par

⁷⁷ Voir ci-dessus à la note 74.

d'autres moyens le problème de l'attribution à un imprimeur, que les ressources actuelles de l'informatique et d'internet rendent dès maintenant la bibliographie matérielle plus facile.

1.2.2. Bibliographie matérielle virtuelle

Nous tentons ici d'identifier l'officine qui était en charge de l'impression du livret en consultant dans Gallica, tous les livres publiés par Rouillé vers 1549. Le matériel typographique utilisé dans ces ouvrages sera comparé à celui de *La magnificence*.

Dans Gallica, cette recherche est plus malaisée que dans un OPAC du type de celui de la Bibliothèque municipale de Lyon ou dans BN-Opale Plus, dans la mesure où aucun champ « Editeur » n'est prévu dans la page permettant les requêtes. Le fait de devoir saisir la requête dans le champ « Recherche libre » met en évidence les limites d'une recherche plein texte, puisqu'il n'y a pas eu d'indexation du nom de l'éditeur sous la forme d'une vedette normalisée, à laquelle renverraient les formes rejetées (sur le modèle des notices d'autorité de la BNF, qui sont bien utilisées dans Gallica pour la vedette auteur). C'est ici que se fait sentir le problème posé par l'emploi d'une norme de catalogage visant à saisir dans la zone de l'adresse les noms propres tels qu'ils se présentent sur la page de titre. Ainsi, en posant la requête « Guillaume Rouillé », nous n'obtenons que sept réponses (sous la forme « Rouille », mais le logiciel ne tient pas compte des accents), et il est nécessaire de la reformuler avec d'autres graphies, dans plusieurs variantes françaises (Rouillé / Roville), italiennes (Rouillio / Rovillio) et latines (Rouillium / Rouilium / Rovillium / Rovilium / Rovillius)⁷⁸. Comme on n'a jamais pu établir la forme exacte du nom de ce libraire, les catalogueurs ont transcrit le V majuscule des pages de titre soit par un v soit par un u. On doit donc considérer ici qu'il y a régression par rapport à un fichier manuel comme celui des « Imprimeurs lyonnais » de la BML — qui comprend également les éditeurs et pour lequel toutes les formes ont été ramenées à une seule — mais au moins la recherche par éditeur est-elle possible dans Gallica, comme recherche

⁷⁸ Le même problème se pose dans BN-Opale Plus. Rouillé ne publie pas encore en espagnol dans les années qui nous intéressent. Au demeurant, il hispanisera son nom en Guillelmo ou Guillermo Rouillio.

libre, ce qui n'est pas le cas pour l'imprimeur, dont le nom n'a pas été saisi dans les notices.

Etant donné le grand nombre de réponses pertinentes sur Guillaume Rouillé dans Gallica, nous limitons notre enquête aux livres publiés jusqu'à deux ans avant et deux ans après *La magnificence*, soit de 1547 à 1551, date à laquelle on observe d'ailleurs un début de renouvellement des ornements typographiques de certains imprimeurs de Rouillé. La date d'édition n'apparaissant pas dans la liste des réponses, qui ne mentionne que l'auteur et le titre, il est nécessaire d'ouvrir la page web de chaque notice, pour vérifier qu'il s'agit bien de notre éditeur (le nombre de réponses pertinentes figure dans la troisième colonne du tableau 3) et pour connaître la date de l'ouvrage. Cette opération réclame bien plus de temps que la consultation d'un fichier manuel organisé en notices d'autorité, mais elle présente l'avantage de pouvoir se faire à distance. Le tableau ci-dessous, avec ses neuf requêtes pertinentes pour trouver une seule entité, rendra compte du manque de commodité de la recherche dans le catalogue de Gallica.

**Tableau 3. — Interrogation dans Gallica :
ouvrages publiés par Guillaume Rouillé de 1547 à 1551**

Requête	Réponses	Rouillé	1547	1548	1549	1550	1551
Guillaume							
Rouillé	7	7			2		
Roville	19	5			1		1
Rouillio	7	7			1		2
Rovillio	1	1					
Rouillium	73	70	6	6	13	4	3
Rouilium	3	3		1		1	
Rovillium	5	5					1
Rovilium	1	1	1				
Rovillius	3	2					

En définitive, Gallica nous offre quarante-et-un ouvrages à comparer avec les livrets de l'entrée de Henri II⁷⁹. Rouillé, proportionnellement aux autres éditeurs contemporains, semble très bien représenté dans cette bibliothèque numérique, notamment grâce au versement d'un important fonds

⁷⁹ Site visité le 20/03/2002. Ce sont 43 ouvrages publiés par Guillaume Rouillé de 1547 à 1551, y compris les deux livrets de l'entrée.

de médecine par la bibliothèque de Cambridge (Mass.). Dans un premier temps, nous avons feuilleté à l'écran cette quarantaine de documents, pour y découvrir des noms d'imprimeurs, dont nous savons qu'il sont, le cas échéant, portés à la fin des volumes. De plus, nous avons noté la présence de lettrines, qui se trouvent dans la plupart des cas dans les premières pages. Enfin, nous avons observé des pratiques d'atelier différentes quant à l'emploi des réclames⁸⁰, qui servent à l'assemblage des cahiers. Ces différentes recherches ont été faites par approximations successives, en cliquant au hasard sur des numéros de page, de manière plus lente que si nous avions eu les exemplaires en main — ce qui est assez gênant. Cependant, il convient de souligner que l'ouvrage requis est mis plus rapidement à la disposition du lecteur dans Gallica que dans une bibliothèque réelle.

Les résultats de cette enquête sont exprimés dans le tableau C (en annexe 3). D'emblée, l'usage des réclames apparaît comme un élément déterminant pour notre recherche, en ce sens que chaque imprimeur aurait utilisé un système propre, du moins d'après le corpus d'une quarantaine d'œuvres disponibles dans Gallica (tableau 4). Naturellement, ce critère aboutit à une périodisation proche de celle que suggérerait la succession des imprimeurs employés par Rouillé (tableau 2). Dans presque tous les cas, il est fait un usage normalisé des signatures, portées sur les rectos de la première partie des cahiers (1-5 pour les cahiers de 8 feuillets, 1-3 pour ceux qui en ont 4). Le système dénommé ci-après A, utilisé avec des cahiers de 8 feuillets (A8), fait une large place aux réclames, qui sont présentes sur les faces de tous les feuillets. Il est utilisé par Roussin et Ausoult en 1547. Une variante (A') se caractérise dans des impressions anonymes de 1547 et 1549 par la présence de la signature jusqu'au f. 4 inclus. Rollet et Frein en 1548-1549 puis Rollet seul en 1550-1551 sont toujours associés au système B (en l'occurrence B8), qui ne juxtapose pas signature et réclame. C8, qui n'emploie qu'une réclame par cahier (f. 8v), pour annoncer au relieur le cahier suivant, est représenté par une impression anonyme de 1549 (Bernieri). Enfin C'8, d'où les réclames sont absentes, semble être le propre de Macé Bonhomme, mais il convient de signaler que ses impressions pour Rouillé — divers tirages des emblèmes d'Alciat — ont une mise en page particulière, qui n'est peut-être pas caractéristique de sa pratique courante.

⁸⁰ Annonce du ou des premiers mots de la page suivante.

Tableau 4. — Indications pour le pliage des cahiers

	Recto	Verso	Imprimeur
A8	f. 1-5 : signature + réclame f. 6-8 : réclame	f. 1-8 : réclame	Roussin & Ausoult
A'8	f. 1-4 : signature + réclame f. 5-8 : réclame	f. 1-8 : réclame	
B8	f. 1-5 : signature f. 6-8 : réclame	f. 1-8 : réclame	Rollet et Frein Rollet
B4	f. 1-3 : signature f. 4 : réclame	f. 1-4 : réclame	
C8	f. 1-5 : signature f. 6-8 : ø	f. 1-7 : ø f. 8 : réclame	
C'8	f. 1-5 : signature f. 6-8 : ø	f. 1-8 : ø	Bonhomme

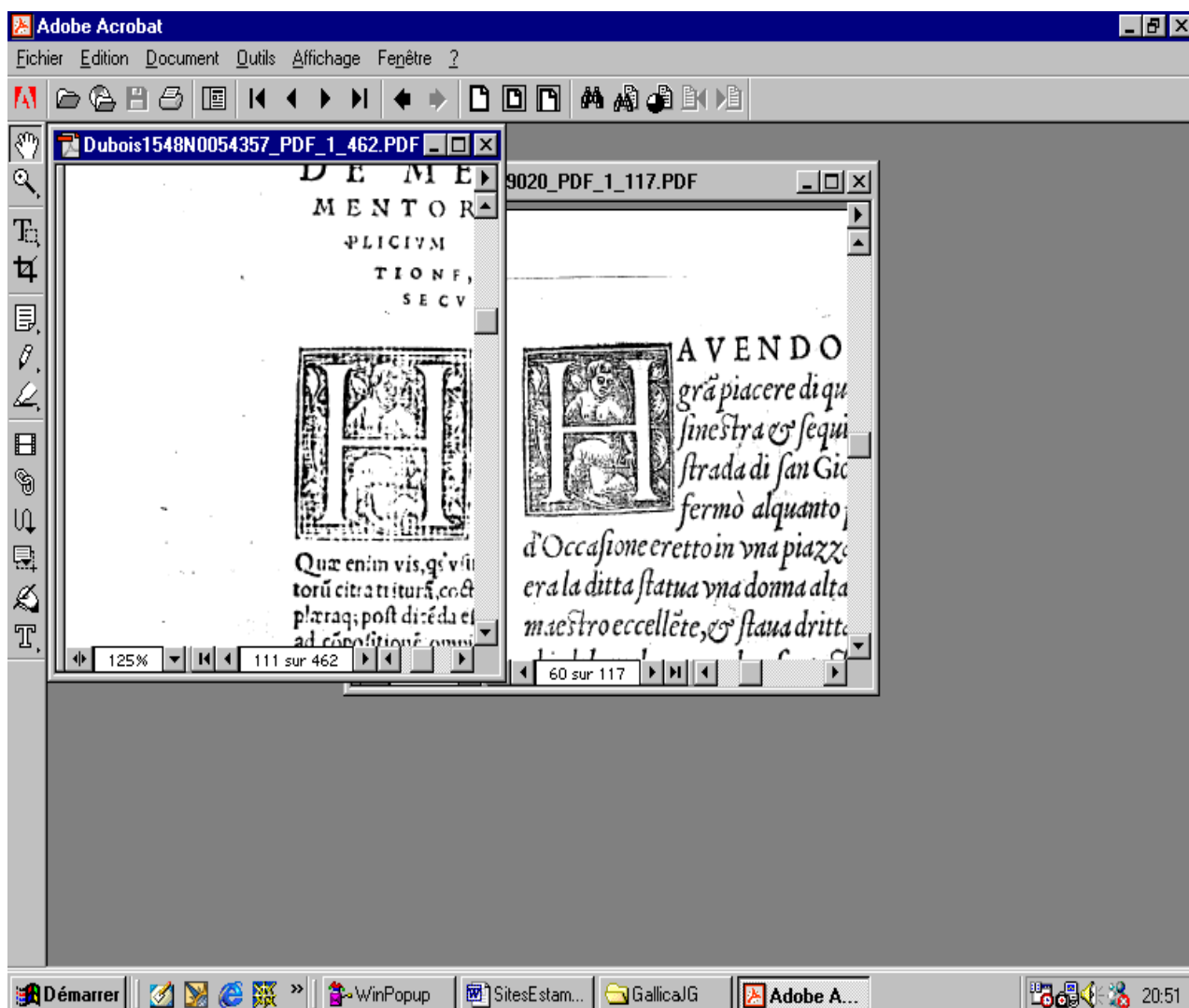
Les trois systèmes (A-C) sont représentés en 1549, l'année de *La magnificence*, livret pour lequel on a employé le système B (variante B4, puisqu'il est composé de cahiers de quatre feuillets), ce qui le rapproche de la production de Rollet et Frein. Cela posé, il restait à faire des vérifications en comparant les ornements typographiques. Dans cette perspective, nous nous sommes constitué notre bibliothèque numérique spécialisée en téléchargeant de Gallica un certain nombre d'ouvrages imprimés entre 1547 et 1551, afin de pouvoir les consulter commodément. Les vérifications ont ensuite montré que *La magnificence* avait été imprimée par Rollet et Frein. La même méthode permet d'attribuer d'autres anonymes à ces imprimeurs, ainsi les *Institutions chirurgiques* de Jean Tagaut (1549), qui sortent bien de la même officine que *La magnificence* puisque nous retrouvons non seulement des lettrines, comme le C de 33 mm de côté provenant de l'imprimerie de Dolet (cf. tableau 1), mais aussi des fleurons. Ce dernier élément s'impose pour esquisser la chronologie relative de la production de cette officine en 1549, dans la mesure où l'on ne retrouve pas dans le livre de Tagaut le fleuron dont le bois avait claqué dès l'impression des premiers feuillets du livret de l'entrée⁸¹. D'après cet indice, l'ordre a été le suivant : *La magnificence*, puis *La magnifica*, enfin les *Institutions chirurgiques*.

⁸¹ Bois entier aux f. Av et C3v de *La magnificence*, dégradé aux f. D, H4v, I2v.

La bibliographie matérielle virtuelle est donc possible. Grâce à la bibliothèque numérique de la BNF, nous avons pu attribuer à Philibert Rollet et Barthélemy Frein cet imprimé anonyme qu'est *La magnificence*, et il suffirait de peu de vérifications supplémentaires pour rendre à leur atelier ou à celui de Macé Bonhomme la plupart des autres ouvrages édités par Guillaume Rouillé dans les mêmes années et dont l'imprimeur n'est pas connu. C'est le fait de disposer avec Gallica de nombreuses éditions de Rouillé numérisées en mode image qui nous a permis d'une part d'établir des critères pertinents pour une première sélection des ouvrages à comparer (étude de l'usage des réclames), d'autre part de rapprocher les ornements typographiques de livres signés ou non. Le corpus actuellement disponible est donc réellement appréciable. Cependant, Gallica est difficilement utilisable par les chercheurs. Comme nous l'avons vu, le catalogue n'est pas conçu de façon à permettre de localiser rapidement tous les documents recherchés. Il faut encore du temps pour télécharger des pages dont rien ne permet de savoir à l'avance si elles vont être utiles. Enfin, lorsque les pages pertinentes ont été ouvertes, aucun outil n'en permet l'exploitation scientifique. En définitive, c'est de manière assez artisanale que l'on va rapprocher les pages de deux documents à comparer visuellement, en réduisant les fenêtres à l'écran. Il est certes possible de se créer, par exemple, une base personnelle de fleurons, à condition de télécharger en format TIFF tout le document sur son disque dur, et de travailler avec les outils d'Imaging®, mais ce collage s'apparente à celui de photocopies et l'exploitation d'un tel outil est extrêmement limitée⁸².

Gallica est donc une bibliothèque numérique que le chercheur consultera avec profit, tout en regrettant le manque de fonctionnalités avancées pour l'étude non seulement des images mais aussi du texte. Voyons maintenant dans quelle mesure DEBORA répond à ses attentes.

⁸² Cf. la copie d'écran ci-dessous.



