

Sommaire

INTRODUCTION.....	4
PARTIE 1 : MÉTHODOLOGIE	7
1. PRÉSENTATION DES AXES DE RECHERCHE.....	7
1.1. <i>Les bibliothèques numériques</i>	7
1.2. <i>Les études sur les usages réels et induits</i>	8
1.3. <i>Le travail en réseau et les fonctionnalités collaboratives</i>	8
2. MÉTHODOLOGIE DE LA RECHERCHE BIBLIOGRAPHIQUE.....	9
2.1. <i>Recherche et localisation</i>	9
2.2. <i>Critères de tri des articles et présentation</i>	11
3. MÉTHODOLOGIE DE L'ANALYSE DES SITES	12
3.1. <i>Choix du corpus</i>	12
3.2. <i>Grille d'analyse</i>	13
4. MÉTHODOLOGIE DES DESCRIPTIONS DÉTAILLÉES.....	15
PARTIE 2 : ETAT DE LA LITTÉRATURE. QUELLE PLACE EST ACCORDÉE À L'USAGER DANS LA RECHERCHE SUR LES BIBLIOTHÈQUES NUMÉRIQUES ?.....	16
1. INTRODUCTION	16
2. CONCEVOIR UNE INTERFACE À PARTIR DE BESOINS INITIAUX	18
2.1. <i>Les postes de travail et sites de bibliothèques numérisées</i>	19
2.1.1 Objectifs et fonctionnalités communs	19
2.1.2 Présentations de projets	20
2.2. <i>La voix des chercheurs</i>	27
2.2.1 L'IRHT	27
2.2.2 Le COSH.....	29
3. PRENDRE EN COMPTE LES RETOURS DES USAGERS POUR FAIRE ÉVOLUER UNE INTERFACE.....	33
3.1. <i>Les sites ayant évolué avec le feed-back</i>	33

3.2. Les études sur les usages : méthodes employées et équipes de recherche.....	36
3.2.1 Les méthodes appliquées	37
3.2.2 Les équipes de recherche	39
3.2.3 Synthèse sur les pratiques observées.....	43
4. CONCLUSION	47
PARTIE 3 : TABLEAUX STATISTIQUES ET ANALYSE DES SITES DE BIBLIOTHÈQUES NUMÉRIQUES	49
1. ANALYSE DES SITES VISITÉS	49
1.1. Les éditeurs et la date de création des sites	49
1.2. Typologie des sites.....	50
1.3. Les fonctionnalités proposées	56
1.3.1 Outils de visualisation.....	56
1.3.2 Recherche.....	57
1.3.3 Travail personnel sur le document	58
1.4. Travail collaboratif et documents associés	59
2. DESCRIPTION APPROFONDIE DE MENASSEH BEN ISRAEL, STUDIOLUM ET THE WILLIAM BLAKE ARCHIVES.....	61
2.1. Menasseh Ben Israel.....	61
2.2. Studiolum	63
2.3. The William Blake Archives.....	66
3. SYNTHÈSE DE L'ANALYSE DES SITES	68
3.1. Classement des sites visités.....	68
3.2. Conclusions sur l'analyse des sites	69
CONCLUSION.....	72
BIBLIOGRAPHIE	78
TABLE DES ANNEXES	83

Introduction

Dans la dynamique actuelle de création de bibliothèques numériques, le projet *DEBORA* tient une place toute particulière. *DEBORA*¹, pour Digital accEss to BOOks of the RenAissance, est un projet financé par l'Union européenne. Il a rassemblé dans un partenariat international, du 1^{er} janvier 1999 au 1^{er} mai 2002, sur une durée de trente mois, les universités lyonnaises, l'INSA de Lyon, les Universités de Lancaster (Grande-Bretagne) et Lisbonne (Portugal), des bibliothèques possédant des fonds anciens importants : la Bibliothèque municipale de Lyon, la Bibliothèque Casanatense de Rome, et la Bibliothèque générale de l'Université de Coïmbra (Portugal), et enfin les sociétés Xerox et SIGB-Lyon. *DEBORA* a pour mission de « faciliter l'accès des chercheurs » aux collections de livres de la Renaissance. Cette période a été choisie pour l'importance qu'elle revêt aux yeux des historiens du livre : « c'est celle où le livre imprimé acquiert, dans toute l'Europe, ses caractéristiques modernes avec la structuration de la page et la normalisation typographique, tout en conservant un mode de production artisanale »².

Dans le cadre de ce projet, de nouvelles normes de numérisation en mode image et une interface de consultation ont été élaborées. D'une part, les ouvrages ont fait l'objet d'une numérisation en mode image, et une transcription (semi-automatique) a été obtenue grâce à un logiciel de reconnaissance de caractères adapté aux polices de caractères du XVI^e siècle. D'autre part, la volonté de rester au plus près des utilisateurs finaux s'est, elle, traduite par une réflexion sur les besoins et usages des chercheurs, qui sont le public principalement visé par *DEBORA*, et ce de façon à concevoir des fonctionnalités adaptées. Une première enquête sur les pratiques documentaires de spécialistes du livre du XVI^e siècle a permis de savoir quelles fonctionnalités étaient attendues, et d'implémenter une interface. *DEBORA* voulait également se positionner en pionnier, et proposer des fonctionnalités

¹ Voir le site Internet de *DEBORA*, disponible sur Internet <http://debor.enssib.fr/description.htm> (consulté le 15.05.2002)

capables de suivre l'évolution des modes de travail, notamment dans le domaine de la collaboration et du travail sur écran. De nombreuses fonctionnalités ont donc été implémentées sur l'interface. La conception de fonctionnalités collaboratives est cependant freinée par l'absence même d'usages établis dans ce domaine chez les seiziémistes.

C'est dans ce contexte que s'inscrit cette étude. Le sujet de cette recherche a pour intitulé précis : « *Interface usagers finals de DEBORA : interactivité et travail collaboratif en recherche d'informations.* » ; son objectif est le suivant : « *Dans le cadre du projet de recherche européen qui s'achève, une interface usagers finals a été implémentée. Elle fait le pari que les chercheurs utiliseront des fonctionnalités d'enrichissement des textes numériques, notamment en matière de métadonnées. Elle veut promouvoir également le travail collaboratif entre chercheurs. Il s'agit d'une part d'expérimenter et d'observer ces pratiques éventuelles et d'autre part de réfléchir sur le rôle de médiation et de modération que doit assumer le bibliothécaire* ».

