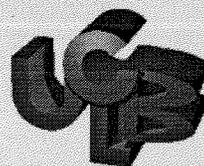


Ecole Nationale Supérieure
des Sciences de l'Information
et des Bibliothèques



Université Claude Bernard
Lyon 1

DESS en INGÉNIERIE DOCUMENTAIRE

Rapport de Stage

Le livre du 16^{ème} siècle à l'aube du numérique

**Production et enquête à la Bibliothèque publique et
universitaire de Genève
dans le contexte de DEBORA**

Enriketa KALLDRËMXHIU BARBEY

Sous la direction de
M Alexis RIVIER
Conservateur Nouvelles Technologies
à la Bibliothèque publique et universitaire de Genève

BIBLIOTHEQUE DE L'ENSSIB

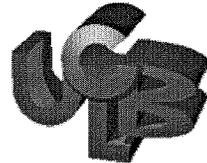


8140346

2000



Ecole Nationale Supérieure
des Sciences de l'Information
et des Bibliothèques



Université Claude Bernard
Lyon 1



DESS en INGÉNIERIE DOCUMENTAIRE

Rapport de Stage

**Le livre du 16^{ème} siècle
à l'aube du numérique**

**Production et enquête à la Bibliothèque publique et
universitaire de Genève
dans le contexte de DEBORA**

Enriketa KALLDRËMXHIU BARBEY

Sous la direction de
M Alexis RIVIER
Conservateur Nouvelles Technologies
à la Bibliothèque publique et universitaire de Genève

2000

Enriketa Kalldrëmxiu Barbey - Le livre du 16^{ème} siècle à l'aube du numérique: Production et enquête à la Bibliothèque publique et universitaire de Genève dans le contexte de DEBORA

Résumé

Les ouvrages du 16^{ème} siècle sont difficiles d'accès. Leur numérisation permet aux bibliothèques qui les possèdent d'ouvrir ces fonds à un public très large et sans limite de distance.

L'objectif du projet européen DEBORA est plus ambitieux encore: la mise en place des postes de travail permettant à l'utilisateur final de travailler interactivement et collaborativement sur ces ouvrages numérisés du 16^{ème} siècle, période à laquelle le livre imprimé acquiert ses caractéristiques modernes.

Ce rapport présente le travail fait à la BPU de Genève, site-test dans le cadre de DEBORA, soit: une enquête sur les besoins et attentes des usagers, une contribution à l'évaluation des coûts et une validation de la chaîne de production du numériseur DigiBook et du serveur TransVision. Ce serveur stocke les images, organise et se présente l'information numérique dans une base de données structurée respectant les métadonnées développées pour DEBORA.

Mots-clés:

Projet, DEBORA, numérisation, bibliothèque numérique, numériseur, livre ancien, métadonnées, base de données, Z39.50, enquête.

Abstract

The 16th century books are not easily accessible. Digitizing them makes it possible for the libraries to open their collections to a very wide readership regardless of distance.

The goal of the European project DEBORA is even more ambitious: the setting up of workstations allowing the user to work in interactive and collective ways on

these digitized 16th century books - this being the period where printed books acquired their modern characteristics.

This report describes the work done at the Geneva Public and University Library, test-site in the framework of DEBORA. It consists of: a survey of the users' needs and expectations, a modest contribution to evaluating the cost and a validation of the production line of th DigiBook and of the TransVision server. This sever keeps in stock the images, organizes and presents the digital information as a structured database, respecting the metadata developed for DEBORA.

Keywords:

Project, DEBORA, digitizing, digital library, digitizer, old book, metadata, database, Z39.50, investigation

Remerciements

Je tiens à remercier toute l'équipe de la Bibliothèque publique et universitaire de Genève, pour l'accueil chaleureux qu'ils m'ont réservé et en particulier:

Alain Jacquesson, son directeur, qui m'a permis de faire mon stage dans cette prestigieuse institution;

Alexis Rivier, mon directeur de stage, pour son aide et son soutien durant l'ensemble du stage;

Marianne Tsioli Bodenman pour sa collaboration sur les métadonnées des livres du 16^{ème} siècle.

L'équipe des chercheurs du Musée historique de la Réformation et de l'Institut d'histoire de la Réformation (MHR/IHR) de Genève et les bibliothécaires de la BPU qui ont répondu au questionnaire.

Les personnes avec qui j'ai travaillé pour la numérisation et la description des ouvrages avec DigiBook et TransVision, notamment:

Patrick Desfarge, administrateur de la base DEBORA sur TransVision;

Jean-Luc Bouchier, photographe de la BM Lyon et parfait connaisseur du numérisateur DigiBook.

Ce stage n'aurait pas eu lieu sans l'aide et les conseils de MM. Bouché et Dalhoumi auxquels je tiens à adresser mes sincères remerciements.

Sommaire

AVANT PROPOS - DE "TOUS LES OUVRAGES" À EN "TOUT UN OUVRAGE"7

I. DEBORA, LA BPU DE GENÈVE ET LE LIVRE DU 16^{ÈME} SIÈCLE 8

1. DEBORA et ses objectifs 8
2. Les partenaires du projet et la Bibliothèque publique et universitaire de Genève 9
3. La mission de la BPU dans ce projet 10
4. Pourquoi numériser les livres anciens? 11
5. Et pourquoi les livres du XVI^{ème}? 11
6. GENEVE au 16^{ème} siècle 12
7. La BPU de Genève 14
8. Le Livre au XVI^{ème} siècle 14
9. L'enthousiasme d'une bibliothécaire à propos de DEBORA 17
10. Critères de choix des livres à numériser pour DEBORA 19
 - 10.1. Les critères de la BPU20

II. LA DESCRIPTION DU LIVRE DU 16^{ÈME} SIÈCLE ET LES MÉTADONNÉES 23

1. Définition 23
2. De l'ouvrage aux images 24
3. Accès aux bases distantes 24
4. Les métadonnées de DEBORA 25
 - 4.1. Les métadonnées du 1^{er} niveau25
 - 4.2. Les métadonnées du 2^{ème} niveau29
 - 4.3. Les métadonnées du 3^{ème} niveau32
 - 4.4. Les métadonnées de 4^{ème} niveau : les annotations36

III. LA CHAÎNE DE PRODUCTION DEBORA - ACQUISITION 38

1. La numérisation 38
 - 1.1. Qu'est ce que la numérisation?38
 - 1.2. Pourquoi numériser?39

2.	DigiBook: Scanner à caméra sur portique	40
2.1.	Qualité minimale requise	41
2.2.	Erreurs constatées	45
2.2.1.	Manque de pages	45
2.2.2.	Pages blanches	45
2.2.3.	Les pages noires et les demi-pages	46
2.3.	Recommandations techniques pour la qualité des images DEBORA	46
2.3.1.	Résolution	46
2.3.2.	Nombre de nuances	46
2.3.3.	Stockage des images	47
2.3.4.	Sauvegarde d'images en niveaux de gris	47
2.3.5.	Format d'image TIFF	47
2.4.	Choix des normes	48
2.4.1.	Normes de numérisation	48
2.4.2.	Normes de stockage	48
3.	Différents avis sur le support de numérisation	50
4.	Autres pratiques de numérisation et apport de DEBORA	52
5.	Autres projets de bibliothèques numériques	53
6.	Analyse des coûts	54
IV. LA CHAÎNE DE PRODUCTION DEBORA - TRAITEMENT		58
1.	Le logiciel TransVision	58
1.1.	De l'imagerie archéologique à un système de gestion de banques d'images	58
1.2.	Utilisateurs, administrateurs, producteurs de TransVision	58
1.3.	Logiciels clients "natifs"	59
1.4.	Le ROBOT	61
1.4.1.	Fonctionnement du Robot	62
1.4.2.	Evaluation du temps de traitement au moyen du robot	65
1.5.	L'architecture client-serveur de TransVision	65
1.6.	L'interface de TransVision	65
1.7.	Saisie et mise à jour des données	66
1.8.	Entité et sous-entités	66
1.9.	Pré-indexation	69
1.10.	Autres opérations dans TransVision	70
1.11.	Le travail sur le livre : les métadonnées du 2 ^{ème} niveau	70
1.11.1.	Etude du contenu	70
1.11.2.	Création des sous-entités	72
1.11.3.	Sélection des pages dans le panorama et transfert dans les sous-entités	72
1.11.4.	Vérification des pages et corrections éventuelles	72
1.11.5.	Copie de la page du titre dans le panorama	72
1.12.	La consultation en mode W3	73
2.	Le client DEBORA	75
2.1.	La norme Z39.50	75
2.1.1.	Protocole d'interrogation client-serveur	76
2.1.2.	Rôle du logiciel client	77
2.1.3.	Rôle du logiciel serveur	77

2.2. Interactions entre composants logiciels	79
V. RÉSULTATS D'ENQUÊTE	82
1. La population	82
2. Objectifs de l'enquête et hypothèses préalables	83
3. Méthode: une enquête par questionnaire	83
4. Analyse	84
4.1. Partie "profil de l'utilisateur" (questions 1 à 14)	84
4.2. Partie "usage du livre ancien" (questions 15 à 23)	85
4.3. Partie "usages du livre du 16 ^{ème} siècle" (questions 24 à 45).....	86
4.4. Partie "livres en ligne et technologie de l'information et de la communication" (questions 46 à 70)	89
4.5. Partie "communications et échanges autour du texte/livre numérique du 16 ^{ème} siècle" (questions 71 à 83)	90
CONCLUSION	93

Avant propos - De "tous les ouvrages" à en "tout un ouvrage"

La base de données MédiuM, conçue dans les années 1970 à l'IRHT, voulait atteindre le rêve utopique de cumuler dans une même base toutes les informations sur les manuscrits médiévaux. Le rêve de tout chercheur est d'avoir accessible sur son poste l'ensemble de tous les catalogues de manuscrits du monde.

Elisabeth LALOU. La numérisation des manuscrits médiévaux à l'Institut de recherche et d'histoire des textes. Document numérique. Volume 3. n. 1-2/1999. p. 29 à 38.

Aujourd'hui, avec la base de données **DEBORA**, les chercheurs ont la possibilité, avec quelques questions habilement posées, de trouver immédiatement des parties d'ouvrage (chapitre, paragraphe, etc.), des illustrations (lettrines, bandeaux, ...) ou encore des sujets sur lesquels porte son intérêt, au lieu d'une longue et fastidieuse recherche dans les fichiers et les ouvrages eux-mêmes.

Le rêve d'un catalogue exhaustif ou d'une base de données universelle de tous les ouvrages anciens n'est pas mort. Il a simplement évolué, vers un objectif encore plus ambitieux. On imagine aujourd'hui le chercheur posséder sur son écran non seulement une liste de *tous les ouvrages*, mais *tout un ouvrage*, soit le texte lui-même, l'objet ultime de son intérêt.

I. DEBORA, la BPU de Genève et le livre du 16^{ème} siècle

1. DEBORA et ses objectifs

En réponse à un appel d'offre de l'Union européenne -DGXIII-, l'Enssib est chargée depuis début 1999 de piloter le projet **DEBORA** (**D**igital **accE**ss to **BO**oks of **RenA**issance). Son objectif est de développer des outils permettant d'accéder à distance à des collections numérisées du XVI^e siècle appartenant à plusieurs bibliothèques européennes.

Conçu dans le cadre d'un vaste programme européen d'application télématique dans les bibliothèques, le projet a pour ambition d'offrir un accès distant à un certain nombre d'ouvrages du XVI^e (peut-être 100), entièrement numérisés pour un nombre d'utilisateurs, et de catégories d'utilisateurs croissants.

Les objectifs particuliers de ce projet sont multiples: produire une analyse des besoins et des usages générés par cet accès distant à des collections patrimoniales et, parallèlement, de mettre au point une définition de la chaîne de production des documents numérisés dans ses aspects techniques et économiques afin de fournir aux bibliothèques des indicateurs précis concernant les coûts et des éléments d'aide à la décision. Sur le plan technique DEBORA réalisera le prototype sur lequel vont être testés la chaîne de numérisation, les procédures d'accès et le poste client.

La mission du projet DEBORA est de faciliter l'accès des chercheurs à ces collections. On peut estimer que l'offre constituée par cet accès distant aux collections numérisées va élargir le nombre de ses utilisateurs avec la possibilité offerte aux historiens de comparer à distance les caractéristiques graphiques liées à l'activité des différents imprimeurs dans plusieurs pays. Mais l'enjeu de DEBORA

est aussi de rendre ces collections accessibles à un public plus large que les seuls chercheurs et érudits.

DEBORA est conduit sous la responsabilité de Richard Bouché, professeur à l'Essib et de l'équipe de recherche SII (Systèmes d'information et interfaces). La réalisation finale est prévue en Juin 2001 ; budget : 500 000 euros.

2. Les partenaires du projet et la Bibliothèque publique et universitaire de Genève

Ce projet est mené en partenariat avec les universités lyonnaises, l'INSA (Institut national des sciences appliquées), les Universités de Lancaster et de Lisbonne, ainsi que les bibliothèques possédant les fonds à numériser. Les questions de numérisation bénéficient de la coopération des sociétés Xérox et de SGBI-Lyon. La Biblioteca municipal de Zaragoza, la Bibliothèque publique et universitaire de Genève (BPU), , ainsi que la Stadtbibliothek München testeront auprès de leurs usagers les résultats obtenus.

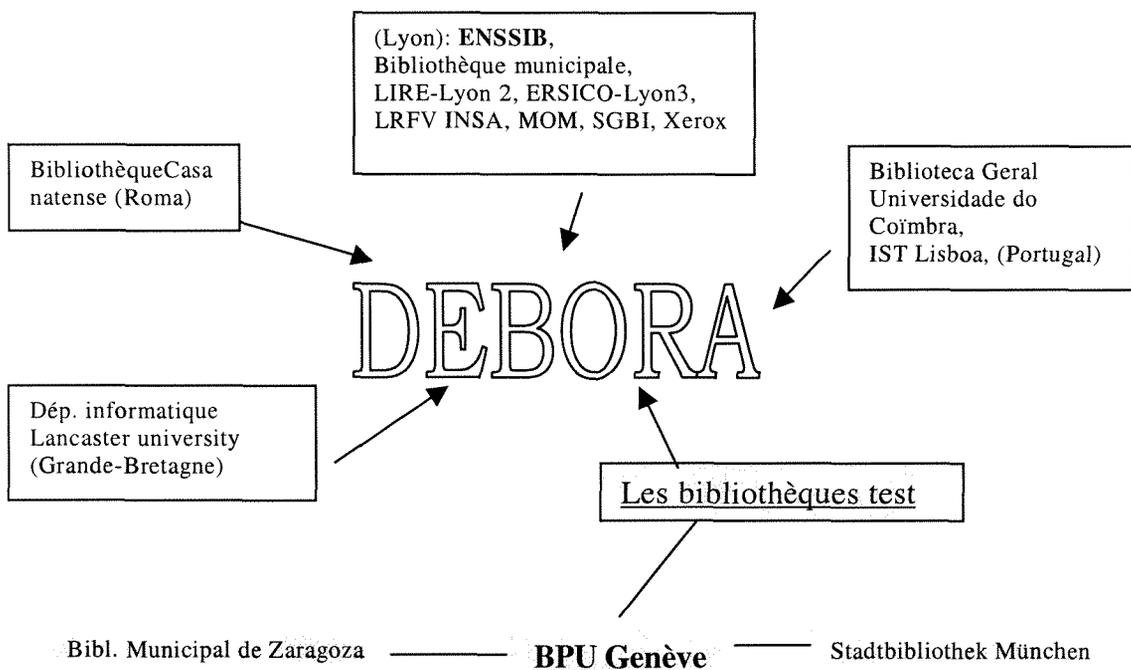


Fig. 1: La BPU de Genève et les partenaires de DEBORA

3. La mission de la BPU dans ce projet

Dans le cadre de DEBORA, la Bibliothèque publique et universitaire de Genève est donc un site test au même titre que deux autres bibliothèques:

Bibliothèque Municipale de Zaragoza
Stadtbibliothek München.

Sa mission dans ce projet en cours de réalisation, est d'appliquer et mettre en place l'interface de consultation d'ouvrages numérisés du 16^{ème} siècle prévue par le projet, de manière adaptée aux besoins de ses chercheurs, ainsi que d'enquêter sur leurs attentes d'un tel système.

La réalisation de cette mission passe par les points suivants:

- Numérisation, description et indexation de 3 ouvrages de la BPU du 16^{ème} siècle;
- Installation du prototype de consultation sur des postes de travail publics;
- Enquêtes auprès des usagers genevois du livre ancien, au moyen d'un questionnaire avant la mise en place du prototype, et d'entretiens approfondis après l'installation de ce prototype pour:
 - Vérifier les besoins généraux des utilisateurs du livre ancien, et leurs pratiques sur les documents électroniques et numérisés;
 - Initier les chercheurs au prototype de consultation de DEBORA et recueillir leurs réactions.

4. Pourquoi numériser les livres anciens?

Tout le monde sait, en effet, que la consultation des livres anciens dans les bibliothèques est réservée aux experts du livre, aux bibliophiles connus, aux spécialistes de diverses disciplines, souvent professeurs et chercheurs d'université, aux doctorants accrédités par lettre de recommandation de leurs enseignants. Tout cela pour garantir la conservation d'un patrimoine précieux, souvent fragile, abîmé, et même exposé au risque de vol du document entier ou partiel: quelques pages ou certaines parties intéressantes comme les illustrations ou les ornements. La disponibilité d'un corpus de livres entièrement numérisés pourra être élargie à chaque étudiant de tout type d'école, lycée ou université, qui lui était jusqu'à présent refusé. Il pourra accéder directement aux sources de sa culture, connues de lui jusqu'ici seulement au travers des seuls noms d'auteurs, de titres d'ouvrages ou de leur contenu général. Pour les experts, déjà usagers ou fervents partisans de l'accessibilité en ligne il donnera corps au rêve de la consultation à distance de tous les documents, où qu'ils se trouvent, tout en s'affranchissant des coûteuses reproductions photographiques ou sur microfilms, et en bénéficiant dans le cas de Debra de nombreuses possibilités de recherche nouvelles.

5. Et pourquoi les livres du XVIème?

Le choix de numériser en mode image les ouvrages du XVIème siècle, en plus d'être une indiscutable valorisation des fonds anciens des bibliothèques, a été fait parce que c'est à partir de ce siècle que le livre imprimé s'est distingué du livre manuscrit et que l'art typographique, tout en restant artisanal, s'est normalisé. Enfin c'est pendant la Renaissance que chaque langue et chaque culture européenne s'est instituée de manière originale et a produit l'immense patrimoine d'ouvrages en langues vernaculaires qui représente les sources de notre savoir moderne.

Les collections choisies dans le cadre de DEBORA sont d'une grande valeur patrimoniale, mais pour des raisons de conservation, elles ne peuvent être communiquées qu'à un nombre restreint de chercheurs. Il s'agit de fonds appartenant à la Bibliothèque municipale de Lyon, la Bibliothèque Casanatense de Rome et la Bibliothèque générale de l'université de Coïmbra (Portugal) et des trois bibliothèques test (citées plus haut).

6. GENEVE au 16^{ème} siècle

Une cité éprise de liberté

Au XVI^e siècle, Genève est une ville prospère. Carrefour commercial sur les routes d'Europe, elle constitue un lieu de rencontre où des marchands allemands font connaître les idées de la Réforme luthérienne dès 1525.

On prête une oreille attentive à ces discours, non pas seulement à cause de leur contenu religieux, mais aussi parce qu'ils véhiculent un message libérateur sur le plan politique.

Genève, ville du Saint Empire romain germanique (l'aigle de son blason), n'a d'autre seigneur que son évêque (la clef sur ses armoiries). Ce dernier, allié au duc de Savoie, a perdu sa popularité : on le considère en ennemi, car le duc nourrit l'ambition de devenir maître à Genève.

Des patriotes, comme Philibert Berthelier et François Bonivard, ont ravivé un esprit d'indépendance dans la cité. Menacés par le duc de Savoie, trompés par leur évêque, les Genevois cherchent des alliés du côté des villes suisses. Tourner ses regards vers la Suisse, c'est se rapprocher de la Réforme luthérienne et repousser la menace de mainmise savoyarde.

L'urgence d'une réforme

A ces raisons locales particulières de s'intéresser à la Réforme, s'ajoutent des motifs plus généraux.

Au début du XVI^e siècle, le malaise était devenu général au sein de l'Eglise régnante : elle avait glissé sur une pente dangereuse. Ses hauts dignitaires, ambitieux et despotes, n'avaient plus devant les yeux l'exemple de Jésus-Christ, dont tous auraient dû s'inspirer. L'urgence d'une réforme apparut plus clairement à tous ceux qui lisaient la Bible et les Pères de l'Eglise. La Renaissance avait, à cette époque, créé un regain d'intérêt pour la littérature ancienne. Facilitée par l'invention de l'imprimerie, l'approche des livres anciens amena de nombreux intellectuels à étudier non seulement des ouvrages profanes mais les Ecritures saintes, l'Ancien et le Nouveau Testaments dans leur texte original.

C'est ainsi que des circonstances religieuses et politiques ont joué un rôle important et ouvert Genève à des influences nouvelles.

L'importance du livre à Genève

Le livre est pour Genève, plus que pour toute autre ville, d'une importance capitale, d'abord parce que c'est lui qui a permis à la Réforme de s'imposer, ensuite parce qu'une ville d'études ne saurait se passer de livres.

A partir de la Réforme, Genève inonde la France, l'Angleterre, les Pays-Bas de bibles, de psautiers et de livres de théologie. Les imprimeurs genevois éditent encore des œuvres célèbres au Siècle des Lumières mais ne retrouvent pas l'activité qui fut la leur aux 16^e et 17^e siècles.



7. La BPU de Genève

La Bibliothèque publique et universitaire de Genève est fondée sous l'influence de Jean Calvin dans la mouvance du Collège et de l'Académie inaugurés en 1559. Mais ses origines remontent à la Réforme, adoptée par le peuple en 1536. Légiférant en matière d'imprimerie, la République nouvellement créée ordonna en 1539 aux imprimeurs de remettre le premier exemplaire de toute publication à la Maison de Ville, donnant ainsi naissance au dépôt légal et à un premier fonds de livres. Le dépôt légal genevois est chronologiquement le second au monde après la France (1537)¹. Le premier catalogue manuscrit de la Bibliothèque de Genève, établi en 1572 recense 723 ouvrages en 554 volumes. Il montre que le fonds s'est accru essentiellement par des dons et par achat de livres provenant notamment des bibliothèques personnelles de Calvin, de l'humaniste florentin réformé Pierre Martyr Vermigli et du patriote, chroniqueur et bibliophile François Bonivard.

Le 5 août 1557, le célèbre imprimeur Robert *Estienne* offre au Conseil de Genève un exemplaire de tous les livres qu'il avait imprimés depuis 1551 en "suppliant de prendre en bonne part pour faire une librairie".

Dès 1562, la bibliothèque est installée au Collège Calvin pour y rester jusqu'en 1872, année de son déménagement dans une aile de la nouvelle Université, à la promenade des Bastions où elle se trouve actuellement.

8. Le Livre au XVIème siècle

Mise au point par Gutenberg vers 1450, l'imprimerie s'est très vite diffusée dans l'Europe entière. La nouvelle technique a transformé **l'aspect du livre** : si certains incunables imitaient encore les manuscrits, le moulage des caractères typographiques en métal a permis de mieux individualiser les lettres, d'harmoniser et d'aérer la mise en page.

¹ Les bibliothèques à Genève : petit essai de chronologie, par Alain Jacquesson et Etienne Burgy, septembre 1997 (non publié)

Au *début du XVIème siècle*, le livre n'a pas réussi à se dégager complètement de l'imitation du manuscrit. Si, par un bonheur merveilleux, il est parvenu presque du premier coup à la perfection typographique, perfection rendue possible par les recherches de plusieurs générations de calligraphes, sa décoration reste empreinte d'une certaine inexpérience; elle se suffit rarement à elle-même et demande encore l'intervention du miniaturiste ou du rubricateur.

Par la combinaison de tirets, de bandeaux, de vignettes, d'encadrements mobiles les éditeurs peuvent varier à l'infini l'aspect de la page imprimée et renouveler à peu de frais les éditions antérieures.

Après une trentaine d'années de recherches, apparaissent des hommes d'un goût délicat qui vont transformer l'aspect du livre, le débarrasser de tout ce qui rappelait le travail de la plume et lui donner les caractères qu'il gardera jusqu'à nos jours.

C'est donc dans le *second tiers du XVIème siècle* que se place le point culminant de l'art de la décoration du livre. La page imprimée réunit alors tous les éléments de perfection: qualité du papier, netteté d'impression, élégance des caractères, proportions des marges et des divers éléments décoratifs, lettres "fleuries", bordures, vignettes et fleurons, jointes à la beauté du dessin dans l'illustration.

Bientôt, la réussite n'est plus qu'accidentelle. La recherche de caractères nouveaux aboutit à des fautes de goût, la qualité des papiers décline, l'illustration se complique, les titres s'alourdissent, le sens des proportions se perd.

Dans le *dernier tiers du XVIème siècle*, de nombreux graveurs au burin vulgarisèrent cette technique qui, par sa rapidité d'exécution, rivalisait avec la gravure en relief sur bois et répondait mieux aux besoins croissants de l'imprimerie.

Durant tout le cours du XVIème siècle, le livre n'aura cessé de se transformer.