Comparaison de deux lettrines dans des ouvrages téléchargés de Gallica :

- fenêtre de gauche : Dubois (Jacques), *Pharmacopoeae... libri tres*, Lyon, 1548, imprimé par Rollet et Frein pour Rouillé et Constantin, p. 110 [Gallica N 054357, p. 111 du document téléchargé]
- fenêtre de droite : *La magnifica*, Lyon, 1549, édité par Rouillé, sans nom d'imprimeur, f. Hv [Gallica N 079020, p. 60 du document téléchargé]

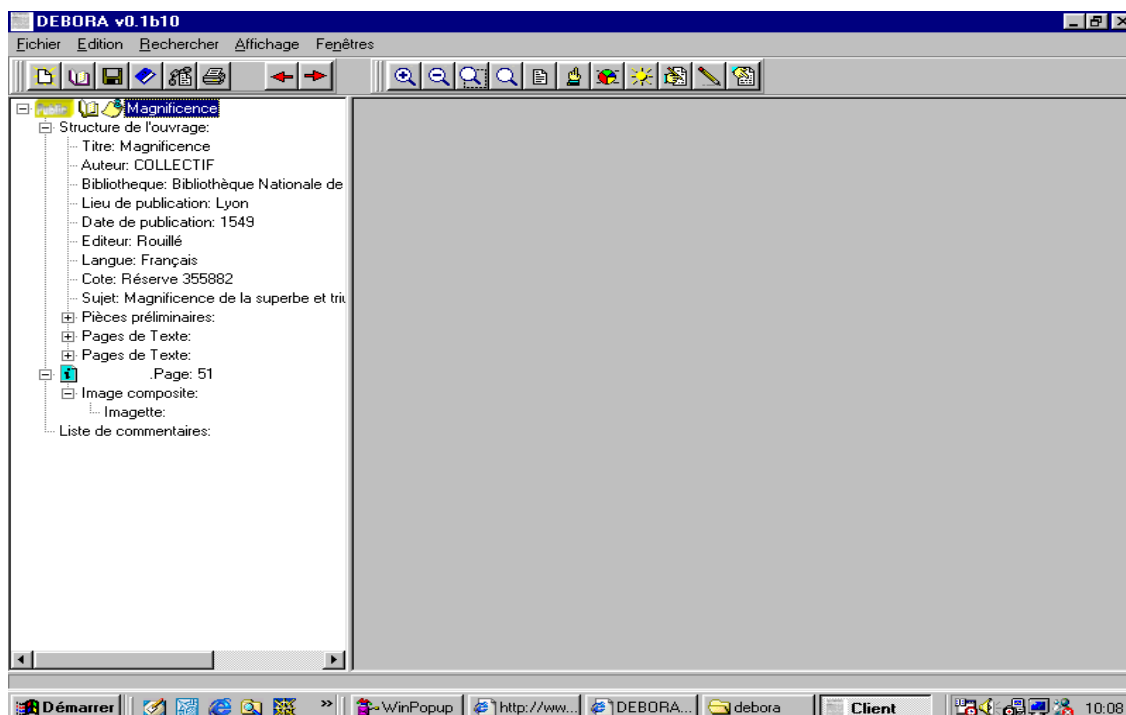
1.3. Les outils proposés par DEBORA

Avant de présenter notre évaluation des fonctionnalités du logiciel DEBORA, il convient de préciser que l'étude de bibliographie matérielle menée précédemment et qui a permis d'établir de manière à peu près certaine le nom de l'imprimeur de *La magnificence* est absolument impossible dans le cadre de la station de travail actuellement disponible. La bibliographie matérielle étant fondée sur des méthodes de comparaison entre exemplaires, il est évident que DEBORA n'est pas pour l'instant adapté à ce type d'usages. Cependant, dans la perspective d'une reprise du projet et d'un accroissement de la base, les outils actuellement disponibles pourraient sans aucun doute rendre des services importants aux spécialistes du livre ancien.

Il convient maintenant de présenter ces outils à travers l'étude de *La magnificence*, en reprenant la typologie esquissée en 2.3. Ce test sera l'occasion d'une démonstration du logiciel, parfois complétée par des propositions d'améliorations qui nous semblent de simples prolongements des fonctionnalités existantes et qui seraient susceptibles d'être intégrées, à notre avis, assez aisément.

1.3.1. Les outils de visualisation

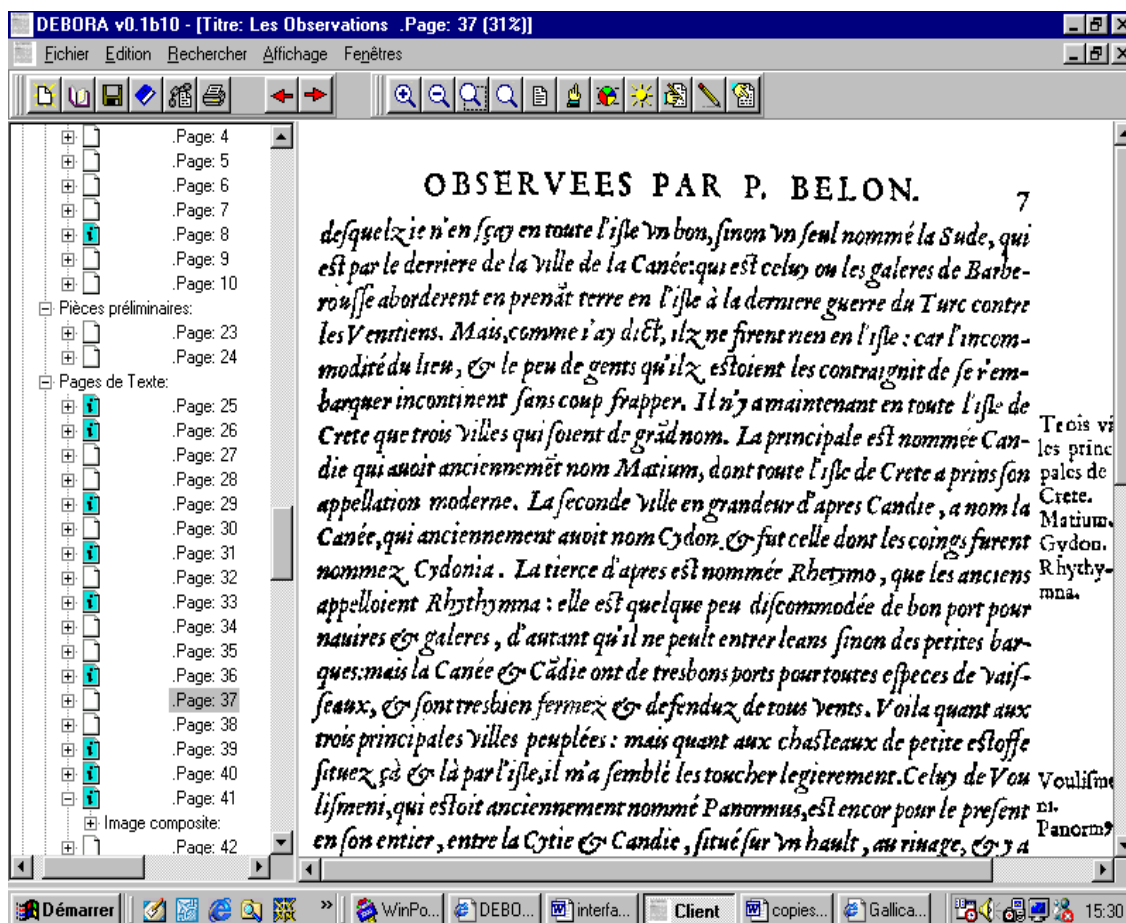
L'illustration de la page suivante représente une copie de l'écran du navigateur DEBORA. On remarque à gauche l'arborescence qui décrit la structure de l'ouvrage : métadonnées (titre, auteur...), pages de texte, images. Derrière l'intitulé "pages de texte" se trouve l'ensemble des pages de l'ouvrage numérisé, regroupées en deux blocs, la page 51 étant rejetée – sans qu'aucune raison logique ne l'explique – à la fin de l'arborescence.



On peut regretter que cette structuration par pages ne rende pas compte de la réalité de l'ouvrage imprimé, contrairement à ce que propose l'interface de Gallica, où le numéro de la page web est identique à celui de la page imprimée. L'exemple suivant⁸³ montre bien le décalage entre la pagination de l'ouvrage (ici folio 7) et celle de l'arborescence (page 37). Plutôt que de paginer à nouveau le livre, ne faudrait-il pas simplement reporter ce qui est imprimé, c'est-à-dire structurer *La magnificence* avec les signatures⁸⁴ ? Dans cette perspective, la collaboration entre les informaticiens mettant en place "l'arborescence virtuelle" et les historiens connaissant la structure originelle du livre imprimé paraît indispensable.

⁸³ Nous avons ici pris le livre de Pierre Belon (*Les observations*) car *La magnificence* n'est pas paginée.

⁸⁴ La signature est le numéro du feuillet dans le cahier (A, Av, A2, A2v etc.). A titre de comparaison, l'interface de Gallica permet de rendre compte de la structuration par cahiers d'un ouvrage ni paginé ni folioté comme *La magnificence*, où la séquence est interprétée comme suit : 1, 1, 2, 2, 3, 3, NP (non paginé), NP, 1, 1, etc... L'URL de *La magnificence* dans Gallica est <http://gallica.bnf.fr/scripts/ConsultationTout.exe?E=0&O=N076015>.



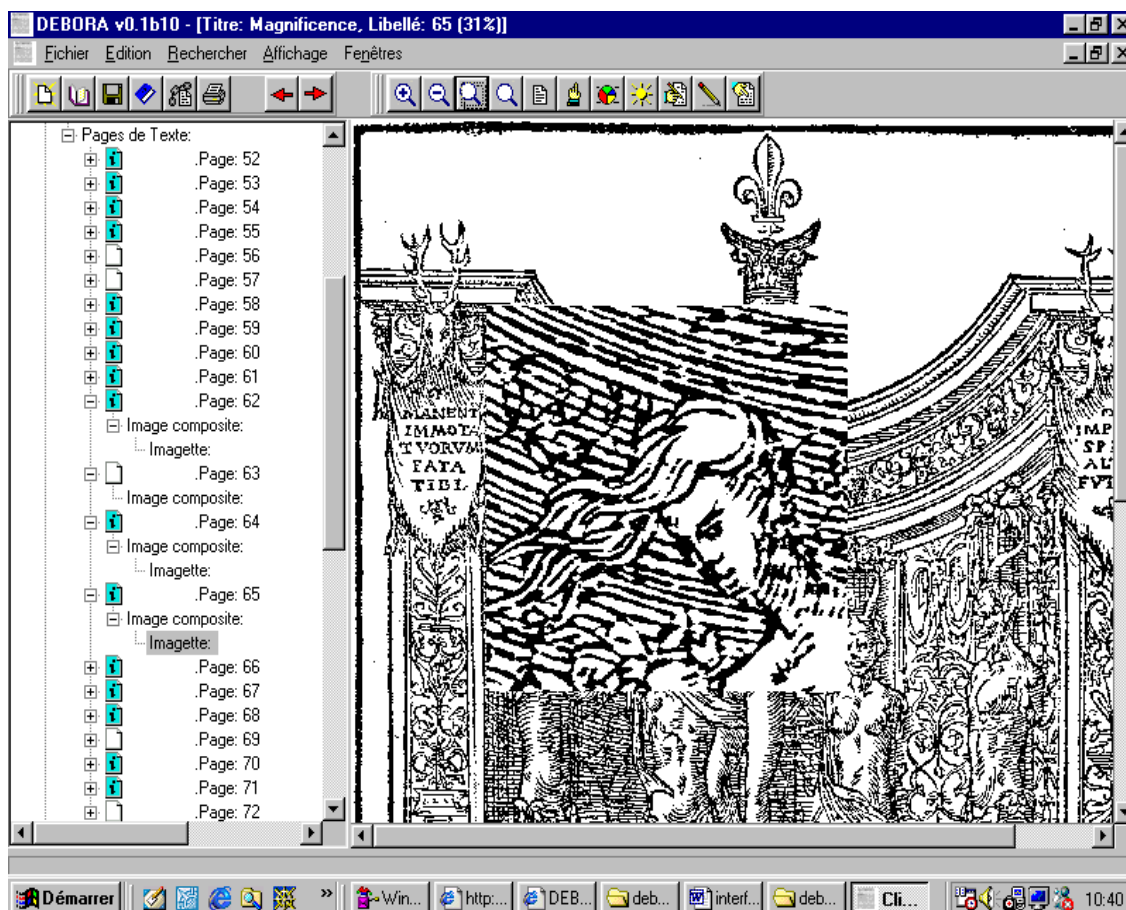
Une telle synergie des compétences est tout aussi fondamentale au moment de l'adaptation de l'arborescence à chaque livre: c'est ainsi que le terme "image composite", s'il était conservé⁸⁵, ne devrait être employé que dans le cas d'une page comportant à la fois du texte et de l'illustration et non appliqué à toutes les pages sans distinction.

L'illustration suivante montre bien la limite du terme d'« image composite » puisque la page 65 que l'on a affichée à l'écran est composée d'une gravure pleine page. Or, cette page apparaît dans l'arborescence de gauche comme structurée en image composite puis en une imagerie. En

⁸⁵ En réalité, on peut se demander dans quelle mesure ce terme est approprié à la description des pages d'un ouvrage, à tout le moins du point de vue de l'utilisateur. En effet, ce terme s'appliquant à toutes les pages sans distinction, il reflète davantage la structure de l'image numérisée (c'est une image composée de plusieurs types de fichiers, à la fois texte et image) que celle de la page imprimée, seule entité qui finalement importe au chercheur.

cliquant sur ce dernier terme, on ouvre uniquement l'illustration dans le plan de travail.

Nous ajoutons ici la démonstration de l'option de « loupe baladeuse »⁸⁶, assez intéressante grâce à un agrandissement de bonne qualité permettant aux spécialistes de la gravure de voir, sans pixellisation, n'importe quel détail.