Le sujet de cette recherche est donc de réfléchir, dans la perspective de la diffusion du logiciel *DEBORA*, sur trois points : la bibliothèque numérique (quelles fonctionnalités, quel public, quel corpus ?), le travail sur écran (les usages qu'il induit, les besoins qu'il crée) et enfin le travail collaboratif en recherche d'information (existe-t-il ? est-il possible et comment ?)

Pour traiter ces trois problématiques, deux pistes de travail étaient possibles : une piste prospective et une piste analytique. La première piste consistait à évaluer quels sont les besoins des usagers d'une bibliothèque numérique telle que *DEBORA*, quels seront leurs usages et comment imaginer un travail collaboratif en fonctions de ces usages et de ces besoins. L'approche analytique, elle, consiste à étudier les bibliothèques numériques existantes, notamment leurs interfaces et les fonctionnalités qu'elles proposent, et de là à les considérer dans la perspective des usages en matière de travail sur écran et de travail collaboratif.

² Richard BOUCHE, présentation de *DEBORA*. Disponible sur Internet <http://debora.enssib.fr/description.htm> (consulté le 15.05.2002)

L'attitude prospective a été rapidement écartée. En effet, si l'on définit les usages comme des pratiques socialement établies dans lesquelles des individus d'un même groupe se reconnaissent, il ne saurait en être déjà question pour *DEBORA*, qui n'a encore fait l'objet d'aucune diffusion : le projet, s'il s'achève, n'a pas atteint sa phase finale de déploiement. De plus, si l'étude des usages est impossible dans l'état actuel des choses, une étude des besoins des chercheurs seizièmistes a déjà été faite³. Il est impossible d'aller plus avant dans cette direction : comment interroger un chercheur sur ses besoins alors même que *DEBORA* fait le pari que le travail sur écran changera ses pratiques et induira de nouveaux besoins ? Les résultats d'une telle enquête ne pourraient refléter qu'une réalité de travail sur papier ou sur imprimé et non un travail tel qu'il est induit par les fonctionnalités proposées par le logiciel *DEBORA*.

Par conséquent, c'est l'approche analytique qui a été choisie pour mener à bien cette recherche. Les objectifs de cette dernière ont été clairement définis : il s'agit d'établir une typologie des bibliothèques numériques sur des critères précis : statut de l'établissement porteur du projet, corpus numérisé, mode de numérisation, fonctionnalités proposées. Cette typologie doit permettre de réfléchir sur les usages induits par ces fonctionnalités en matière de travail sur écran et de travail collaboratif. Dans cette étude sur les usages des bibliothèques numériques, deux axes de réflexion ont été privilégiés : la problématique de l'usage induit par rapport à l'usage constaté et la problématique qui en découle, sur la politique de l'offre et de la demande.

Cette étude des bibliothèques numériques doit permettre, dans une seconde phase, de comprendre comment *DEBORA* s'inscrit dans cette typologie et d'aider à répondre aux deux questions : quelles sont les perspectives de *DEBORA* en matière de travail collaboratif et comment envisager une véritable étude des usages du travail sur écran ?

³ « Enquête : les pratiques actuelles des chercheurs » in *DEBORA*, WP2. *Rapport final d'évaluation de l'usage des outils et interface DEBORA*. 2001.

Partie 1 : Méthodologie

1. Présentation des axes de recherche

La réflexion a été menée dans trois directions : sur les bibliothèques numériques, sur les usages réels (étudiés) et induits (fonctionnalités des interfaces), et pour finir sur les fonctionnalités collaboratives et le travail en réseau. Pour traiter ces questions, deux axes de recherche ont été choisis : l'état de la littérature et l'analyse de sites de bibliothèques numériques. Mais tout d'abord, il faut donner quelques définitions nécessaires aux trois termes utilisés : bibliothèque numérique, usages et travail collaboratif.

1.1. Les bibliothèques numériques

La terminologie des bibliothèques numériques recouvre des réalités diverses. La multiplicité des termes reflète la confusion qui règne : bibliothèque numérique, bibliothèque électronique, bibliothèque numérisée ou encore bibliothèque virtuelle, les adjectifs s'accumulent sans que souvent des définitions précises ne soient posées. De même en anglais, où « digital library », « electronic library » et « virtual library » coexistent, bien que « digital library » semble s'imposer. Tout d'abord, le critère commun à ces diverses expressions est que toutes ces bibliothèques permettent un accès à l'information primaire, par opposition aux catalogues ou aux bibliographies en ligne.

Au delà de ce critère commun, il convient de distinguer des réalités différentes. Deux distinctions sont fondamentales pour établir une typologie des bibliothèques numériques : le statut de la collection et le mode de numérisation. Dans le premier cas, il s'agit de savoir si la bibliothèque numérique est l'image d'une collection existante dans un établissement précis, ou si elle est une construction virtuelle. Dans le second, la numérisation en mode texte se distingue de la numérisation en mode image. Un des enjeux principaux de *DEBORA* étant cette numérisation en

mode image et les usages qui peuvent en découler, c'est cette approche qui a été ici privilégiée.

Précisons tout d'abord que l'expression « bibliothèque numérique » sera employée de manière générique dans cette recherche. Le choix d'une numérisation en mode image implique l'existence de documents physiques : on parlera alors de « bibliothèque numérisée ». A l'inverse, on dira qu'une « bibliothèque électronique » propose des documents en mode texte, indépendamment d'un support physique : l'emploi de cet adjectif fait référence à l'édition électronique. Dans ces deux cas que sont les bibliothèques numérisées et les bibliothèques électroniques, le second niveau de définition concerne donc le statut de la collection : soit celle-ci existe comme entité matérielle et intellectuelle dans un établissement, et l'on parlera d'une « collection existante », soit les documents ont été rassemblés de manière artificielle, et l'on parlera de « collection virtuelle ». N'ont pas été reconnus comme bibliothèques numériques des portails donnant accès à d'autres sites, des expositions virtuelles de livres numérisés, ou encore des sites, institutionnels ou personnels, présentant quelques pages numérisées sans aucun travail d'indexation, de communication et de réflexion d'ensemble.