Au sujet de **l'illustration du livre**, des problèmes délicats se posent. Les livres du XVIème siècle dont on connaît avec certitude l'illustrateur, ceux surtout dont les gravures sont signées, sont bien peu nombreux en regard de tous ceux, souvent remarquables, dont la décoration reste anonyme. Ce qui complique encore les choses, c'est qu'au XVIème siècle, les artistes gravaient rarement eux-mêmes leurs dessins.

Avec la diffusion croissante des livres imprimés, les impératifs de vente entraînent la création de pages particulières, portant l'état-civil du livre : vers 1550, **la page de titre** comporte tous ses éléments, titre, nom de l'auteur, "marque" puis adresse et nom de l'imprimeur-libraire, année d'impression.

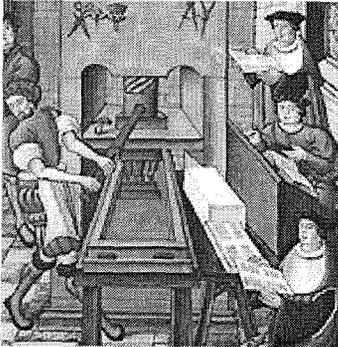
Le sort de **l'humanisme** fut lié à l'apparition puis au développement du livre.

Au siècle précédent, les humanistes avaient recherché fiévreusement les manuscrits anciens, surtout grecs et latins. Les rassembler, les déchiffrer et les commenter avait constitué une bonne part de leur activité.

Ce que diffuse d'abord l'imprimerie, ce sont les textes anciens; du XVème au XVIème siècle, les catalogues des bibliothèques montrent bien leur augmentation. Durant la première moitié du XVIème siècle, l'imprimerie rend accessible à un vaste public, dans toute l'Europe, l'antiquité latine, puis la grecque et - dans une certaine mesure également - l'hébraïque.

Parmi les idées nouvelles que servit l'imprimerie, il faut mettre au premier rang **la Réforme**. Celle-ci est préparée d'abord par la traduction en langue vulgaire de l'Écriture Sainte; Lefèvre d'Étaples, au début du siècle, met à la disposition du peuple chrétien un texte sûr et accessible; outre les Évangiles et les Épîtres, le Psautier devient un livre de piété, où chaque psaume est accompagné d'un bref argument pour prier et comprendre la prière. Quand la Réforme devient militante, l'imprimerie répand un très grand nombre de libelles, de pamphlets, d'exposés des principes et des applications de la foi. La page imprimée fut le premier de ces moyens de communication de masse, qui ont fini maintenant par dominer nos vies.

L'appel anonyme, adressé davantage à l'émotion qu'à la raison, devint possible. Le placard imprimé devint une arme d'activité politique et joua son premier grand rôle lors de la propagation de la Réforme.



A Paris ou à Lyon, centres principaux de l'édition en France, naissent de véritables dynasties **d'imprimeurs-libraires**. Simon de Colines, et les familles Estienne, Bade, De Tournes, etc., sont à l'origine de productions souvent luxueuses ou savantes qui concevaient la typographie comme un art, apportant un soin tout particulier au dessin des caractères.

Fig 2 : Imprimeurs, Chants royaux sur la conception couronnés au puy de Rouen, 1519-1528

9. L'enthousiasme d'une bibliothécaire à propos de DEBORA²

Ada Corongiu, bibliothécaire spécialiste du fonds du XVIème siècle à la Bibliothèque Casanatense de Rome, avec qui j'ai eu le privilège de travailler sur la description des livres de cette bibliothèque pour corpus DEBORA, exprime l'enthousiasme que lui procure ce projet:

"Il y a deux raisons à cet enthousiasme : la première est d'avoir la possibilité de sortir d' une atmosphère très particulière d'appartenance à un monde séparé, presque aristocratique, celui des anciennes bibliothèques ; la seconde, qui me concerne personnellement, est la perspective d'avoir à disposition des instruments de travail sûrs, à l'aide desquels il est enfin possible de vaincre le sentiment du bibliothécaire face à tout type de description du livre ancien: celui de frustration dû à la conscience de l'impossibilité de rendre parfaitement compte de chaque caractéristique et élément du livre décrit dans une fiche normale de bibliothèque. Je ne peux pas oublier qu'au début de ma carrière de bibliothécaire (il y a 20 ans),

² Ada Corongiu est coauteur du premier chapitre de DEBORA, juin 2000.

après avoir décidé avec le directeur Alfredo Serrai - très attentif aux problèmes de la bibliographie matérielle en vue de la parfaite identification des éditions - de commencer un catalogage du fond ancien du XVI^{ème} siècle, suivant des règles particulières tirées de l'ISBD(A), dans la perspective d'une future mise en mémoire des données, nous nous sommes rendu compte tout de suite que toute description est incomplète (même celle des annales typographiques) et que, par exemple, la transcription et l'abrègement du titre et d'autres éléments de la page de titre, constituent une espèce de violence: souvent on se disait que la seule photo du frontispice, unie à la description bibliographique, aurait rendu justice au livre et identifié l'édition.

Un problème très important qui concerne la conservation du livre se pose : s'il est vrai qu'un livre ancien, exclu de la large consultation, gît comme une chose morte dans la poussière d'une bibliothèque, il est aussi vrai qu'au moment où il est soumis à n'importe quel genre de reproduction, il est en danger, qu'il s'agisse de microfilmage ou de photographie, parce que ces deux techniques nécessitent un fort éclairage et une ouverture forcée et violente du volume. Mais rien n'est pire que la photocopie, même si dans ce domaine on observe des améliorations. Le projet DEBORA prévoit d'utiliser un appareil projeté de façon à limiter les risques et les contacts avec le livre. Même si cela est dur à dire, je pense qu'il vaut la peine de prendre un minimum de risques si le patrimoine du livre ancien devient ainsi accessible à tout le monde sans nécessité de déplacement et avec une fonctionnalité aussi vaste que le permet la recherche.

Je pense que la tâche du bibliothécaire expert en livres anciens et conscient des exigences des usagers, sera de favoriser une numérisation en mode image «intelligente», en indiquant le nombre le plus large de champs nécessaires pour renseigner sur les parties et les caractéristiques du livre (soit comme édition, soit comme exemplaire). Ceci concerne tout genre de recherche, et surtout détermine ce qui est le plus apte à figurer dans un corpus en fait d'images particulières, à consulter séparément : p. ex. un corpus des portraits, gravures sur cuivre ou bois, blasons, cartes géographiques, marques typographiques, initiales, reliures, etc. En

particulier un regroupement par matière ou genre littéraire rendrait le plus grand service aux chercheurs de tout type[COR 99]."

10. Critères de choix des livres à numériser pour DEBORA

Le corpus de DEBORA est constitué d'un échantillon de 30 à 40 ouvrages pris dans chacune des trois bibliothèques participantes, c'est-à-dire la Bibliothèque municipale de Lyon, la Bibliothèque Universitaire Geral de Coimbra, et la Bibliothèque Casanatense de Rome, en ajoutant ici les livres des bibliothèques tests, notamment les trois de la BPU de Genève.

Le but est de numériser des ouvrages significatifs des collections de ces bibliothèques et de valoriser leur patrimoine.

Les **principaux critères**³ retenus pour le choix des ouvrages sont:

Des critères culturels : l'intérêt de l'ouvrage quant à son contenu et à sa forme:

- Langue
- Contenu thématique
- Esthétique;

Des critères d'usage : l'intérêt de l'ouvrage pour différents usagers

- Objet de l'activité de l'utilisateur;

Des critères de technique documentaire : l'intérêt de l'ouvrage quant à ses caractéristiques techniques;

Des critères économiques : l'intérêt de l'ouvrage quant à son impact économique;

Des critères techniques pour étudier les contraintes de la numérisation.

³ Pour ceux qu'intéressent plus particulièrement les critères de sélection on renverra à :

- <http://debora.enssib.fr>
- DEBORA : projet européen : n°LB 5608 A, sous la coordination de R.Bouché. [Lyon], juin 2000. 167 p.

Un autre élément a été considéré comme important : la présence de tables des matières et /ou d'index, toujours dans la perspective d'expérimenter, à travers la reconnaissance de la forme d'un mot ou d'un groupe de mots de l' index, la possibilité pour le lecteur d'accéder directement à une page particulière, à une partie du livre ou à un argument précis.

10.1. Les critères de la BPU

En tant que bibliothèque test, il a été décidé de numériser trois ouvrages seulement; le choix était donc plus difficile, vu la richesse de la BPU en ouvrages de cette époque.

La motivation principale était d'offrir en consultation au moins quelques ouvrages concernant directement la problématique du XVIIe s. étudiée à Genève. Du fait que ces ouvrages sont numérisés à l'extérieur, en l'occurrence à la Bibliothèque municipale de Lyon (avec le système DigiBook), ce sont des *doublets* qui ont été choisis pour minimiser les risques de pertes.

Pour le critère de "langue" la BPU, contrairement à la BM-Lyon et à la bibliothèque de Rome qui ont privilégié leurs langues vernaculaires, a plus "osé". Pour les trois ouvrages, trois langues apparaissent : *le français, le latin et le grec* (pour voir les contraintes de numérisation qui se posent avec le grec).

Un des critères historico-culturels pour la BPU était l'édition: les trois ouvrages ont été *édités à Genève* (voir: Genève au 16^{ème} siècle). En plus, la présence d'un *auteur genevois* de l'époque: Théodore de Bèze (selon les enquêtes réalisées, une partie des chercheurs actuels travaillent sur la correspondance de Th. de Bèze).

Les ouvrages sélectionnés intéressent potentiellement un grand nombre d'utilisateurs.

Pour ce qui est de *l'aspect esthétique*, la BPU ne l'a pas pris en considération (ces ouvrages ne sont pas aussi riches en images que ceux d'autres bibliothèques).

Pour les critères techniques, l'accent a été mis sur *le format*. Trois formats différents ont d'ailleurs été retenus afin de voir les différences de durée dans la manipulation et l'incidence sur les coûts de la taille et du nombre de pages.

BIBLIOGRAPHIE

[BRU30] BRUN, Robert.

Le livre illustré en France au XVIIe siècle.- Paris : F. Alcan, 1930.-336 p.

[COR99] CORANGIU, Ada.

Description des lignes générales du Projet Debora. Choix des livres. Possibilités de recherche. in: Actes du colloque Vers une nouvelle érudition : numérisation et recherche en histoire du livre, Rencontres Jacques Cartier, Lyon, décembre 1999
http://histoire.enssib.fr/6archives/Colloques_journees/Nouvelle_erudition/corongiu.html (visité le 04.08.2000)

[DUR85] DUREAU, Jeanne-Marie, MERLAND, Marie-Anne.

La fabrication du livre. Villeurbanne : ENSB, 1978. 34 p. 29 cm

[GUE98] GUEDON, Jean-Claude.

La bibliothèque virtuelle: une antinomie? : conférence prononcée à la National Library of Medicine, Washington. 1998
<http://linux.univ-lyon2.fr/nlm-fr.html#anchorintro> (visité le 30.08.2000)

Introduction à la journée d'étude sur la documentation informatisée au sein du ministère de la culture et de la communication. 16 novembre 1998
http://www.culture.gouv.fr/culture/mrt/numerisation/fr/seminaire_du_161198/introduction.htm (visité le 05.06.2000)

[LAL99] LALOU, Elisabeth.

La numérisation des manuscrits médiévaux à l'Institut de recherche et d'histoire des textes. *Document numérique*. Volume 3. n. 1-2/1999. p. 29 à 38.

[MEN68] MENAGER, Daniel.

Introduction à la vie littéraire du XVI^e siècle.- [Paris] : Bordas, 1968
<http://gallica.bnf.fr/themes/LivXVI.htm> (visité le 07.08.00)

[ORI87]ORIEUX, Madeleine

Imprimeurs/libraires XVIe-XVIIIe siècles : guide pour la rédaction des notices d'autorité en format InterMarc / [document établi par Madeleine Orioux...et al.]. [Paris] : Bibliothèque nationale, Dép. des livres imprimés, Dép. des entrées étrangères, 1987. IV, 66 p., [5] leaves of plates. (Etudes, guides et inventaires. No. 8)

[TSI00] TSIOLI BODENMANN, Marianne.

Henri Estienne : 1531-1598. *Info Gutenberg, mars 2000*, n.14 Fribourg : Gesellschaft der Freunde des Gutenberg Museums (Association des amis du musée Gutenberg).

II. La description du livre du 16^{ème} siècle et les métadonnées

1. Définition

Les metadata (ou métadonnées) peuvent être définies comme étant des données relatives à d'autres données (data about data : données sur des données). Par conséquent, une notice catalographique classique peut être considérée comme un ensemble de métadonnées. Le terme metadata est surtout utilisé pour désigner l'information concernant des fichiers de données "lisibles par machine", donc ce terme désigne en quelque sorte une information référentielle sur des données électroniques.

Utilisée dans le contexte de l'information numérique géospatiale qui est à l'origine du terme, *metadata* est conçue comme *l'information fondamentale qui décrit le contenu, la qualité, les conditions et autres propriétés et caractéristiques des données.*

Le choix des métadonnées est une phase importante dans tout développement d'une bibliothèque numérique. Il préfigure les modes d'accès et de navigation qu'on offre aux usagers.

Dans DEBORA, deux niveaux de descriptions sont choisis:

- un catalogage au niveau de la collection en format MARC et respectant la norme ISBD (A); une seule notice permet d'accéder à plusieurs pages-images;
- une deuxième description plus fine à l'unité pour exprimer les particularités de chaque page-image (lettrines, colophons, mots-clés, etc.)

Une relation est établie en stockant le nom du fichier dans un champ de la notice descriptive MARC. Cette relation sera gérée par encodage de la zone 856 de ce format. Ce choix de description s'explique par la nécessité d'une numérisation en mode image des livres du 16^{ème} siècle.

2. De l'ouvrage aux images

Dans le cas des catalogues en ligne, la création de la zone 856 du format USMARC a fortement facilité l'accès aux documents numériques. Cette zone établit une relation entre la notice et un document qui peut se trouver en dehors du catalogue.

Suivant le format de stockage choisi (ici MARC) , il est nécessaire de déterminer la structure du document. La consultation d'un document devient beaucoup plus aisée lorsqu'il est possible, à partir du sommaire, d'accéder directement à la page souhaitée. La structure peut être représentée linéairement ou de façon hiérarchique. Il est bien évident que la consultation restera plus aisée avec une structure hiérarchique. Cette structuration offre les mêmes avantages dans DEBORA que l'utilisation de langages de balisage, par exemple pour la mise à disposition de la base sur un serveur W3.

3. Accès aux bases distantes

Le protocole Z39.50 implémenté au-dessus du protocole de communication TCP/IP, permet l'interrogation simultanée de banques de données par l'envoi d'une requête unique à *plusieurs serveurs distants* sans se préoccuper de l'interface propre au module de recherche utilisé par les usagers directs de ces banques de données. Le protocole Z39.50 et les *services* qu'il fournit prévoient non seulement la navigation dans les bases de données, mais également le parcours de listes d'index sur un serveur. L'annexe 1 est un exemple de conversion entre les différents champs d'un format Marc et ceux supportés par Z39.50. Une conversion des champs de DEBORA vers ceux de Z39.50 Bib1 a été effectué.

Par exemple :

Debora	Unimarc	Z39.50
Authors	\$700 ; \$701 ; \$702	Authors name (1003)
Title	\$200	Title (4)
Date of publication	\$100	Date-publication (31)
Subjects	\$600;601;602;604;605;606;608	Subject (21)
Language	\$101	Code--language (54)
Note	\$316;317;321	Note (63)
Exemplaires	\$140; 141	Note (63)

Exemple de conversion entre les champs de DEBORA et ceux de Z39.50 ⁴

4. Les métadonnées de DEBORA

Les métadonnées choisies pour DEBORA se répartissent en quatre niveaux:

4.1. Les métadonnées du 1^{er} niveau

Elles reprennent celles que l'on trouve habituellement dans le catalogage classique effectué dans les bibliothèques:

Author (Auteur) : La personne ou l'organisation principale responsable de la création du contenu intellectuel du document.

Title (Titre) : Nom donné au document par l'auteur ou l'éditeur.

Place of publication (Lieu de publication) : Nom de la ville où a eu lieu la publication.

Publication date (Date de publication) : Date à laquelle le document a été mis à disposition dans sa forme actuelle.

Publisher (Editeur – Imprimeur) : L'entité qui a mis à disposition du public le document dans sa forme actuelle.

Language (Langue) : Langue du texte du document.

⁴ Entre les champs de DEBORA et ceux de Unimarc voir Annexes

Collation (Collation) : Description physique du livre qui détaille la pagination, la numérotation des feuillets, les signatures, le format, et les planches, les illustrations, les bandeaux, etc. Elle permet ainsi de s'assurer que le document est bien complet.

Call number (Cote et localisation) : Identificateur dans la classification locale de la bibliothèque où se trouve l'ouvrage.

Fingerprint (Empreintes) : Caractères du livre permettant d'identifier une édition par référence à une base-source.

Notes (Notes) : Toutes les remarques soit sur l'édition, sur les caractéristiques de l'exemplaire, soit sur le contenu.

Secondary author (Auteurs secondaires) : Tous les auteurs en dehors de l'auteur principal (préfacier, collaborateur, traducteur, etc.).

Subject (Sujet) : Le thème du document, exprimé habituellement par des mots-clés ou des phrases qui décrivent le sujet ou contenu du document.

Les métadonnées sont écrites en anglais en vue des traitements informatiques qui sont faits sur ces champs. Ainsi le choix d'une langue commune et internationale est primordial pour assurer également une bonne exécution des programmes de traitement des données contenues dans les différents champs.

Voici les trois ouvrages de la BPU avec leurs métadonnées du 1^{er} niveau :

1

Author	Bèze, Théodore de
Title	Theodori Bezae Annotationes majores in Novum Dn. Nostri Jesu Christi Testamentum : in duas distinctae partes, quarum prior explicationem in quatuor Evangelistas et Acta Apostolorum, posterior vero in Epistolas et Apocalypsin continet ; quibus etiam adjuncti sunt indices rerum ac verborum locupletissimi
Place of publication	[Genève]
Publisher	[Jerémie des Planches]
Publication date	1594
Collation	570 [i.e. 574], [1bl.], [34], [1bl.], 648, [52] p. : ill. ; 8° (20 cm)
Notes	Marque employée, entre autres, par J. Des Planches Texte disposé en colonnes Signatures : A-Z@8, Aa-Nn@8, t-v@8, x@@2, a-z@8, aa-rr@8, ss@4, [croix]@8, BB-CC@8, DD@@2

	Empreinte : o-e- n-l- i-i- ve<te> (3) M.D.XCIIII. Références : CDM 139 ED./VERSION : Nova autem haec editio multo correctior e emendatior priore, nunc primum hac minore forma in gratiam et utilitatem studiosorum in lucem prodit
Subject	Bible, Nouveau Testament
Language	Latin
Call number	S 14789 – Bibliothèque publique universitaire Genève

2

Author	Gentillet, Innocent
Title	Discours sur les moyens de bien gouverner et maintenir en bonne paix un royaume, ou autre principauté : divisez en trois parties, asavoir, du conseil, de la religion, et de la police que doit tenir un prince : contre Nicolas Machiavel Florentin : à très-haut et très-illustre prince François duc d'Alençon, fils et frère de roy
Place of publication	[Genève]
Publisher	[François Estienne?]
Publication date	1576
Collation	[15], 639, [17] p. ; 8° (17 cm)
Notes	Imprimé à Genève par François Estienne d'après E. Droz " Sig. [stilet]@8, A-Z@8, Aa-Ss@8 Références: CDM et article de E. Droz dans BHR t. 23(1961), p. 576
Subject	Politique
Language	Français
Fingerprint	tscy s,n- l,de ense (3) M.D.LXXVI
Call number	S 6403 Bibliothèque publique universitaire Genève

3

Author	Anonyme
Title	Virtutum encomia sive gnomae de virtutibus : ex poetis et philosophis utriusque linguae / graecis versibus adjecta interpretatione Henrici Stephani
Place of publication	[Genève]
Publisher	Henri Estienne
Publication date	1573
Collation	[32], 255, [1 bl.] p. ; 16° (14 cm)
Notes	Sig. [oeil de mouche]-2[yeux de mouche]@8, a-q@8 Références: Renouard, A.A. (Annales de l'imprimerie des Estienne), 2° éd., p. 139 Ex-libris Bibliothèque Mr. Dupan, avocat
Subject	Morale
Language	Grec, Latin
Fingerprint	s,re <n.n.> <phy>S. MiIn (3)ANNO M.D.LXXIII
Call number	cd 308 w – Bibliothèque publique universitaire Genève

Les recherches dans la base de données portent d'abord sur les métadonnées du 1^{er} niveau. Au moment de l'affichage des pages numérisées du livre, l'utilisateur a tous les éléments du modèle pour la page de titre. Par contre, au niveau des autres pages, on ne conserve que les trois éléments suivants : titre, auteur et date de publication.

Les figures ci-dessous illustrent comment ces types de pages se présentent dans Transvision, le logiciel de structuration et de présentation de bases de données utilisé dans le cadre de DEBORA

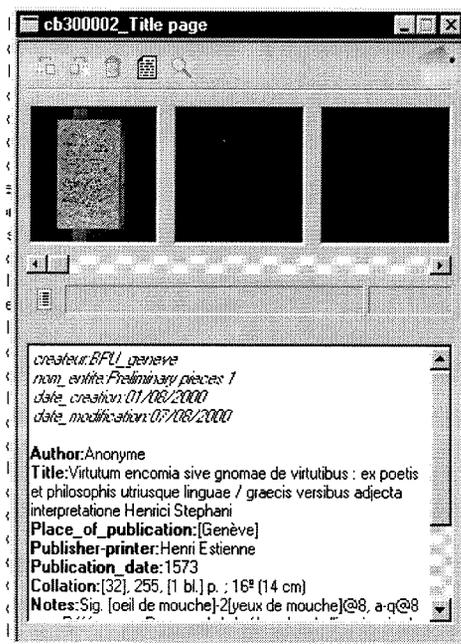
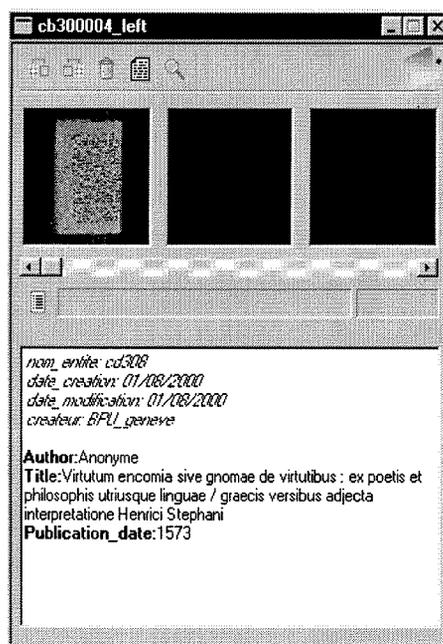


Fig. 3 Page de titre avec les métadonnées du 1er niveau



Une page du corps du texte contenant seulement les champs Auteur, Titre et Date de publication

4.2. Les métadonnées du 2^{ème} niveau

Les métadonnées du 2^{ème} niveau permettent de représenter la structure du contenu du livre.

Même si tous ne sont pas présents dans un livre, nous énumérons ci-dessous tous les éléments que DEBORA prend en compte pour ce deuxième niveau. L'ordre de ces éléments peut varier d'un ouvrage à l'autre.

Pages de titre

Elles contiennent trois types d'information:

-Titre, sous-titre

-Auteur (s)

-Edition : nom, lieu et marque typographique (iconographique, de type varié) de l'éditeur-libraire et parfois de l'imprimeur avec la date.

Frontispice

-Composition gravée ou imprimée comprenant généralement aussi la mention

abrégée du titre et placée généralement face à la page de titre (définition française).

- page de titre (définition italienne)

Pièces préliminaires

-Approbation : Accord de publication donné par un censeur

-Avertissement : Information sur l'ouvrage ou bien sur l'auteur

-Avis aux lecteurs

-Catalogue

-Dédicace : Texte en début d'ouvrage exprimant l'hommage qu'un auteur fait de son œuvre à quelqu'un : comporte le nom du dédicataire, le texte, puis le nom de l'auteur de la dédicace, parfois le lieu et la date de la dédicace

-Épitaphe

-Ex-libris

-Floraison : Petits textes sur l'ouvrage ou sur l'histoire littéraire

-Lettre

-Pièce de vers : Hommage à une personne vivante

-Portrait : Représentation gravée ou imprimée de l'auteur

-Préface : Texte placé en tête d'un livre et qui sert à le présenter au lecteur

-Privilège : Texte faisant état de la permission et propriété des droits d'imprimer attribuée par une autorité pour une période donnée à un personnage donné (typographe, éditeur, auteur) parfois avec mention de rénovation du privilège et/ou de son partage

-Prologue

-Autres textes variés d'auteurs variés (incipit, sonnet, quatrain, poème)

Pages de texte

L'ensemble du texte, y compris les pages avec illustration, bandeaux ou lettrines.

L'illustration peut être sur bois ou sur cuivre. Elle peut remplir toute la page. Au-dessous et au-dessus d'une illustration, il peut y avoir un titre et une nomenclature (ou autre chose) qui peut être imprimé ou faire partie de l'illustration. Il peut y avoir des ornements typographiques, des fleurons.

Hors texte

-Planches, cartes repliées, illustrations collées à la main.

Index

-Récapitulation par ordre alphabétique des sujets abordés dans l'ouvrage, ou des annotations.