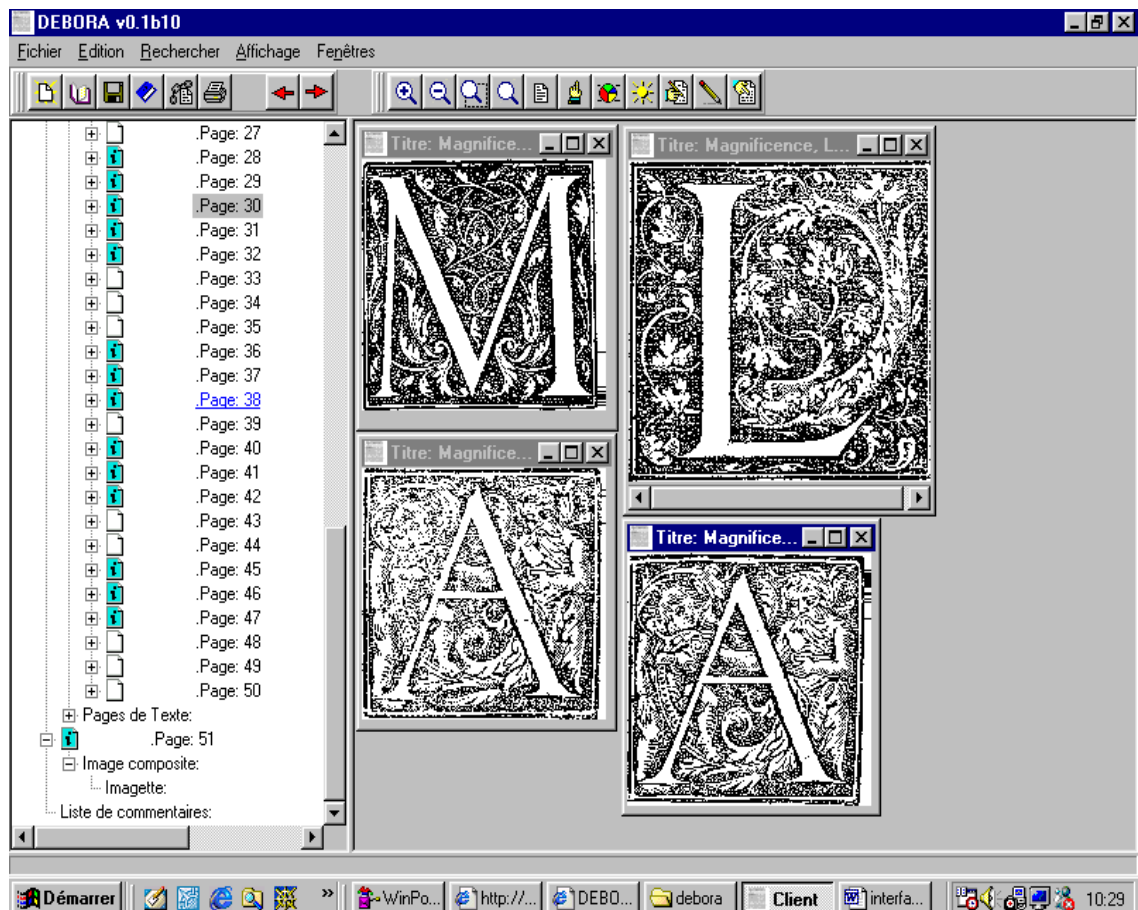
- Affichage multi-fenêtre

Dans la perspective d'une recherche comparative, sur les lettrines par exemple, l'affichage multi-fenêtre est extrêmement précieux. Notons en outre qu'il s'agit là d'une fonctionnalité nouvelle par rapport au fac-similé ou à Gallica⁸⁷. Il est cependant dommage de ne pas disposer d'une règle permettant

⁸⁶ Notons que cette fonctionnalité se trouvait à titre de prototype sur le site de la Bibliothèque vaticane (http://www.vatican.va/library_archives/vat_library/index_it.htm, visité le 19/04/2002).

⁸⁷ A comparer avec la copie d'écran d'après Gallica à la fin du paragraphe 1.2 de la troisième partie.

de mesurer les gravures extraites des ouvrages⁸⁸. En outre, les pourcentages de grossissement de l'image ne correspondent pas à l'ouvrage original : en effet, lorsque l'agrandissement est à 100 %, le résultat à l'écran ne représente pas l'ouvrage à l'échelle 1 ce qui est gênant pour des travaux de bibliographie matérielle. Le même problème se pose dans Gallica.



En outre, plutôt que de faire apparaître la mention "Titre: Magnificence" dans l'en-tête de la fenêtre, il serait plus utile pour le chercheur de connaître la page (ou le feuillet) dont est issue la lettrine.

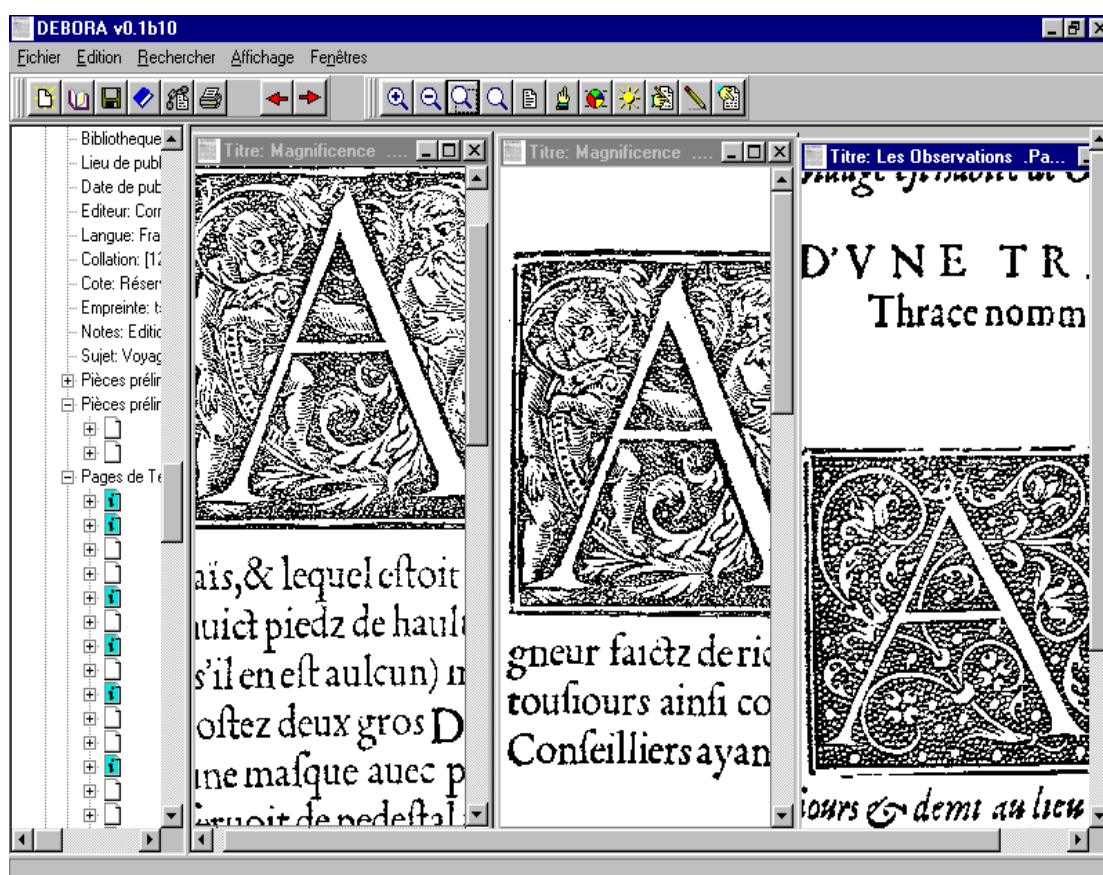
Dans l'exemple que nous présentons ici, on voit un jeu de quatre lettrines à fond criblé extraites de *La magnificence*: un L à rinceaux et figures provenant de l'imprimerie d'Etienne Dolet (54 x 54 mm), un M à rinceaux (42 x 42 mm) et deux A à rinceaux et figures (42 x 42 mm) issus de deux pages

⁸⁸ Le système de coordonnées n'est pas d'un usage très facile pour définir les dimensions des gravures et, de plus, il nous semble hors échelle.

différentes pour lesquels on peut remarquer, malgré la perte d'information due à la copie d'écran puis au tirage sur l'imprimante laser, la différence d'encrage au moment de l'impression en 1549. Pour la typologie des différents types d'alphabets de lettrines utilisés dans le livret, nous renvoyons à l'étude et au tableau 1 (1.2.1 de la partie III).

Grâce à cette possibilité d'ouverture simultanée de plusieurs pages, la recherche effectuée précédemment pour identifier l'imprimeur par la comparaison avec un matériel déjà connu serait possible sous DEBORA, dans l'hypothèse d'une base regroupant un certain nombre d'ouvrages du XVI^e siècle lyonnais. En effet, il serait alors possible d'ouvrir côte à côte des lettrines issues d'ouvrages distincts et de les comparer, plus aisément peut-être qu'on ne le ferait avec deux fac-similés voire deux originaux.

La copie d'écran suivante est présentée à titre purement indicatif⁸⁹, pour faire comprendre que les similitudes entre deux lettrines apparaissent au premier coup d'œil lorsque les gravures sont ainsi mises en regard.

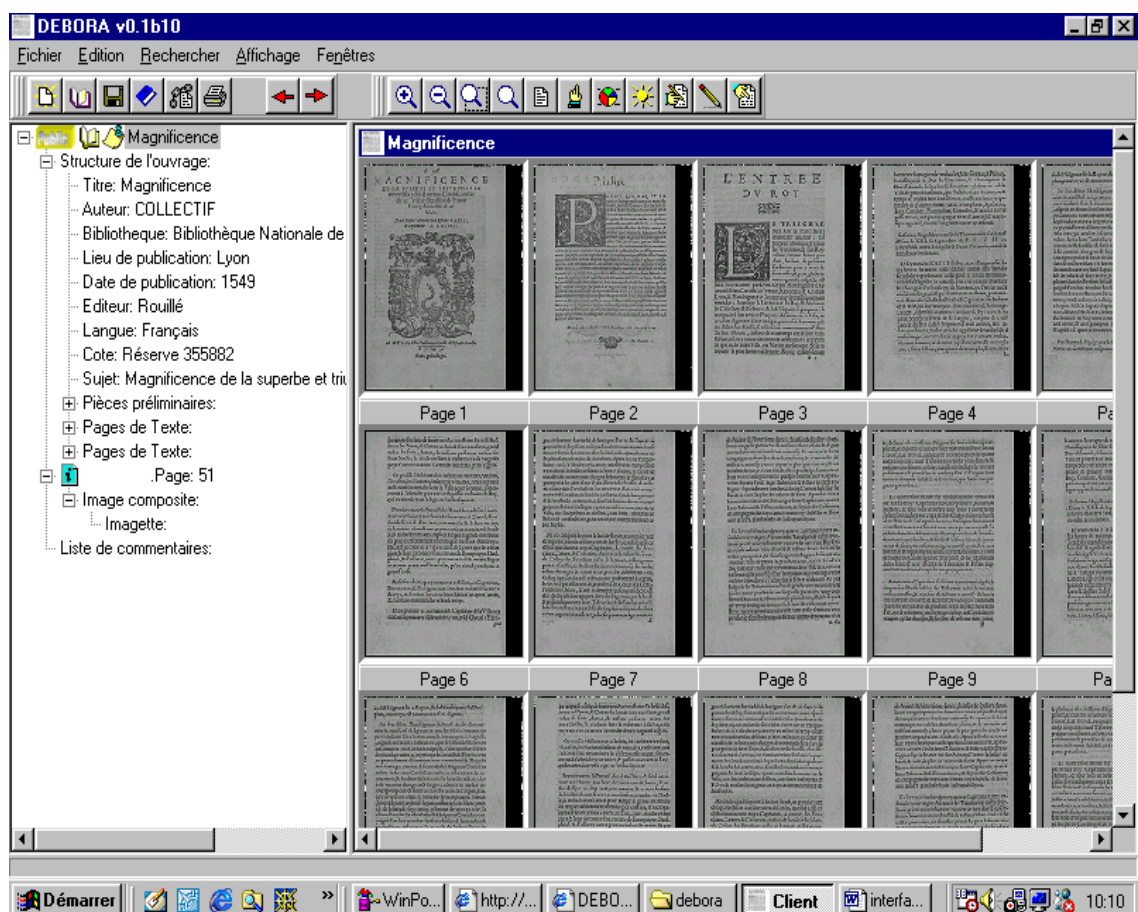


⁸⁹ Le matériel utilisé pour *Les observations...* (fenêtre de droite de la copie d'écran), ouvrage de Pierre Belon imprimé à Paris en 1554, n'a bien sûr aucun rapport avec celui de *La magnificence*.

Si, grâce à cet outil, les premières comparaisons se révélèrent fructueuses, il serait alors intéressant de parvenir à recréer – à l'écran – l'alphabet des lettrines dont disposait l'imprimeur lorsqu'il tira les ouvrages. C'est alors que les outils de DEBORA deviendraient particulièrement utiles. Cependant, si le chercheur veut recomposer cet alphabet en le disposant tel qu'il pouvait apparaître à l'imprimeur, il se heurte pour l'instant à la taille réduite et non extensible de la zone d'affichage des pages (plan de travail). Pour une recherche plus aisée, il faudrait pourvoir cette zone d'une barre de défilement.

- Le vignettage

Cette fonctionnalité permet de faire défiler l'ensemble des pages du livre sous forme d'un aperçu ce qui permet notamment de repérer les lettrines, les gravures pleine page, les fleurons, etc.

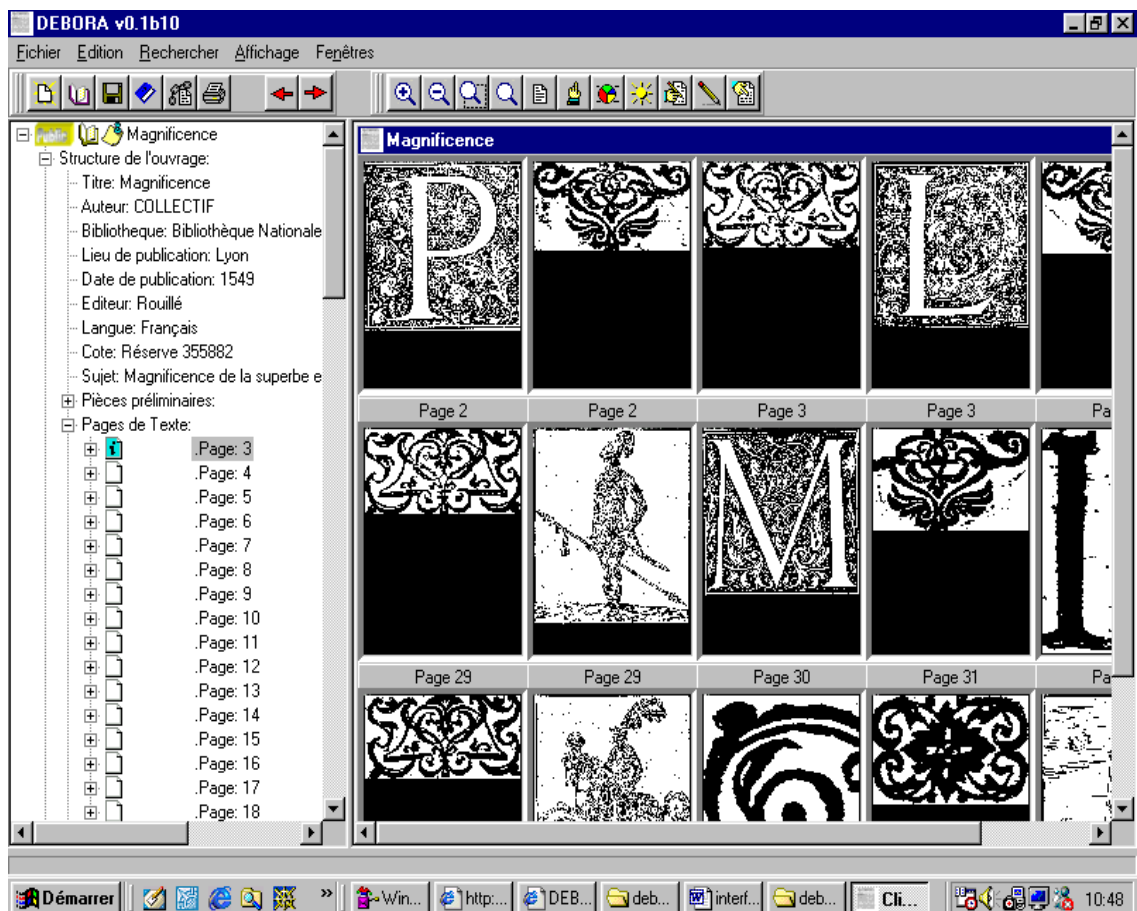


1.3.2. Les outils de recherche

- L'outil de recherche dans les éléments graphiques de l'ouvrage

Cet outil apporte un gain de temps appréciable au chercheur travaillant sur les gravures. En effet, il permet de faire défiler toutes les illustrations d'un ou de plusieurs ouvrages et facilite ainsi la navigation au sein d'un livre illustré.