1.2. Les études sur les usages réels et induits

On entend par usage réel un usage constaté « in vivo » par enquête ou par observation. Les articles rapportant de telles enquêtes et leurs résultats sont traités dans la synthèse bibliographique.

Par opposition, un usage « induit » est une hypothèse concernant un usage, fondée sur l'étude des fonctionnalités proposées par une bibliothèque numérique. C'est la méthode appliquée ici à l'analyse des sites.

1.3. Le travail en réseau et les fonctionnalités collaboratives

Le travail en réseau s'est développé grâce aux technologies informatiques, notamment dans les entreprises (knowledge management). Dans le domaine des bibliothèques, il s'applique en particulier au catalogage partagé et dans les catalogues collectifs, ce qui permet de gagner en efficacité et en temps de travail.

Selon l'équipe du projet *DEBORA*, ce principe pourrait s'appliquer également aux bibliothèques numériques, et c'est dans cette perspective que les fonctionnalités collaboratives du logiciel client ont été créées.

Cependant, le travail collaboratif est différent du travail en réseau. Il ne s'agit pas seulement de rationaliser le travail, ni même de récupérer des informations communiquées par autrui (ce que permettent les messageries, ou encore les newsgroups), mais de recréer les conditions de travail permises par la proximité réelle des personnes. Le travail collaboratif doit permettre la confrontation des idées, les dialogues autour d'un document et l'enrichissement collectif d'une recherche. Ce qui est rendu possible, dans le monde de l'entreprise, par des techniques telles que la visioconférence, l'archivage partagé, ne l'est évidemment pas dans les bibliothèques numériques. Comment recréer un environnement permettant, par exemple, d'enrichir une base de données, de corriger un texte, d'ajouter une note ou un commentaire, voilà ce sur quoi a réfléchi l'équipe du projet *DEBORA*.

Qu'en est-il des autres sites de bibliothèques numériques ? Où en est la réflexion scientifique et technique à ce sujet ? Les deux axes de cette recherche, synthèse bibliographique et analyse des sites, veulent répondre à ces questions.

2. Méthodologie de la recherche bibliographique

Cette recherche bibliographique a deux objectifs : il s'agit dans un premier temps de comprendre quelle place occupe le thème des fonctionnalités et usages dans la recherche sur les bibliothèques numériques, et d'autre part de synthétiser les résultats obtenus par les différentes équipes de recherche qui se sont penchées sur cette problématique.

2.1. Recherche et localisation

Pour répondre à ces questions, plusieurs bases de données de dépouillement d'articles de périodiques ont été interrogées :

- LISA : Library and Information Science Abstracts de Bowker-Saur. Cette base disponible en ligne sur abonnement donne accès aux références et résumés des

articles de 440 titres de périodiques en bibliothéconomie et sciences de l'information et à des ressources électroniques, pour plus de 68 pays, et en 20 langues. Elle couvre la période de 1969 à nos jours.

- FRANCIS : base de données produite par l'INIST, Francis couvre le champ des sciences humaines et sociales.
- Les bases de données en sciences de l'information du serveur Dialog.
- La base de périodiques en ligne Elsevier, pour les articles relevant de l'informatique et des sciences sociales.

Ces bases dépouillent évidemment un certain nombre de titres en commun, et des doublons ont pu être relevés dans les résultats obtenus, mais il s'agissait de réduire autant que faire ce peut le risque de silence.

L'absence d'unité dans le vocabulaire employé pour parler des bibliothèques numériques, décrit dans l'introduction, est à l'origine de la multiplicité des mots-clés utilisés pour lancer les recherches. On a privilégié l'utilisation des index propres à chaque base de données, pour réduire là encore autant que possible le silence, et l'emploi des troncatures a permis de faire porter la recherche aux termes au singulier comme au pluriel. La combinaison des différents termes permettent dans un second temps de réduire le nombre de références retrouvées. Les termes suivants ont donc été retenus :

Digital / virtual / electronic library

Bibliothèque numérique / électronique / virtuelle

Reproduction / fac-similé / fonds anciens / special collections / old and rare material

Numérisation / digitization

Usage / usager / enquête / user survey / use / evaluation

Des critères de dates ont également été introduits lors des interrogations.

Les bibliographies des articles pertinents relevés après ces premières interrogations ont permis de lancer ensuite des recherches par auteur. L'Internet s'est alors révélé précieux, pour cerner précisément les groupes de recherche, et situer les axes de travail.

La localisation des références trouvées s'est faite grâce aux catalogues de bibliothèques, SUDOC notamment, ou directement par un moteur de recherche sur

l'Internet : nombre de chercheurs proposent un site personnel où leurs publications sont disponibles. Enfin, certaines références trouvées n'ont pu être localisées ni donc consultées, notamment pour des comptes-rendus de colloques dont les programmes avaient été trouvés sur l'Internet.

2.2. Critères de tri des articles et présentation

Dans cette seconde étape de la recherche bibliographique, la sélection des références puis articles pertinents s'est faite selon des critères chronologiques et thématiques.

Chronologiques, car il s'agissait de se placer dans la continuité des recherches déjà menées par l'équipe de recherche de *DEBORA*, et de proposer un état de la littérature actuelle : seuls les articles postérieurs à 1998 ont été retenus.

Thématiques, car les références retrouvées portaient pour des raisons déjà évoquées sur des bibliothèques numériques très variées ; il convenait donc de restreindre la recherche aux sites comparables à *DEBORA* ou se situant dans une problématique proche. Il va toutefois sans dire que les bibliothèques numériques proposant des documents anciens en mode image ont été privilégiées, et que les recommandations des chercheurs n'ont été retenues que s'il s'agissait de personnes travaillant sur des manuscrits ou ouvrages anciens.

Enfin, le choix a été fait de ne pas se limiter aux seules études sur les pratiques : les articles décrivant les bibliothèques numériques en termes de fonctionnalités et besoins des publics potentiels ont également été retenus, de même que ceux émanant de chercheurs habitués à travailler avec des documents anciens, susceptibles de donner des indications intéressantes quant à leurs attentes.

Les articles finalement retenus traitent soit de projets ciblés, soit de réflexions et enquêtes menées dans le cadre de la mise en place d'une bibliothèque numérique : dans ce dernier cas, la plupart des articles émanent de groupes de recherche anglo-saxons.