Tables

-Table alphabétique

-Table des auteurs

-Table des cartes

-Table des chapitres

-Table des choses

-Table des citations

-Table des épigrammes

-Table des livres

-Table des matières : Liste, avec indication de pages, de noms, de sujets,

- Tables des noms
- Table des œuvres
- Table des sommaires

Pièces finales

- Approbation** : Accord de publication donnée par un censeur
- Catalogue** : Liste de livres écrits par le même auteur ou édités par le même imprimeur
- Colophon** : Indications du lieu, du nom de l'imprimeur et de la date de l'impression
- Élégie** : une sorte de poésie
- Errata** : Erreurs à corriger
- Explicit** : Rappel du titre de l'ouvrage et de l'auteur (à vérifier)
- Hors-texte** : Élément non compris dans les signatures ; peut être dépliant
- Marques typographiques** : marque de l'éditeur ou de l'imprimeur
- Privilège** : Texte faisant état de la permission et propriété des droits d'imprimer attribuée par une autorité pour une période donnée à un personnage donné (typographe, éditeur, auteur) parfois avec mention de rénovation du privilège et/ou de son partage
- Registre** : Table des signatures (indication d'assemblage des feuillets) ou des mots
- Autres textes variés d'auteur variés** (excipit, quatrain, poème)

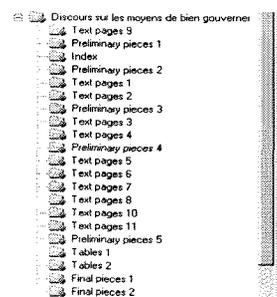


Fig. 4 : Les métadonnées du 2^{ème} niveau dans TransVision

4.3. Les métadonnées du 3^{ème} niveau

Il s'agit ici des éléments de description au niveau de la **page** d'un document ancien. DEBORA a prévu⁵ une description au niveau des pages de chaque livre, qui correspond aux métadonnées du 3^{ème} niveau. Au cours de cette étape, il y a identification, soit manuellement par le producteur de la base, soit automatiquement, (par un algorithme de reconnaissance des formes), de certains éléments, par exemple, les bandeaux, la foliation, les lettrines etc. Les éléments de ce niveau sont des éléments de base pour aider les chercheurs, les "seizièmistes", à étudier ces livres et à identifier avec une certaine précision la période de leur impression.

Bandeau

-Pour séparer les chapitres, marquer le début d'un paragraphe

-Élément décoratif de forme rectangulaire, qui orne le haut des pages. Il est en général gravé sur bois et parfois signé d'initiales



Fig.5: Bandeau de chapitres dans "Discours sur les moyens de bien gouverner" (BPU Genève, S 6403)

Colonnes : Une ou plusieurs colonnes

Cul-de Lampe : Motif décoratif ou figure gravée ou typographique centrée en fin de chapitre.



Fig. 6 : Cul-de Lampe (BPU Genève, S 6403)

⁵ La description du 3^{ème} niveau était encore en cours de développement en été 2000. Les laboratoires Reconnaissance des Formes et Vision de l'INSA à Lyon et l'Instituto Superior Tecnico de Lisbonne travaillent sur des processus de traitement, de compression et d'indexation de l'image.

Dorure : Décoration et titrage de la reliure d'un livre par l'application à chaud, à l'aide de fers et de poinçons, d'une fine pellicule dorée

Enluminure : Lettre ornée et de petite dimension illustrant les feuillets d'un livre ou d'un manuscrit

Estampe : Toute image réalisée au moyen d'un élément d'impression (en creux, en relief, etc.)

Exlibris : Vignette que les bibliophiles collent au revers des reliures de leurs livres et qui porte leur nom ou leur devise

Filet : Trait imprimé, simple ou décoratif

Foliotation : Numérotation des feuillets d'un livre. Seul le recto du feuillet est numéroté. On parle donc de feuillet recto – verso. Les incunables comme les manuscrits sont foliotés (ils peuvent également ne pas porter de numérotation)

Frontispice : Illustration ou une gravure qu'on trouve généralement en regard de la page de titre

Garnitures : Bandes de diverses largeurs servant à l'espacement en impression typographique. Elles constituent les marges et le blanc de "tête" ou de "pied"

Illustration : Images, gravures, cartes, médaillons, etc.

Incipit : Terme qui désigne les premiers mots d'un manuscrit ou d'un livre

Lettre d'attente : Lettre minuscule mise en "attente" dans un manuscrit pour que le rubricateur dessine une initiale plus importante. On trouve aussi des lettres d'attente dans les premiers imprimés où les initiales étaient également rajoutées à la main. Ensuite ces initiales furent gravées sur bois et imprimées en même temps que le texte

Lettrine : Lettre décorée, ornée, historiée

-Grande capitale souvent ornée en début de paragraphe

-Lettre sur bois : lettre gravée sur bois le plus couramment

-Lettre sur cuivre : on les appelle lettres grises reconnaissables à la cuvette



Ligne de pied : Abrégé du titre, souvent de type Tome II

Manchette : Elles peuvent être sous forme de références ou de notes explicatives brèves placées en marge du texte. Dans les manuscrits, ces manchettes étaient constituées souvent du dessin d'un bras ou d'un doigt attirant l'attention sur un passage, d'où ce nom de manchettes.

Marque typographique : Une composition emblématique ou héraldique adoptée par un éditeur, un imprimeur ou un libraire comme marque commerciale. Ces signes graphiques se trouvent généralement sur la page de titre, au verso du dernier feuillet ou à ces deux endroits. Parfois, la marque du libraire est présente sur la page de titre et celle de l'imprimeur au colophon.

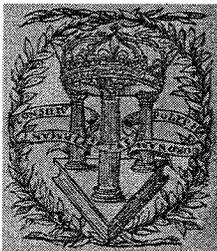


Fig. 7 : Marque typographique (BPU Genève, S 6403)

Miniature : Terme qui signifiait à l'origine, toute lettre ornée et colorée en rouge au minium⁶ (d'où le mot "miniature"). Ensuite cette dénomination s'est généralisée pour désigner tout ornement écrit ou peint avec une encre de couleur. C'est un terme qui pourrait également s'appliquer aux peintures fines de petite taille, dessinées sur les manuscrits anciens.

Notes : Texte écrit en bas de page ou dans la marge

Pagination, numérotation : Jusqu'en 1550, deux façons de paginer coexistent. La première, celle des signatures en deux parties en bas à droite, est formée d'une ou de plusieurs lettres indiquant la place du cahier dans la suite des cahiers formant le livre, et d'un chiffre indiquant la place de la feuille dans le cahier. Le deuxième mode de pagination, l'utilisation des chiffres en haut à droit (le verso ne porte pas de chiffre), apparaît progressivement au cours du XVIème siècle.

⁶ Peinture rouge, à l'oxyde de plomb

Panier de fruits ou fleurs ou encore fleurons : éléments décoratifs utilisés pour signaler la fin des chapitres. Ils peuvent également être sur la page de titre à la place de la marque.



Fig. 8: Fleuron (BPU Genève, S 6403)

Pied de mouche⁷ : Signe de ponctuation équivalent au signe de paragraphe (le même signe que celui des claviers modernes). Il était utilisé pour attirer l'attention sur un passage et pour délimiter un paragraphe. Il est également employé parfois par le rubricateur à des fins décoratives.



Fig. 9: pied de mouche



œil de mouche (BPU Genève, S 6403)

Réclame : Inscription au bas d'une page du mot ou du début du mot de la page suivante. Cet élément qui se trouve déjà dans certains manuscrits a été utilisé jusqu'au début du 19^{ème} siècle

-On peut avoir des réclames à chaque page, ou le plus fréquemment à la fin de chaque cahier. La réclame est un repère qui permet de collationner les feuillets et le cahier d'un livre avant de le brocher ou de le relier

Réglure : Lignes tracées sur la page par le copiste avant d'écrire son texte manuscrit.

Signature : Lettre ou signe que l'on met au bas de la page imprimée au-dessous de la dernière ligne et qui est destinée à faire connaître l'ordre des cahiers et des pages qui le composent au relieur. Les cahiers liminaires sont notés soit avec des signes, soit par la succession des voyelles indiquées en minuscules. Pour le reste du volume on utilise l'alphabet latin majuscule (le I et le J et le U et le V étant

⁷ Sur indication de Mme M. Tsioli-Bodenman, spécialiste du livre ancien à la BPU, on a ajouté ce signe, manquant dans les éléments de description du 3^{ème} niveau de DEBORA, mais caractéristiques des ouvrages de cette époque-là.

confondus), s'il y a plus de 24 cahiers on redouble les lettres (AA ou Aa; BB ou Bb), parfois il faut même tripler la lettre. Les lettres sont suivies d'un chiffre qui indique la place de la page dans le cahier. Généralement, seule la première moitié du cahier est signée, en comptant le nombre des feuillets d'un cahier signé.

Texte : Le corpus textuel du livre.

Titre courant: Titre de l'ensemble du livre ou du chapitre porté au haut de chaque page et se répétant d'une page à l'autre, dans la marge de tête ou la marge de pied. Certains titres courants ont parfois la forme d'un bandeau.

Titre de départ : Titre se trouvant dans la marge de tête de la première page du texte d'un livre.

Typographie : Le type de lettres s'exprime en hauteur de 20 lignes, du bas de la queue au haut de la haste, suivie de G = gothique, R = romain et I = italique, capitales ou bas de casse.

Vignette : Illustration gravée, qu'on trouve généralement sur la page de titre.

Le problème des métadonnées de 3^{ème} niveau se pose en terme d'usage et de coût. Le niveau de détail de la description pouvait être très variable suivant le type d'usager. Pour un historien du livre qui veut identifier avec précision un objet livre, plus les indices seront nombreux plus il sera satisfait. Pour d'autres, le niveau de détail est beaucoup moins important. D'autre part produire de l'information pour chaque page, parfois avec une description très fine par des experts, est complètement irréaliste au niveau du temps et du coût.

4.4. Les métadonnées de 4^{ème} niveau : les annotations⁸

De nombreux seiziémistes, au cours de leurs travaux, sont conduits à écrire des "notes" relatives aux ouvrages qu'ils étudient; et, ainsi, ils peuvent apporter des connaissances nouvelles utiles pour eux-mêmes ou pour d'autres lecteurs.

⁸ Dans le cadre de DEBORA, l'Université de Lancaster est chargée de la conception du poste de travail ou client Debora, intégrant également une gestion des annotations des chercheurs.

Par suite, DEBORA a été amené à définir un niveau 4 de métadonnées qui, ne seront pas fournies par la bibliothèque mais par les usagers. Ce niveau de description fonctionne par le moyen des annotations.

Grâce à une gestion des annotations, il est possible à plusieurs usagers travaillant ensemble ou de façon isolée d'enrichir la description des documents au niveau de la page, de se constituer un corpus personnel local et structuré.

Cette approche fait de DEBORA un outil de recherche en particulier sur l'histoire du livre et le livre du 16^{ème} siècle.

BIBLIOGRAPHIE

[AME00]AMEROUALI, Youcef ; IHADJADENE, Majid.

Métadonnées et bibliothèque numérique. Lyon : UCBL, ENSSIB ; Paris : Paris 10.
DEBORA : D 6.3
<http://debor.enssib.fr/publications.htm> (visité le 24.07.2000)

[BEL00] BÉLISLE, Claire; BOUANZI, Latifa.

Métadonnées et livres du XVII^e siècle : Debora, WP6 / Version 20.05.2000.
[Lyon], 2000.

ISBD(CF) : description bibliographique internationale normalisée des documents électroniques = international standart bibliographic description for computer files / Fédération internationale des associationd de bibliothécaires et des bibliothèques. Paris : Bibliothèque nationale de France, 1996. 96 p. ; 30 cm ISBN 2-7177-1983-0

[MAN99] Manuel Unimarc : format bibliographique; version française / [Fédération Internationale des Associations de Bibliothécaires et des Bibliothèques, IFLA Universal Bibliographic Control and International MARC programme, Deutsche Bibliothek, Frankfurt am Main]. Trad. par Marc Chauveinc. 3^{ème} édition. München : Saur, 1999. 427 p. : 24 cm. (UBCIM publications : Programme CBU-MI)

III. La chaîne de production DEBORA - Acquisition

L'objectif de DEBORA étant de développer des outils permettant l'accès , à partir de postes de consultation distants, à des collections de documents du 16^{ème} siècle des bibliothèques par la numérisation des ouvrages, une chaîne de production de documents numérisés numérisés spécifiques a été définie et mise en place.

1. La numérisation

Avec Gutenberg, l'humanité a vécu la révolution de l'imprimerie. Au cours du siècle précédent, microformes et photographies ont facilité l'accès à la consultation. La numérisation constitue aujourd'hui une alternative, et ouvre de nouveaux horizons pour le livre et l'écrit.

1.1. Qu'est ce que la numérisation?

Une numérisation est une transformation de l'image en une information basée sur un ensemble fini de nombres "compréhensibles" par un ordinateur. On passe alors d'un espace continu à un espace discret (l'image est composée d'un ensemble fini de points nommés *pixels*). Plus l'image a de pixels, plus celle-ci sera nette, plus elle véhiculera d'informations. Chaque pixel représente soit une valeur de noir ou de blanc, soit un niveau de gris ou encore une nuance de couleur. La résolution représente la densité de pixels dans l'image, exprimée en points par pouces (notée dpi ou ppp). Naturellement, plus l'image est dense, plus elle occupe de place dans la mémoire de l'ordinateur, plus elle est donc coûteuse à stocker; par contre elle est alors beaucoup plus précise et fidèle.

1.2. Pourquoi numériser?

Pour sauvegarder le patrimoine

Parce qu'il existe des ouvrages et documents anciens conservés dans les bibliothèques et archives qui sont menacés de disparaître, rongés par l'acidité du papier et par l'humidité.

Pour consulter

La numérisation, associée à l'archivage et à l'indexation offre désormais aux professeurs, étudiants, chercheurs la possibilité d'accéder au patrimoine à l'aide de PC multimédia qui appartiennent désormais à notre environnement quotidien.

Pour imprimer

L'implantation des réseaux, le développement d'Internet, le fait que l'on puisse s'affranchir des contraintes de lieu, de temps et d'espace, permettent dès lors de réaliser sur papier des fac-similés, et de partager ainsi le "Patrimoine du Futur".

Pour les livres anciens, la numérisation peut poser quelquefois des problèmes du fait de leur fragilité. Cette numérisation dépend du choix du scanner, du mode de numérisation, du format de stockage. Bien des problèmes sont résolus grâce à des stations de numérisation spécialisées. Celles-ci sont prévues pour éviter de casser la tranche du livre et pour respecter le document afin de le garder intact.

2. DigiBook: Scanner à caméra sur portique

Le choix s'est porté sur le système DigiBook 5600⁹ pour la numérisation de livres de la Bibliothèque municipale de Lyon et de la Bibliothèque publique universitaire de Genève.

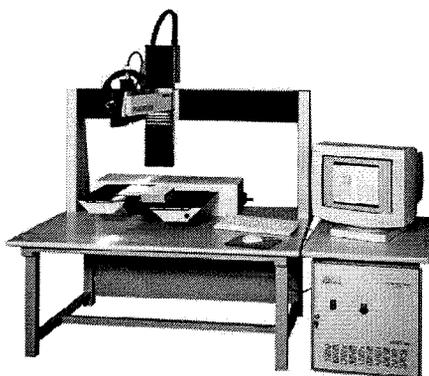


Fig. 1 : DigiBook 5600

Il s'agit d'une station de numérisation de haute productivité pour ouvrages à reliure rigide. Le DigiBook 5600 accepte des ouvrages dont la taille maximum ne dépasse pas 840 mm de largeur et 500 mm de hauteur (600 mm en option). Il assure le respect des ouvrages par deux caractéristiques techniques:

- les documents sont posés sur des plateaux autocompensés par le poids du livre, à plat ou à 120°, évitant de "casser" la reliure des livres très fragiles,¹⁰
- tout contact avec l'original est évité, notamment par l'absence de vitre.

Ce système est particulièrement bien adapté à la numérisation des livres reliés malgré son coût important.

Différentes chaînes de numérisation ont été mises en place dans le cadre du projet sur les sites des bibliothèques de Rome, Lyon et Coimbra pour tester tous les

⁹ Le DigiBook 5600 a été développé par la société I2S en partenariat avec Xerox.

problèmes rencontrés sur la qualité des images. La BPU de Genève, dans le cadre de DEBORA, a bénéficié d'une des stations de numérisation, ses trois livres ayant été numérisés avec une station DigiBook 5600 installée à la Bibliothèque municipale de Lyon.

Avec le DigiBook, les réglages optiques de caméra se font automatiquement. A ce moment, il n'est pas toujours évident de trouver la hauteur juste de caméra; l'opération prend du temps (ce qu'il faut calculer dans l'évaluation du temps et des coûts).

Un équilibrage de pages est nécessaire: pour toutes les 15 double-pages (celle de gauche et celle de droite en même temps) un réglage du plateau est préférable.

Au début du projet, M. J.-L. Boucher a fait ses expériences avec les premiers livres de la BM Lyon en numérisant chaque page à part. Par la suite il a été décidé de numériser deux pages à la fois (une séparation en page de gauche et page de droite s'effectuant après dans Transvision). Une double-page est considérée comme un seul fichier-image, d'une taille d'environ 27 Mo. Chaque nom de fichier est une chaîne numérique,

2.1. Qualité minimale requise

Afin d'assurer une longévité aux documents numérisés, il est nécessaire d'établir une qualité minimale pour la numérisation. L'objectif premier de cette numérisation est de pouvoir travailler sur le document numérique, et ainsi de protéger l'original. Il faut ainsi s'assurer de conserver toutes les informations utiles, car toute information perdue lors de l'acquisition ne sera jamais récupérée par traitement d'image¹¹.

Malgré la rigueur avec laquelle la numérisation a été effectuée, un certain nombre d'erreurs ont été observées et un travail non négligeable de correction a dû être

¹⁰ Au cours de la numérisation d'une Bible avec le photographe de la BM de Lyon, M. Bouchier qui manipulait la station, il a été constaté que l'ouvrage s'ouvrait mieux par le bas plutôt que par le haut. M. Bouchier a préféré respecter cette ouverture "capricieuse" pour ne pas casser l'ouvrage. Par conséquent, des zones floues peuvent apparaître sur l'image.

¹¹ Voir le phénomène des pages manquantes, blanches, noires ou demi-pages.

effectué. C'est dans ce cadre de mise en évidence de ces erreurs, et des corrections éventuelles, que j'ai travaillé pendant le mois de juin à la Maison de l'Orient méditerranéen (MOM) à Lyon sur les 25 livres de la BM de Lyon numérisés en format TIFF¹².

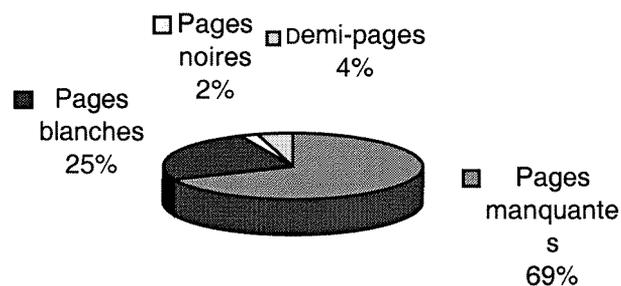
Les tableaux et les graphiques ci-dessous parlent du volume du travail de numérisation en *format TIFF*, effectué par DigiBook installé à la BM Lyon pour les deux bibliothèques: *la BM Lyon* et *la BPU Genève*, et le taux d'erreurs dû à la numérisation sur cette station.

Titre	Nb. de pages	Nb. pages manquantes	Nb pages blanches	Nb. pages noires	Nb demi-pages
1-La légende des Flamens	320				
2-Le Propriétaire des choses	570		6		
3-Les Exemplaires d'écriture	48				
4-Arithmetica	128	1		1	
5-L'Histoire de la nature...	412				
6-De honneste volupté	330		2		
7-Historia disputes et discours...	944	4			
8-Les Oeuvres & melanges...	644	7			
9-La Magnificence de la superbe...	96				
10-La Déclaration de l'instrument	98				
11-Les Oeuvres	962				
12-Excellent & moulutile Opuscule	242	4			
13-Les Amours	1138 (?)	9	2		
14-Prespectiva communis	36				

¹² Neuf autres livres ont été numérisés en format JPEG pour cette bibliothèque.

15-Traicté de la nature des viandes	268	7			
16-Le Mirouer d'or de l'ame pécheresse	136				
17-Le Kalendrier des bergiers	192				
18-Les Annales Dacquoise	450	1	16		4
19-Le livre des prouffitz ...	366	14			
20-Les Observations de plusieurs sing...	456	42		2	
21-Les fleurs et manières du temps	196				
22-Histoire merveilleuse et ...	574		2		2
23-La Pyrotechnie	340				
24-Les Première oeuvres poétiques	368	4			
25-Le Pegme	424				
Total	9738	93	34	3	6

Tableau 1 : BM Lyon (volume des pages numérisées et erreurs)

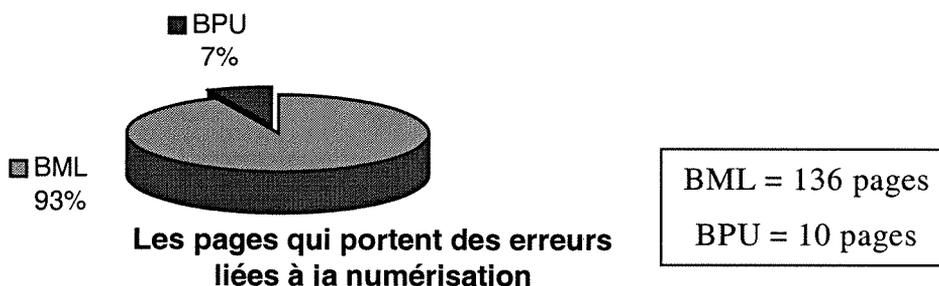
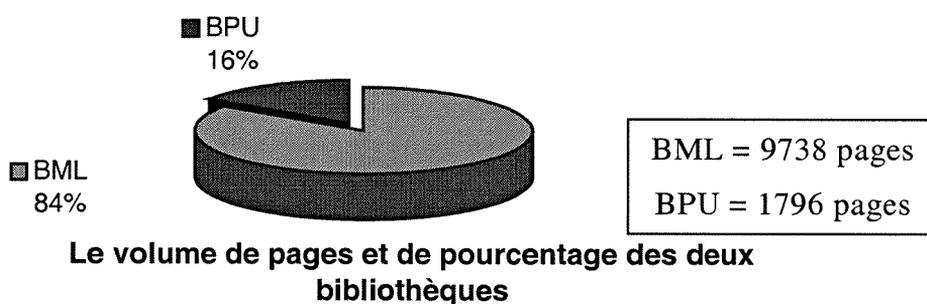


Les erreurs pour la BM Lyon

Titre	Nbr. de pages	Nbr. de pages manquantes	Nbr. de pages illisibles
1-Virtutum encomia	145		
2-Theodori Bezae Annotationes majores in Novum Dn	1314	4	10
3-Discours sur les moyens de bien gouverner	337		
total	1796	4	10

Tableau 2 : BPU Genève (le volume des pages numérisées et les erreurs)

Les deux graphiques ci-dessous permettent de voir que le nombre d'erreurs lié à la numérisation a baissé plus de deux fois pour les livres de la BPU Genève par rapport à ceux de la BM Lyon. L'expérience acquise sur le DigiBook est à l'origine de ce progrès.



2.2. Erreurs constatées

2.2.1. Manque de pages

L'erreur appelée "manque de pages" recouvre deux phénomènes:

- A. Lorsque la page est numérisée, mais qu'elle est "perdue" après coup (peut-être pendant le transfert), elle ne figure donc plus parmi les pages, un "trou" a lieu dans la numérotation.

Le phénomène de "trou" dans la numérotation des fichiers se produit quand une image est enregistrée par erreur deux fois. On doit alors supprimer manuellement l'un des fichiers à double, ce qui entraîne un saut dans la numérotation.

BML: 37 pages manquantes pour 25 livres

- B. Le "phénomène trou" se produit pendant la numérisation aussi, quand une image est enregistrée deux fois. On la supprime, mais des corrections manuelles doivent être ensuite apportées.

BML: 37 pages manquantes pour 25 livres

Dans ce cas deux solutions sont envisageables:

- soit on numérise les pages oubliées après coup; des chiffres bis apparaîtront dans la suite de la numérotation;
- soit, si la numérisation ultérieure n'est plus possible (comme c'est le cas pour la BPU), une note informant l'utilisateur de ce manque est ajoutée dans la zone de notes des métadonnées.

2.2.2. Pages blanches

- A. Des pages sont représentées par des images complètement blanches, soit des fichiers vides mais portant bien un nom de fichier. Dans ce cas, le "trou" concerne donc le contenu.
- B. Des images de nouveau blanches, mais où apparaissent des explications portées par l'opérateur. Dans ce cas le phénomène est dû au manque réel de la page du document physique. En s'apercevant de cette absence, l'opérateur numérise une feuille blanche portant une explication sur la raison de ce manque.

2.2.3. Les pages noires et les demi-pages

On parle du phénomène "pages noires ou demi-pages" dans le cas où la page est complètement noire ou à moitié blanche et qu'elle suive malgré tout la numérotation. Dans ce cas-là on parle plutôt de problèmes de numériseur.