On pourrait cependant préférer une possibilité de recherche par type de décoration : recherche dans les lettrines du livre, dans les gravures pleine page etc. Ceci nécessiterait au préalable une structuration plus fine des métadonnées sur les décorations (distinction entre lettrines et fleurons par exemple).



La copie d'écran ci-dessus permet de montrer que le fait même de faire apparaître à *la suite* l'ensemble des décorations d'un ouvrage permet, sinon

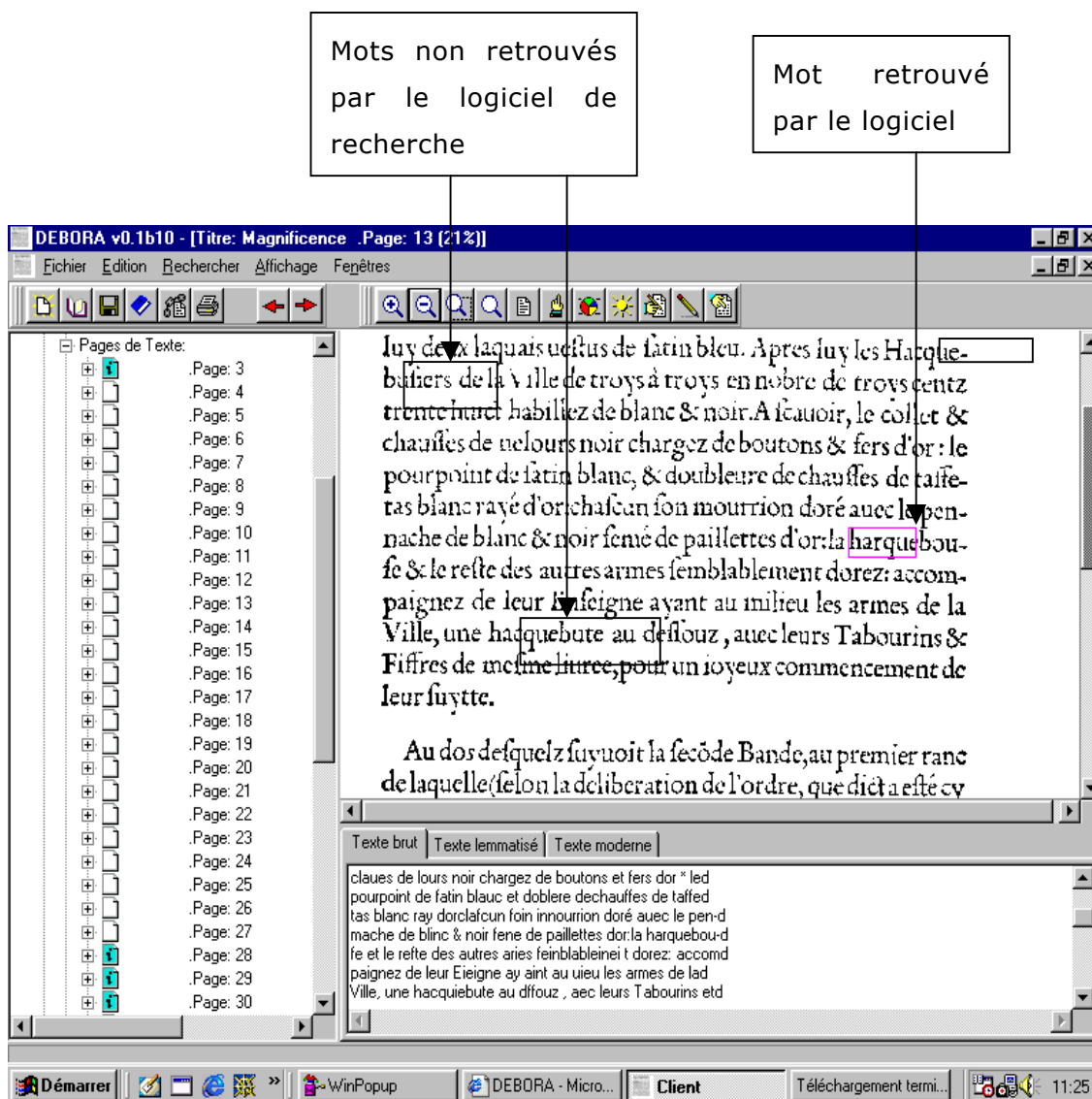
d'amorcer une étude de bibliographie matérielle, du moins de pointer quelques détails qui peuvent avoir leur importance. En effet, en observant les fleurons du début du livre, on remarque l'utilisation de deux bois différents: le fleuron A en page 2, le fleuron B en page 3. On retrouve ce dernier à la page 29 et le premier à la page 31. Cependant, on peut relever que le bois utilisé à la page 31 est en partie cassé, dans le coin supérieur gauche de la gravure. Le bois du fleuron A a claqué à l'impression et il fut remplacé par l'imprimeur comme on le voit dans la troisième ligne des décorations de l'ouvrage : on retrouve le fleuron B des pages 3 et 29, mais le A a disparu au profit d'une autre forme, C (cf. la 4^e colonne de la ligne 3).

- L'option de recherche de mots dans un ou plusieurs ouvrages

Si cette fonctionnalité peut rendre de nombreux services aux chercheurs, notamment aux linguistes, elle est cependant limitée pour les ouvrages du XVI^e siècle à cause du caractère non unifié de l'orthographe à cette époque.

En effet, pour trouver le mot arquebuse, il faudrait interroger par les diverses graphies du mot: harquebouse, hacquebute etc. Dans l'exemple suivant, notre requête par "harque" ramène le mot « harquebouse » mais pas « hacquebute » ni « hacquebusiers ». On prend conscience ici des limites d'une telle recherche. La solution serait de créer un dictionnaire recensant les mots de l'ouvrage dans des graphies normalisées et présentant les formes rejetées associées. On pourrait aussi envisager une lemmatisation du texte⁹⁰ transcrit, laquelle permettrait de retrouver tous les mots d'une même famille.

⁹⁰ En effet, si le champ du texte lemmatisé est prévu dans la fenêtre de la transcription, il n'est pas actuellement renseigné.

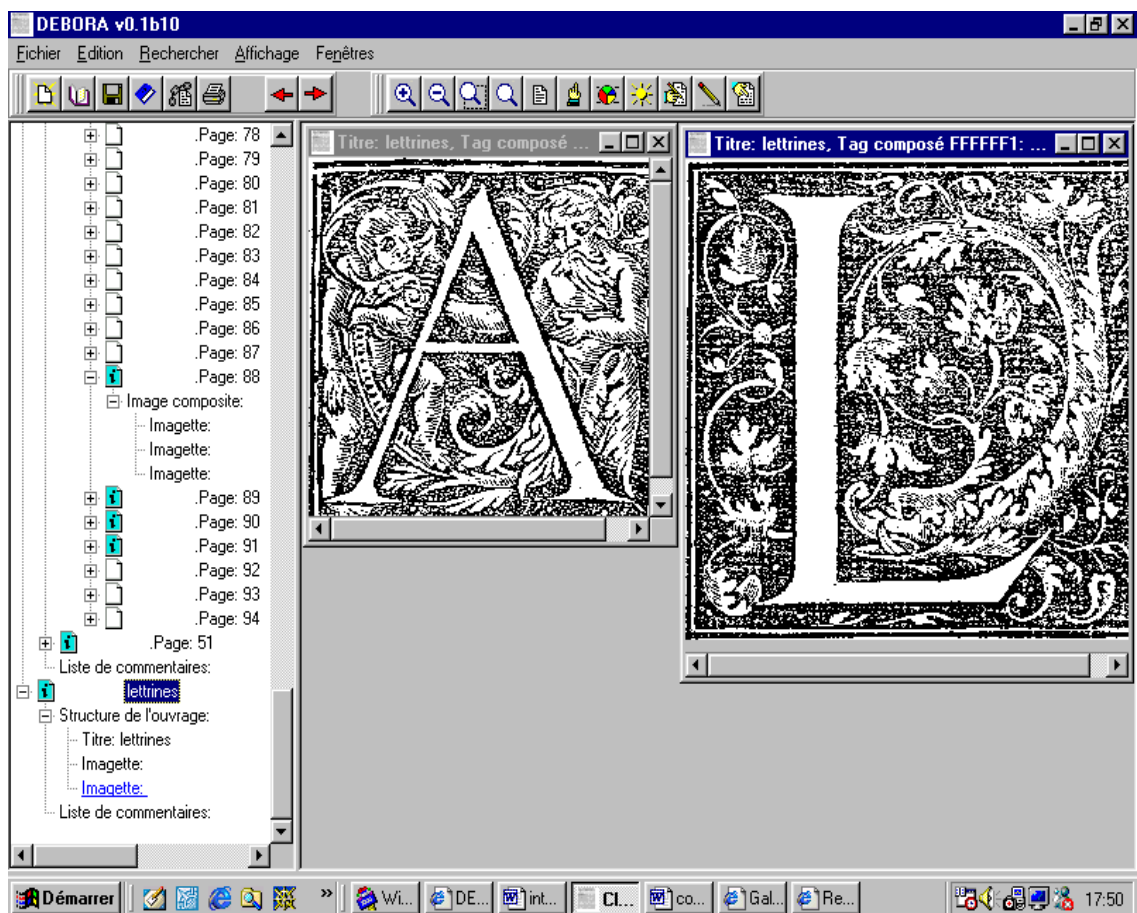


La possibilité d'interroger plusieurs ouvrages en même temps pourrait alors permettre de dater l'apparition d'un mot et de suivre ses différentes graphies comme son évolution au cours du temps.

Bien évidemment, l'efficacité d'un tel outil repose sur la qualité de la transcription. Comme on peut le voir ci-dessus, cette dernière demeure, en raison des difficultés techniques évoquées par ailleurs, difficilement exploitable, sans une fastidieuse correction.

1.3.3. Les outils d'exploitation personnelle

La création d'un dossier personnel est une fonctionnalité inédite qui permet au chercheur de se constituer sa propre base avec les éléments dont il a besoin, par exemple les lettrines, comme le montre l'exemple ci-dessous. Ce dossier personnel est sauvegardé sur le disque dur et peut être ouvert indépendamment des ouvrages.



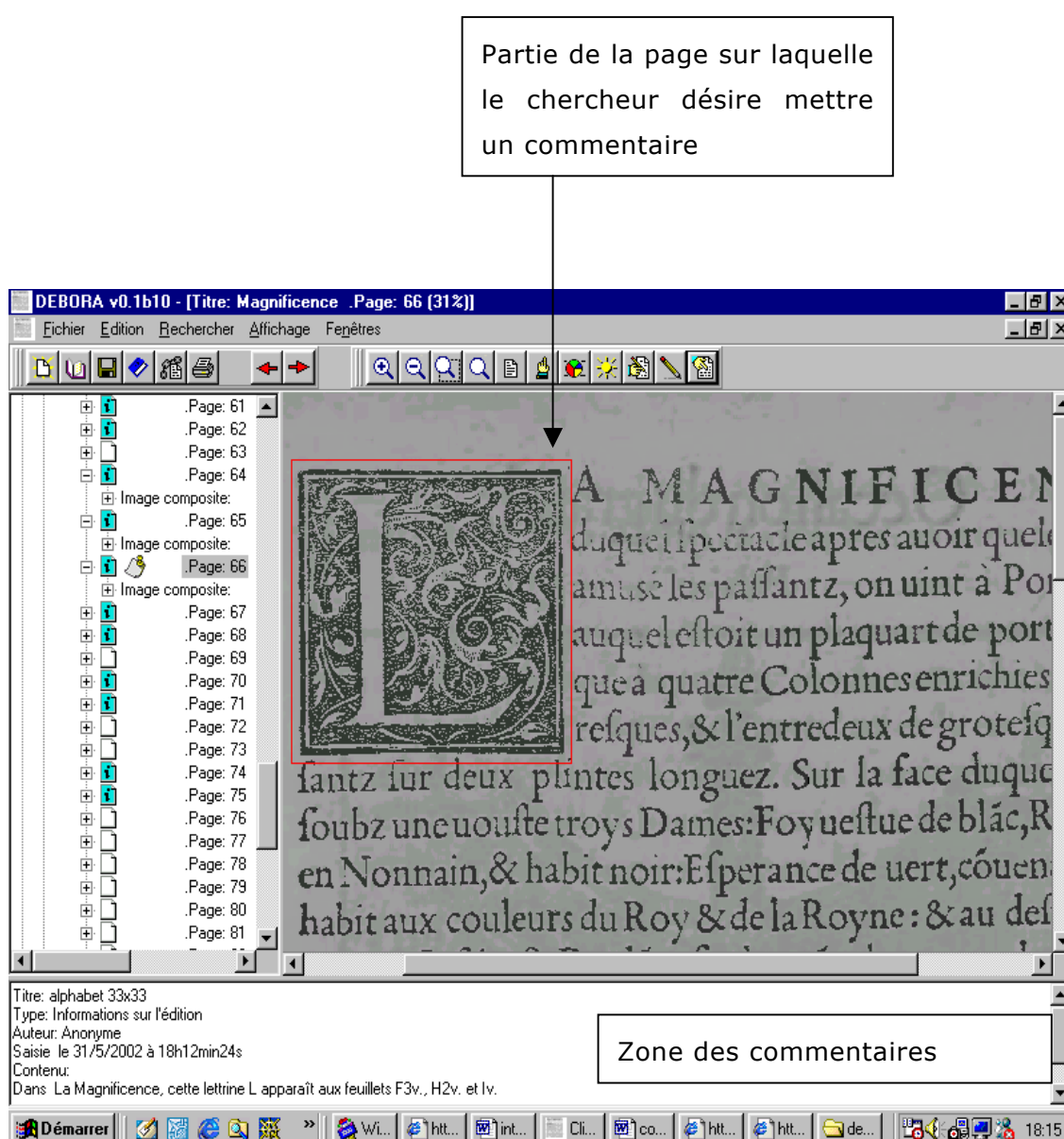
Deux lettrines extraites de *La magnificence* et placées dans un dossier personnel nommé

« lettrines »⁹¹

⁹¹ Celui-ci apparaît sur fond bleu dans l'arborescence de gauche.