La présentation des résultats de cet état de la littérature suit un ordre thématique, mais conserve l'unité des différents groupes de recherche rencontrés : il a semblé autant intéressant de situer entre elles les préoccupations et pratiques des

chercheurs que de maintenir la cohérence des réflexions menées dans les différents groupes de recherche.

La présentation des références suit les règles de description bibliographique internationale, éditées en France par l'AFNOR⁴.

3. Méthodologie de l'analyse des sites

Cette analyse de sites de bibliothèques numériques vise à comparer leurs fonctionnalités de recherche d'information et de visualisation des documents avec celles de *DEBORA*.

3.1. Choix du corpus

Le corpus des sites visités a été établi à partir d'une liste de sites de bibliothèques numériques extraite d'un répertoire de ressources⁵, de liens disponibles sur certains sites de bibliothèques numériques (*Studiolum*) et dans des articles de revues. Seuls les sites correspondant à la définition proposée en 1.1. ont été retenus, et surtout ceux proposant des livres numérisés en mode image, ou mieux, qui allient l'image et une transcription.

Dans la perspective d'une comparaison avec *DEBORA*, qui propose à la fois l'image du livre et une transcription, les bibliothèques numériques en mode texte seul sont peu pertinentes, mais quelques-unes ont été retenues néanmoins pour les fonctionnalités de recherche qu'elles proposent.

Ont notamment été écartés les projets de bibliothèques numériques, les expositions numériques, les répertoires de ressources, les portails⁶. Les sites de bibliothèques numériques ne comportant qu'un seul ouvrage, sans aucune fonctionnalité, n'ont pas été retenus.

⁴ AFNOR. *Références bibliographiques : contenu, forme et structure*. ISO 690 1987. Paris : AFNOR, 1987.
AFNOR. *Références bibliographiques : partie 2, documents électroniques, documents complets ou parties de documents*. ISO 690-2 1997. Paris : AFNOR, 1998.

⁵ Documentation du projet *DEBORA*.

⁶ Voir en annexe 2 la liste des sites visités et non retenus.

3.2. Grille d'analyse

Elle a été élaborée à partir des fonctionnalités proposées sur le logiciel de visualisation développé par l'INSA de Lyon pour *DEBORA*⁷.

Le premier bloc permet d'identifier le site :

- o Nom du site
- o Editeur (institutionnel et/ou la personne responsable)
- o Adresse URL
- o Date de création lorsqu'elle apparaît

Le deuxième bloc décrit la base de données :

- o Type de documents numérisés : livres imprimés, manuscrits ou périodiques, époque, langue, lieu de conservation
- o Nombre de livres ou de pages/feuilletts. Le nombre de données annoncé ne correspond pas forcément à la réalité de l'avancement du projet de numérisation
- o Mode de numérisation : texte, image, texte et image
- o Couleurs : numérisation en couleurs ou en noir et blanc
- o Mise à jour lorsqu'elle apparaît
- o Base de données enrichie : présence de documents annexes au texte principal comme des bibliographies, des index, des cartes, des plans, des commentaires historiques ou biographiques
- o Niveau de description des métadonnées⁸ : niveau 1 (élaborée et précise, en particulier d'après les recommandations de la Text Encoding Initiative), niveau 2 (format MARC), niveau 3 (description « maison », élaborée mais sans référence aux normes et aux métadonnées), niveau 4 (inexistante ou très limitée)
- o Format des fichiers images et logiciels de lecture éventuellement nécessaires

⁷ Version 10, disponible sur le site Internet de l'INSA : <http://rfv6.insa-lyon.fr/debora/client.htm> (consulté le 12.02.2002). Cette définition des métadonnées concerne les livres et peut donc s'appliquer à l'ensemble des bibliothèques numériques, contrairement à la définition technique des 4 niveaux de métadonnées proposée dans le Rapport final sur *DEBORA*, qui concernait la structure même de l'interface *DEBORA*, et qui était par conséquent trop spécifique.

Le troisième bloc décrit les outils de visualisation :

- o Vignettage : les pages numérisées apparaissent en réduction les unes à côté des autres. (Cette présentation du livre s'oppose à celle du sommaire et des numéros de page comme autant de liens permettant d'aller directement à la page souhaitée, ou à la possibilité d'inscrire le numéro de la page souhaitée)
- o Mosaïque : les pages numérisées apparaissent en réduction les unes empilées sur les autres
- o Affichage des annotations : affichage à côté du document numérisé d'annotations concernant le document lui-même ou l'auteur
- o Fonds : supprimer les effets de transparence, améliorer la lisibilité
- o Luminosité : éclaircir ou assombrir
- o Grossissement : loupe

Le quatrième bloc décrit les outils de recherche proposés :

- o Recherche par titre
- o Recherche en plein texte : signalement des occurrences d'un mot ou d'une expression ; ou simple moyen de retrouver un document qui n'a pas encore été indexé
- o Illustrations : extraction des illustrations par rapport au texte
- o Autres : auteur, sujet, éditeur

Le cinquième bloc décrit les fonctionnalités de travail personnel sur le document :

- o Corpus personnel : téléchargement en enregistrement de tout ou partie d'un document numérisé (par exemple *Gallica*)
- o Fichier personnel : possibilité d'introduire des annotations, soit dans le corps du document lui-même, soit sur un fichier (par exemple le « *shadow file* » de *Studiolum*)
- o Impression

Le sixième bloc décrit les fonctionnalités de travail collaboratif :

- o Messagerie électronique
- o *Newsgroup* (forum, *Frequently Asked Questions*) ou listes de discussion

⁸ d'après R. Bouché, *RIST*, p. 56-57.

- o Banque de données permettant de gérer le travail collaboratif
- o Autres : rubriques permettant de suggérer des corrections, des commentaires

Le septième bloc décrit les documents associés :

- o Enquêtes : enquête en cours sur le site ou résultats d'enquête
- o Liens : liste ou répertoire de liens vers d'autres bibliothèques numériques
- o Webmaster, mél

Ce tableau a été conçu en fonction de *DEBORA*. Il doit permettre de comparer les différents sites, mais également de les situer par rapport à *DEBORA*. Cette méthode n'est pas sans comporter un inconvénient : chaque site a été développé selon une logique propre, et il est parfois difficile de faire correspondre une fonctionnalité observée à une case du tableau.