2.3. **Recommandations techniques pour la qualité des images DEBORA**

2.3.1. Résolution

Contrairement aux documents actuels standardisés sur les formats A4, A5 et format poche, les formats du XVI^{ème} siècle peuvent aller de 50x80 cm (grands livres) à 8x6 cm (petits livres). Par conséquent, il ne faut pas tenir compte de la résolution du scanner (utilisées pour les scanners actuels à plats en nombres de points par pouces ou dpi) mais du ratio avec le corps (hauteur moyenne d'une ligne) du texte.

DEBORA a défini alors :

$$\text{La résolution moyenne en pixels des caractères} = \frac{\text{Résolution optique}}{\text{Corps du texte}}$$

2.3.2. Nombre de nuances

Afin d'assurer une meilleure durée de vie de l'image, son acquisition doit être faite avec un minimum de 256 niveaux de gris. Une image couleur (24 bits minimum) est cependant très fortement recommandée dans les configurations où elle est nécessaire (documents en couleur ou dégradés...)

En effet, plus l'image est riche en informations, plus sa durée de vie est longue, et plus les traitements, qui y seront effectués, seront pertinents.

Pour DEBORA une image en 256 niveaux de gris est requise.

2.3.3. Stockage des images

Une fois que l'image du document est numérisée, il est nécessaire de la sauvegarder afin de pouvoir s'en resservir ultérieurement. Pour cela, Debora a décidé de stocker ses images une fois sur le réseau en passant par des stockage "temporels" sur le disque dur. La capacité du disque dur pour le stockage initial est importante, mais la nécessité du transfert sur la base commune d'image est inévitable.

Remarque: Au cours du travail nous avons remarqué que ce transfert se fait mieux par groupe de 5 fichiers et pas davantage.

Ces fichiers-images sont stockés sur CD-ROM (qui vont être utilisés notamment à l' Institut National des Sciences Appliquées de Lyon - INSA pour la reconnaissance de forme). Au début du projet la BM Lyon a sauvegardé ces images sur des DVD-RAM. Plus tard ce support de stockage a été "abandonné", compte tenu de sa longévité douteuse.

Les livres de la *BPU Genève* ont eu pour support de stockage le *CD-ROM*.

2.3.4. Sauvegarde d'images en niveaux de gris

Ce mode était conseillé pour DEBORA. Les images en niveaux de gris permettent des traitements sophistiqués pour segmenter les caractères dégradés, séparer le texte des images par analyse de texture, réparer et restaurer les images des caractères.

2.3.5. Format d'image TIFF

Les deux formats: TIFF et JPEG sont utilisés pendant la numérisation des livres de la BM Lyon (25 livres en format TIFF et 9 livres en format JPEG).

Pour la BPU Genève, ses trois livres ont été numérisés en format TIFF. Ce format est capable de décrire des images binaires, en niveaux de gris, en couleur ainsi qu'en « full-color » (utilisation du nombre maximum de couleurs disponibles).

	LYON	GENEVE	ROME	COIMBRA
Scanner	DigiBook 5600	DigiBook 5600	OmniScan 7000	—————
Nuances	256 niveaux de gris			
Résolution	600 dpi	600 dpi	300 à 600 dpi	300 à 600 dpi
Format d'image	TIFF	TIFF	TIFF	TIFF
Support de sauvegarde	DVD-RAM	CD-ROM	CD-ROM	DVD-RAM

Tableau 3 : Configuration utilisée par les différentes bibliothèques partenaires de DEBORA

2.4. Choix des normes

Les normes choisies ont pour objectif de permettre l'archivage des images numérisées du document afin d'assurer une base de travail stable avec une durée de vie maximale. C'est pourquoi, on s'assurera que **l'image sera brute**, exempte de tout risque de modifications susceptibles de détruire des informations à partir desquelles il ne sera plus possible de revenir en arrière sans recourir à l'original.

2.4.1. Normes de numérisation

La compensation des courbures est prise en compte par le logiciel qui pilote la station. 256 niveaux de gris différents sont traités, ce qui garantit une bonne qualité de numérisation des manuscrits ou des imprimés.

Grâce à une résolution importante (250 à 600 dpi), la qualité de restitution du document original est largement suffisante.

Le passage à la couleur (prévu dans le courant 2000) permettra une numérisation encore plus complète et fidèle.

2.4.2. Normes de stockage

On notera que les supports informatiques à hautes capacités n'ont jamais une durée de vie illimitée et même très inférieure aux ouvrages dont ils vont devoir assurer la pérennité! Néanmoins ces copies de substitution sur support informatique sont des

images hautement fidèles à l'original. Le support physique doit donc répondre à de fortes contraintes de longévité ainsi que de maniabilité.

Il doit garantir une durée de vie minimale fixée à 15 ans afin d'éviter les copies de sauvegarde trop fréquentes. Une durée légèrement plus importante impliquerait un choix de support plus réduit pour un coût beaucoup plus élevé. Sont conseillées aussi fortement les sauvegardes sur supports hétérogènes (disques optiques, magnétiques...).

Remarque : un support inscriptible une seule fois est suffisant.

Les informations suivantes sur la numérisation seront indiquées:

- Paramètres utilisés par le numériseur (éclairage, temps d'exposition, profondeur de champ ...);
- Notes éventuelles de l'opérateur.

Les normes qui ont été définies ici sont minimales, c'est-à-dire que les résolutions, nombres de couleurs, etc. données sont les valeurs minimales à utiliser. Nous devons nous rappeler que, en règle générale, plus une image est riche en informations, plus elle restera intéressante à long terme.

Les réglages de l'optique sont automatiques, et mémorisables par format de livre.

La correction de l'éclairage en temps réel, des marges programmables ainsi que les fonctions de correction de courbure, de détournage automatique, de détection des pages blanches, de redressement du texte associés à la retouche d'image (tramage, modification du contraste, gommage) garantissent des images de très haute qualité.

Des images peuvent être affichées soit en 256 niveaux de gris et archivées en JPEG, soit en bitonal et archivées en TIFF G4.

L'automatisation de la chaîne de numérisation, visualisations, traitements, sauvegarde est réalisée par la fonction auto-scan sur DigiBook 5600 garantissant ainsi des gains de productivité et un grand confort pour l'opérateur.

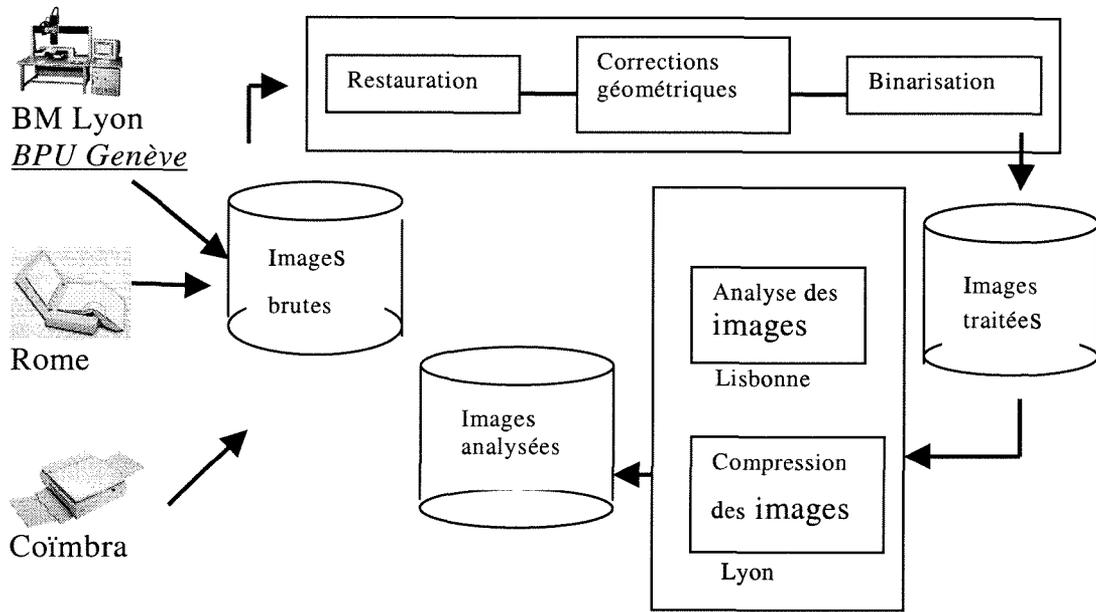


Fig. 2 : La chaîne de numérisation de DEBORA

3. Différents avis sur le support de numérisation

1. "La numérisation directe à partir des documents est peu pratiquée dans les bibliothèques. Il est préférable pour l'instant de passer par un support intermédiaire, quand c'est possible, pour des raisons qui peuvent être liées aux caractéristiques physiques des originaux (taille), à l'état ou à la valeur des documents qui interdisent la sortie de la bibliothèque, ou à la nécessité d'avoir constamment ce document à disposition du public, etc." [BEQ00]

Cet auteur, G. Bequet cite quelques exemples:

- Pour la numérisation des livres imprimés, la Bibliothèque nationale de France a utilisé des exemplaires sans valeur qui ont été massicotés, ou des microfiches.
- Les universités de Yale de Floride numérisent des journaux à partir de microfilms.

- Un programme de numérisation directe de journaux avec production ultérieure de microfilms 35 mm (Computer Output Microfilm) a été réalisé de 1993 à 1995 à l'université Cornell aux Etats-Unis.
2. Par contre à l'INSA de Lyon, l'accent est mis sur l'origine de l'image, donc la version papier. La capture de l'image est l'étape primordiale qui déterminera ses qualités intrinsèques¹³.

Des exemples d'acquisition sont données:

Mauvais exemples:

- document papier → microfilm → scanner
- document papier → photographie → scanner
- document électronique → imprimerie → scanner
- document papier → photocopie → scanner

Bons exemples:

- document papier → scanner
- document papier ou objet → photographie numérique

En règle générale il faut donc éviter de créer des étapes intermédiaires inutiles!

Les images obtenues doivent être, non seulement, le plus fidèle possible aux documents originaux, mais elles doivent aussi contenir suffisamment d'informations pour pouvoir être analysées automatiquement.

Certes, aujourd'hui, les algorithmes de reconnaissance de formes sont encore limités pour le traitement du contenu des documents anciens. Mais dans un avenir proche, les progrès dans ce domaine permettront l'analyse automatique du contenu par l'informatique.

¹³ E.Trinh; F.Lebougeois; N. Martin – DEBORA: Telematics Applications, Programme: WP 4.2

4. Autres pratiques de numérisation et apport de DEBORA

Les normes utilisées par la Bibliothèque nationale de France et par la Bibliothèque du Congrès à Washington pour la numérisation de leurs documents sont les suivantes:

A- Bibliothèque nationale de France

Page de texte imprimée: résolution de 400 dpi, format TIFF Aldus monopage 5.0, compression CCITT groupe IV; photographie en noir et blanc: résolution de 2000 par 3000 dpi, codage du pixel sur un octet, compression JPEG (1:10); photographie en couleurs: résolution de 2000 à 3000 dpi, codage du pixel sur trois octets, compression JPEG (1:10).

B- Bibliothèque du Congrès

Le imprimés sont traités en mode caractère (codage des documents selon SGML) ou en mode image (format TIFF, pixel codé sur un octet, résolution différente selon que le document est destiné à la consultation – 300 dpi – ou à l'archivage – 300 à 1200 dpi -) ; photographies en noir et blanc ou couleurs: 3 formats sont sélectionnés. Imagette: pixel codé sur un octet, format GIF, résolution 200 par 200 dpi; image de consultation: codage sur un ou trois octet, format JFIF, compression JPEG, résolution de 500 par 400 à 4000 par 3000 pixels; image d'archivage: codage sur un ou trois octets, format TIFF, pas de compression, résolution de 500 par 400 à 4000 par 3000 pixels. [BEQ00]

La très grande majorité des projets de rétro-conversion de documents considèrent l'image brute comme suffisante. Ceci a de nombreux inconvénients. D'une part, l'image numérisée du document est entachée de défauts (taches, caractères dégradés...), ce qui rend l'image parfois désagréable à regarder. D'autre part, aucune récupération automatique des informations contenues dans ces images n'est effectuée. Dans ce cas, les indexations sur le contenu des images ne peuvent être effectuées que manuellement.

D'autres projets se contentent seulement de conserver le texte contenu dans le document (projet Gutenberg). Ces projets ne permettent pas la conservation de l'aspect visuel de l'édition originale. Par contre, toute indexation du contenu est rendue plus simple.

D'autres groupes de projets encore apportent les solutions les plus efficaces en matière de rétro-conversion pure du document. Par contre, aucune solution intéressante n'a été proposée pour utiliser au mieux les informations qui sont contenues dans l'image de l'ouvrage.

L'intérêt du projet DEBORA est la prise en compte intégrale de ces lacunes. DEBORA se charge de toute la chaîne de diffusion du document, de la numérisation par un numériseur à la mise à la disposition du public sur le Web. Il consiste à définir le processus de numérisation et à corriger l'image après acquisition, de manière à ce quelle soit le plus conforme possible à l'original et agréable à regarder. Mais aussi, il effectuera des traitements complexes d'indexation par le contenu même de l'image (recherche de mots clés, indexation des lettrines...). Finalement, il assurera toutes les indexations nécessaires dans une base de données et la conversion des données dans un format utilisable sur Internet.¹⁴

5. Autres projets de bibliothèques numériques

MALVINE

MALVINE a pour principal objectif d'élaborer en Europe un réseau des catalogues de ces diverses institutions ; réseau qui, indépendamment de toute configuration technique particulière (cf. Informations techniques détaillées), et accessible partout dans le monde via le Web, fonctionne comme une base de données unifiée et homogène. Une interface utilisateur multilingue, utilisant une terminologie commune, sera mise à la disposition de l'utilisateur, qui pourra également obtenir des copies numérisées des documents originaux. Sa particularité est le fait que, non

¹⁴ DEBORA : Projet européen : n° LB 5608 A/ coordination R. Bouche. juin 2000. 167 p [20p.]. : ill. ; 30 cm

seulement elle facilite l'accès aux manuscrits, mais elle optimise aussi l'échange des données entre systèmes hétérogènes, tout en permettant de perpétuer les traditions locales de catalogage.

Son but ultime est d'offrir à tous les pays un accès commun à tout type de document autographe. En dépassant les différences entre les langues et les particularismes des équipements techniques, MALVINE favorise ainsi le développement d'une Europe unifiée.

MASTER (Manuscript Access through Standards for Electronic Records)

MASTER est un projet de l'Union Européenne qui vise à réaliser un catalogue général et en ligne des manuscrits médiévaux conservés dans les bibliothèques d'Europe.

MOSAIC (Museums Over States and Virtual Culture)

MOSAIC est également un projet financé par l'Union Européenne, implanté au sein de Ten-Telecom. Il se focalise sur le développement de pôles de services qui fournissent en ligne des informations concernant le patrimoine culturel. Ces Centres de Services Culturels Européens (CSC) encourageront la recherche et favoriseront la diffusion du patrimoine culturel européen en offrant des services et des outils de mise en réseau internationale.

6. Analyse des coûts

S. Lenga et A. Silem ont souligné¹⁵ qu': "il semble qu'il n'existe aujourd'hui aucune étude portant sur les coûts de la numérisation d'intégration dans une base de données et de diffusion par l'Internet permettant d'offrir les résultats utilisables et transposables tels quels. Notre propos [dans le cadre de DEBORA] se veut donc exploratoire [...] Néanmoins, l'ambition demeure technologiquement limitée, car il sera difficile sinon impossible d'évoquer les effets d'expérience, de prendre en

¹⁵ DEBORA : Projet européen : n° LB 5608 A/ coordination R. Bouche. juin 2000. 167 p [20p.]. : ill. ; 30 cm.

compte des coûts de plus en plus faibles de matériels et de logiciels de meilleure qualité susceptibles d'apparaître sur les marchés dans les prochaines années".

Evaluation du temps passé pour la numérisation

En suivant de près la numérisation des trois livres de la BPU de Genève avec la station de numérisation DigiBook à la BM Lyon j'ai constaté que la cadence moyenne de numérisation est de

86 pages-images par heure.

Plusieurs opérations sont comprises dans cette évaluation. Il faut ainsi tenir compte des éléments suivants:

- pour chaque ensemble de 25-30 pages (cela varie en fonction de l'épaisseur de la page), un réglage du plateau est recommandé pour obtenir une meilleure qualité;
- les images sont stockées une première fois sur le disque dur (d'où l'importance d'un disque dur de haute capacité), avant leur transfert sur la base. Pour effectuer ce transfert, l'expérience a montré qu'un fractionnement des fichiers par groupes de 5 (pas davantage!) était préférable;
- les téléphones auxquels il faut répondre;
- le contrôle d'image et la correction d'erreurs de numérisation constatées;
- la préparation et le collage des étiquettes d'identification des CD-Roms etc.

Evaluation des coûts

L'évaluation des coûts dépendent de la catégorie du personnel qui procède à la numérisation¹⁶. Pour la BM de Lyon la personne qui a fait la numérisation était photographe de profession. Elle estimait également que certaines connaissances photographiques sont appréciables pour ce travail, connaissances qu'un assistant photographe possède par ailleurs.

¹⁶ Voir les catégories respectives dans les annexes

BIBLIOGRAPHIE

ARKHËNUM / Patrimoine du Futur

<http://www.gutenberg21.asso.fr/pages/acteurs/arkhenum.htm> – visité le 10.08.2000

[BEQ00]BEQUET, Gaëlle.

La numérisation des documents patrimoniaux " Bibliothèques et informatique"
publié par la Direction du livre et de la lecture.

<http://www.culture.fr/culture/mrt/numerisation/fr/technique/dll.htm> (visité le
01.09.2000)

Les Bibliothèques Virtuelles/par: Marie-Hélène Balland, Alexandre Delavanne,
David Fortino, Sylvie Loore.

ENSSIB/INSA, février 1999.

<http://www.enssib.fr/bibliotheque/documents/travaux/bibvirt/rapp3.html>

[CAZ99]CAZABON,Marie-Renée.

UNIMARC : manuel de catalogage, monographies, publications en série, musique
imprimée, documents électroniques / par Marie-Renée Cazabon avec la
participation de Maria Witt ; préfaces de Marcelle Beaudiquez et Marc Chauveinc.
Cercle de la Librairie, 1999. 205p. (Collection Bibliothèques).

[FOU98]FOUREL, Frank.

La station DigiBook 5600 de I2S, 1998. Article du journal MOS (Mémoire
optiques et systèmes) N 166, p. 9.

[JAC99]JACQUESSON Alain, RIVIER, Alexis.

Bibliothèques et documents numériques : concepts, composantes, techniques et
enjeux. Paris : Electre-Edition du Cercle de la Librairie,1999. 377 p. : ill.
(Bibliothèques)

[LEC98]LE CROSNIER Hervé

Les bibliothèques numériques. *Document numérique*. Vol. 2 - n° 1/1998.

[PLA98]PLAISANT, Catherine.

La recherche d'interfaces utilisateurs pour la bibliothèque du Congrès américain.
Catherine Plaisant - Gary Marchionini - Anita Komlodi . Document numérique.
Vol. 2 – n 1/1998.

[SET98]SETA, Frédérique.

Le rôle de la numérisation dans la mise en valeur des fonds patrimoniaux:
l'exemple de la bibliothèque interuniversitaire Cujas : Mem. DCB . Enssib, 1998.
65 f, Annexes (20 f.)

IV. La chaîne de production DEBORA - Traitement

1. Le logiciel TransVision¹⁷

1.1. De l'imagerie archéologique à un système de gestion de banques d'images

Après la numérisation avec le DigiBook, la réalisation d'une banque d'images de pages numérisées a été effectuée au moyen du système TransVision. Cette base est dotée de moyens d'accès puissants qui permettent à la fois une validation de l'outil par les usagers et une reconnaissance internationale du travail accompli.

TransVision est un système de gestion d'images remplissant aussi la fonction de serveur d'images. Conçu dans un laboratoire d'archéologie de la Maison de l'Orient Méditerranéen (MOM), il a été adapté au projet DEBORA.

1.2. Utilisateurs, administrateurs, producteurs de TransVision

On distingue deux types d'utilisateurs/producteurs:

A. Utilisateurs déclarés comme administrateurs.

Ces administrateurs ont la possibilité de :

- créer une nouvelle base de données,
- gérer les droits d'accès des utilisateurs en lecture ou en écriture sur la banque d'images,

17 Le projet SGBI (Système de Gestion de Banques Images) a été mis au point par des chercheurs du CNRS, de l'Université Lumière Lyon 2 avec des collaborations extérieures. Evalué par la société France Innovation Scientifique et Technique (FIST), il est aujourd'hui développé et commercialisé par SGBI Entreprise, société fondée par les auteurs du projet. A l'initiative de Bruno Helly et Patrick Desfarges (Maison de l'Orient méditerranéen, CNRS-Université Lyon 2), le travail préliminaire à la mise au point du système de gestion de banques d'images (SGBI) a commencé en 1989, dans le cadre d'un thème de recherche de la Maison de l'Orient méditerranéen (CNRS-Université Lyon 2), sur l'archéologie comme système d'information scientifique.

- définir la structure de la base de données textuelles associées à la banque d'images,
- gérer l'espace disque,
- etc.

B. Utilisateurs déclarés comme producteurs d'informations.

L'administrateur peut rattacher un utilisateur à une banque d'images dont il a la gestion soit en lecture (dans ce cas l'utilisateur ne peut que consulter les éléments de la banque) soit en écriture, ce qui lui permet de produire des informations.

Dans le cas de DEBORA, le bibliothécaire ou le producteur de la base de livres numérisés sont des utilisateurs déclarés.

TransVision est doté d'un client natif qui assure l'interface graphique entre l'utilisateur et le serveur. Ce client est disponible sur des postes Macintosh (Power PC) et Windows (95-98-NT).

1.3. Logiciels clients "natifs"

TransVision se compose de deux logiciels clients "natifs" différents:

A. Le client administrateur, dont les fonctions sont:

- la création et l'administration des utilisateurs d'une banque d'images TransVision
- la création et la gestion des éléments composant le modèle de références textuelles associées à la banque,
- la gestion de la liste des serveurs disponibles

B. Le client de production, dont les fonctions sont:

- la création d'une entité,
- l'intégration d'un objet image dans la base,
- la constitution de collections

- l'interrogation multi-critères des références textuelles associées aux éléments images.

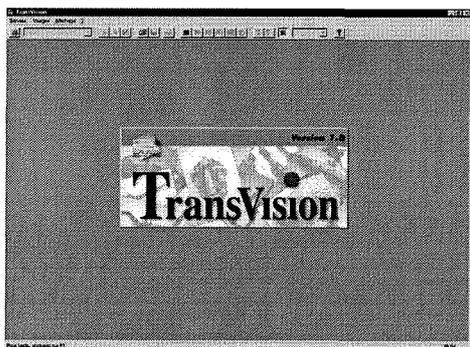


Fig. 1. Ecran d'accueil de TransVision

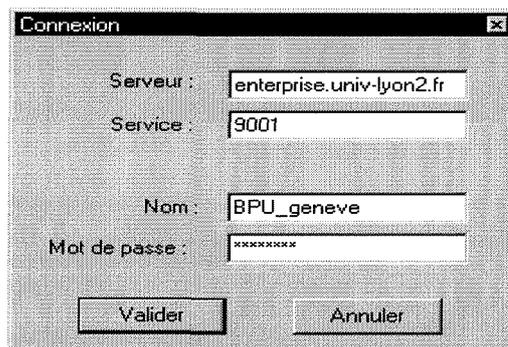


Fig. 2. Accès à la base protégé par mot de passe

Le travail sur la base TransVision a débuté à la MOM de Lyon, sous la direction des administrateurs de cette base, et plus directement de M. Desfarge, pour continuer, à la BPU de Genève, expérimentant ainsi pour la première fois le travail à distance avec ce progiciel.

Personnellement, j'ai eu l'occasion de faire des expériences en travaillant d'abord, (toujours dans le cadre de DEBORA), sur les livres de la Bibliothèque municipale de Lyon (BML) et de la bibliothèque Casanatense de Rome.

En numérisant une trentaine de livres du fond du 16^{ème} siècle au moyen d'une station DigiBook installée sur place, la BML a été la première bibliothèque à alimenter la base DEBORA, acquérant de ce fait une certaine expertise dans le domaine de la numérisation¹⁸. Les fichiers obtenus ont été stockés sur un support provisoire (DVD-Ram) qui permet un échange facile entre les différentes institutions devant traiter ensuite l'information, telles que La MOM pour l'intégration dans la base de données, le laboratoire de connaissance des formes de l'INSA – Lyon et l'Instituto Superior Tecnico de Lisbonne pour le traitement et l'indexation des images¹⁹.

¹⁸ DEBORA: Projet européen: n° LB 5608 A. Coordination : R. Bouché, juin 2000.

¹⁹ A l'heure actuelle cette tâche est dans sa phase expérimentale.

Initialement la chaîne de traitement des livres (BML et bibliothèque Casanatense), était répartie entre plusieurs opérations séparées, chacune prenant beaucoup de temps (fig. 3).