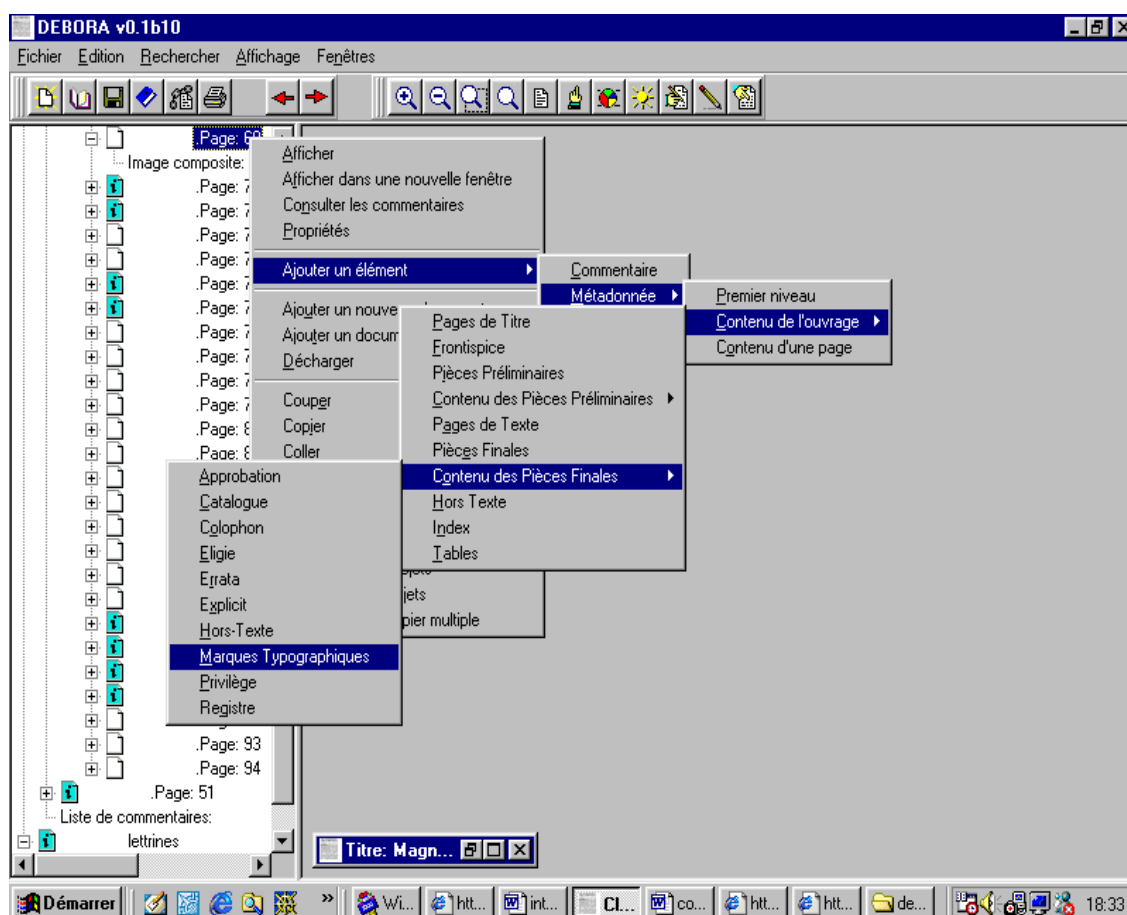
Le logiciel DEBORA offre en outre la possibilité d'annoter l'ouvrage et de visualiser les commentaires. Ceux-ci sont signalés dans l'arborescence de gauche (fiche punaisée) et apparaissent lorsqu'on passe le curseur sur l'élément annoté (ici la lettrine L entourée).

Petit à petit, les ouvrages s'enrichiraient de ces commentaires validés par un coordonnateur. Ils pourraient ainsi être lus par tous les utilisateurs de la base ou réservés au groupe de recherche de l'annotateur, la note étant clairement datée et signée pour préserver le droit de l'auteur sur celle-ci.



De plus, si un chercheur souhaite consulter les commentaires existants portant par exemple sur les lettrines de l'ouvrage, il doit utiliser le menu « Recherche dans les textes », sélectionner le champ « Annotations » puis interroger par le sujet idoine. Le moteur de recherche lui ramène alors tous les commentaires présents dans l'ouvrage et contenant ce mot.

Précisons également que l'historien du livre est en mesure d'ajouter un élément de type métadonnée sur quelque page que ce soit, afin d'en préciser le contenu (cf. ci-dessous). Si les champs concernant le texte sont développés avec un luxe de détail, en revanche il est impossible actuellement de préciser le statut des illustrations. Ceci pourrait être une extension à ajouter.



Les fonctionnalités que nous avons brièvement décrites et analysées renouvellent la façon d'appréhender le livre ancien. DEBORA étant un prototype, ces outils ne sont actuellement pas parfaitement opérationnels et gagneraient, pour certains d'entre eux du moins, à être développés dans une version future du navigateur.

À l'issue de nos recherches et dans l'éventualité d'un nouveau projet plus global, nous avons envisagé un certain nombre de pistes d'évolution qui font l'objet de la partie suivante.

1.4. Les perspectives de développement : pour une nouvelle architecture de la base

Pour l'instant, le site DEBORA, hébergé par l'INSA de Lyon, propose une juxtaposition de cinq livres téléchargeables. Ceux-ci sont proposés dans le format développé par le laboratoire RFV de l'INSA et pour certains possèdent des métadonnées et la transcription partielle du texte. Pour un travail portant sur plusieurs ouvrages, on doit d'abord télécharger les différents fichiers sur son disque dur, puis les ouvrir dans le navigateur DEBORA. Il est alors possible de mener diverses recherches à partir de cet ensemble, par exemple extraire les décorations des différents ouvrages, ou trouver des occurrences de mots dans plusieurs textes. Dans le cadre d'un nouveau développement du projet DEBORA, la base devra naturellement s'enrichir d'un corpus important d'œuvres numérisées, pour en faire une véritable bibliothèque. Il convient désormais de réfléchir à son architecture.

1.4.1. Conception du catalogue

Si un historien s'intéresse à un seul ouvrage contenu dans la « base », DEBORA peut lui rendre un certain nombre de services. En revanche, si un spécialiste des arts graphiques travaille sur un sujet plus large, l'évolution de l'illustration dans les livres de la Renaissance lyonnaise par exemple, ou encore dans le cas de travaux de bibliographie matérielle portant sur la production d'un atelier typographique - toutes études fondées sur la comparaison - la

question se posera de savoir *a priori* quels livres télécharger. Avant d'augmenter sensiblement le nombre de livres contenus dans la base, il est en premier lieu nécessaire de concevoir une interface de recherche.

Celle-ci, pour donner le meilleur accès au catalogue des ouvrages présents dans la base, ne devra pas reprendre le principe retenu pour Gallica dont nous avons souligné les limites dans le paragraphe 3.2.2. En effet, pour pouvoir répondre de manière pertinente à tous les types de requêtes des chercheurs et des non spécialistes, cette interface avec le catalogue devra offrir le plus grand nombre de champs d'interrogation. Actuellement, Gallica propose les possibilités de recherche suivantes : mots du titre, auteur, sujet et recherche libre. Ce choix minimaliste ne saurait satisfaire le chercheur, amené à s'intéresser à toute autorité secondaire, comme l'imprimeur, l'éditeur, l'illustrateur, l'auteur de pièces liminaires, mais aussi aux illustrations, à leur sujet, leur technique. La recherche par sujet pêche par une indexation fautive et incomplète. L'exemple de *La magnificence* le montre clairement, car le sujet « Entrées royales » ramène six réponses⁹², mais pas le livret de 1549 relatant celle d'Henri II à Lyon. Puisqu'il est indispensable de disposer de renseignements bibliographiques avant d'interroger le catalogue de Gallica, il apparaît que cette bibliothèque numérique ne peut pas être considérée comme s'adressant à tous les publics. Quant à la recherche libre, elle est peut-être paradoxalement « trop libre » et les requêtes par ce champ amènent du bruit et ne se révèlent pas réellement exhaustives⁹³.

Ce que nous proposons pour les recherches dans la future base DEBORA, c'est une interface la plus fine possible permettant aux spécialistes, notamment les historiens du livre, d'accéder aux ouvrages pertinents pour leur recherche et à ceux-ci seulement. Voilà qui ne devrait d'ailleurs pas poser de grands problèmes informatiques, puisque le format AdHoc a structuré les ouvrages en utilisant un système de balises prenant appui sur les métadonnées des ouvrages⁹⁴ (voir 3.3). Il serait donc possible d'imaginer une interface avec le catalogue permettant non seulement les recherches par titre et auteur, mais aussi par éditeur, imprimeur, graveur etc., par section du document, comme les épîtres dédicatoires, ou par le sujet des illustrations, ces données permettant aux spécialistes, quel que soit leur domaine, d'interroger avec le

⁹² A savoir une médaille du règne de Henri IV, trois estampes représentant l'entrée du roi et de la reine à Paris en 1660, et deux études historiques datant du XIX^e siècle.

⁹³ Voir le paragraphe 1.2 de la troisième partie et l'exemple de la recherche sur le libraire Rouillé.

⁹⁴ Cf. 1.3 de la troisième partie.

maximum de pertinence le catalogue des ouvrages présents dans la base. Pour reprendre l'exemple traité dans les paragraphes précédents, il devrait être possible de poser une requête telle que : lister tous les ouvrages publiés par Guillaume Rouillé entre 1547 et 1551, imprimés par Philibert Rollet et Barthélemy Frein et dont l'illustration est attribuée à Bernard Salomon. Il est évident qu'un gros travail d'indexation (avec vedettes et formes rejetées) est à prévoir en amont, pour éviter à l'utilisateur de devoir formuler par exemple neuf requêtes afin de trouver une seule entité⁹⁵. On envisagera même de renseigner des champs jusque-là négligés par les catalogueurs, comme ce qui concerne les indications pour le pliage des cahiers. Par exemple, on pourrait savoir, avant même de télécharger des documents au hasard, quels sont les livres de la base imprimés à Lyon entre 1547 et 1551 dont les rectos portent la signature sur les feuillets 1-4 et une réclame sur les feuillets 5-8 (dans le cas de cahiers de 8 feuillets), sachant que tous les versos ont une réclame (dispositif codé A'8 dans les tableaux 4 et C).

On voit combien il est nécessaire de penser l'architecture des notices en termes d'arborescence. Peut-être conviendrait-il ici de s'inspirer de la norme AFNOR Z 44074 pour le catalogage du livre ancien, qui définit plusieurs zones (présentation simplifiée) :

- la vedette auteur
- le pavé de transcription du titre et de l'adresse
 - o titre
 - o lieu
 - o éditeur
 - o date
- la zone de notes sur l'édition
 - o les pièces liminaires
 - o le matériel typographique
 - o l'illustration
 - o les états
- la zone de notes sur l'exemplaire
 - o la reliure
 - o les annotations
- les renvois d'autorités secondaires
 - o préfacier

⁹⁵ Cf. le cas de la recherche sur le libraire Rouillé dans Gallica, tableau 3.

- o illustrateur
- o possesseur

1.4.2. Corpus thématiques et bases associées

- Des corpus cohérents

Il apparaît évident que pour être utile aux usagers, une bibliothèque numérique doit offrir un catalogue abondant. Cependant, si une certaine « masse critique » doit être atteinte, il ne suffit pas de remplir la base avec comme seul objectif de présenter un nombre impressionnant de livres : encore faut-il que les choix répondent à une logique.

Le projet initial de DEBORA se voulait destiné à la fois aux spécialistes du livre du XVI^e siècle et au grand public. Or, il est incontestable que les besoins des premiers étant très ciblés et requérant la mise en place d'outils spécifiques, la reprise du projet doit passer par une définition claire du public visé. Nous nous plaçons résolument dans le cadre d'une bibliothèque numérique destinée aux chercheurs, permettant des travaux de bibliographie matérielle virtuelle, de linguistique ou d'histoire de la gravure. En effet, le format développé par l'INSA permettant des recherches très poussées sur les ouvrages du XVI^e siècle, on tâchera d'en tirer toutes les potentialités en vue d'un nouveau développement du projet.

Aussi, plutôt que de multiplier les ouvrages issus de différents fonds de bibliothèques européennes et présentant les particularités des imprimeries nationales à la Renaissance (option choisie à la base dans le projet européen DEBORA), nous proposons d'élaborer des corpus qui puissent répondre efficacement aux besoins de la recherche. Ils seraient mis en place en collaboration avec des seiziémistes envisageant tous les aspects de l'étude des livres anciens. Il s'agirait de corpus cohérents. L'un d'eux mettrait par exemple à la disposition des chercheurs la production de tous les imprimeurs lyonnais dans les années 1530-1560, ce qui devrait permettre d'établir avec certitude la provenance des impressions anonymes, de saisir, du point de vue de la mise en page, le passage du livre encore médiéval au livre moderne, enfin d'offrir le meilleur de la production littéraire lyonnaise de la Renaissance, depuis le séjour de la cour en 1536 – période de l'apogée de la poésie néo-latine –

jusqu'au sac de la ville par les protestants en 1562, qui mit fin à cette brillante période.

A l'instar de Gallica, ces corpus cohérents pourraient reposer sur un thème fédérateur. Par exemple, une liste d'ouvrages du XVI^e siècle relatant les entrées royales dans les villes du royaume pourrait être proposée.

- Des bases associées

A côté de ces corpus et en guise de complément, on pourrait envisager la constitution de bases associées à la bibliothèque numérique, notamment une base regroupant les gravures des ouvrages. Celle-ci pourrait reprendre le principe de bases de lettrines existant déjà pour le XVIII^e siècle⁹⁶. Une fois les livres de la base DEBORA codés au format AdHoc, et grâce à un balisage fin mettant en œuvre des tags tels que « lettrine », « fleuron », « gravure pleine page » par exemple (et non le simple terme « imagerie »), il serait possible d'extraire ces différents types d'ornements et de les basculer dans une base à part. Ces bases spécialisées étant constituées, elles fonctionneraient indépendamment de la bibliothèque numérique, ce qui permettrait un gain de temps substantiel pour les chercheurs ne travaillant que sur les illustrations des ouvrages, car il ne serait plus nécessaire de télécharger tous les livres dont les ornements seraient susceptibles d'être intéressants, ni d'utiliser la fonction vignettage ou enfin de se constituer une base personnelle de gravures.

Comme pour les bibliothèques numériques, il importe dans une base d'images numérisées de favoriser l'accès de l'utilisateur aux informations et donc de ménager divers types de requêtes. L'interrogation par mots-clefs semble la première étape. Cette recherche passe par l'établissement d'un thésaurus exhaustif permettant de décrire de manière précise tout type d'ornement. Cependant, cette méthode, si elle était la seule possible jusqu'à présent (emploi du thésaurus Garnier, par exemple), pourrait être avantageusement remplacée par les nouveaux outils fondés sur la reconnaissance de formes. En effet, l'indexation des images pose des problèmes insolubles, même lorsqu'on utilise un vocabulaire normalisé : comme le rappelle Silvio Corsini, l'un des

⁹⁶ Mòriane (<http://www.ulg.ac.be/moriane/>, visité le 13/03/2002) ou Passe-Partout (http://www.unil.ch/BCU/docs/collecti/res_prec/fr/todai_intro.html, visité le 13/03/2002) par exemple.

responsables de la base Passe-Partout⁹⁷, « la subjectivité est importante » quand il s'agit de décrire des images et le « risque est donc grand de passer à côté d'un ornement qui a été indexé avec des termes différents de ceux proposés pour la recherche. Par ailleurs, certains ornements non franchement représentatifs sont difficiles voire impossible à indexer »⁹⁸. C'est ici que l'on attend des logiciels de reconnaissance des formes des performances rendant caduque l'indexation la plus subtile. En effet, le projet TODAI (Typographical Ornaments Database and Indexation), qui a été intégré à la base Passe-Partout, permet une recherche automatisée à partir d'un ornement digitalisé (scan ou photo numérique) sans faire intervenir de description textuelle.