4. Méthodologie des descriptions détaillées

Pour compléter l'analyse quantitative des sites de bibliothèques numériques, il a semblé intéressant de choisir trois d'entre eux, pour en établir une description détaillée et les comparer à *DEBORA*.

Les critères de choix ont été établis par rapport à *DEBORA* :

- Corpus de textes anciens numérisés
- Interface de consultation élaborée, ayant manifestement fait l'objet d'une réflexion lors de sa conception
- Possibilité de travail sur les images ou textes numérisés
- Diffusion de l'interface et des documents numérisés

Pour la description proprement dite, le tableau utilisé pour les analyses de sites n'a pas été retenu : dans le cadre d'une description détaillée, les inconvénients du tableau figé sont encore exacerbés, et se baser sur les fonctionnalités de *DEBORA* aurait pu fausser la compréhension du lecteur.

Partie 2 : Etat de la littérature. Quelle place est accordée à l'utilisateur dans la recherche sur les bibliothèques numériques ?

1. Introduction

L'équipe de *DEBORA* a déjà produit en juin 1999 un rapport intitulé *Etat de l'art des recherches sur les pratiques et sur les usagers des bibliothèques virtuelles*⁹, qui fait le point sur les enquêtes publiées à la date de rédaction, et souligne la pauvreté de la recherche en ce domaine. Les attentes des utilisateurs, qui font l'objet d'une synthèse, sont peu développées. Elles concernent les fonctionnalités des logiciels de lecture, l'accès à l'information et les nécessaires adaptations personnelles. Le présent état de la littérature se situe donc à la suite de ce rapport, et constitue une mise à jour des connaissances, selon les critères énoncés dans la méthodologie.

La méthode d'approche est toutefois un peu différente, dans la mesure où le choix a été fait ici de ne pas se limiter aux seules études sur les pratiques, mais de considérer également l'expression des besoins et attentes des usagers .

Deux types de bibliothèques numériques se dégagent : des bibliothèques numérisées, qui émanent d'établissements reconnus et desservant un large public et qui ont pour but autant d'assurer une meilleure diffusion à leurs documents que d'en faciliter la consultation ; et des postes de travail destinés aux spécialistes des textes anciens, que ce soit en histoire ou en littérature. Ces postes destinés aux

⁹ HEMBISE, C., et al. *Etat de l'art des recherches sur les pratiques et sur les usagers des bibliothèques virtuelles*, 1999. Disponible sur Internet <<http://debora.enssib.fr/de121.doc>> (consulté le 12.03.2002)

chercheurs sont pour certains d'entre eux élaborés en étroite collaboration avec leurs utilisateurs finaux, dans des laboratoires de recherche. Ils se distinguent des bibliothèques numérisées par les nombreux outils logiciels qu'ils proposent, contre souvent un faible nombre de titres disponibles. A l'inverse, dans le cadre d'une politique de diffusion, les sites de bibliothèques numériques proposent à la consultation, par ordinateur interposé, de nombreux documents, mais avec des interfaces de consultation simplifiées.

De fait, les besoins des chercheurs sont différents de ceux des bibliothèques, et selon que l'un ou l'autre se trouve à l'origine d'un projet, les résultats et méthodes de mises en oeuvre sont assez différents, notamment pour ce qui est de la prise en compte de l'utilisateur : les bibliothèques numériques élaborées par et pour les chercheurs intègrent la phase d'analyse des besoins en amont, les bibliothèques commencent par réfléchir aux questions techniques de numérisation et de constitution d'une collection virtuelle, ne se tournant qu'après vers les pratiques de lecture.

Pour ces deux types de sites, on trouve donc un public spécifique, et la politique de numérisation tout comme l'élaboration de l'interface de consultation s'en trouve affectée : à un public défini correspondent des besoins propres, et la prise en compte de ces derniers est le résultat d'une approche tout à fait différente.

L'investissement du public n'est pas lui non plus le même, selon qu'il s'agit d'un public de professionnels ou non : les spécialistes qui s'intéressent aux reproductions électroniques de documents se sont investis dans la conduite de certains projets, ou formulent dans des revues spécialisées leurs besoins et attentes, en amont du travail de numérisation. A l'inverse, pour des bibliothèques destinées à un public plus large, qui peuvent toutefois inclure également des chercheurs, les responsables de la bibliothèque numérique vont à la rencontre des pratiques de leurs usagers dans un second temps, après avoir conçu une première interface et l'avoir testée en situation réelle. Les usagers, s'ils acceptent de participer à des enquêtes, semblent ne donner que peu de retours volontaires.

Les enquêtes sur les pratiques des usagers des bibliothèques numériques sont récemment devenues présentes dans le paysage de la recherche sur les

bibliothèques numériques, et une évolution qui va vers l'intégration progressive de la problématique des pratiques dans la conception de bibliothèque numérique se dessine.

Ces deux démarches de prise en compte des besoins et pratiques des usagers sont présentées ici successivement.

2. Concevoir une interface à partir de besoins initiaux

Avec la multiplication des projets de numérisation de documents anciens, manuscrits ou imprimés, une réflexion s'est mise en place sur les fonctions des interfaces de consultation à développer. Le point de vue de l'utilisateur est souvent présent dans ces travaux, sans qu'aucune étude sur les pratiques réelles ne soit citée. C'est en terme de besoins et attentes que les usagers sont traités, dans la mesure où le projet est souvent mené en collaboration avec des chercheurs spécialisés en histoire ou sciences humaines, et de ce fait habitués à la fréquentation des documents anciens : on peut donc voir dans les descriptions des postes l'expression indirecte des besoins des chercheurs.

Cette collaboration ne saurait être l'équivalent d'une enquête sur les pratiques des chercheurs en matière de documentation électronique, car les besoins exprimés par les chercheurs relèvent pour l'essentiel de leurs usages des documents traditionnels, transposés dans le domaine du numérique. Cette méthode, dont la valeur ne saurait être contestée dans le cadre de projets pionniers, n'est pas à l'abri d'une évolution des usages des chercheurs liée à l'évolution des supports.

Les fonctions développées dans les différents logiciels et sites proposant des documents anciens sont donc issues de l'expression des besoins des chercheurs, sans refléter leurs pratiques réelles qui peuvent se développer a posteriori, lors de la mise à disposition du produit fini. Il y a ainsi dans cette démarche un aspect très théorique, car la représentation que les chercheurs ont de leurs besoins en matière de documents numérisés peut être différente de l'usage qu'ils en feraient réellement, dans le cadre de leur travail quotidien.