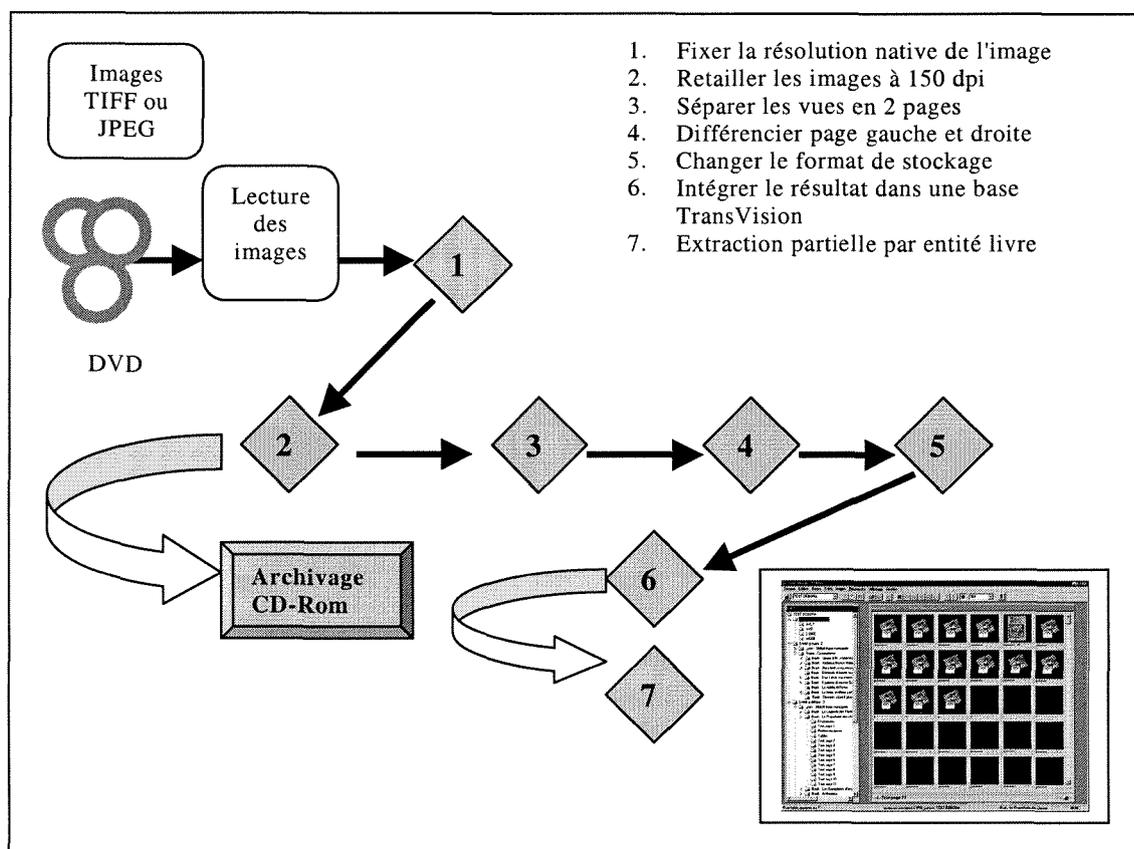


Fig. 3: Chaîne de traitement initial DEBORA

1.4. Le ROBOT

Les livres de la BPU de Genève ont pu bénéficier des avantages d'un robot qui a été mis en place à la MOM en juillet 2000. Ce robot permet de regrouper en une seule les quatre opérations de la chaîne de traitement initial, ce qui représente un gain de temps trois fois supérieur (fig. 4).

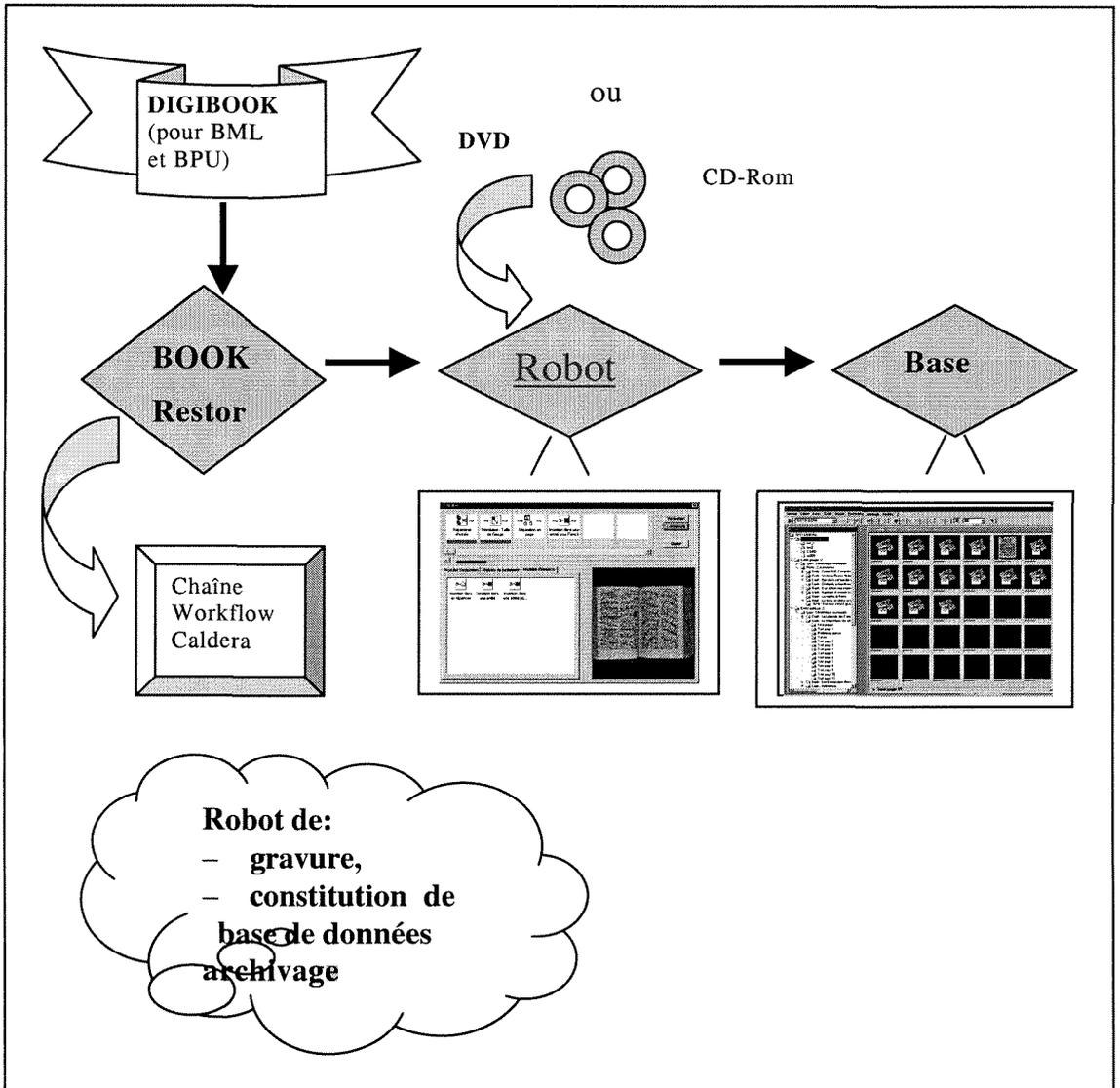


Fig. 4 : Chaîne de production DEBORA

1.4.1. Fonctionnement du Robot

Le robot fait le lien entre le serveur, la base et l'entité. Il se compose de trois modules:

1. Module d'acquisition

On définit le répertoire d'entrée des images stockées sur CD-ROM ou autre support.

2. Modules de traitement

On y définit:

1. la résolution soit la *Taille de l'image*, exprimée par ses dimensions (largeur ou hauteur) maximum en pixel;
2. la séparation en pages (page de gauche et page de droite)

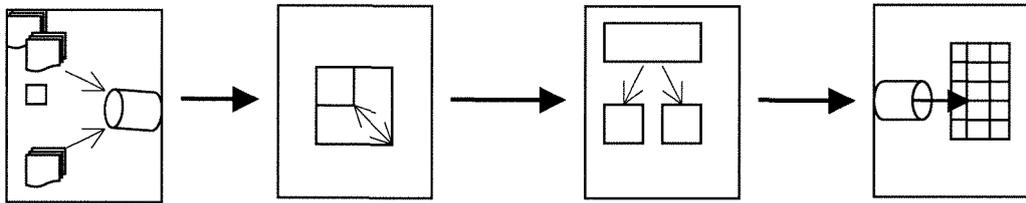
3. Module d'insertion

Il permet l'insertion de ce répertoire dans une entité: c'est dans ce module que sont établis les liens entre l'entité, le serveur et la base.

Avant de mettre le robot en route, on décide des paramétrages, qui sont propres à chaque livre. Mais même pour chaque CD-ROM d'un livre, il faut bien réajuster le curseur qui commande le découpage des pages au bon endroit; des tests sont recommandés chaque fois. L'avantage lié à cette méthode est la possibilité de faire des contrôles immédiats de l'image, et si l'on détecte un découpage incorrect (Fig. 7), on arrête le robot pour recommencer. L'inconvénient est qu'il faut alors reprendre le paramétrage depuis le début, ce qui prend beaucoup de temps. Il n'y a malheureusement pas d'autre manière d'apporter une correction pour la fonction "séparation en page". C'est la raison pour laquelle on découvre parfois des pages mal coupées). Le pire est bien sûr de s'en apercevoir dans les dernières pages du traitement.

Il existe pourtant une autre façon de corriger uniquement les pages mal coupées: remonter jusqu'à la racine, où toutes les pages se trouvent encore réunies sous forme de répertoires: cocher les répertoires corrects que l'on veut garder, et corriger uniquement le ou les répertoires qui posent problèmes en les traitant à nouveau avec le robot. Toutefois ce mode d'intervention nous a été montré trop tard pour l'appliquer aux documents de la BPU.

Ces problèmes s'expliquent par le fait que le logiciel n'est pas initialement conçu pour le découpage des pages.



Choix du répertoire
d'entrée

Résolution
Taille de l'image

Séparation en pages
(left/right)

Insertion dans une
entité X

Fig. 5 : Représentation schématique des opérations du Robot

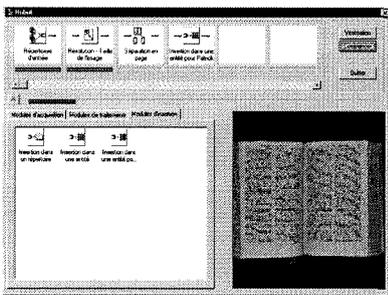
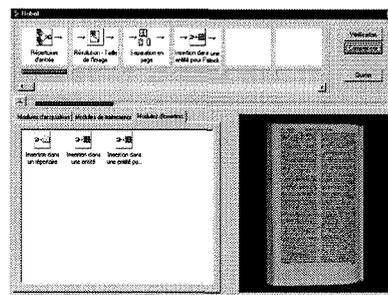


Fig. 6 : Deux pages non séparées



une seule page après découpage

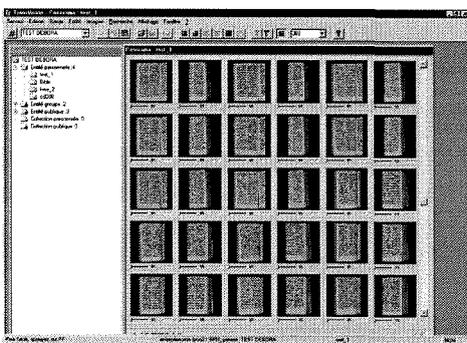


Fig. 7 : Des pages mal coupées

1.4.2. Evaluation du temps de traitement au moyen du robot

Paramétrage des CD-ROM pour le robot:

2-3 min pour chaque CD-ROM.

Traitement du robot sur les 23 fichiers-images du CD-ROM, 46 pages images séparées:

Autour de **8-10 min**

Remarque

Plus le livre devient épais, plus on met le curseur pour le découpage doit être placé vers le centre.

1.5. L'architecture client-serveur de TransVision

TransVision possède cinq serveurs:

- un serveur de base proprement dit;
- un serveur d'indexation;
- un serveur d'administration;
- un serveur de listes hiérarchiques;
- une librairie se comportant comme serveur W3.

Pour chaque serveur, il existe trois librairies:

- une librairie de fonctions: cette librairie regroupe toutes les actions que le serveur est capable d'effectuer;
- une librairie de communication pour le serveur. Cette librairie regroupe toutes les fonctions permettant le transport et l'interprétation des commandes issues des postes clients;
- une librairie de communication pour le client. Cette librairie regroupe toutes les fonctions permettant le transport et l'émission des commandes des postes clients.

1.6. L'interface de TransVision

Le serveur TransVision propose essentiellement deux modes d'accès à l'information: un mode d'accès de type arborescent, et un mode d'accès par

recherche sur les éléments textuels de la base. Ces deux méthodes d'accès à l'information, sont parfaitement complémentaires : l'une privilégiant l'accès purement visuel à l'information, l'autre permettant de mener des requêtes plus précises sur des contenus textuels.

L'accès de type arborescent a été le plus utilisé pour le travail sur les livres des trois bibliothèques, dont je me suis occupée.

1.7. Saisie et mise à jour des données

Le logiciel TransVision offre les fonctions suivantes:

- la production répartie d'informations sur un réseau,
- la navigation arborescente dans une banque d'images,
- l'interrogation multi-critères des éléments qui la composent,
- et bien sûr la gestion des droits d'accès à l'information, directement assurée au niveau des producteurs de la base, l'ensemble mis à jour en temps réel.

Cette gestion, établie en mode client/serveur, est transparente pour l'utilisateur et lui permet d'être en permanence renseigné sur l'évolution des éléments qui lui sont accessibles.

Pour constituer les éléments de la banque d'image, le producteur d'information a à sa disposition un certain nombre d'éléments primaires qui sont l'objet-image et ses références textuelles, l'entité, la collection

1.8. Entité et sous-entités²⁰

L'entité est un ensemble dans lequel, en tant que producteur, on dépose nos objets-image, sources d'information regroupées par thème sur lesquelles les utilisateurs vont pouvoir travailler. C'est également au niveau de l'entité qu'on définit les champs des références-textes qu'on souhaite associer à chaque objet-image. Ce modèle de référence est défini par l'administrateur de la base TransVision.

²⁰ voir Fig. 8

Une entité peut elle-même contenir d'autres entités, que l'on appelle des *sous-entités*. Il n'y a pas de limites dans le nombre de niveaux hiérarchiques. Le nombre d'éléments compris à l'intérieur d'une entité peut être illimité, mais il est recommandé aux administrateurs de base TransVision de limiter ce nombre d'éléments à 100, car au-delà il devient fastidieux à un utilisateur de retrouver un élément. A partir de différentes entités, on peut ensuite créer des collections virtuelles.

Le producteur a la possibilité de créer des sous-entités en arborescence, à l'identique, avec un modèle de références textuelles différent ou non. Par définition, au moment de sa création, les droits d'accès à l'entité sont restreints à son créateur. Le producteur peut à tout moment décider de la diffusion d'une entité vers d'autres utilisateurs, suivant des règles définies.

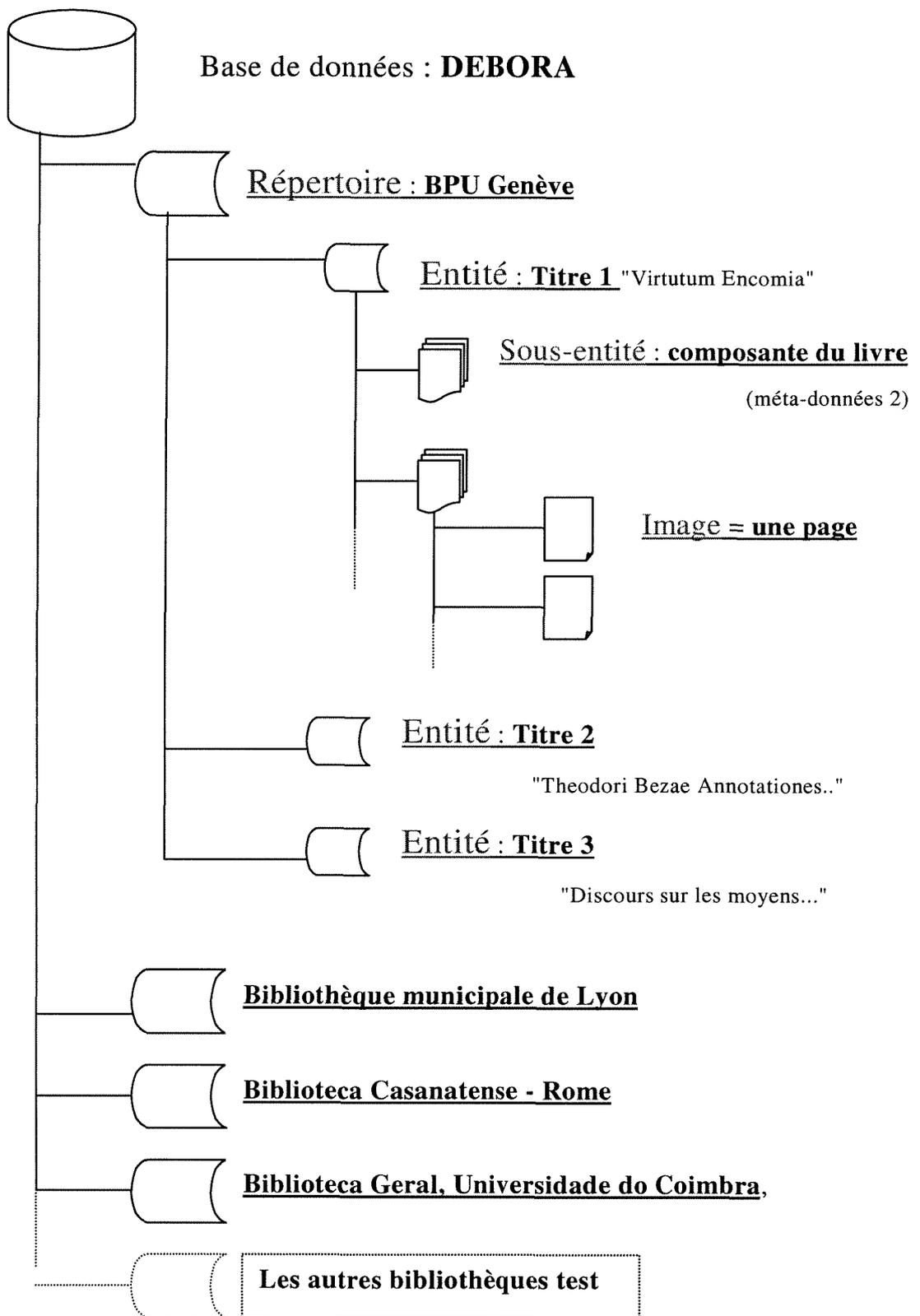


Fig. 8. L'arborescence de DEBORA

1.9. Pré-indexation

Pour l'utilisation de DEBORA avec le poste-client de consultation, un système de recherche efficace des documents et des pages-images est nécessaire. Les informations textuelles saisies en mode production doivent être indexées par le serveur d'indexation de TransVision. Chacun des champs qui la composent peut contenir un maximum de 32 Ko.

Une version des références textuelles sous forme de fichiers HTML est également conservée. Elle permet de reconstituer automatiquement la base de données en cas de crash du serveur d'indexation.

Lorsque les objets-images sont créés, une *pré-indexation* automatique est effectuée sur les références textuelles. Dans le cadre de Debora, les trois champs "Auteur", "Titre" et "Date de publication" sont renseignés, ainsi que toutes les métadonnées pertinentes du 1er niveau pour la page de titre. Les utilisateurs - producteurs ou non -, peuvent d'emblée mener des interrogations sur les entités qui leur sont accessibles.

Pour le *Theodori Besae Annotationes* (Genève BPU, S 14789) le temps de pré-indexation s'est élevé à

4 heures pour 1314 pages

Cette pré-indexation est longue, mais indispensable. D'ailleurs son élaboration manuelle n'est pas envisageable, vu l'ampleur que cette tâche impliquerait.

Suggestion pratique

Il peut être intéressant de disposer de deux PC de production simultanément. L'opérateur lance le processus de pré-indexation sur l'une, tout en continuant de travailler normalement sur l'autre.

1.10. Autres opérations dans TransVision

Un double-clic sur une *page-image du panorama*, active la fonction de loupe et entraîne l'affichage de la partie textuelle de cette page (Fig. 9), avec donc l'ensemble des métadonnées du 1^{er} niveau, pour la page du titre, et les seuls champs auteur, titre, et date de publication, pour les autres pages. La modification des références textuelles est possible à cette étape.

Un double-clic *sur l'image réduite d'une page* (Fig. 9), entraîne l'affichage de la partie image agrandie (Fig. 10). La présence de l'outil main donne la possibilité de se "promener" confortablement sur cette page.

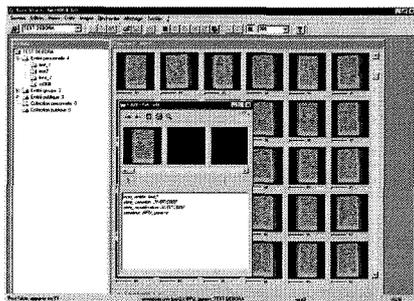


Fig. 9. Mode panorama avec les références textuelles d'une page

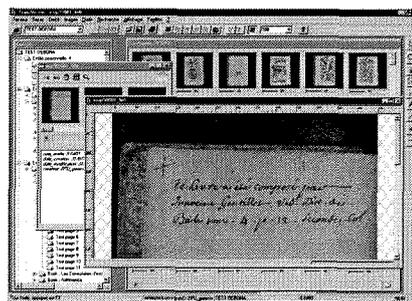


Fig. 10. Affichage de l'image taille normale

1.11. Le travail sur le livre : les métadonnées du 2^{ème} niveau

Initialement toutes les pages-images d'un livre insérées dans la base TransVision se trouvent ensemble au même niveau hiérarchique. Pour les répartir dans les sous-entités (métadonnées du 2^{ème} niveau, voir ch. II.4.2) selon le schéma proposé par le groupe de travail DEBORA, la procédure suivante est appliquée:

1.11.1. Etude du contenu

Le livre entier est passé en revue afin de déterminer la manière dont il sera structuré avec les métadonnées du 2^{ème} niveau.

La durée de cette tâche dépend du livre, lui-même : d'abord du nombre de pages, puis de la visibilité des pages-images dans le panorama, de l'écriture: dense ou pas. Pour certains livres, la fonction de loupe (Ctrl+clic) fait apparaître une image très claire, du point de vue du contenu: on y voit tout de suite le commencement ou la fin de chapitres. Pour d'autres il faut agrandir considérablement l'image (deux, voire trois fois) pour visualiser "de plus près" la page, ce qui exige beaucoup plus de temps. Mais celui-ci dépend aussi des performances du poste de travail et de la connexion avec le serveur.

Mon expérience personnelle sur la base des livres des trois bibliothèques (Lyon, Rome et Genève) me montre que cette opération s'effectue à une moyenne de:

5 minutes pour 78 pages

Dans un cas extrêmes l'examen d'un livre de 155 pages a demandé 75 minutes, toutes les pages ayant dû être agrandies deux fois.

Remarque

Dans la version utilisée de TransVision, l'ascenseur du panorama ne fonctionnait pas. Seuls les touches de direction étaient disponibles pour naviguer dans le panorama et visualiser les images, ce qui est très fastidieux.

Décisions

Après avoir acquis une certaine expertise sur TransVision, le groupe de travail de DEBORA a pris les décisions suivantes:

- A. Le contenu de l'ouvrage doit être respecté, soit la répartition en chapitres ou autres subdivisions.
- B. Chaque entité contiendra une cinquantaine de pages au maximum, dans les cas où le contenu ne peut pas être déterminé plus précisément. Cette limitation facilite le téléchargement, trop lourd autrement.

1.11.2. Création des sous-entités

La création des sous entités se fait selon une structure prédéfinie correspondant aux métadonnées du 2^{ème} niveau. Cette opération est relativement rapide, et dure en moyenne:

1 minute pour une sous-entité.

1.11.3. Sélection des pages dans le panorama et transfert dans les sous-entités

Pour la sélection des pages dans le panorama et leur transfert dans les sous-entités (soit les métadonnées du 1^{er} niveau), il faut compter

1 heure pour 350 pages.

Suggestion

Pour gagner du temps pendant le transfert des pages d'un livre dans les sous-entités créées, il est préférable passer par "couper – coller" plutôt que par "copier – coller".

Décision

Pour les chapitres qui commencent au milieu de la page, le groupe de travail de DEBORA a décidé de laisser la page "physiquement" au début de la sous-entité suivante et d'établir un lien, appelé copie-renvoi, de la même page à la fin de la sous-entité précédente.

1.11.4. Vérification des pages et corrections éventuelles

Afin d'assurer que le traitement est correct, on contrôle systématiquement que toutes les pages-images appartiennent bien aux bonnes sous-entités. Le cas échéant, les corrections nécessaires sont effectuées.

1.11.5. Copie de la page du titre dans le panorama

Si la page du titre ne contient pas encore toute l'information nécessaire, (les métadonnées du 1^{er} niveau, on la complète.

Une copie de cette page (copie-renvoi) est introduite au niveau de l'entité "livre", afin que son image apparaisse d'emblée lorsqu'on "ouvre" le livre.

1.12. La consultation en mode W3

La plupart des clients W3 actuels étant capables de reconnaître les "frames" et supportant la norme HTML 3.0, peuvent accéder en consultation au serveur TransVision par une interface W3. L'URL du serveur est de type "http://www.cei-sgbi.insa-lyon.fr:9001", où "9001" représente le service utilisé et "cei-sgbi.insa-lyon.fr" le nom du serveur sur le réseau. Le client W3 affiche alors la liste des bases accessibles. A chaque base correspond la liste des entités rendues publiques par les créateurs.