Cette base de gravures numérisées pourrait à terme rendre de grands services aux spécialistes de l'histoire du livre en leur permettant de reconstituer virtuellement le stock de bois gravés des imprimeurs, d'établir leur transmission etc. Dans le schéma que nous avons décrit de constitution de corpus homogènes d'ouvrages, la base des gravures augmenterait au rythme du nombre de livres numérisés.

1.4.3. Environnement de recherche

Après avoir mis en place une bibliothèque numérique comprenant un nombre important de titres se rapportant à un même sujet, ou permettant de mener des recherches fondées sur la comparaison des formes (grâce au transcripteur DEBORA, le contenu serait automatiquement doublé en mode texte), il conviendrait de fournir aux chercheurs les outils dont ils peuvent avoir besoin lors de leurs travaux.

Ces usuels, que le spécialiste consultait jusqu'à présent dans les salles des fonds anciens, pourraient faire l'objet de liens intégrés au navigateur DEBORA. Quelques exemples permettront de saisir l'ambition d'un éventuel nouveau projet DEBORA.

Dans le cadre d'une bibliothèque numérique d'ouvrages du XVI^e siècle, il paraît indispensable de fournir un accès aux dictionnaires de langue facilitant

⁹⁷ Banque internationale d'ornements d'imprimerie du XVIII^e siècle fédérant plusieurs bases. http://www.unil.ch/BCU/docs/collecti/res_prec/fr/todai_intro.html#banques (visité le 03/06/2002).

⁹⁸ CORSINI (Silvio), *La base Passe-Partout (ornements typographiques des livres anciens) dans l'optique de l'identification des imprimeurs clandestins*. [en ligne] Disponible sur : <<http://aanla.free.fr/corsini.htm>> (visité le 03/06/2002).

la compréhension de certains mots de moyen français. On pense ici à des liens vers les dictionnaires historiques de Huguet ou de Godefroy⁹⁹ qui, s'ils étaient en ligne, complèteraient les informations données par ceux des contemporains Robert Estienne (1552)¹⁰⁰ et Jean Nicot (1606)¹⁰¹. On pourrait ainsi vérifier l'apparition d'un terme donné, discuter la date avancée par les lexicographes, voire les corriger. Toutes ces recherches lexicographiques portant notamment sur l'emploi plus ou moins fréquent de certains mots dans les ouvrages du XVI^e siècle, sur les néologismes etc. étaient bien sûr déjà réalisées par les chercheurs. DEBORA ne serait pas, dans ce domaine, révolutionnaire, mais le site pourrait leur apporter un confort dans la recherche : d'un simple clic, on vérifierait le sens ou la date de première apparition d'un mot...

Cette aide sémantique pourrait se doubler d'un accès à des dictionnaires biographiques permettant de situer dans le contexte les personnages historiques apparaissant dans un texte. On le voit, le but ultime de cette bibliothèque numérique est de mettre à la disposition du chercheur directement sur son écran tout ce dont il pourrait avoir besoin. Cette perspective étant sans bornes, on peut imaginer des liens vers tout type de sites susceptible d'aider le spécialiste étudiant les ouvrages du XVI^e siècle¹⁰². En reprenant l'exemple d'un corpus d'ouvrages lyonnais du XVI^e siècle, on pourrait envisager la numérisation du répertoire de Baudrier recensant les productions des imprimeurs lyonnais de cette époque. Ce « Baudrier numérique » est sans doute souhaité par de nombreux chercheurs. On pourrait à cette occasion imaginer la numérisation d'un exemplaire particulier auquel chaque bibliothèque ajouterait les annotations sur les exemplaires conservés dans les fonds du monde entier, fruit du travail des bibliothécaires locaux. Cette entreprise serait si gigantesque qu'elle pourrait à elle seule faire l'objet d'un projet particulier.

Cet environnement des ressources disponibles sur le réseau et vers lesquelles pointerait DEBORA afin de satisfaire les besoins des lecteurs savants pourrait s'étendre à des liens vers d'autres sites de bibliothèques numériques permettant de compléter le catalogue de la base. Ces liens seraient tenus à

⁹⁹ Qui ont dépouillé *La magnificence*.

¹⁰⁰ <http://www.lib.uchicago.edu/efts/ARTFL/projects/dicos/ESTIENNE/> (visité le 06/06/2002).

¹⁰¹ <http://www.lib.uchicago.edu/efts/ARTFL/projects/dicos/TLF-NICOT/> (visité le 06/06/2002).

¹⁰² Par exemple celui du Centre d'Etudes Supérieures de la Renaissance <http://www.cesr.univ-tours.fr/Epistemon/> (visité en avril 2002) qui héberge depuis décembre dernier celui de la Société Française d'Etudes du Seizième siècle.

jour par un webmestre à qui les chercheurs pourraient soumettre des titres ou des corpus thématiques à numériser et à intégrer à la base DEBORA.

Enfin, on imagine aussi des liens vers des catalogues en ligne de « bibliothèques réelles ». Ainsi, le spécialiste, après avoir mené virtuellement sa recherche de bibliographie matérielle derrière son écran, pourrait *in fine* savoir dans quelle bibliothèque proche de chez lui se trouvent les ouvrages sur lesquels il a travaillé et *peut-être* se rendre dans la salle des fonds anciens pour faire les ultimes vérifications, rien, fût-ce la meilleure numérisation ne pouvant remplacer le document original.

Comme on le voit, ce projet d'évolution de DEBORA s'apparente pour l'instant à ce que l'on pourrait appeler de la recherche-fiction. *Pour l'instant* certainement. Notons bien toutefois que les bibliothèques numériques sont appelées à se développer et que les outils et le format créés dans le cadre du projet DEBORA apportent à celles-ci des avantages non négligeables et susceptibles d'être encore améliorés. Nous concluons sur un des aspects réellement novateurs : le travail collaboratif. Si pour l'instant, celui-ci semble balbutiant, si les chercheurs sont encore réticents à partager leurs informations avant toute publication, il est possible que la tendance déjà à l'œuvre dans les sciences dures se transpose dans des disciplines telles que l'histoire du livre. La possibilité ménagée dans une bibliothèque numérique comme DEBORA de faire des annotations sur des ouvrages¹⁰³, tout en préservant les droits de l'auteur sur sa découverte, pourrait accompagner, voire initier, un changement de mentalité dans le groupe des chercheurs. En effet, si la publication en ligne d'une information acquérait la même degré de reconnaissance qu'une publication papier, nul doute que les travaux des chercheurs sur les ouvrages via les bibliothèques numériques se développeraient. C'est alors que des bibliothèques numériques « avancées » offrant des outils pertinents et inédits verraient leur excellence reconnue.

¹⁰³ Voir la mise en perspective de cet aspect du sujet par Wagneur 1999, notamment p. 81 pour la « critique génétique » et, d'une manière plus générale, pour la mise « en abyme » des bibliothèques virtuelles, afin de tenter de se représenter ce que fut « l'univers documentaire d'une époque ».

Conclusion

La recherche sur DEBORA menée durant les cinq derniers mois a été pour nous particulièrement riche d'enseignements. Outre l'apprentissage du travail en équipe – et des contingences inhérentes à ce genre d'exercice –, elle nous a permis d'allier notre goût pour le livre ancien à la découverte d'un aspect du web : les bibliothèques numériques. Nouvelles technologies et histoire du livre se conjuguèrent ainsi, offrant une extension intéressante au sein de notre cursus à l'ENSSIB. Cet exercice d'école fut aussi l'occasion de rencontrer à la fois des responsables de fonds anciens – de futurs collègues – et des spécialistes d'autres domaines, comme les informaticiens de l'INSA, dont les réflexions et les points de vue nous ont apporté un éclairage nouveau sur les questions du livre ancien. Le cloisonnement entre ces diverses professions, pourtant amenées à travailler ensemble, nous est alors apparu. Notre statut à part (ni encore vraiment professionnels des bibliothèques ni vraiment connaisseurs des problèmes techniques et informatiques posés par la numérisation) s'est finalement révélé être un atout. Ce sujet de recherche nous a fait prendre conscience du fait que la collaboration entre ces deux professions, pour l'instant trop rare, constitue un élément essentiel : les fruits de la mise en commun de ces compétences complémentaires, si elle voyait le jour, ne manqueraient pas de rejaillir sur le monde des bibliothèques numériques.

Dans l'environnement actuel du livre ancien en ligne, DEBORA apparaît comme un prototype atypique. En effet, laissant de côté les expériences existantes en matière de document numérique, les concepteurs du projet ont d'emblée voulu repenser entièrement la question de l'accès au livre du XVI^e siècle. Il s'agissait en fait de généraliser cet accès et de le faciliter. En outre, au-delà de cette prise en compte des attentes présagées du « grand public », le défi à relever était de fournir aux chercheurs des documents structurés dans lesquels il eût été plus facile de retrouver l'information quel que soit son contenu ou son statut au sein du livre. L'objectif était en somme de « rendre vivant le document », selon l'expression d'Hubert Emptoz, directeur du laboratoire RFV de l'INSA. Une telle conception était novatrice par rapport à celles qui ont présidé à la création de la plupart des bibliothèques numériques actuelles qui ne proposent guère d'outils avancés capables d'aider le chercheur.

L'ampleur de la tâche et le temps imparti pour le projet européen ont conduit à la création d'un logiciel d'expérimentation inadapté à la mise en ligne. Sorte de vitrine de la technologie actuelle en matière de compression des images et de structuration des données, DEBORA propose des outils qui pourraient se révéler à l'avenir très profitables aux chercheurs.

Cependant ces fonctionnalités - indéniablement performantes - testées sur le livret de l'entrée de Henri II à Lyon ne permettent pas, dans l'état actuel des choses, de répondre à l'une des questions qui nous ont été posées, à savoir l'identification de l'imprimeur de cet ouvrage. En effet, si les outils sont riches de potentialités, ils ne peuvent être réellement utiles à une recherche de ce type dans la mesure où ils ne sont pas associés à une véritable bibliothèque numérique comportant un nombre important de livres. L'histoire du livre s'appuie largement sur des méthodes comparatives d'exemplaire à exemplaire. De ce fait, DEBORA, avec sa large gamme d'outils sophistiqués mais sans corpus auquel les appliquer, ne pouvait pas nous aider à répondre au problème posé.

En revanche, Gallica, la bibliothèque numérique de la BNF, grâce au vaste ensemble d'ouvrages proposés, et malgré la relative indigence des outils disponibles, nous a permis, par recoupements successifs de donner un nom à l'imprimeur du livret. Ceci ne fut d'ailleurs possible qu'en combinant les ressources numériques, les ouvrages conservés à la bibliothèque municipale de Lyon et une solide connaissance du livre ancien.

Cette question de bibliographie matérielle a été envisagée comme représentative des attentes des chercheurs à l'égard des bibliothèques numériques. La conclusion qui s'impose et qui devrait servir de *vade-mecum* aux futurs concepteurs de bibliothèques numériques peut paraître d'une banalité évidente : des outils perfectionnés sans livres ne servent à rien, et des masses d'ouvrages numérisés sans possibilité de recherches fines sont tout aussi inutiles, en tout cas pour un public de chercheurs comme les historiens du livre ou de la gravure. La véritable révolution des bibliothèques numériques à l'usage du public savant réside donc dans la combinaison de ces deux idées... ce qui n'a jusqu'à présent pas été réalisé. En fait, cette bibliothèque numérique d'un genre nouveau, souhaitée pour ne pas dire attendue par les spécialistes, pourrait tout à fait naître dans le sillage de DEBORA. En effet, les outils testés ayant été globalement validés, il faudrait désormais associer au navigateur une véritable bibliothèque virtuelle. Grâce à un choix judicieux et cohérent des ouvrages à intégrer dans la base et moyennant quelques améliorations à

apporter au logiciel, nul doute que les spécialistes trouveraient là une aide précieuse pour leurs recherches.

Parmi les chantiers qui s'ouvrent aujourd'hui pour les bibliothèques patrimoniales, la numérisation des livres anciens est certainement l'un des plus importants. Cette technique a de grands atouts car le document sous forme numérique peut être soumis à tous les traitements logiques susceptibles d'aider le chercheur. Toutefois, elle ne saurait être une fin en soi et il serait de peu de profit de « numériser pour numériser ». Ce transfert de l'information d'un support vers un autre doit s'accompagner d'un enrichissement. Cette plus-value ajoutée au livre ancien paraît aujourd'hui techniquement réalisable. Sa mise en œuvre reste néanmoins subordonnée à une collaboration étroite entre informaticiens, chercheurs et bibliothécaires.

Bibliographie

LES BIBLIOTHÈQUES NUMÉRIQUES

ANDRE (Jacques), CHABIN (Marie-Anne) (dir.), *Les documents anciens*, Document numérique vol. 3, n° spécial 1-2/1999, Paris, Hermès, 169 p.

Bibliothèques numériques, cours INRIA (Institut national de recherche en informatique et en automatique), 9-13 octobre 2000, Paris, ADBS éditions, 246 p.

CORSINI (Silvio) , *La base Passe-Partout (ornements typographiques des livres anciens) dans l'optique de l'identification des imprimeurs clandestins*. [en ligne] Disponible sur : <<http://aanla.free.fr/corsini.htm>> (consulté le 3 juin 2002).

CUBAUD (Pierre), GIRARD (Didier), « ABU : une bibliothèque et son public », [en ligne]. Disponible sur : <http://jasmin.cnam.fr:8081/RR9813/> (visité le 11/06/02)

CUBAUD (Pierre), TOPOL (Alexandre), « A WWW-based digital library for antiquarian collections », [en ligne]. Disponible sur : <http://jasmin.cnam.fr:8081/RR9909/> (visité le 22/03/02)

DHERENT (Catherine), *La numérisation dans les archives de France*, Document numérique, 1999, vol .3, n° spécial 1-2, pp. 13-27.

JACQUESSON (Alain), RIVIER (Alexis), *Bibliothèques et documents numériques : concepts, composantes, techniques et enjeux*, Paris, éd. du Cercle de la Librairie, 1999, 377 p. (collection Bibliothèques)

GIRAUD (Yves), « Protocole pour l' édition des textes imprimés en moyen français », Bulletin de la Société Française d'Étude du Seizième Siècle, mars 1997, n°42, p. 37-40.

LALOU (Elisabeth), *La numérisation des manuscrits médiévaux à l'Institut de recherche et d'histoire des textes*, Document numérique, 1999, vol .3, n° spécial 1-2, pp. 29-38.

LE BARS (Fabienne), MUNCK (Gilles), *La communication des documents patrimoniaux*, Protections et mise en valeur du patrimoine des bibliothèques de France : recommandations techniques, Paris, 1998, p. 45-50.

TOURNON (André), « Propositions et discussions », [en ligne]. Disponible sur : <http://193.52.215.193/sfdes/sfdes/Bull45.html#tournon> (visité le 10/05/02).

TRUDEAU (Danielle), « L'apport des nouvelles technologies à l'édition critique ». in : *Les études françaises valorisées par les nouvelles technologies de l'information et de la communication*, 12-13 mai 2000, Toronto [en ligne]. Disponible sur : http://www.sjsu.edu/depts/foreign_lang/Constant/apport.html .