Enfin, l'expression de leurs besoins et attentes par les chercheurs ne se fait pas toujours dans le cadre de l'élaboration d'une interface : certains s'expriment sur le sujet, par le biais d'articles dans des revues spécialisées.

2.1. Les postes de travail et sites de bibliothèques numérisées

Les divers projets qui ont vu le jour dans le milieu et la fin des années 90 présentent un certain nombre de points communs, notamment dans leurs objectifs et les fonctions logicielles proposées, mais ont également chacun leurs particularités.

2.1.1 Objectifs et fonctionnalités communs

Les objectifs de ces projets relèvent de deux ordres : d'une part faciliter l'accès aux documents anciens pour les chercheurs, et donc essayer de résoudre le grand dilemme souvent posé au bibliothécaire, entre conservation et communication du document. D'autre part, permettre le travail des philologues, en leur donnant les outils informatiques dont ils peuvent avoir besoin : le travail de recherche devient moins répétitif et plus scientifique. Ce dernier objectif est plus particulièrement à l'origine des fonctionnalités implémentées dans les logiciels de consultation.

Bon nombre de ces travaux se sont en effet attachés à associer une image numérisée du document original, et une transcription codée en ASCII, ceci afin de ne pas perdre d'informations sur la mise en page tout en rendant possible les manipulations sur le texte, notamment la recherche¹⁰. La technique de l'hypertexte permet de lier étroitement une partie d'une image à un texte, et donc de passer rapidement de l'un à l'autre, ou encore de les afficher simultanément sur l'écran. La transcription couplée par ce biais à l'image de l'original ne peut que difficilement être automatisée pour les documents manuscrits : un système de transcription semi-automatique est conçu, qui soulage le travail du chercheur sans accroître les risques d'erreurs.

¹⁰ Cette double numérisation concerne le site consacré à *Yvain ou le Chevalier au Lion*, le projet *Charrette*, le projet *Cervantes*, *BAMBI* et les postes de travail « Philological Workstation », le projet *Philectre*, et les *William Blake Archives*.

Un système permettant d'annoter une partie du document est également commun à de nombreux projets¹¹ : annoter le texte, la transcription, une partie de l'image, ou encore créer des liens vers des sources externes sont autant de possibilités données aux utilisateurs. Ces annotations, affichées simultanément avec le texte auquel elles se réfèrent, peuvent constituer tout de même un corpus séparé du texte du document travaillé, et être gérées séparément. Des bases de données d'analyses linguistiques ou historiques sont ainsi disponibles sur certains sites, en complément des deux versions (texte et image) des textes anciens.

Comme l'hypertexte rendait possible le couplage de l'image et du texte, l'utilisation de documents structurés a ses conséquences : l'introduction de métadonnées et l'application de DTD améliorent les possibilités de recherche, et rendent plus aisée l'identification d'un titre ou d'un exemplaire.

A ces caractéristiques communes s'ajoutent les particularités de chaque projet.

2.1.2 Présentations de projets

Le périodique *Document numérique*¹², publié par les éditions Hermès, a consacré un numéro spécial aux documents anciens et à leur numérisation en 1999, présentant plusieurs projets de numérisation de documents anciens et d'élaboration de stations de travail. Chaque projet est unique, notamment parce qu'il s'attache à la numérisation d'un corpus défini : chaque type de document présente des caractéristiques propres, son public particulier, et appelle des traitements spécifiques.

Le rôle des chercheurs spécialistes en histoire des textes ou en littérature est lui plus ou moins important selon le laboratoire qui est à l'origine du projet. Dans les quatre cas présentés ici en premier, il s'agit de laboratoire de sciences humaines : l'objectif fondamental est l'élaboration d'un outil de travail répondant aux besoins, et non la recherche en informatique.

¹¹ Les annotations concernent *Don Quixote*, *BAMBI* et « Philological workstation », *Philectre*, *Xlibris* et *William Blake Archives*, avec le logiciel Inote.

¹² *Document numérique*, vol. 3, n° spécial 1-2, 1999.

Les projets décrits ont été regroupés selon les fonctions qu'ils avaient le plus développé : transcription, annotations, métadonnées.

L'élaboration d'un site Internet¹³ consacré au roman *Yvain ou le Chevalier au Lion* de Chrétien de Troyes a été l'occasion d'une réflexion sur les besoins des chercheurs en histoire ou sciences humaines en matière de documentation électronique : les professeurs d'ancien français de deux universités canadienne et américaine sont à l'origine de ce projet à visée philologique, mené depuis 1991. Les éditions imprimées et critiques des oeuvres du Moyen-Age ne favorisent pas d'après les concepteurs la connaissance de la langue originale et l'étude des manuscrits. Proposer gratuitement en ligne une image numérique de tous les manuscrits disponibles, assortie de deux transcriptions de niveaux critiques différents permettra un retour à ces domaines. A ces documents sont jointes les bases de données des analyses lexicale et morpho-syntaxique menées par les chercheurs. Soulignons que le choix d'une présentation sur l'Internet permet à ces chercheurs de rejoindre le projet Charrette, et de travailler par pallier, en tenant compte « du feed-back des lecteurs »¹⁴. Si les chercheurs ont eu une part active à l'implémentation du logiciel, aucune enquête d'usage postérieure n'est organisée.