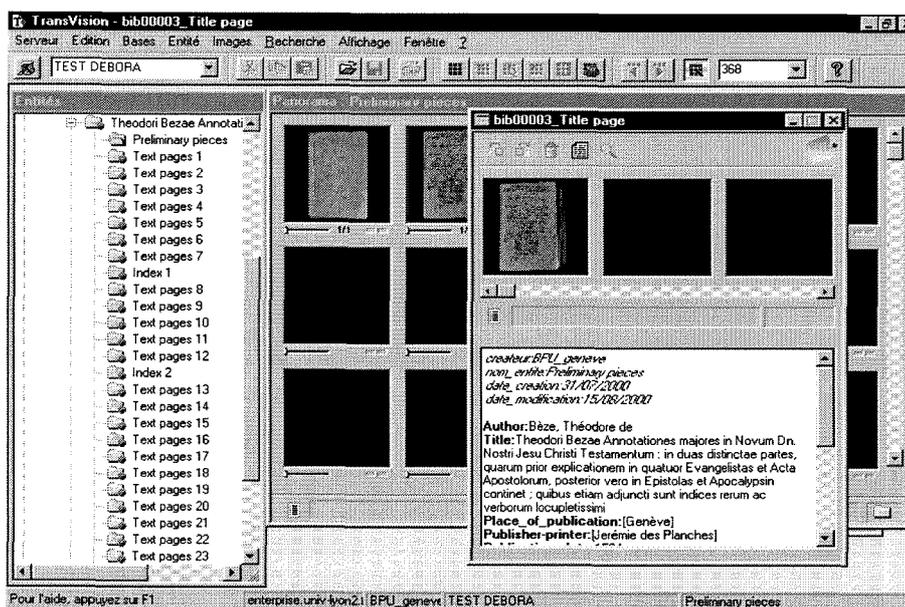


Fig. 10. Interface du client natif. Partie textuelle de la page de titre de "Theodori Bezae" (Genève BPU, S 14789)

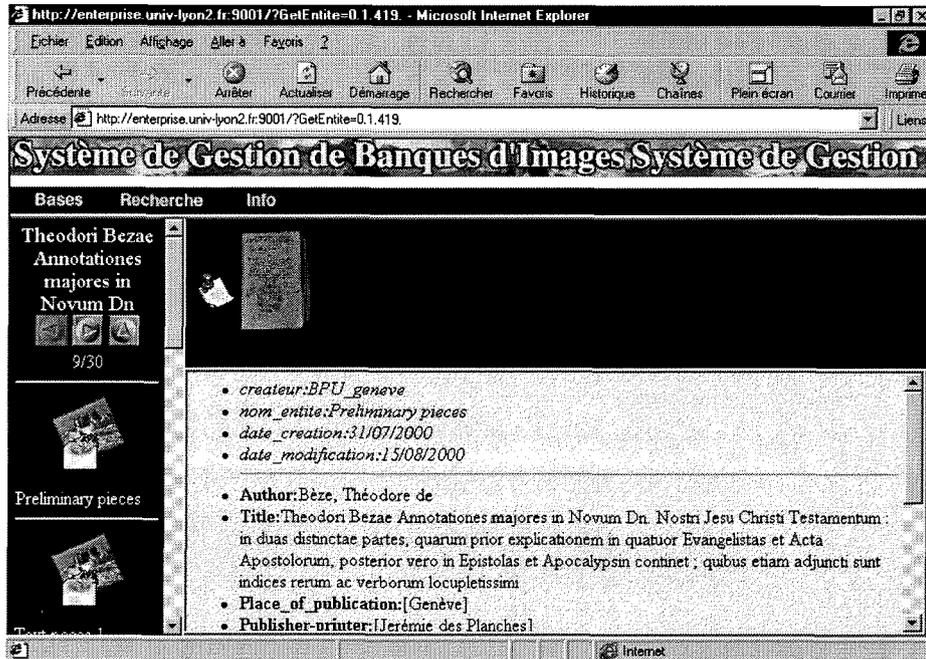


Fig. 11 : Interface Web. Partie "textuelle" de la page de titre de "Theodori Bezae..."

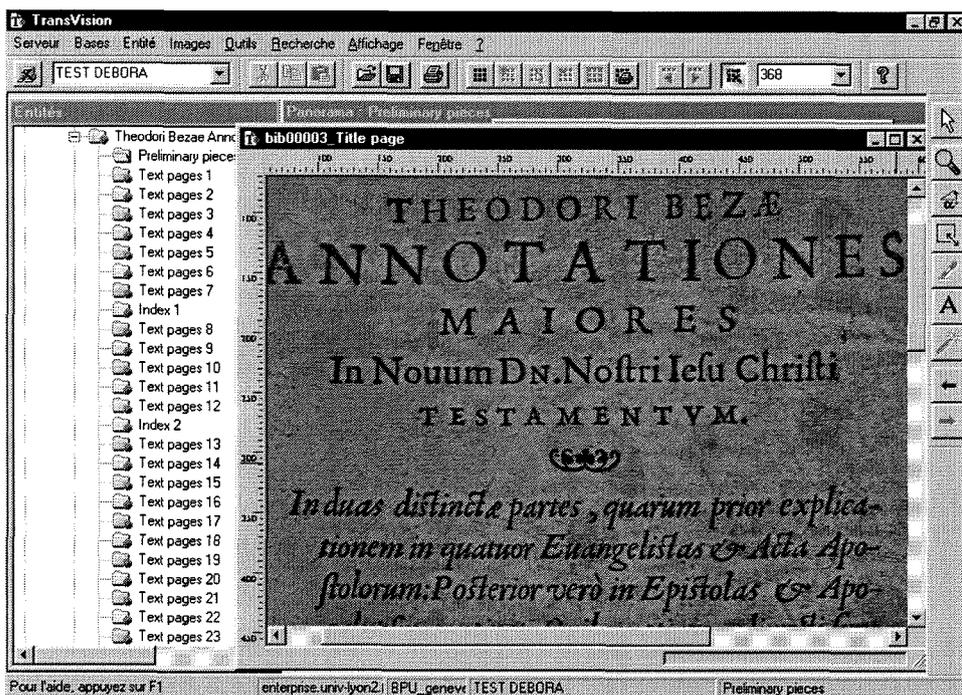


Fig. 12 : Interface du client natif. Partie image de la page de titre de "Theodori Bezae..."

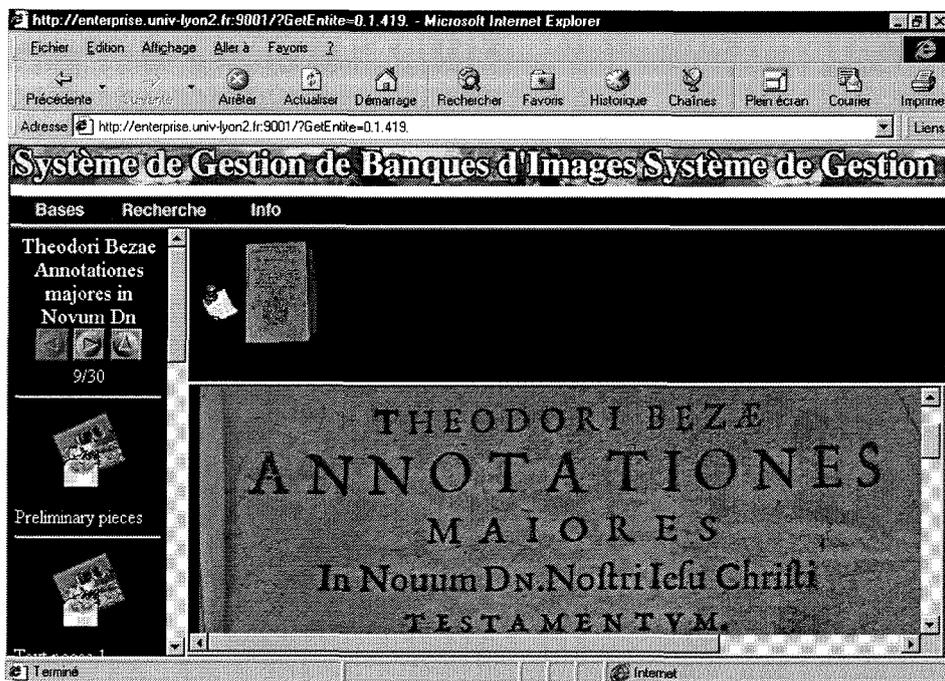


Fig. 13 : Interface Web. Partie image de la page de titre de "Theodori Bezae..."

2. Le client DEBORA

Le but de DEBORA n'est pas uniquement de rendre les ouvrages accessibles par le Web. Ce projet comprend la mise en place d'un poste de travail permettant à l'utilisateur final de pouvoir non seulement accéder à distance aux corpus de livres numérisés, mais aussi de travailler interactivement et collaborativement sur les documents électroniques. L'adaptation de l'outil coopératif ARIADNE pour le projet DEBORA est pris en charge par Lancaster.

Pour garantir l'interopérabilité entre le client et l'ensemble des systèmes de serveurs, le protocole réseau utilisé pour la communication client-serveur est basé sur la norme Z39.50. Ceci permet au client de se transposer facilement si nécessaire.

2.1. La norme Z39.50

Z39.50 est un protocole de communication dédié aux applications de recherche documentaire. Il permet de spécifier les procédures nécessaires pour effectuer des

recherches dans des bases de données distantes et le rapatriement sur un poste local du résultat de ses recherches. Z39-50 est une norme nationale américaine conçue initialement pour l'environnement OSI. Si l'on se réfère à la division en couches utilisée par l'OSI "Open Systems Interconnection", Z39.50 se situe dans la couche application, mais qui a été finalement implantée au-dessus de TCP / IP (Transfer Control Protocol), protocole de communication utilisé sur Internet. Ce protocole ne s'applique pas à l'interface avec les usagers; il porte sur la partie de communication, de recherche et d'application. L'implémentation de Z39-50 est encore une opération longue et coûteuse pour l'instant.

Z39.50 est basé sur un modèle d'architecture client-serveur. Ce protocole définit un protocole d'interrogation et de réponse suivant un dialogue entre un client et un serveur. Son avantage est que tout utilisateur disposant d'un logiciel client Z39-50 peut interroger toute base d'information accessible sur un serveur conforme Z39-50. Aujourd'hui Z39-50 s'appuie sur les normes de communication TCP/IP et se trouve de ce fait parfaitement adapté à Internet. Il est donc facile d'imaginer que l'usage d'un logiciel client conforme à Z39-50 et fonctionnant sur un système relié à Internet puisse interroger toute base bibliographique disponible sur Internet et mettant en œuvre un logiciel serveur conforme à Z39-50. Ces informations bibliographiques peuvent non seulement servir pour la consultation, mais également pour d'autres besoins, notamment dans le cadre d'échange et de récupération de notices

2.1.1. Protocole d'interrogation client-serveur

Z39-50 définit un protocole d'interrogation entre un client posant une requête d'interrogation à un serveur effectuant la recherche dans ses bases et renvoyant le résultat trouvé. Le protocole fournit ainsi la nature et le format d'un ensemble de questions et de réponses - appelés services - sans aucune hypothèse sur la nature, la structure, le format bibliographique et le jeu de caractères employés dans les bases bibliographiques ou documentaires consultées.

Ce protocole applique le principe de fonctionnement client-serveur : un logiciel client Z39-50 s'adresse à un logiciel serveur Z39-50.

2.1.2. Rôle du logiciel client

Le logiciel client doit prendre en charge les fonctions suivantes :

- dialogue homme- machine avec l'utilisateur;
- présentation, interface d'interrogation, ergonomie générale, formatage et présentation des résultats, fonctionnalités complémentaires à l'interrogation;
- impression, mémorisation, export, traduction en Z39-50 des requêtes utilisateurs d'interrogation, l'adressage des requêtes vers les serveurs interrogés, le décodage des résultats retournés en Z3950.

Le logiciel client reçoit toutes les informations constituant le résultat dans des formats structurés de stockage ou d'échange classiques ; il a ensuite la charge de traiter ces informations " brutes " en local pour les présenter à l'utilisateur.

2.1.3. Rôle du logiciel serveur

Le logiciel serveur doit prendre en charge les fonctions suivantes :

- le décodage Z39-50 des requêtes reçues;
- l'interprétation de cette requête, puis son exécution sur les bases interrogée (la requête est généralement adressée en interne sur le serveur au logiciel de gestion de la base de données);
- le formatage puis le codage du résultat trouvé en Z39-50;
- puis l'envoi du tout au client Z39-50.

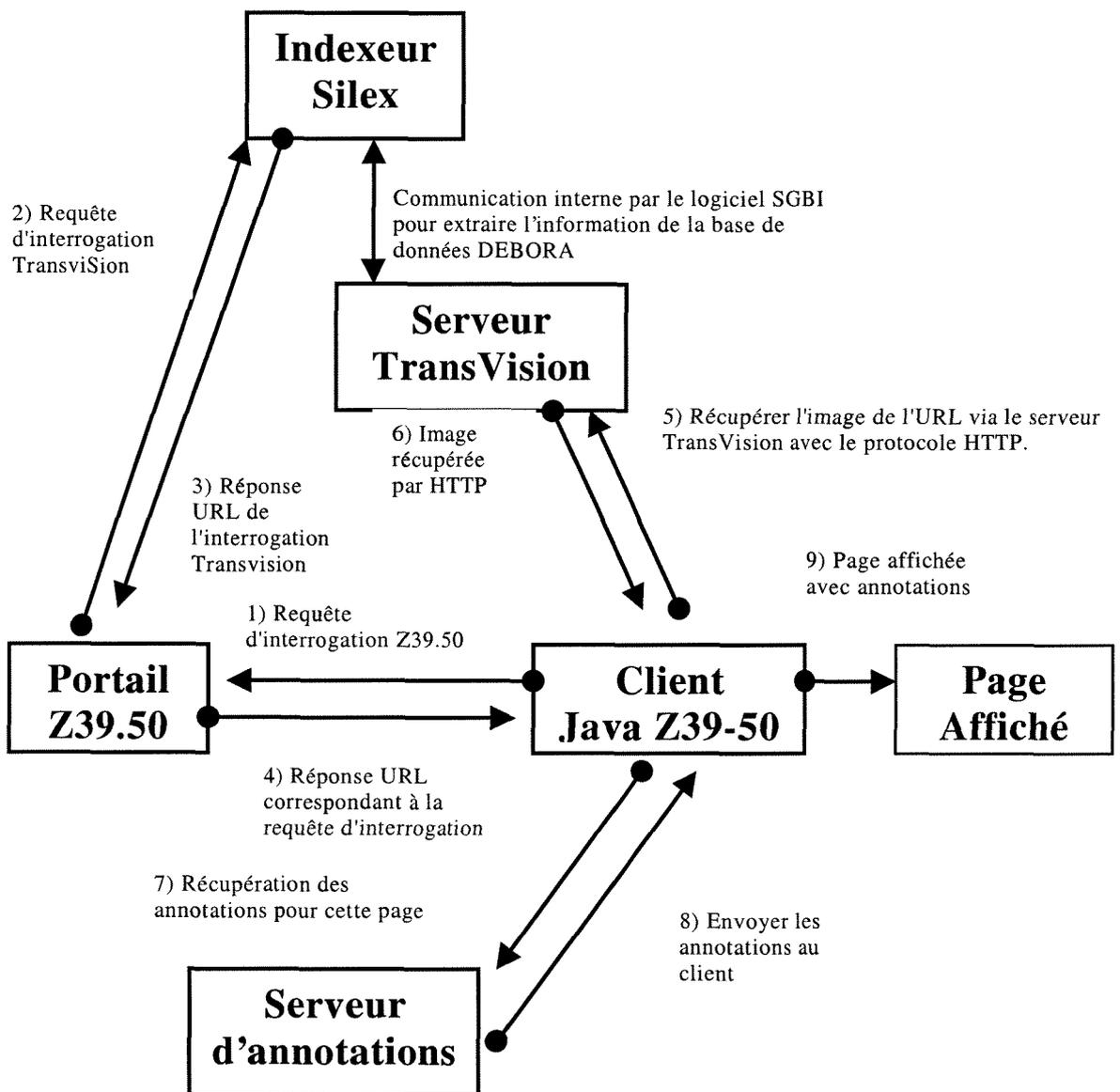


Fig. 14. Architecture client-serveur DEBORA

2.2. Interactions entre composants logiciels

1. Requête d'interrogation Z39.50

Le client Java envoie une requête d'interrogation de type – 1 RPN Z39.50 au portail SGBI Z39.50. Celle-ci est ensuite acheminée dans une deuxième étape à l'indexeur Silex qui fait partie de l'application SGBI.

2 Requête d'interrogation native TransVision

Le portail Z39.50 traduit la requête Z39.50 venant du client Java dans un format compréhensible en interne par le logiciel SGBI. Cette requête est ensuite envoyée à l'indexeur Silex pour l'interrogation de la base de données.

3 Réponse URL à la requête TransVision

Après une communication de l'indexeur Silex au serveur TransVision, la réponse à la requête est envoyée au portail Z39.50.

4 Réponse URL correspondant à la requête d'interrogation

L'URL de l'image qui correspond à la requête des étapes précédentes est retournée au client Java.

5 Requête d'image HTTP

Le client Java demande alors l'image numérique de la page dans la base de données du serveur TransVision en utilisant le protocole HTTP et cette URL.

6 Téléchargement de l'image HTTP

Les données image de la page (e.g. GIF, TIFF, JPEG, etc.) sont renvoyées au client via une réponse HTTP.

7 Récupération des annotations pour la page affichée

Avant l'affichage pour l'utilisateur final, le client Java récupère toutes les annotations existantes pour la page affichée. Ces annotations peuvent être faites par d'autres utilisateurs via les mécanismes CSCW d'annotation intégrés au client.

8 Annotations envoyées au client

Le serveur d'annotations, basé sur le langage Java, retourne toute annotation correspondant à la page demandée.

9 Page affichée et annotations chez le client

La page précédemment demandée est maintenant affichée sur l'application client ainsi que toute annotation d'utilisateur s'y référant.

BIBLIOGRAPHIE

CZAPLINSKI, Jean Marc

L'avenir des formats de communication : Z39.50 et l'échange de données bibliographiques en ligne. 8 octobre 1996
en ligne :(<http://www.acctbief.org/avenir/z3950.htm#par2>) visité le 11.08.2000

DEBORA

Developing an Interface to Support Collaboration in a Digital Library.
(<http://www.comp.lancs.ac.uk/computing/research/cseg/projects/ariadne/docs/>)
visité le 23.08.2000

PEMBERTON Duncan

DEBORA client functional requirements / by Duncan Pemberton, David Nichols, Majid Ihadjadene. 23.08.2000. (DEBORA Telematics Applications : Programme - project no. 5608)
(<http://www.comp.lancs.ac.uk/computing/research/cseg/projects/debora/>) visité le 23.08.2000

PEMBERTON Duncan

DEBORA : Digital Acces to Books of the Renaissance / by Duncan Pemberton, David Nichols, in Exploit Interactive, Issue 3, octobre 1999
(<http://www.exploit-lib.org/issue3/debora/>) visité le 5.08.2000

ROLE, François

Deux outils pour les bibliothèques distribuées : la norme Z39.50 et le protocole HTTP; in BBF. Paris, T.40, n.%,1995. p. 50-53

SITES INTERNET

<http://frantiq.mom.fr/>

<http://www.univ-lyon2.fr/SGBI/SGBImenu.html>

V. Résultats d'enquête

Parallèlement au développement des techniques informatiques, le groupe de travail de DEBORA a conduit une enquête auprès des utilisateurs du livre du 16^{ème} siècle afin de mieux envisager les outils à mettre en place, leurs fonctionnalités et leur pertinence d'usage pour un accès numérisé à ces documents.

L'enjeu de cette enquête est important, les pratiques d'accès numérisé restant peu fréquentes dans ce domaine de recherche. C'est également ce qui en fait la difficulté, par manque de points de comparaison.

Un des objectifs de mon stage à la BPU de Genève était la réalisation d'une enquête auprès des utilisateurs du livre du 16^{ème} de cette ville qui a connu un si grand essor à cette époque (voir: "Genève au 16^{ème} siècle").

Les résultats obtenus à Genève pourront être comparés avec les enquêtes déjà menées dans le cadre de DEBORA et sur la base du même questionnaire à Lyon, Rome et Coimbra.

1. La population

La pré-enquête a été réalisée auprès d'une population-test de chercheurs-historiens, bibliothécaires et étudiants, supposée avoir une pratique régulière de la recherche documentaire informatisée et du travail sur les livres anciens.

Cette population est composée de 9 chercheurs de l'Institut et du Musée d'histoire de la Réformation (IHR/MHR) dépendant de l'Université de Genève, de 2 bibliothécaires de la BPU de Genève, ainsi que d'1 étudiante.

2. Objectifs de l'enquête et hypothèses préalables

Selon le préambule du questionnaire, cette enquête vise, dans le cas de DEBORA, à recueillir les habitudes d'usage du livre, de préférence auprès des seiziémistes, pour une meilleure prise en compte de leurs besoins spécifiques et pour l'accès distant au livre numérique du 16^{ème} siècle, ainsi qu'à étudier leurs pratiques de travail collaboratif.

L'examen du questionnaire lui-même nous a permis de construire quelques hypothèses. Il s'agirait de cerner les différents usages du texte numérique (modalité, outils d'accès aux documents et de travail sur les documents) pour évaluer l'influence de ces usages sur la recherche et notamment sur le travail collaboratif. Les hypothèses pourraient donc être formulées de la manière suivante:

- Le livre du 16^{ème} siècle est intéressant à étudier, à cause de sa complexité.
- Le travail sur le texte numérique du 16^{ème} siècle favoriserait une diversification des pratiques de recherche et générerait des méthodes de travail particulières pour les seiziémistes.
- Le travail sur le texte numérique favoriserait le travail collaboratif.

3. Méthode: une enquête par questionnaire

L'enquête réalisée à Genève au cours de l'été 2000, auprès d'utilisateurs du livre du 16^{ème} siècle, n'a pu se faire que sous la forme d'un questionnaire établi dans le cadre de DEBORA.

Pour des questions techniques (indisponibilité de la base DEBORA), il n'a pas été possible d'expérimenter un scénario pour définir les attentes et les besoins de l'utilisateur face aux livres numérisés, tels qu'ils se présentent dans le client DEBORA.

Ce questionnaire a été rédigé par l'équipe de DEBORA et divisé en cinq parties principales:

1. Profil de l'utilisateur interrogé
2. Usages du livre ancien
3. Usages du livre du 16^{ème} siècle
4. Livres en ligne et technologie de l'information et de la communication
5. Communications et échanges autour du texte/livre numérique du 16^{ème} siècle

4. Analyse

4.1. Partie "profil de l'utilisateur" (questions 1 à 14)

Les questions regroupées dans cette partie (7 questions fermées, 5 questions ouvertes et 2 questions numériques) proposent un certain nombre de variantes utiles à la définition du profil socio-professionnel des personnes interrogées. 3 types d'indicateurs sont identifiables: les indicateurs "biologiques" (sexe, âge), les indicateurs relatifs au capital social, culturel et économique (langue, diplôme, profession) et des indicateurs plus spécifiquement adaptés à la population ciblée (les "chercheurs") devant servir à les catégoriser avec plus de précision.

- Notre échantillon est constitué principalement d'enseignants-chercheurs (70%), tous francophones, mais parlant plusieurs langues;
- la population questionnée est à majorité masculine (75%);
- on note une dispersion des âges: de 25 ans (étudiante) à 72 ans (retraité) avec une moyenne d'âge de 50,7 ans;
- du point de vue de la nationalité, l'échantillon comprend 7 Suisses, 3 Français et 3 Italiens, qui habitent tous en Suisse à l'exception d'une personne;
- 6 personnes mènent une recherche à la fois collective et individuelle, 4 travaillent individuellement et 2 collectivement. L'écart entre l'approche individuelle et collective est donc minime.
- dans leurs recherches actuelles ces personnes se trouvent :
 - pour la plupart au milieu de leur activité de recherche (46%)

- pour une minorité, au début (23 %) et pour le reste à la fin (31%). A noter qu'une personne mène plusieurs recherches de front et se trouve, selon l'objet de sa recherche, au début, au milieu et à la fin!

4.2. Partie "usage du livre ancien" (questions 15 à 23)

Cette partie concerne les questions 15 à 23, soit 7 questions fermées et 3 questions ouvertes. Elle traite des pratiques de recherches et des modalités d'accès aux livres anciens.

- Toute personne interrogée a travaillé ou travaille actuellement sur des livres imprimés du 15^{ème} au 18^{ème} siècles;
- la préférence va clairement aux éditions de l'époque, la photocopie papier recueillant également beaucoup de suffrages. Par contre, la copie électronique est actuellement très peu utilisée (une personne seulement: 3%);
- la majorité des personnes questionnées utilise les services des bibliothèques dépositaires des originaux (11 sur 12), tout en opérant aussi sur leurs lieux de travail (8 sur 12);
- toute la population est familiarisée avec les catalogues de bibliothèques et les bibliographies spécialisées. Il s'agit des instruments les plus utilisés pour les recherches bibliographiques;
- une seule personne n'est pas familiarisée avec les outils électroniques; pour le reste, toutes les utilisent, avec une préférence pour les catalogues de bibliothèques et bases de données en ligne;
- le texte lui-même arrive en tête des centres d'intérêt (11 sur 12), au contraire de la couleur qui n'intéresse qu'une seule personne.