VENTABERT (Gusnard de), *Réflexions sur la représentation des documents anciens : le projet Philectre*, Document numérique, 1999, vol .3, n° spécial 1-2, pp. 57-73.

WAGNEUR (Jean-Didier), *La bibliothèque électronique de la BNF ou Quand un lecteur a le droit d'écrire sur les livres*, Revue de la Bibliothèque nationale de France, n° 2, juin 1999, p. 78-81.

Sites de bibliothèques numériques (cf. annexe 1) :

<http://irht.cnrs-orleans.fr><http://www.promo.net/pg/> ou <http://www.gutenberg.net>

<http://abu.cnam.fr/>

<http://www.bmlisieux.com/>

<http://gallica.bnf.fr/>

<http://cnum.cnam.fr/>

<http://www.cesr.univ-tours.fr/Epistemon/Hyperb.htm>

<http://www.cesr.univ-tours.fr/Epistemon>

<http://zeus.inalf.cnrs.fr>

<http://un2sg4.unige.ch/athena/html/athome.html>
<http://www.chass.utoronto.ca/epc/langueXIX/lexotor/>
<http://www.chass.utoronto.ca/~wulfric/rentexte/fouillou/>
<http://inalf.ivry.cnrs.fr/ccrti>
<http://www-ellug.u-grenoble3.fr/ellug/>
http://www.mpiwg-berlin.mpg.de/Galileo_Prototype/MAIN.HTM
http://www.unil.ch/BCU/docs/collecti/res_prec/fr/todai_intro.html#banques

DEBORA :

BOUCHE (Richard) (dir.), *DEBORA : projet européen n° LB 5608 A*, juin 2000, dactylog., 179 p.

BOUCHE (Richard), MAJID (Ihadjadene), *Le Projet DEBORA*, Revue d'Information Scientifique et Technique, vol. 10, n° 1-2, 2000, p. 55-61.

DEBORA, *Etat de l'art sur les pratiques et sur les usagers des bibliothèques virtuelles*, Délivrable 2.1 juin 1999.

DEBORA, Délivrable 4.2 [en ligne]. Disponible sur: <http://debora.enssib.fr/del42.rtf> (visité le 20 avril 2002).

DEBORA, *Numérisation*, Délivrable 4.5, [en ligne]. Disponible sur : http://debora.enssib.fr/Format_AdHoc.doc (visité le 27/04/2002).

DEBORA, *Critères choisis*, Délivrable 6.1, avril 1999.

DEBORA, *Collection*, Délivrable 6.3, avril 1999.

BIBLIOGRAPHIE MATÉRIELLE :

Aristotelis ethicarum, sive de moribus, ad Nicomachum filium libri decem [...], Lyon, G. Rouillé, 1548, [16]-299-[21] p. [= ARISTOTE 1548].

BAUDRIER (Henri), BAUDRIER (Julien), *Bibliographie lyonnaise : recherches sur les imprimeurs, libraires, relieurs et fondeurs de lettres de Lyon au XVI^e siècle*, Lyon, L. Brun / Paris, A. Picard, t. IX, 1912.

COOPER (Richard), *Maurice Scève, the Entry of Henri II into Lyon (september 1548) : a Facsimilé with an Introduction*, Tempe, Arizona Board of Regents for Arizona State University, 1997, 326 p. (« Medieval and Renaissance Texts and Studies », 160) [inclut trois réimpressions anastatiques entre les p. [160] et [300], dont SCÈVE 1549a et F.M. 1549].

F.M., *Particolare descrizione della comedia fatta recitare in Lione la natione fiorentina a richiesta di sua maesta christianissima*, sig. M-O4, P2 [= F.M. 1549, imprimé avec SCÈVE 1549b].

GALIEN, *De sanitate tuenda libri sex [...]*, Lyon, G. Rouillé, 1548, 375-[37] p.

Bern. Gordonii opus lilium medicinae inscriptum, de morborum prope omnium curatione, septem particulis distributum [...], Lyon, G. Rouillé, 1550, 910-[14] p.

IACONO (Giuseppe), FURONE (Salvatore Ennio), *Les marchands banquiers florentins et l'architecture à Lyon au XVI^e siècle*, Paris, Publisud, 1999, 285 p. (« La France au fil des siècles »).

LONGEON (Claude), *Bibliographie des œuvres d'Etienne Dolet écrivain, éditeur et imprimeur*, Genève, Droz, 1980, 213 p. (« Travaux d'humanisme et Renaissance », 174).

MAROT (Clément), *Les oeuvres de C.M. de Cahors, valet de chambre du roy*, Lyon, E. Dolet, 1543, 272 f.

MARTIN (Henri-Jean), *Mise en page et mise en texte du livre français : la naissance du livre moderne (XIV^e-XVII^e siècle)*, [Paris], Ed. du Cercle de la librairie, 2000, 491 p.

PARADIN (Guillaume), *Mémoires de l'histoire de Lyon*, Lyon, A. Gryphe, 1573, [16]-444-[14] p. (réimpr., Lyon, Dioscor, 1985).

SAULNIER (Verdun-Louis), *Le prince de la Renaissance lyonnaise, initiateur de la Pléiade, Maurice Scève, italianisant, humaniste et poète (ca. 1500-1560) : les milieux, la carrière, la destinée*, Paris, C. Klincksieck, 1948-1949, 2 vol. (th. de doct. lettres).

[SCÈVE (Maurice)], *La magnificence de la superbe et triumpante entree de la noble et antique cité de Lyon faicte au tres chrestien roy de France Henry deuxiesme de ce nom, et a la royne Catherine son espouse, le XXIII de septembre MDXLVIII*, Lyon, G. Rouillé, 1549, sig. A-L4 [= SCÈVE 1549a].

[SCÈVE (Maurice)], *La magnifica et triumphale entrata del christianiss. re di Francia Henrico secondo di questo nome fatta nella nobile et antiqua città di Lyone a luy et a la sua serenissima consorte Chaterina alli 21 di septemb. 1548*, Lyon, G. Rouillé, 1549, sig. A-L4 [= SCÈVE 1549b, imprimé avec F.M. 1549].

[SCÈVE (Maurice)], *La magnificence de la superbe et triumpante entree de la noble et antique cité de Lyon faicte au tres chrestien roy de France Henry deuxiesme de ce nom, et a la royne Catherine son espouse, le XXIII de septembre MDXLVIII : relations et documents contemporains*, éd. Georges Guigue, Lyon, Société des bibliophiles lyonnais, 1927, 373 p.

Tabulae de schematibus et tropis Petri Mosellani ; In rhetorica Philippi Melanchthonis ; In Erasmi Roterodami libellum De duplici copia, Lyon, T. Payen, 1548, 62-[2] p.

TAGAULT (Jean), *De chirurgica institutione libri quinque [...]*, Lyon, G. Rouillé, 1549, [64]-572 p.

TIRAQUEAU (André), *Commentariorum de utroque retractu, et municipali et conventionali, secunda editio [...]*, Lyon, G. Rouillé, 1549, [136]-1023 p.

Table des annexes

ANNEXE 1	I
Tableau d'analyse des sites web de bibliothèques numériques	I
ANNEXE 2	IX
Tableau A. — Alphabet sur fond rayé, avec rinceaux et putti	X
(24 x 24 mm)	X
Tableau B. — Alphabet sur fond criblé, avec rinceaux	XI
(33 x 33 mm)	XI
ANNEXE 3	XII
Tableau C. — Consultation dans Gallica :	XII

Annexe 1

Tableau d'analyse des sites web de bibliothèques numériques

	PROJECT GUTENBERG	EARLY CANADIANA ON LINE	GALILEO GALILEI
Etat civil du site	Visité le 27.03.02	Visité le 27.03.02	Visité le 28.03.02
Adresse	www.gutenberg.net	www.canadiana.org	www.mpiwg-berlin.mpg.de/Galileo_Prototype/INDEX.HTM
Editeur	Michael Hart	Institut Canadien de Microreproductions Historiques	Biblioteca Nazionale Centrale (Florence), Istituto e Museo di Storia della Scienza (Florence), Max Planck Institute for the History of Science (Berlin)
Accessibilité du site	Accès libre	Accès sur abonnement	Accès libre
Date de création	1971	1997	
Date de mise à jour	2002	2002	10.05.99
Corpus			
Nombre de documents numérisés	Inconnu	890.000 pages pour les membres, plus 633.000 pages accessibles pour les non membres	Un volume (ms. Gal. 72 de la BNC de Florence)

Types de documents numérisés	Tout ce qui est exempt de droit d'auteur (avant 1923). Classiques de la littérature, comme Shakespeare, Poe, Rabelais, Montaigne	Imprimés depuis les origines (ex. Voyages de Champlain, Paris, 1613)	Manuscrit scientifique
Texte intégral/Extraits	Texte intégral	Texte intégral	Extrait, fol. 33-196
Type d'approche des textes numérisés	Philosophie du gratuit. Bibliothèque encyclopédique en anglais	Patrimoine imprimé du Canada	Histoire des sciences. Etude génétique des textes
Public visé	Diffusion large	Diffusion large	Chercheurs
Principes éditoriaux	Transcription à partir d'éditions plus ou moins récentes (Rabelais dans une traduction anglaise du XVII ^e s. imprimée en 1838)	Sources de l'histoire nationale, dossiers pédagogiques	Prototype. Interactivité, collaboration entre chercheurs
Textes accompagnés d'un appareil critique	Introduction tirée de l'édition papier	Notices bibliographiques	Longue analyse codicologique (sur la foliotation, les différentes mains, les possesseurs). Précis biographique, histoire du ms.
Notice bibliographique présentant l'ouvrage			Bibliographie du ms.
Aspects techniques			
Mode de numérisation	Mode texte	Mode image	Mode image et mode texte
Qualité de l'image		Bonne	Très bonne restitution du texte
Outils d'exploitation des textes			
Accès au texte et aux illustrations	Recherche par auteur, titre, mots du titre	Texte intégral, titre, auteur, sujet, éditeur, bibliographie	Présentation page par page du ms. Trois accès possibles : liste des folios, indices, « proposition » des <i>Discorsi</i> .
Recherche sur tout le site ou dans toute la base	Non	En mode texte intégral, la recherche d'un mot renvoie aux notices des œuvres avec le n° de page sous forme de liens. Le lien renvoie à une édition du XIX ^e ou du XX ^e siècle (donc usage de l'OCR)	
Recherche dans le(s) texte(s) ou les tables	Texte « au kilomètre », sans pagination. En version .txt, lecture seule ; en version .zip (WordPad), recherche plein texte possible.		Par « proposition », mot ou chiffre
Possibilité de commenter en ligne les textes	Non	Non	Non, mais on peut envoyer des e-mails aux éditeurs

	Athena	BODLEIAN LIBRARY : Browse images of manuscripts	Base Enluminures
Etat civil du site	Visité le 18.04.02	Visité le 28.03.02	Visité le 18.04.02
Adresse	http://un2sq4.unige.ch/athena/html/	www.bodleian.ox.ac.uk/depts/scwmss/wmss/medieval/browse.htm	http://www.bm-lyon.fr (site général avec lien à la base)
Editeur	Pierre Perroud (informaticien et enseignant à Genève)	Bodleian Library / University of Oxford	BM de Lyon
Accessibilité du site	Accès libre	Accès libre	Accès libre
Date de création	1986	Déc. 1999	
Date de mise à jour	La dernière en 2002	25.07.00	
Corpus			
Nombre de documents numérisés	Le site annonce 10000 liens, en partie vers d'autres bibliothèques numériques, pour des livres de disciplines et langues diverses. Le corpus français compte plus de 2000 titres + le corpus suisse + celui des sciences.	Petit nombre de manuscrits.	12000 images réparties en 4200 fiches 50 documents pour le 16 ^e siècle
Types de documents numérisés	Des classiques francophones, des textes genevois, des textes universitaires non publiés	Sélection d'images de mss. du XI ^e au XVII ^e siècle. Pas de textes.	Enluminures du 6 ^e au 16 ^e siècle
Texte intégral/Extraits	Texte intégral	Extraits	Images ou lettres retirées de leur contexte
Type d'approche des textes numérisés	Littérature	Enluminures en ligne	Approche iconographique, histoire de l'art
Public visé	Diffusion large	Diffusion large	Plutôt le chercheur dans ces domaines
Principes éditoriaux	Offrir un large panorama littéraire par accès aux éditions Athena sous forme numérique et catalogage des ressources documentaires des bibliothèques en ligne	Trésors de la bibliothèque	Pas de volonté bibliographique : seuls sont indiqués, pas de description précise du manuscrit ou du livre imprimé
Textes accompagnés d'un appareil critique	Non mais propose des références d'éditions critiques	Non	Non
Notice bibliographique présentant l'ouvrage	Notice avec informations sur l'auteur et références de l'édition numérisée. Rappel des principes de transcription	Notices succinctes	Réduite à sa plus simple expression : auteur, titre, la date et la cote du livre
Aspects techniques			
Mode de numérisation	Mode texte (fichier html ou rtf)	Mode image	Mode image
Qualité de l'image		Faible	Bonne mais texte pas lisible
Outils d'exploitation des textes			
Accès au texte et	Accès au texte par table	Liste des manuscrits.	Résultats en tableau ou

aux illustrations	des matières ou grille listant les numéros de page	dans la table des matières	mosaïque avec liens vers les documents
Outils de visualisation	Non	Vignettage. Loupes	Image en vignette sur la fiche ou en agrandissement pleine page
Recherche sur tout le site ou dans toute la base	Oui, par corpus (suisse, français, œuvres étrangères traduites en français ou l'inverse, œuvres illustrées), auteur, mot-clé ou mot du titre		Oui, par tout ou un ou plusieurs critères suivants : date, cote, titre, auteur, légende, descripteur
Recherche dans le(s) texte(s) ou les tables	Pas de recherche spécifique dans un texte		
Possibilité de commenter en ligne les textes	Non	Non	Non

	CNUM	GALLICA	BM de Troyes
Etat civil du site	Visité le 18.04.02	Visité le 18.04.02	Visité le 18.04.02
Adresse	http://cnum.cnam.fr	http://gallica.bnf.fr	http://www.bm-troyes.fr
Editeur	CNAM	Bibliothèque nationale de France avec compléments de certains éditeurs ou bibliothèques	Bibliothèque municipale de Troyes
Accessibilité du site	Accès libre	Accès libre	Accès libre
Date de création	2001	Octobre 1997	
Date de mise à jour		Fait régulièrement semble-t-il	
Corpus			
Nombre de documents numérisés	Inconnu	80000 documents numérisés pour 15 millions de pages en ligne	La présentation annonce des manuscrits de Clairvaux, des livrets de la Bibliothèque bleue, des photos mais dans les faits que deux ouvrages
Types de documents numérisés	Sciences et techniques, économie et sociologie appliquée	Divers du Moyen-Age au début du XX ^e siècle (416 livres du 16 ^e en mode image, 2 en mode texte)	Un livre d'heure du XV ^e et un livre de chœur du XVI ^e siècle
Texte intégral/Extraits	Texte intégral	Texte intégral	Texte intégral pour l'un, que les enluminures pour l'autre
Type d'approche des textes numérisés	Histoire des sciences et techniques	Approche très souple, pour la détente aussi bien pour le chercheur	Histoire du livre, esthétique
Public visé	Diffusion large	Diffusion large	Diffusion large
Principes éditoriaux	Donner largement accès aux collections spécialisées du CNAM avec un objectif de familiarisation et de sensibilisation aux caractéristiques du numérique	Démocratiser les collections patrimoniales et simplifier l'accès au texte	Mise en valeur d'acquisitions de la BM