Le projet Charrette qui vient d'être cité est né de l'association du groupe de recherche du Centre d'études supérieures de civilisation médiévale, de l'Université de Poitiers, au projet qui vient d'être décrit¹⁵. En mars 1998, un site miroir de celui d'*Yvain, Lancelot* est installé sur les serveurs de Poitiers. Il devient indépendant en mars 2000. Les sites proposent 7 des 8 manuscrits disponibles en mode image, une transcription en mode texte, deux index, une traduction en français moderne, un

¹³ KUNSTMANN, Pierre, MARTINEAU, France. Chrétien de Troyes sur le Web: présentation, analyse et interprétation des manuscrits d'*Yvain ou le Chevalier au Lion*. Editors of *CHWP* 2000. Disponible sur Internet <<http://texttechnology.mcmaster.ca/kunstmannmartineau.html>> (consulté le 15.04.2002)
Le site est disponible sur Internet <<http://www.uottawa.ca/academic/arts/lfa/activites/textes/chevalier-au-lion/P/Ppresentation.html>> (consulté le 18.04.2002)

¹⁴ PELLEN, René. Le « Projet Charrette » à Poitiers. *Le Médiéviste et l'ordinateur*, n° 40, automne 2001, *La numérisation des manuscrits médiévaux : actes de la journée d'étude, Paris, 13 octobre 2000*. Disponible sur Internet <<http://irht.cnrs-orleans.fr/meto/mo40-07.htm>> (consulté le 20.05.2002)

¹⁵ PELLEN, René. Le « Projet Charrette » à Poitiers. *Le Médiéviste et l'ordinateur*, n° 40, automne 2001, *La numérisation des manuscrits médiévaux : actes de la journée d'étude, Paris, 13 octobre 2000*. Disponible sur Internet <<http://irht.cnrs-orleans.fr/meto/mo40-07.htm>> (consulté le 20.05.2002)

essai sur Chrétien de Troyes. Les objectifs de l'Université de Poitiers sont les suivants : « pouvoir *analyser, fragmenter, manipuler* l'image comme du texte (ou presque), c'est-à-dire en extraire des portions, localiser les zones sombres par rapport aux zones claires, identifier des figures, des tracés, des « ductus », des caractères ; aller vers une *comparaison partiellement automatique* entre transcription(s) et image, vers une transcription également semi automatique (validée par le chercheur) ; *coupler image et texte* dans les bases de données »¹⁶.

Dans le même esprit que les sites consacrés à *Yvain* ou à *Lancelot*, les chercheurs de l'Université du Texas ont pour leur part mis en ligne une édition électronique de *Don Quichotte* de Cervantès¹⁷ : ce texte, transmis à la postérité par plusieurs éditions incunables, présente de nombreuses variantes. L'édition électronique doit permettre de présenter au lecteur ces différents états du texte, assortis de possibilités de personnalisation de l'affichage, de constitution d'un état personnel du texte, d'annotations et de liens vers d'autres ressources en ligne. Le projet souligne l'importance de la création d'une interface usager : il s'agit de rendre plus largement accessibles les documents traités, et de permettre au lecteur de créer ses propres documents par la fusion de parties des documents proposés.

La station de travail *BAMBI*¹⁸, projet développé par l'Institute for computational linguistics de Pise depuis 1995, est conçu pour rendre possible à partir de l'image du document original une transcription et une indexation semi-automatique, avec correction par le chercheur : cette fonction le distingue des trois projets précédents. L'article signale que « la station de travail *BAMBI* a été évaluée par plusieurs spécialistes du domaine (...) et a donné globalement satisfaction à ses utilisateurs »¹⁹, mais cette expérimentation de laboratoire, si elle a permis de

¹⁶ Ibid.

¹⁷ HU, Shuen-Cheng, FURUTA, Richard, URBINA, Eduardo. An electronic version of *Don Quixote* for humanities scholars. *Document numérique*, 1999, vol. 3, n° spécial 1-2, p. 75-91.
Site disponible sur Internet <http://www.csd.tamu.edu/cervantes/english> (consulté le 26.05.2002)

¹⁸ CALABRETTO, Sylvie, PINON, Jena-Marie, BOZZI, Andrea. *BAMBI* : système de gestion de manuscrits anciens pour historiens. *Document numérique*, 1999, 2, 3-4, p. 31-50.

¹⁹ Ibid., p. 47.

conclure que les fonctionnalités implémentées correspondaient aux besoins exprimés par les chercheurs, ne saurait être considérée comme l'équivalent d'une enquête d'usage.

Toujours à l'Institut de linguistique assistée par ordinateur de Pise, le projet « Philological Workstation for digital libraries »²⁰, développé en 1997-1998, devait aboutir à un logiciel permettant la constitution de « stemma » (arbre généalogique des manuscrits et éditions). Lié au projet BAMBI précédemment décrit, il a également fait l'objet d'articles séparés. Les concepteurs voulaient par là convertir les chercheurs en sciences humaines à l'informatique, pour renverser la tendance observée : « on note la permanence d'un certain manque d'intérêt pour des logiciels innovateurs de la part des personnes auxquelles ces logiciels sont adressés » (p. 94). Précisons également que « le logiciel décrit [...] a été testé par des philologues qui travaillent sur des textes latins, italiens et occitans »²¹, mais ce test en laboratoire est la seule étude menée sur ce logiciel quant à l'usabilité du système.

Autre poste d'édition et de lecture hypermédia développé dans le cadre d'un projet : Philectre²², qui vise à implémenter des « postes de travail hypertextuels »²³ pour les philologues. Développé notamment par le GIS science de la cognition du CNRS, en collaboration avec des chercheurs en histoire des textes de l'IRHT et des chercheurs en informatique, de 1994 à 1997, il utilise le langage structuré XML qui rend possible la création de liens entre les différentes parties du document, voire avec des informations externes (liens hypertextes). Il se situe également dans le courant de la TEI. Ce poste de travail a été conçu en étroite collaboration avec des chercheurs « spécialistes des oeuvres littéraires anciennes ou modernes »²⁴, et les concepteurs ont « tenté de mettre en évidence certaines tâches et besoins qui

²⁰ BOZZI, Andrea. Pour un système de philologie numérique. *Document numérique*, 1999, 3, 1-2, p. 93-101.

²¹ Ibid., p. 95.

²² VENTABERT, Gusnard de. Représentation et exploitation électroniques de documents anciens. *Document numérique*, 1999, 3, 1-2, p. 57 à 73.

²³ Ibid., p. 58.

²⁴ LECOLINET, Eric, ROBERT, Laurent. Conception d'un poste d'édition et de lecture d'hypermédiat littéraires. *Document numérique*, 1999, 3, 1-2, p. 114.

leur sont spécifiques afin de proposer des solutions adaptées »²⁵, mais aucune évaluation poussée sur l'usabilité du système n'a été menée. Le projet Philectre, abandonné pour des raisons financières notamment, prévoyait également de proposer au chercheur une vue contextuelle du manuscrit, dans un environnement tridimensionnel : la vue contextuelle inclue, outre l'image du manuscrit et une vue de la transcription, une vue de l'emplacement hypermédia du document consulté. Ce logiciel²⁶, qui a fait l'objet d'articles indépendamment du projet Philectre, propose une vue du bureau en trois dimensions : il est développé par Eric Lecolinet et Laurent Robert, du laboratoire du CNRS cité. Nommé ZoomTree, il a pour but de réduire la désorientation du lecteur en représentant l'ensemble des informations disponibles. Les deux auteurs affirment que leur travail sur les interfaces a le mérite de lier trois axes de recherche : hypertexte, document structuré et visualisation. Ce travail sur l'interface fait la particularité du projet Philectre.