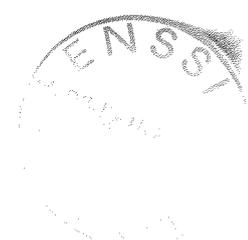
Les intérêts de pour le livre ancien (en particulière du 16^{ème} siècle) sont différents selon les spécialités des chercheurs, comme le montre le tableau comparatif ci-dessous:

spécialités	l'aspect original du livre	un exemplaire particulier	le format du livre	la couleur	les illustrations	les imperfections le texte	les différents parties du contenu	autres
bibliographie-histoire	1	1	1		1	1	1	
bibliothécaire						1		
histoire religieuse	3						1	
archiviste-paléographe	1						1	1
histoire	2	3	2		3	1	1	1
sc. politiques	1					1		
édition critique	1							
histoire de l'art	1	1	1		1	1	1	
philosophie-théologie-droit	1		1			1	1	

4.3. Partie "usages du livre du 16^{ème} siècle" (questions 24 à 45)

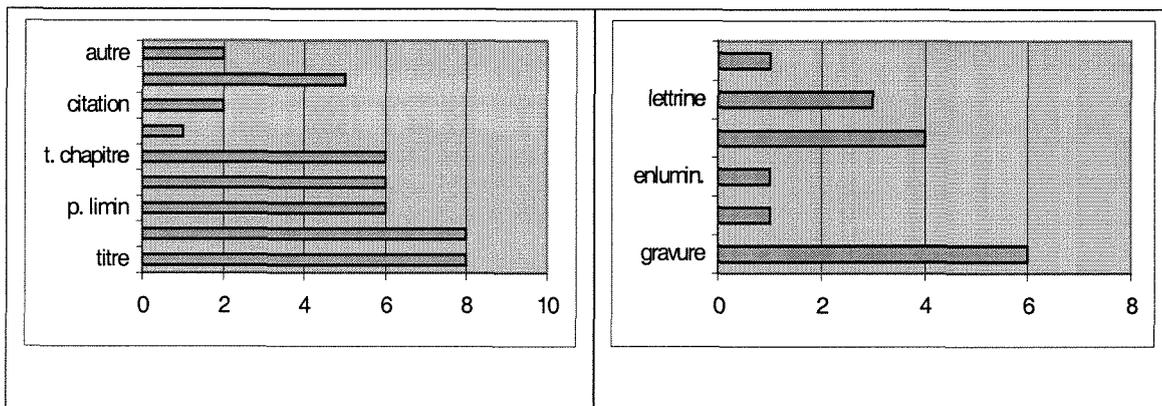
Cette partie concerne les questions 24 à 45, soit 14 questions fermées, 7 questions ouvertes et 1 question numérique. Elle traite directement des pratiques de recherches et des modalités d'accès aux livres du 16^{ème} siècle, et étudie les éléments les plus importants de ces livres au niveau de leur structure et de leur contenu.

- Sur 12 personnes questionnées, 10 travaillent avec/sur les livres du 16^{ème} siècle;
- les bibliothèques de Genève sont les plus sollicitées, très peu de chercheurs se déplaçant;
- la moitié seulement (5 sur 10) des personnes consultent des bases bibliographiques propres aux livres du 16^{ème} siècle, et parmi elles 4 seulement citent des noms de bases bibliographiques qu'elles utilisent; pour ce qui est des bases textuelles propres à ces livres, elles sont moins nombreuses: sur 8 personnes, une seule les consulte, mais même cette dernière ne mentionne aucune des bases utilisées;



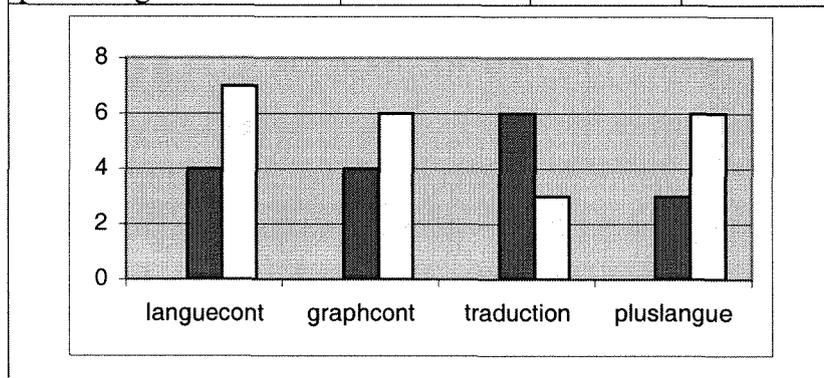
- les bibliographies spécialisées viennent en tête des sources bibliographiques utilisées par les chercheurs pour le repérage des livres du 16^{ème} siècle;
- parmi ceux qui travaillent avec les livres du 16^{ème} siècle, certains le font de longue date (depuis 1948 pour la mention la plus ancienne), d'autres depuis peu de temps (début 2000 pour la mention la plus récente). La moyenne se situe aux alentours de 1975, ce qui représente globalement une longue expérience;
- pour 50% de ceux qui travaillent avec les livres du 16^{ème}, il s'agit de leur activité principale;
- pour tous l'accès aux fonds des livres du 16^{ème} est généralement facile; s'il y a des difficultés, elles concernent plutôt l'obtention des reproductions (6 sur 7 personnes, soit 28% des difficultés);
- s'agissant de la structure du texte numérique du 16^{ème} siècle on peut dire qu'elle est complexe : les deux parties qui intéressent particulièrement les chercheurs dans ces livres sont: le titre et le texte (6 sur 8); et parmi les éléments des illustrations, ce sont les gravures (6 sur 9);

structextliv16		illustliv16	
S'agissant de la structure du texte numérique, quelles sont les parties qui vous intéressent particulièrement dans le livre du 16 ^e ?		S'agissant des illustrations, quels sont les éléments qui vous intéressent dans le livre du 16 ^e ?	
non-réponse 2		non-réponse 3	
titre	17,5%	gravures	38%
texte	17,5%	bandeaux	6%
pièces liminaires	14%	enluminures	6%
tables des matières	14%	matériel typographique	25%
tables des chapitres	14%	lettrines	19%
notes	2%	autres	6%
citations	4%		
index	13%		
autres	4%		
TOTAL	100%	TOTAL	100%



- concernant la langue des livres du 16^{ème}, la majorité refuse:
 - d'utiliser la translation en langue contemporaine (7 sur 10),
 - d'utiliser la transcription en graphie contemporaine (6 sur 10),
 - d'utiliser des textes en plusieurs langues (6 sur 9),
 la préférence va à la langue usuelle (6 sur 10).

	toujours	parfois	jamais
langue contemporaine	0	4	7
graphie contemporaine	0	4	6
traduction	0	6	3
plus. langues	0	3	6



- les chercheurs privilégient surtout la prise de notes manuscrites (7 sur 10), par rapport à la constitution de corpus (5 sur 10) ou de dossiers et d'extraction de mot.

4.4. Partie "livres en ligne et technologie de l'information et de la communication" (questions 46 à 70)

Cette partie concerne les questions 46 à 70, soit 21 questions fermées et 3 questions ouvertes. Elle traite plus précisément les habitudes des personnes qui travaillent avec les livres du 16^{ème} siècle et qui utilisent un ordinateur pour leurs recherches (méthodes, environnement informatique, outils, etc.).

A propos des questions concernant le mode d'utilisation du texte numérique, on peut faire les remarques suivantes:

- L'ensemble de la population interrogée utilise un ordinateur pour ses travaux de recherches (67 sur 10).
- Elle n'est pas très habituée à la terminologie " mode texte" et " mode image", la moitié des personnes (6 sur 12) n'ayant pas répondu à cette question (n°68). Pour le reste, 2 personnes disent clairement qu'elles ne font pas la différence entre les deux modes sur un écran; en fin de compte 5 personnes sont capables de faire la différence et toutes les cinq utilisent le mode texte.
- A propos du matériel de visualisation du texte numérique (4 réponses seulement sur 12 questionnaires), une personne a répondu par un signe d'interrogation (soit elle ne comprend pas la question, soit elle n'a aucune connaissances de la taille d'un écran). En revanche tous veulent travailler sur leurs écrans soit en taille réelle de l'image, soit en agrandissant avec une loupe. Personne ne souhaite la réduction de l'image.
- L'ensemble des personnes interrogées travaille pour le moment principalement avec le mode texte, mais estime que celui-ci n'est pas indispensable.
- Les recherches sur le web ne sont pas très familières: 6 personnes sur 7 ont déjà consultés des bibliothèques numériques sur le web, 3 sur 6 utilisent des moteurs de recherche (les principaux cités sont Altavista, Yahoo, Copernik, Ask), mais ils ont habitude de se servir des mots-clés et des listes de liens préétablie (6 sur 7).
- On peut remarquer aussi que 71% des personnes qui ont répondu utilisent des signets (bookmarks) pour ce type de travail et que tous ceux qui ont répondu

avouent rencontrer des difficultés dans ce type de recherche, surtout parce qu'ils trouvent peu d'informations réellement utilisables.

- L'attachement au papier se fait encore bien sentir car 4 personnes sur les 7 qui ont répondu préfèrent travailler sur un document original papier; une seule sur une reproduction imprimée et 2 directement sur l'écran. Ce fait est confirmé par la question n°58 qui montre que tous impriment des fragments de texte numérique pour réaliser un corpus thématique. Ceci rejoint le constat fait à la question n°45 à propos de l'habitude des prises de notes manuscrites et constitution de corpus.

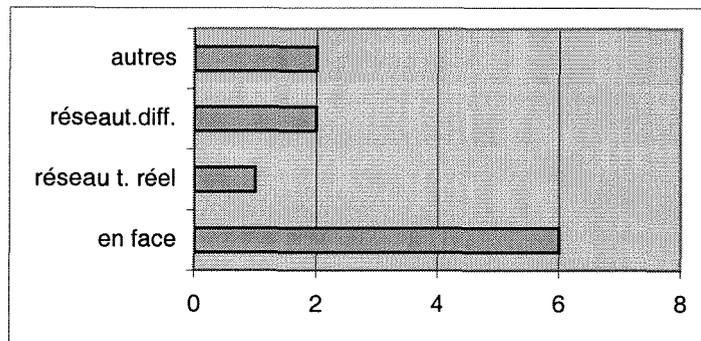
4.5. Partie "communications et échanges autour du texte/livre numérique du 16^{ème} siècle" (questions 71 à 83)

Cette dernière partie concerne les questions 71 à 83, subdivisées en 7 questions fermées (choix multiple) et 5 questions ouvertes. Ces questions touchent au cœur de l'objectif visé par le projet DEBORA, soit la détermination des pratiques de travail collaboratif des seizièmistes et des besoins qui les conduisent à cette collaboration.

Les questions concernant le travail collaboratif autour des livres numériques du 16^{ème} siècle, montrent que les personnes qui ont répondu:

- travaillent plutôt seules (8 sur 10), ou en équipe (4 sur 10), ou à la fois seul et en équipe (2 sur 10), ou encore seules et au sein d'un réseau (2 sur 10);
- celles qui acceptent l'échange des informations (6 sur 9) sont plutôt d'avis que cet échange passe d'abord par le biais de discussions informelles (5 sur 6), et ensuite par le biais des messages électroniques (3) et des réunions de travail (3); personne ne préfère (ou ne connaît) les publications électroniques;
- celles qui acceptent de collaborer préfèrent le faire en face à face (6 sur 9);

travcollabmaniere		
non-réponses 3		
en face à face	6	55%
par réseau informatique en temps réel	1	9%
par réseau informatique en temps différé (48 heures maximum)	2	18%
par réseau informatique à plus long terme	0	
autres	2	18%
TOTAL	9	100%



- la majorité des personnes souhaite partager leurs notes avec d'autres usagers du livre numérique du 16^{ème} (6 sur 9), pour autant qu'elles se limitent à des informations sur le texte lui-même;
- ces personnes reconnaissent le besoin de collaboration avec les autres, pour l'obtention d'informations sur les documents (6 sur 9), ainsi que d'adresses Internet pertinentes (5 sur 9);
- pour les découvertes (ou des idées originales), les chercheurs sont plutôt d'avis de les soumettre à des personnes ressources pour validation (6 sur 8), seule une minorité affirme catégoriquement garder le secret jusqu'à la publication (3);
- curieusement, les personnes interrogées ne sont pas très intéressées à disposer d'outils d'échanges avec d'autres chercheurs intégrés à leur poste de travail (5 sur 6); la plupart n'a pas répondu à cette question (6 sur 12), une personne est hésitante (elle a répondu à la fois oui et non!), une seule personne est entièrement pour la possession de ces outils.

BIBLIOGRAPHIE

[**ARB99**] ARBORIO, Anne-Marie; FOURNIER, Pierre; SINGLY, François de
L'enquête et ses méthodes: l'observation directe. Nathan, 1999. 127 p (Collection
128. Sciences sociales). N. 216

[**BER99**] BERLOTTIER, Anne Sereine

L'action culturelle dans les bibliothèques de la Seine-Saint-Denis. Enssib 1999.
Mémoire DCB. 105p. IX f.

[**ENQ95**] Enquête DCB1 et DCB2 / [Ecole nationale supérieure des sciences de
l'information et des bibliothèques]. [1] Résultats, questions ouvertes; [2]
Traitements statistiques Enssib 1995. 2 vol. 137, 41f.

[**KHA00**] KHAMKHAM Laurence

Les initiateurs nouvelles technologies face aux besoins des utilisateurs : une
double enquête à la BM de Lyon. DCB : Mémoire d'étude. Enssib 2000. 109 +
26 p.

[**LES94**] LESCARBEAU, Robert

L'enquête feed-back. Montréal : Les Presses de l'Université de Montréal, 1994. XI,
155 p. (Intervenir)

[**MOS90**] MOSCAROLA, J.

Enquêtes et analyses de données avec le Sphinx. [Paris]: Vuibert, 1990. 307 p. : ill.
(Gestion).

Conclusion

DEBORA est un projet pilote financé par l'Union européenne en cours de réalisation et plusieurs équipes y travaillent actuellement. L'INSA de Lyon et l'Instituto Superior Tecnico de Lisbonne se penchent sur les questions de reconnaissance de forme, l'Université de Lancaster développe le poste de consultation de DEBORA. Ces recherches concernent directement les niveaux de métadonnées 3 (données sur des éléments du livre) et 4 (annotations et échanges). Cependant l'état d'avancement du projet ne nous a pas permis de travailler sur ces deux derniers niveaux de métadonnées prévues dans le concept de DEBORA.

Le bilan de notre stage peut se résumer de la manière suivante:

3 ouvrages, totalisant 1796 pages, ont été intégralement numérisés, stockés sous forme électronique et organisés logiquement dans le serveur TransVision ainsi que mis à disposition de manière publique sur le Web.

Parallèlement, une enquête a été entreprise dans le but de reconnaître les pratiques et les attentes des chercheurs qui travaillent sur le livre du XVIème s. ou plus généralement, sur le livre ancien, fréquentant la Bibliothèque publique et universitaire.

Mais il n'a pas été possible de mettre ce chercheur en face du livre numérisé, pour observer ses réactions, son enthousiasme - ou son appréhension - face à cette forme nouvelle d'approche. En effet des questions techniques – transfert de la base TransVision sur support magnétique, état encore expérimental des développements déjà souligné – ainsi que le temps limité, n'ont pas permis l'installation d'un poste de travail équipé du client DEBORA. Ceci reste encore à faire, afin que la BPU puisse remplir sa mission de bibliothèque-test dans le cadre de ce projet.

ANNEXES

Sommaire

ANNEXE 1: CORRESPONDANCE ENTRE Z39.50 ET LE FORMAT MARC	II
ANNEXE 2 – DOCUMENTS DE SAISIE DES COÛTS.....	IV
ANNEXE 3. USAGES ET USAGERS DU LIVRE DU 16^E SIECLE	VII
1. PROFIL	VIII
2. II USAGE DU LIVRE ANCIEN	XI
3. USAGES DU LIVRE DU 16 ^{ème} SIECLE	XIV
4. Livre en ligne et technologie de l'information et de la communication	XXI
5. COMMUNICATIONS ET ECHANGES AOUTOUR DU TEXTE/LIVRE NUMERIQUE DU 16 ^e SIECLE	XXVIII

Annexe 1: Correspondance entre Z39.50 et le format MARC

Bib-1 Attribute Set used	OCLC MARC
Not a Bib-1 attribute	File:
1031 Material type	655 or 516
4 Title	245\$a
1032 Doc ID	856\$u (if URL)
43 Title abbreviated	246
1 Name-Personal 2 Name - Corporate 3 Name - Conference 1000 Author-name-and- title 1002 Name 1003 Author-Name 1004 Author-name - personal 1005 Author-name- corporate 1006 Author Name- conference	100, 110, 111, 700, 710, 711, or 245\$c
1000 Author-name-and- title 1002 Name 1003 Author-Name 1004 Author-name - personal 1005 Author-name- corporate 1005 Author-Name- conference	100, 110, 111, 700, 710 or 711
63 Note	500
63 Note	538
62 Abstract	520
63 Note	500
63 Note	524
63 Note	500
1018 Name-Publisher	260\$b
63 Note	500
31 Date-Publication	260\$c
63 Note	500
21 Subject	653
Not a Bib-1 attribute	250

Not a Bib-1 attribute	538
Not a Bib-1 attribute	256
54 Code-Language	Lang:
63 Note	500
7 Identifier-ISBN	020
8 Identifier-ISSN	022
1012 Date/Time last modified	260\$c
(As below)	
13 Classification-Dewey 14 Classification-UDC 16 Classification-LC 20 Classification-Local	082, 080, 050 or 084

Annexe 2 – Documents de saisie des coûts¹

RELEVÉ DE MESURE DU TEMPS DE TRAVAIL		
Travail Courant et Travail de numérisation sur DEBORA		
FONCTION/POSTE OCCUPE : SALAIRE GLOBAL BRUT : SALAIRE HORAIRE BRUT :	VOLUME HORAIRE GLOBAL POUR L'ACTIVITE	EN POURCENTAGE S SUR LA PERIODE
PERIODE DU AU		
OPERATIONS COURANTES : ORGANISATION & GESTION ADMINISTRATION & ANIMATION ACQUISITIONS, COURRIER, DEPOUILLEMENT DOSSIERS DOCUMENTAIRES BASE DE DONNEES QUESTIONS PRODUITS DOCUMENTAIRES ETC.		
<i>SOUS TOTAL I</i>		
OPERATIONS SPECIFIQUES DE LA CHAÎNE DEBORA : SELECTION PREPARATION & NUMERISATION INTEGRATION BASE DE DONNEES DISTRIBUTION – WWW		
<i>SOUS TOAL II</i>		
AUTRE(S), PRECISEZ		
TOTAL HEURES DE TRAVAIL (I + II)		100%

¹¹ [DEB-00] DEBORA, juin 2000: chap. 4 : Analyse des coûts de la numérisation, d'intégration et de distribution par l'Internet de documents anciens et modèles économiques / par Serge Lenga et Ahmed Silem.

CATEGORIES DE LA FONCTION PUBLIQUE TERRITORIALE²

CATEGORIE	CADRE D'EMPLOI	DIPLOMES ET FORMATIONS	INDICE
Catégorie A	CONSERVATEUR	Bac + 3 ans d'études supérieures 18 mois à l'ENSSIB	
	BIBLIOTHECAIRE	Bac + 3 ans d'études supérieures 12 mois à CNFPT	
Catégorie B	ASSISTANTS QUALIFIE DE CONSERVATION	Bac + diplôme technico-professionnel (exemple : DUT métier du livre, etc.) 6 mois CNFPT	
	ASSISTANT DE CONSERVATION	Bac	
Catégorie C	AGENT QUALIFIE DU PATRIMOINE	Diplôme homologué du niveau IV	
	AGENT DU PATRIMOINE	Ni diplôme, ni concours	

² pour la France

Matrice des coûts de sélection des documents

	Éléments / tâches	Nombre(s) (1)	Temps de travail moyen/ Coût unitaire de l'élément (2)	Coût des opérations (3)=(1)×(2)
COUTS DE SELECTION	Ouvrages retenus			
	Pages			
	Contrôle de la qualité des ouvrages			
	Acquisition d'ouvrages			
	Emprunts à des collections externes			
	Droits d'utilisation (copyright)			
AUTRES ELEMENTS DE COUTS ADDITIONNELS	Equipements			
	Ordinateurs			
	Scanners			
	Dispositifs de stockage			
	Logiciels, précisez			
	1.			
	2.			
	3.			
	4.			
	Intervention de tiers			
	Formation			
	Consultants			
	Autres, précisez			
	Consommables			
	CD-ROM, DVD-RAM, etc.			
	Transports			
	Téléphone			
Electricité				
Papiers, encre, etc.				
Autres, précisez				
	COUT GLOBAL DE SELECTION			

Annexe 3. USAGES ET USAGERS DU LIVRE DU 16^e SIECLE

1. Plan de l'enquête

L'enquête comporte 83 analyse(s) réparties comme suit:

Questions fermées

1. Sexe; 4. Pays; 5. Nationalité; 6. Langue; 10. Activreche; 13. Indivgroup; 14. Rechactuelle; 15. Livrimprim; 16. Edition; 17. Lieuxconsult; 18. Rechbibliogr; 20. Rechbibelctr; 21. Aspectinters; 23. Sessiontrav; 24. Travliv16ème; 26. Bibliogrpropre16; 28. Basliv16; 30. Sourcebibliogr; 33. Domcherch16e; 34. Acceliv16; 35. Difficaces; 37. Structextliv16; 39. Illustliv16; 41. Languecont; 42. Graphcont; 43. Traduction; 44. Pluslangue; 45. Opératext16; 46. Ordinateur; 47. Taillecran; 48. Bibnumweb; 49. Rechwebliv16; 50. Rechweb; 52. Difficrechweb; 53. Difficrechweb_Oui; 54. Bookmark; 55. Motrech; 57. Travailleur; 58. Consultecran; 60. Metadonecran; 61. Métadonestrav; 62. Métadonesref; 63. Corpusnumtrav; 64. Classtextnum; 65. Fonctionnalités; 67. Modtextimage; 68. Modtextimage_Oui; 69. Modtext; 70. Tailletexte; 71. Travcollab; 72. Travcollab_Equipe; 75. Partagenotes; 76. Contenuchange; 77. Besoincollab; 79. Attitudedecouv; 81. Outilechange.

Questions ouvertes

7. Langue_Autre; 8. Dernidiplôme; 9. Profession; 11. Domainreche; 12. Sujeteche; 19. Rechbibliogr_Autres; 22. Aspectinteres_Autres; 25. Bibliothprinc; 27. Nombibliogr; 29. Basetextprinc; 31. Sourcebibliog_Autres; 36. Difficaces_Autres; 38. Structextliv16_Autres; 40. Illustliv16_Autres; 51. Rechweb_Autres; 56. Motrechsouv; 59. Consultecran_Autres; Fonctionnalité_Autres; 74. Travcollabmaniere; 78. Besoincollab_Autre; 80. Typetravlivnum; 82. Outilechange_Autre; 83. Suggestcomment.

Questions numériques

2. Age; 3. Codepostal; 32. Daterech.

2. Analyse

2.1. PROFIL

1. Sexe

Sexe:

- SEXE	Nb. cit.	Fréq.
Homme	9	75 %
Femme	3	25%
TOTAL OBS.	12	100%

2. Age

Age ?

âge	Nb. cit.	Fréq.
de 20 à 30	2	17%
de 31 à 40	1	8%
de 41 à 50	2	17%
de 51 à 60	4	33%
plus de 61	3	25%
TOTAL OBS.	12	100%

Minimum = 25 ans, Maximum = 72 ans

Somme = 608 ans, Moyenne = 50,7 ans

La question est à réponse ouverte numérique. Les observations sont regroupées en 5 classes d'égale amplitude.

3. Codepostal

Code postal de votre lieu d'habitation?

CH 1294; CH 1228, CH 1228, CH 1204 (3 fois), CH 1211, F 74100, CH 1227, CH 1205 (2 fois), CH 1231.

4.Pays:

Pays:

Suisse (11 fois), France

5.Nationalité

Nationalité

Suisse (8 fois), Française (2 fois), Italienne (2 fois).

6.Langue

Langue(s) pratiquées ?

Vous pouvez cocher plusieurs cases.

Langue	Nb. cit.	Fréq.
Français	12	25%
Allemand	8	17%
Anglais	9	19%
Espagnol	2	4%
Italien	4	21%
Latin	10	8%
Portugais	1	2%
Autres	2	4%
TOTAL OBS.	48	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (8 au maximum).

7.Langue_Autre

Si "autres langues", précisez:

grec (2 fois), hébreu

8.Derndiplôme

Quel est votre dernier diplôme obtenu?

Doctorat es Lettres (3 fois), Doctorat Théologie, Licence es Lettres (3 fois), Licence sc. politiques, Archiviste-paléographe, Habilitation, DEA, $\frac{3}{4}$ Licence.

9. Profession

Quelle est votre profession?

Professeur UNI (4 fois), Chercheur, Chercheur en Histoire, Historien, Bibliothécaire (2 fois), Conservatrice/Bibliothèque, Retraité, Etudiante.

10. Activreche

Votre activité de recherche actuelle concerne plutôt :

actvreche	Nb. cit.	Fréq.
Non- réponse		
une préoccupation de recherche et/ou d'enseignement	9	70%
une activité professionnelle (autre qu'enseignement et recherche)	2	15%
une activité privée (amateur, érudition)	2	15%
TOTAL OBS.	13	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (3 au maximum).