Textes accompagnés d'un appareil critique	Bibliographie	Non	Non
Notice bibliographique présentant l'ouvrage	Notice du type catalogue de bibliothèque avec notamment collation et cote CNAM	Notice comportant l'édition numérisée et par qui, la cote, la taille du fichier, le mode de numérisation, la discipline, le numéro de la notice, avec lien vers le document	Petite notice rédigée avec quelques éléments historiques et descriptifs pour l'édition et l'exemplaire (mais pas normalisée)
Aspects techniques			
Mode de numérisation	Numérisation en mode image	Numérisation en mode image ou en mode texte (extrait de Frantext)	Numérisation en mode image
Qualité de l'image	Bonne	Bonne	Bonne restitution du texte
Outils d'exploitation des textes			
Accès au texte et aux illustrations	Accès au texte par chapitre grâce à table des matières, abrégée ou intégrale : une fois dans un chapitre, accès direct aux première et dernière page, feuilletage, choix numérique d'une page Table des illustrations	Accès par table des matières, différents accès par pagination (en tout cas en mode image)	
Outils de visualisation	Divers réglages de la taille de l'image : originale, réduction, agrandissement, automatique	En mode image, barre d'outils avec différents outils de visualisation : masquer la barre, zoom, diverses tailles Possibilité de stocker des signets, vues miniatures et articles	
Recherche sur tout le site ou dans toute la base	Oui, par auteurs ou anonymes	Oui, simple ou croisée, par fonds, titre, auteur, sujet avec possibilité choisir un type de document + les navigations thématiques	
Recherche dans le(s) texte(s) ou les tables	Par mots (un ou plusieurs, début ou fin de mot) mais que dans les tables et les notices	Oui, par le champ Recherche libre (dans les notices, les oeuvres en mode texte et les tables des matières) mais aussi fonction recherche dans un texte (options : mot entier, respect de la casse, recherche en arrière)	Non
Possibilité de commenter en ligne les textes	Non	Non	Non

	EPISTEMON Cornucopie	BIUM Collection Medic@	BM de Lisieux
Etat civil du site			Visité le 18.04.02
Adresse	www.mshs.univ-poitiers.fr/adoni/Epistemon.htm	www.bium.univ-paris5.fr/histmed/medica.htm	http://www.lisieux.com
Editeur	Unité de recherche A.DO.NI (Université de Poitiers Université de Tours (CNRS-CESR))	Bibliothèque inter-universitaire de Médecine	BM de Lisieux
Accessibilité du site	Accès libre	Accès libre	Accès libre
Date de création	1998	Octobre 2000	Existait déjà en 1998
Date de mise à jour	Régulière	Régulière	Régulière
Corpus			
Nombre de documents numérisés	11	105 volumes	2000 documents pour le fonds local + les autres
Types de documents numérisés	Textes de la Renaissance dont : <i>Odes</i> de Ronsard <i>La Saulsaye</i> et <i>Microcosme</i> de Maurice Scève	Rééditions de textes anciens et notamment : Ensemble des éditions <i>princeps</i> du XVI ^e siècle des écrits d'Hippocrate	Fonds local, littérature et documentation du 19 ^e siècle essentiellement
Texte intégral/Extraits	Texte intégral ou extraits	Texte intégral	Texte intégral (souvent bref) ou extrait
Type d'approche des textes numérisés		Histoire de la médecine, philologie, un peu d'histoire du livre	Approche littéraire et linguistique (collaboration avec Lexotor, base canadienne sémantique)
Public visé	Diffusion large	Chercheurs, historiens de la médecine, philologues, médecins	Pas de ciblage précis
Principes éditoriaux	Reproduction la plus exacte possible de l'original et notamment : Respect strict des graphies, de la ponctuation, de la mise en page Exceptions : dissimilation des i/j et des u/v, résolution des abréviations, transformation des lettres ornées en lettres ordinaires	Rééditer « un certain nombre d'ouvrages médicaux depuis longtemps inaccessibles » : Editions anciennes des grands textes de l'antiquité (Hippocrate, Galien...) Textes fondateurs d'une discipline	Donner accès à un contenu peu accessible (brochures, etc.) et/ou peu connu
Textes accompagnés d'un appareil critique	Certains textes font l'objet d'une notice détaillée (remarques sur la structure de l'ouvrage, son sujet, son édition)	Chaque ouvrage est présenté par un chercheur qui présente l'auteur, le contexte, le texte, l'édition Ajout souvent d'une bibliographie sur l'auteur ou le sujet	Non
Notice bibliographique présentant l'ouvrage	Non	Non	Mention de l'édition papier de référence, de la personne l'ayant établie

Aspects techniques			
Mode de numérisation	Mode texte (saisie manuelle)	Mode image	Mode texte
Qualité de l'image		Bonne restitution du texte	
Outils d'exploitation des textes			
Accès au texte et aux illustrations	Par chapitrage	Accès par une table des matières : par page ou par chapitrage	Par lien avec le nom de l'auteur, dans la rubrique correspondante ou dans l'index
Outils de visualisation		Non	
Recherche sur tout le site ou dans toute la base	Non	Non	Oui, par un index des auteurs et anonymes
Recherche dans le(s) texte(s) ou les tables	Non	Non	Non
Possibilité de commenter en ligne les textes	Non	Non	Non

	Association des Bibliophiles Universels (ABU)	Laurent Joubert	ELLUG
Etat civil du site			
Adresse	http://abu.cnam.fr	www.bama.ua.edu/~gd_eroche/joubert/index.html	http://www.u-grenoble3.fr/ellug/livres/Nef_des.Folles/s/
Editeur	CEDRIC Equipe MIHM	Grégory de Rocher, University of Alabama	Bibliothèque municipale de Lyon ELLUG Université Stendhal (Grenoble)
Accessibilité du site	Accès libre	Accès libre	Accès libre
Date de création	1993		1995
Date de mise à jour	Régulière	24/03/2000	Edition achevée
Corpus			
Nombre de documents numérisés	288	3	1
Types de documents numérisés	Oeuvres du domaine public francophone De l'Antiquité au XIX ^e siècle	Textes de Laurent Joubert (1529-1582), médecin ordinaire du roi : <i>Le Traitté des Archusades</i> (1574, Jean de Tournes) <i>La Pharmacopée</i> (1588, Antoine de Harsy) <i>Erreurs populaires</i> (1578, S. Millanges)	Edition de 1500 de <i>La Nef des Folles</i> de Josse Bade Texte original numérisé + traduction en mode texte
Texte intégral/Extraits	Texte intégral	Texte intégral	Texte intégral
Type d'approche	Donner accès à tous		Histoire littéraire

des textes numérisés	aux grands textes		Histoire du livre
Public visé	Grand public	Historiens du livre ou de la médecine	Historiens
Principes éditoriaux	Edition en français actuel	Donner accès aux éditions princeps	Sortir de l'oubli un ouvrage qui a une grande valeur littéraire et artistique
Textes accompagnés d'un appareil critique	Non	Non (envisagé)	Présentation et annotations par Charles Béné, professeur émérite à l'Université Stendhal de Grenoble Index des noms Bibliographie
Notice bibliographique présentant l'ouvrage	Oui (métadonnées)	Non	Succincte
Aspects techniques			
Mode de numérisation	Mode texte (saisie manuelle)	Mode image	Modes image et texte (original et traduction placés en regard)
Qualité de l'image		Restitution moyenne	Bonne restitution du texte
Outils d'exploitation des textes			
Accès au texte et aux illustrations	En texte complet ou par paragraphe	Accès au texte par chapitrage par page Accès aux indexes originaux des ouvrages	Accès aux bois gravés par une liste
Outils de visualisation		Non	Possibilité d'isoler le texte numérisé et la traduction ou les disposer en regard mais aussi d'obtenir séparément tous les bois gravés grâce à une liste
Recherche sur tout le site ou dans toute la base	Possibilité de constituer un corpus de textes sur lesquels travailler		
Recherche dans le(s) texte(s) ou les tables	Recherche par mots/expressions sur un texte en particulier ou sur tout le corpus	Non	Non
Possibilité de commenter en ligne les textes	Non	Non	Possibilité d'envoyer des commentaires et critiques au concepteur du site

Annexe 2

Deux alphabets de lettrines de *La magnificence* identifiés :

Tableau A. – Alphabet sur fond rayé, avec rinceaux et putti (24 x 24 mm)

Tableau B. – Alphabet sur fond criblé, avec rinceaux (33 x 33 mm)

**Tableau A. — Alphabet sur fond rayé, avec rinceaux et putti
(24 x 24 mm)**

			Editeur	Imprimeur	Cote BML
ARISTOTE	1548	8°	Rouillé & Constantin	Rollet & Frein	B 509103
GALIEN	1548	16°	Rouillé & Constantin	Rollet & Frein	811856
<i>Tabulae</i>	1548	8°	Payen	Frein & Rollet	B 509667
<i>Magnifica</i>	1549	4°	Rouillé	?	Rés. 355883 = Gallica
TAGAUT	1549	8°	Rouillé	?	Rés. 444280
TIRAQUEAU	1549	4°	Rouillé	?	Rés. 161196
GORDON	1550	8°	Rouillé	Rollet	402706

	A	B	C	D	E	F	G	H	I	L
ARISTOTE	259		253	71	282			274	199	265
GALIEN	10		3							
<i>Tabulae</i>			3		5					
<i>Magnifica</i>						Dv		Hv		
TAGAUT										
TIRAQUEAU			70		35	61			127	
GORDON	140	103								

	M	N	O	P	Q	R	S	T	V	Π
ARISTOTE			1		126		†2			
GALIEN					15					
<i>Tabulae</i>			44							
<i>Magnifica</i>										
TAGAUT		558		aav	428				169	545
TIRAQUEAU				221		A5	61		36	
GORDON										

Tableau B. — Alphabet sur fond criblé, avec rinceaux (33 x 33 mm)

(Longeon 1980, n° A-2)

			Editeur	Imprimeur	Cote BML
MAROT	1543	8°	Dolet	[Dolet]	- Rés. 318246 - Gallica
Magnificence	1549	4°	Rouillé	?	Rés. 355882 = Gallica = DEBORA
Magnifica	1549	4°	Rouillé	?	Rés. 355883 = Gallica
TAGAUT	1549	8°	Rouillé	?	Rés. 444280
TIRAQUEAU	1549	4°	Rouillé	?	Rés. 161196

	A	B	C	D	E	F	G	H	I	K
MAROT										
Magnificence	E2v		G							
Magnifica	E3		D3	F4					I2	
TIRAQUEAU			A3					Av		
TAGAUT									aa2	
	L	M	N	O	P	Q	R	S	T	V
MAROT	2									
Magnificence	F3v							H4		
Magnifica					G4			H4v		Gv
TIRAQUEAU								35		
TAGAUT						462			1	

NB tableaux A et B.

1) seule est indiquée la première occurrence de chaque lettre (p. ou f.)

2) le dépouillement du Tiraqueau n'a pas été fait de manière exhaustive

Annexe 3

**Tableau C. — Consultation dans Gallica :
ouvrages publiés par Guillaume Rouillé de 1547 à 1551**

N° Gallica	Œuvre	Date	Editeur	Imprimeur	Réclame
Rouillium 1 N 053791	Galien	1547	Rouillé	Roussin & Ausoult	A8
Rouillium 9 N 053931	Galien	"	Rouillé	Roussin & Ausoult	A8
Rouillium 52 N 053856	Galien	"	Rouillé	Roussin & Ausoult	A8
Rouillium 15 N 054098	Galien	"	Rouillé	Roussin & Ausoult	A8
Rouillium 7 N 054177	Hippocrate	"	Rouillé	Roussin & Ausoult	A8
Rouillium 69 N 052602	Tagaut	"	Rouillé	Roussin & Ausoult	A8
Rovilium 1 N 054191	Thriveris	"	Rouillé	?	A'8
Rouillium 68 N 052810	Altomare	1548	Rouillé	Rollet & Frein	B8
Rouillium 16 N 054357	Dubois	"	Rouillé & Constantin	Rollet & Frein	B8
Rouillium 44 N 053904	Fuchs	"	Rouillé & Constantin	Rollet & Frein	B8

Rouillium 4 N 054040	Galien	"	Rouillé & Constantin	Rollet & Frein	B8
Rouillium 53 N 053834	Galien	"	Rouillé & Constantin	Rollet & Frein	B8
Rouillium 58 N 053747	Galien	"	Rouillé & Constantin	Rollet & Frein	B8
Rouillium 2 N 054416	Justinien	"	Rouillé & Payen	?	B8
Roville 2 N 072027	Alciat	1549	Rouillé	Bonhomme	C'8
Rouillium 13 N 054257	Bernieri	"	Rouillé	?	C8
Rouillium 17 N 054356	Dubois	"	Rouillé	Rollet & Frein	B8
Rouillium 26 N 054131	Dubois	"	Rouillé	?	A'8
Rouillium 2 N 054138	Galien	"	Rouillé	Rollet & Frein	B8
Rouillium 14 N 054097	Galien	"	Rouillé	?	B8
Rouillium 25 N 054147	Galien	"	Rouillé	?	B8
Rouillium 38 N 053789	Galien	"	Rouillé	?	B8
Rouillium 43 N 053930	Galien	"	Rouillé	Rollet & Frein	B8
Rouillium 46 N 053885	Galien	"	Rouillé	Rollet & Frein	B8
Rouillium 54 N 053790	Galien	"	Rouillé	?	B8
Rouillium 65 N 053412	Galien	"	Rouillé	Rollet & Frein	B8
Rouillium 42 N 053953	Hippocrate	"	Rouillé	Rollet & Frein	B8
Rouillio 2 N 079020	Magnifica	1549	Rouillé	?	B4

Guillaume					
Rouillé 1	Magnificence	1549	Rouillé	?	B4
N 076015					
Rouillium 56 N 053777	Tagaut	"	Rouillé	?	B8
Guillaume Rouillé 5 N 053776	Tagaut	"	Rouillé	?	B8
Rouillium 1 N 054642	Alciat	1550	Rouillé	Bonhomme	C'8
Rouillium 29 N 054095	Galien	"	Rouillé	Rollet	B8
Rouillium 45 N 053886	Galien	"	Rouillé	Rollet	B8
Rouillium 19 N 054069	Gordon	"	Rouillé	Rollet	B8
Rouillium 57 N 053770	Hippocrate	"	Rouillé	Rollet	B8
Rouillio 3 N 058175	Alciat	1551	Rouillé	?	C'8
Rouillio 4 N 054545	Alciat	"	Rouillé	?	C'8
Roville 1 N 054502	Boccace	"	Rouillé	?	B8
Rouillium 21 N 054215	Galien	"	Rouillé	Rollet	B8
Rouillium 34 N 054044	Galien	"	Rouillé	Rollet	B8
Rouillium 61 N 053832	Galien	"	Rouillé	?	B8
Rovillium 3 N 053746	Galien	"	Rouillé	?	B8

NB tableau C : pour l'explication des sigles de la colonne « Réclame », se reporter ci-dessus au tableau 4.