11. Domainreche

Quel est le domaine de votre activité de recherche actuelle?

Histoire religieuse 16^{ème}, 16^e – 17^e , Edition critique, Correspondance de Bèze, Bibliographie-Histoire, Histoire des mentalités du 16^e s., Histoire des Idées politiques, Histoire, Histoire-droit-théologie-philosophie, Philosophie et Théologie début 16^{ème}, Thèse en cours et Expositions dans le musée de la BPU, (une personne n'a pas répondu).

12. Sujeteche

Quel est le sujet ou thème de votre recherche actuelle?

Correspondance Bèze (2 fois), Guillaume d'Orange, Histoire théologie, Perception de l'Afrique noire au 16^{ème} siècle, Définition de la souveraineté chez Gaspar de Real de Corbon (1682-1702), Vie artistique genevoise 18^e-

19^e siècles, Histoire des voyages 16^e-17^{es}., Eglise genevoise au 17^{es}., Les dialoghi d'amore de Léon Hébreux, La tyrannie dans l'histoire occidentale.

13.Indivgroup

Votre activité de recherches actuelle est-elle :

Vous pouvez cocher plusieurs cases

indivgroup	Nb. cit.	Fréq.
Non- réponse		
individuelle	10	56%
collective	8	44%
TOTAL OBS.	18	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (2 au maximum).

14.Rechactuelle

S'agissant de votre activité de recherche actuelle, êtes-vous :

rechactuelle	Nb. cit.	Fréq.
Non- réponse	1	
plutôt au début de votre activité de recherche	3	23%
plutôt au milieu de votre activité de recherche	6	46%
plutôt à la fin de votre activité de recherche	4	31%
TOTAL OBS.	14	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (3 au maximum).

Les non-reponses ne sont pas prise en compte.

2.2. II USAGE DU LIVRE ANCIEN

15.Livrimprim

Avez-vous travaillé ou travaillez-vous actuellement avec/sur des livres imprimés du 15^{ème} au 18^{ème} siècle?

livrimprim	Nb. cit.	Fréq.
non-réponse		
oui	12	100%
non		
TOTAL OBS.	12	100%

16. Edition

Ces livres avec/sur lesquels vous travaillez sont des:

Vous pouvez cocher plusieurs cases

Edition	Nb. cit.	Réf.
non-réponse		
éditions de l'époque	12	34%
éditions postérieures	7	20%
copies microformes	6	17%
photocopies papier	9	26%
copies électroniques (numériques)	1	3%
TOTAL OBS.	35	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (5 au maximum).

17. Lieux consult

Consultez-vous ces documents dans:

Vous pouvez cocher plusieurs cases

Lieux consult	Nb. cit.	Fréq.
non-réponse		
les bibliothèques dépositaires des originaux	11	46%
autres bibliothèques	5	21%
vos lieux de travail habituels	8	33%
TOTAL OBS.	24	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (3 au maximum).

18. Rech bibliogr

Vous utilisez pour vos recherches bibliographiques des documents papiers tels que:

Vous pouvez cocher plusieurs cases

Rech bibliogr	Nb. cit.	Fréq.
non-réponse		
catalogues de bibliothèques	12	29%
annuaires	6	14%
bibliographies spécialisées	12	29%

citations et références d'ouvrages	9	21%
autres	3	7%
TOTAL OBS.	42	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (5 au maximum).

19.Rechbibliogr_Autres

Si "autres" précisez:

"Conseil des collègues" (2 fois), "dictionnaires".

20.Rechbibelectr

Utilisez-vous des outils électroniques pour vos recherches bibliographiques ?

Vous pouvez cocher plusieurs cases.

Rechbibelectr	Nb. cit.	Fréq.
aucun outil électronique	1	5%
catalogues de bibliothèques en ligne bases de données bibliographiques en ligne	11	55%
bases de données sur cédéroms	6	30%
citations d'ouvrages et de références de textes numériques	2	10%
autres		
TOTAL OBS.	20	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (5 au maximum).

21.Aspectinteres

S'agissant du livre ancien, ce qui vous intéresse particulièrement est:

Vous pouvez cocher plusieurs cases.

Aspectinteres	Nb. cit.	Fréq.
non-réponse		
l'aspect originel du livre	6	15%
un exemplaire particulier	4	10%
le format du livre	5	12%
la couleur		
les illustrations	4	10%

les imperfections	2	5%
le texte	11	26%
les différents parties du contenu	7	17%
autres	2	5%
TOTAL OBS:	41	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (9 au maximum).

22.Aspectinteres_Autres

Si "autres", précisez:

"Variantes entre éditions (verif. exemplaire)".

23.Sessiontrav

Au cours d'une session de travail, travaillez-vous habituellement sur :

sessiontrav	Nb. cit.	Fréq.
Non-réponse	1	
un seul ouvrage à la fois		
plusieurs ouvrages simultanément	11	79%
plusieurs ouvrages successivement	3	21%
TOTAL OBS.	14	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (3 au maximum).

Les non-réponses ne sont pas prise en compte.

2.3. USAGES DU LIVRE DU 16^{ème} SIECLE

24.Travliv16ème

Travaillez-vous ou avez-vous travaillé spécifiquement avec/sur des livres du 16^e siècle?

si "non" allez à la question n°71

Travliv16ème	Nb. cit.	Fréq.
non-réponse	2	17%
oui	10	83%
non		

TOTAL OBS.	12	100%
------------	----	------

25. Bibliothprinc

Quelles sont les 3 bibliothèques principales que vous utilisez pour vos travaux avec/sur le livre du 16^e siècle?

"La BPU - Genève (9 fois), MHR/IMR - Genève (5 fois), BnF Paris (3 fois), Biblioth. protestantisme - Paris, Bibl. Arsenal – Paris, Bibl. universitaire et UB-Bâle, BCU Fribourg".

26. Bibliogrpropre16

Consultez-vous des bases bibliographiques propres aux livres du 16^e siècle?

Bibliogrpropre16	Nb. cit.	Fréq.
non-réponse	2	
oui	5	50%
non	5	50%
TOTAL OBS.	12	100%

Les non-réponses ne sont pas prise en compte.

27. Nombibliogr.

Si "oui" citez les 3 principales bases bibliographiques que vous utilisez le plus souvent:

Quatre personnes seulement ont répondu:

VD 16, Adams, Schottenlauer, Cioranescu, VTLS, OPALE, ATLA, PCI (Périodique Context Index), MSL.

28. Basliv16

Consultez-vous des bases textuelles propres aux livres du 16^e siècle?

Basliv16	Nb. cit.	Fréq.
non-réponse	4	33%
oui	1	8%
non	7	59%
TOTAL OBS.	12	100%

29. Basetextprinc

Si "oui" citez les 3 principales bases textuelles que vous utilisez le plus souvent:

Basetextprinc	Nb. cit.	Fréq.
non-réponse	12	100%
TOTAL OBS.	12	100%

30. Sourcebibliogr.

Quelles sont les autres sources bibliographiques que vous utilisez pour repérer les livres du 16^e siècle qui vous intéressent?

Vous pouvez cocher plusieurs cases.

Sourcebibliogr.	Nb. cit.	Fréq.
non-réponse	2	
catalogues de ventes contemporains	5	16%
catalogues de ventes anciens	6	19%
bibliographies spécialisées	9	28%
inventaires privés	3	9%
archives	5	16%
réseaux professionnels	1	3%
autres sources	3	9%
TOTAL OBS.	34	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (7 au maximum).

Les non-réponses ne sont pas prises en compte.

31. Sourcebibliogautres

Si "autres sources", précisez:

(trois personnes ont répondu)

"Catalogues de bibliothèques en ligne", "Conseils des collègues", "Citations or références d'ouvrages".

32. daterech

Depuis quelle année, effectuez-vous des travaux avec/sur les livres du 16^e siècle?

daterech	Nb. cit.	Fréq.
Non- réponse	2	
de 1948 à 1960	3	30%
de 1961 à 1980	3	30%
de 1981 à 1990	1	10%
de 1991 à 2000	3	30%
TOTAL OBS.	12	100%

Les non-réponses ne sont pas prises en compte.

La question est à réponse ouverte numérique. Les observations sont regroupées en 4 classes d'égale amplitude.

Minimum = 1948, Maximum = 2000

Somme = 19757, Moyenne = 1975

33.Domcherch16e

Est-ce que vos travaux actuels avec/sur le livre du 16^e siècle constituent votre domaine de recherche?

Domcherch16e	Nb.cit.	Fréq.
non-réponse	2	
exclusif	3	30%
principal	5	50%
occasionnel	2	20%
autres		
TOTAL OBS.	12	100%

Les non-réponses ne sont pas prises en compte.

34.Acceliv16

Travaillez-vous avec/sur des fonds du 16^e siècle qui vous sont:

Acceliv16	Nb. cit.	F
non-réponse	2	
facilement accessibles	10	71%
difficilement accessibles	4	29%
TOTAL OBS.	16	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (2 au maximum).

Les non-réponses ne sont pas prises en compte.

35.Difficaces

Si vous éprouvez des difficultés pour accéder aux livres du 16^{ème} siècle, ces difficultés sont-elles dues à:

Vous pouvez cocher plusieurs cases.

Difficaces	Nb.cit.	Fréq.
non-réponse	5	24%
l'accès au texte original	4	19%
l'accès distant au texte numérique	2	10%
la localisation des documents	3	14%
l'obtention des reproductions	6	28%
autres	1	5%
TOTAL OBS.	21	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (5 au maximum).

36.Difficaces_Autres

Si "autres", précisez:

Difficaces_AUTRE	Nb. cit.	Fréq.
Non- réponse	12	100%
TOTAL OBS.	12	100%

37.Structextliv16

S'agissant de la structure du texte numérique, quelles sont les parties qui vous intéressent particulièrement dans le livre du 16e?

Vous pouvez cocher plusieurs cases.

structextliv16	Nb. cit.	Fréq.
Non- réponse	2	
titre	8	17,5%
texte	8	17,5%
pièces liminaires	6	14%
tables des matières	6	14%
tables des chapitres	6	14%
notes	1	2%
citations	2	5%
index	5	11%

autres	2	5%
TOTAL OBS.	46	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (9 au maximum).

38.Structextliv16_Autres

Si "autres", précisez:

Une seule réponse:

"Marques de propriété, annotations etc."

39.Illustliv16

S'agissant des illustrations, quels sont les éléments qui vous intéressent dans le livre du 16^e?

Illustliv16	Nb.cit.	Fréq.
non-réponse	3	
gravures	6	38%
bandeaux	1	6%
enluminures	1	6%
matériel typographique	4	25%
lettrines	3	19%
autres	1	6%
TOTAL OBS.	19	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (6 au maximum).

Les non-réponses ne sont pas prises en compte.

40.Illustliv16_Autres

Si "autres", précisez:

Une seule réponse: "les cartes".

41. Languecont

Quand vous travaillez sur un texte du 16^e siècle, aimeriez-vous avoir sa translation en langue contemporaine?

Languecont	Nb.cit.	Fréq.
non-réponse	2	
toujours		
parfois	4	36%
jamais	7	64%
TOTAL OBS.	13	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (3 au maximum).

Les non-réponses ne sont pas prises en compte.

42. Graphcont

Quand vous travaillez sur un texte du 16^e siècle, aimeriez-vous avoir sa transcription en graphie contemporaine?

Graphcon	Nb.cit.	Fréq.
non-réponse	2	
toujours		
parfois	4	40%
jamais	6	60%
TOTAL OBS.	12	100%

Les non-réponses ne sont pas prises en compte.

43. Traduction

Quand vous travaillez sur un texte du 16^e siècle, aimeriez-vous avoir sa traduction dans votre langue usuelle?

Traduction	Nb.cit.	Fréq.
non-réponse	2	
toujours		
parfois	6	60%
jamais	4	40%
TOTAL OBS.	12	100%

Les non-réponses ne sont pas prises en compte.

44.Pluslangue

Quand vous travaillez sur un texte du 16^e siècle, aimeriez-vous l'avoir en plusieurs langue?

Pluslangue	Nb.cit.	Fréq.
non-réponse	3	
toujours		
parfois	3	33%
jamais	6	67%
TOTAL OBS.	12	100%

Les non-réponses ne sont pas prises en compte.

45.Opératext16

Quels types d'opérations, effectuez-vous à partir de texte du 16^e siècle?

Vous pouvez cocher plusieurs cases

opératexte16	Nb. cit.	Fréq.
Non- réponse	2	
prise de notes manuscrites	7	28%
marquage de passages du textes	3	13%
reconstitution de textes	1	4%
reconstitution d'abréviations		
réagencement du texte	1	4%
extraction de mots	3	13%
signets		
constitution de dossiers	3	13%
constitution de corpus	5	21%
autres	1	4%
TOTAL OBS.	26	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (10 au maximum).

Les non-réponses ne sont pas prises en compte.

2.4. Livre en ligne et technologie de l'information et de la communication

46.Ordinateur

Utilisez-vous un ordinateur pour vos travaux de recherches sur le 16^e siècle?

Si "non" allez à la question 71

Ordinateur	Nb.cit.	Fréq.
non-réponse	2	
oui	7	70%
non	3	30%
TOTAL OBS.	12	100%

Les non-réponses ne sont pas prises en compte.

47.Taillecran

Quelle est la taille minimale d'écran que vous souhaitez avoir pour visualiser vos travaux ?

taillecran	Nb. cit.	Fréq.
Non- réponse	8	68%
15 pouces	1	8%
17 pouces	1	8%
19 pouces	1	8%
21 pouces	1	8%
autre		
TOTAL OBS.	12	100%

48.Bibnumweb

Consultez-vous des bibliothèques numériques sur le Web?

Bibnumweb	Nb.cit.	Fréq.
non-réponse	5	42%
toujours	2	17%
parfois	4	33%
jamais	1	8%
TOTAL OBS	12	100%

49.Rechwebliv16

Si vous avez déjà utilisé le Web, avez-vous fait des recherches sur un thème défini concernant les textes/les livres du 16^e siècle?

Rechwebliv16	Nb.cit.	Réf.
non-réponse	5	42%
plusieurs fois	3	25%
une ou deux fois	3	25%
jamais	1	8%
TOTAL OBS.	12	100%

50.Rechweb

Comment procédez-vous lorsque vous cherchez une information sur le WEB à propos des textes/des livres du 16^e siècle?

Vous pouvez cocher plusieurs cases

Rechweb	Nb. cit.	Fréq.
Non-réponse	5	29%
vous utilisez des mots clés	6	35.5%
vous utilisez une liste d'adresses préétablie	6	35.5%
vous partez des liens hypertextes existants dans les documents		
vous procédez autrement		
TOTAL OBS.	17	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (4 au maximum).

Les non-réponses ne sont pas prises en compte.

51.Rechweb_Autres

Si "vous procédez autrement", précisez:

Aucune réponse.

52.Difficrechweb

Avez-vous rencontré des difficultés dans la recherche sur le WEB à propos des textes/des livres du 16^e siècle?

Difficrechweb	Nb. cit.	Fréq.
Non- réponse	5	42%
oui	7	58%
non		
TOTAL OBS.	12	100%

53.Difficrechweb_Oui

Si 'oui', précisez :

Vous pouvez cocher plusieurs cases

Difficrechweb_Oui	Nb. cit.	Fréq.
Non-réponse	5	26%
informations mal organisées	4	21%

peu d'informations réellement exploitables	6	32%
trop d'informations récupérées	2	11%
absence d'aide pour trouver ce que je cherche	1	5%
difficultés pour trouver les mots clés pour formuler ma question	1	5%
autres		
TOTAL OBS.	19	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (6 au maximum).

54.Bookmark

Avez-vous déjà enregistré des adresses dans votre fichier de signets (bookmarks) concernant les textes/les livres du 16ème siècle ?

Bookmark	Nb. cit.	Fréq.
Non- réponse	5	41.5%
oui	5	41.5%
non	2	17%
TOTAL OBS.	12	100%

55.Motrech

Dans vos recherches bibliographiques et textuelles sur le 16^e siècle, avez-vous déjà utilisé un moteur de recherche ?

Motrech	Nb. cit.	Fréq.
Non- réponse	5	42%
oui	4	33%
non	3	25%
TOTAL OBS.	12	100%

56.Motrechsouv

Si " oui", précisez les 3 moteurs que vous utilisez le plus souvent:

Trois réponses: "Altavista", "Yahoo", "Copernik", "Seek", "Ask".

57.Travaillur

Préférez-vous travailler directement sur :

Travail sur	Nb. cit.	Fréq.
Non- réponse	5	
un écran d'ordinateur	2	29%
un document original papier	4	57%
une reproduction imprimée	1	14%
TOTAL OBS.	12	100%

Les non-réponses ne sont pas prises en compte.

58.Consultecran

Quand vous consultez un texte sur écran :

Vous pouvez cocher plusieurs cases

Consultecran	Nb. cit.	Fréq.
Non- réponse	5	
vous prenez des notes manuscrites	4	25%
vous prenez des notes sur ordinateur	5	31%
vous imprimez les fragments de textes que vous consultez en même temps	7	44%
autre		
TOTAL OBS.	21	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (4 au maximum).

Les non-réponses ne sont pas prises en compte.

59.Consultecran_Autres

Si "autres", précisez:

Aucune réponse.

60.Metadonecran

Eprouvez-vous besoin de consulter les renseignements bibliographiques (métadonnées) du texte/du livre numérique que vous visualisez sur écran?

Metadonecran	Nb.cit.	Fréq.
non-réponse	6	50%
oui	5	42%
non	1	8%
TOTAL OBS.	12	100%

61. Metadonestrav

Eprouvez-vous le besoin de consulter les renseignements bibliographiques (métadonnées) du texte/du livre numérique pendant votre travail?

Metadonestrav	Nb.cit.	Fréq.
non-réponse	7	59%
oui	4	33%
non	1	8%
TOTAL OBS.	12	100%

62. Metadonesref

Si "oui", aimeriez-vous que ces métadonnées se réfèrent à (aux):

Vous pouvez cocher plusieurs cases.

Metadonesref	Nb. cit.	Fréq.
non-réponse	7	43%
l'ouvrage	3	19%
l'exemplaire	1	6%
principales parties de l'ouvrage	3	19%
la page en cours de consultation	2	13%
TOTAL OBS.	16	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (4 au maximum).

63. Corpusnumtrav

Regroupez-vous des fragments de textes pour former un corpus numérique de travail?

Corpusnumtrav	Nb.cit.	Féq.
non-réponse	6	50%
toujours		
parfois	4	33%
jamais	2	17%
TOTAL OBS.	12	100%

64. Classtextnum

Si "parfois" ou "toujours", comment classez-vous ces textes numériques?

Vous pouvez cocher plusieurs cases.

Classtextnum	Nb.cit.	Fréq.
non-réponse	8	50%
par titre	1	6%

par thème	4	25%
par date	3	19%
par type de document (image, texte, son)		
autres		
TOTAL OBS.	16	100%

65.Fonctionnalités

Quelles sont les fonctionnalités dont vous auriez besoin et que vous voudriez retrouver sur un écran, pour travailler sur le texte numérique ?

Vous pouvez cocher plusieurs cases.

Fonctionnalités	Nb. cit.	Fréq.
non-réponse	7	35%
annotation	3	15%
aide en ligne		
navigation cartographique	2	10%
surlignage (couleur)	1	5%
surbrillance (contraste)	2	10%
statistiques sur le texte		
recherche d'occurrence	4	20%
historique	1	5%
autres		
TOTAL OBS.	20	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (9 au maximum).

66.Fonctionnalité_Autres

Si "autres", précisez:

Aucune réponse.

67.Modtextimage

Faites-vous la différence entre le mode image et le mode texte sur un écran d'ordinateur ?

modtextimage	Nb. cit.	Fréq.
Non- réponse	5	41.5%
oui	5	41.5%
non	2	17%
TOTAL OBS.	12	100%

68.Modtextimage_Oui

Si "oui", quel mode utilisez-vous ?

modtextimage_Oui	Nb. cit.	Fréq.
Non-réponse	6	50%
mode texte	5	42%
mode image	1	8%
TOTAL OBS.	12	100%

69.Modetext

Le mode texte vous est-il indispensable ?

modetext	Nb. cit.	Fréq.
Non- réponse	7	58%
oui	3	25%
non	2	17%
TOTAL OBS.	12	100%

70.Tailletexte

Sur votre écran, vous souhaitez travailler sur le texte en :

Vous pouvez cocher plusieurs cases.

tailletexte	Nb. cit.	Fréq.
Non- réponse	5	33%
taille réelle 1/1	6	40%
taille réduite		
disposant de loupe grossissante	4	27%
TOTAL OBS.	15	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (3 au maximum).

2.5. COMMUNICATIONS ET ECHANGES AOUTOUR DU TEXTE/LIVRE NUMERIQUE DU 16^e SIECLE

71.Travcollab

Dans le cadre de vos travaux de recherches actuels, travaillez-vous :

Vous pouvez cocher plusieurs cases.

Travcollab	Nb. cit.	Fréq.
Non- réponse	2	
seul	8	57%
en équipe	4	29%
au sein d'un réseau informel	2	14%

au sein d'un réseau formel		
TOTAL OBS.	16	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (4 au maximum).

Les non-réponses ne sont pas prises en compte.

72.Travcollab_Equipe

Si votre réponse est : "en relation avec un réseau informel" ou " en équipe", échangez-vous des informations par le biais de :

Vous pouvez cocher plusieurs cases.

travcollab_Equipe	Nb. cit.	Fréq.
Non- réponse	6	
texte écrit	3	19%
publications électroniques		
messages électroniques	3	19%
réunions de travail	3	19%
rencontres scientifiques	2	13%
discussions informelles	5	30%
autres		
TOTAL OBS.	22	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (7 au maximum).

Les non-réponses ne sont pas prises en compte.

73.Travcollab_Autre

Si "autres", précisez:

Aucune réponse.

74.Travcollabmaniere

Le plus souvent, quand vous collaborez avec quelqu'un, est-ce:

Vous pouvez cocher plusieurs cases.

Travcollabmaniere	Nb. cit.	Fréq.
non-réponse	3	

en face à face	6	55%
par réseau informatique en temps réel	1	9%
par réseau informatique en temps différé (48 heures maximum)	2	18%
par réseau informatique à plus long terme		
autres	2	18%
TOTAL OBS.	14	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (5 au maximum).

Les non-réponses ne sont pas prises en compte.

75.Partagenotes

Accepteriez-vous de partager vos notes avec d'autres usagers du livre numérique du 16^e siècle?

Partagenotes	Nb. cit.	Réf.
non-réponse	3	
oui	6	67%
non	3	33%
TOTAL OBS.	12	100%

Les non-réponses ne sont pas prises en compte.

76.Contenuéchange

Le contenu de ces échanges serait une :

contenuéchange	Nb. cit.	Fréq.
Non- réponse	7	58%
production commune		
information sur le texte	5	42%
information sur la nature de la collaboration		
information sur les compétences techniques		
TOTAL OBS.	12	100%

77.Besoincollab

Quels sont les besoins qui vous amènent à collaborer avec d'autres?

Vous pouvez cocher plusieurs cases.

Besoincollab	Nb. cit.	Réf.
non-réponse	3	
notes	3	13%

renvois à d'autres références	4	17%
indications d'autres lectures	4	17%
information sur les documents	6	27%
adresses Internet	5	22%
autres	1	4%
TOTAL OBS.	26	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (6 au maximum).

Les non-réponses ne sont pas prises en compte.

78. Besoincollab_Autre

Si "autres", précisez:

Une seule réponse: "Réflexion sur l'Histoire".

79. Attitudedécouv

Lors d'une découverte (ou une idée originale) seriez vous tenté plutôt de:

Vous pouvez cocher plusieurs cases.

Attitudedécouv	Nb. cit.	Réf.
non-réponse	4	
garder le secret jusqu'a publication	3	27%
la diffuser dans votre groupe de travail	1	9%
la soumettre à des personnes ressources pour validation	6	55%
la soumettre à votre responsable hiérarchique	1	9%
la diffuser dans un forum ou liste de diffusion		
autres		
TOTAL OBS.	15	100%

Le nombre de citations est supérieur au nombre d'observations du fait de réponses multiples (7 au maximum).

Les non-réponses ne sont pas prises en compte.

80. Typetravlivnum

A quels types de travaux voulez-vous accéder dans le cadre d'un travail de groupe (collaboratif) sur le livre numérique du 16^{ème} siècle?



Que deux réponses: "Edition critiques et recherche d'occurrence", "Pensée politique et religieuse".

81.Outiléchange

Aimeriez vous avoir des outils d'échanges avec d'autres chercheurs qui soient intégrés à votre écran de travail?

Outiléchange	Nb. cit.	Réf.
non-réponse	6	47%
oui	2	38%
non	5	15%
TOTAL OBS.	13	100%

82.Outiléchange_Oui

Si "oui", précisez lesquels?

Une seule réponse: "Contact direct avec les bibliothécaires".

83.Suggestcomment

Vos suggestions et commentaires sont les bienvenus:

Deux suggestions:

"Fournir ou perfectionner les outils concernant les articles des revues (non seulement les livres)",

"L'intérêt majeur de la numérisation est d'abord l'accès au texte lui-même, même s'il accessible par microforme mais surtout, il est la recherche d'occurrences – de mots clés; ça serait vu vrai propres, car l'image à l'écran n'est en fait qu'un autre support (papier, microfilm, microfiche)".