A propos des annotations, aux Etats-Unis, Cathy C. Marshall, qui travaille pour l'entreprise Xerox, mène des études originales sur l'annotation en milieu numérique.

A partir d'une enquête portant sur les annotations d'étudiants américains sur leurs manuels de cours et du rôle que ces mêmes annotations jouent lors des changements de propriétaire des ouvrages, C. C. Marshall développe une typologie des annotations, selon leurs formes et leurs fonctions²⁷. Sa réflexion se développe autour du thème de l'hypertexte et des bibliothèques numériques à venir : pour elle, il faut à terme arriver à concevoir des outils qui rendent possible le transfert des usages liés aux supports imprimés vers les supports électroniques²⁸. Un troisième article publié en 1998 aborde également la question du partage de ces annotations : de l'annotation personnelle en marge d'un texte à des annotations

²⁵ Ibid.

²⁶ LECOLINET, Eric, ROBERT, Laurent, ROLE, François. Text-image coupling for editing literary sources. *Computers and the Humanities*, 2002, 36, p. 49-73. Article disponible en ligne [URL : <http://ipsapp008.lwwonline.com/content/search/4589/22/1/fulltext.pdf>], consulté le 16 mai 2002.

²⁷ MARSHALL, Catherine C. Annotation : from paper books to the digital library. Disponible sur Internet <<http://www.csdl.tamu.edu/~marshall/di97.pdf>> (consulté le 25.04.2002)

²⁸ MARSHALL, Catherine C. Toward an ecology of hypertext annotation. Disponible sur Internet <<http://www.csdl.tamu.edu/~marshall/ht98-final.pdf>> (consulté le 25.04.2002)

partagées appelées à former un autre texte, les outils informatiques devraient accompagner cette évolution²⁹. Au terme de ces études marquées par leur fort aspect théorique, même si elles se fondent sur une première enquête d'usages, C. C. Marshall participe aux expérimentations d'un prototype développé par la firme Xerox : *Xlibris*, un e-book permettant d'annoter les textes et de constituer des corpus d'annotations. Calquée sur le modèle traditionnel du livre imprimé, cette tablette électronique permet des annotations manuscrites libres grâce à un stylet, mais est également dotée d'un système de mémorisation des passages annotés et d'un système d'indexation pondérée des termes soulignés, pour améliorer les performances de la recherche. Il s'agit là d'une « active reading machine » qui permet d'étudier le processus de lecture, la recherche d'information, et à l'avenir le travail collaboratif. Cathy Marshall et ses collaborateurs ont mené une étude sur les pratiques de six chercheurs travaillant ensemble sur différents projets : observation de la lecture sur papier, de la lecture sur *Xlibris*, du processus d'annotation et de lecture des références bibliographiques dans l'un et l'autre cas³⁰. L'auteur conclue que l'annotation est un mode d'appropriation du texte très important, et que « our experiences with the Reader's Notebook suggest that we need to involve readers in defining computational services if they are to be useful to them ».

Enfin, le département Anglais de l'Université de Caroline du nord, aux Etats-Unis, a développé sur l'Internet une bibliothèque numérique consacrée aux oeuvres de William Blake : William Blake Archives. Cette bibliothèque est étonnamment riche en fonctionnalités³¹. Ce projet a été l'occasion d'une réflexion sur la mise en page électronique : travail sur les nuances de gris, la réduction des marges de la

²⁹ MARSHALL, Catherine C. The future of annotation in a digital (paper) world. 1998. Article disponible en ligne [URL : <http://www.csdl.tamu.edu/~marshall/uiuc-paper-complete.pdf>], consulté le 27 avril 2002.

GOLOVCHINSKY, Gene, PRICE, Morgan N., SCHILIT, Bill N. From reading to retrieval : freeform ink annotations as queries. Disponible sur Internet <http://www.fxpal.com/PapersAndAbstracts/papers/go199a.pdf> (consulté le 15.05.2002)

GOLOVCHINSKY, Gene, MARSHALL, Catherine C. Hypertext Interaction Revisited. Disponible sur Internet <<http://www.fxpal.com/PapersAndAbstracts/papers/go12000a.pdf>> (consulté le 20.05.2002)

³⁰ MARSHALL, Catherine C., et al. Introducing a digital library reading appliance into a reading group. In *Proceedings of ACM Digital Libraries 99*, ACM Press, pp. 77-84, 1999. Disponible sur Internet <http://www.csdl.tamu.edu/~marshall/dl99.pdf> (consulté le 22.05.2002)

version imprimée (ce qui permet un gain de volume pour les données et de place à l'écran), la correction du flou issu de la numérisation, le choix d'une qualité d'image permettant des réductions et visualisation de détails, le maintien des couleurs originales. Des fonctions logicielles permettent de gérer agrandissements à l'écran, qualité de l'image, multi-fenêtrage pour l'affichage simultané d'une image et des informations bibliographiques liées : transcription du texte, description fine de l'image, renseignement sur sa provenance, son actuelle localisation, l'établissement qui la conserve. Ces informations textuelles sont conservées dans un espace du fichier image, qui lui est donc lié. L'utilisateur peut également utiliser le logiciel d'annotations Inote, qui permet de créer des annotations relevant d'image ou de parties d'image, ou de lire toutes les annotations se rapportant à l'image qu'il consulte. Un moteur de recherche d'image permet également de rechercher des images, à partir d'une indexation ou des annotations enregistrées. Les besoins des usagers sont donc manifestement pris en compte dans le discours et les réalisations des William Blake Archives, mais là encore, aucune étude sur les pratiques réelles des utilisateurs n'est citée à son propos.

L'élaboration de métadonnées, qui rendent possibles les recherches précises et efficaces est également l'occasion de faire collaborer spécialistes des textes anciens et informaticiens. L'application de la TEI à un corpus de lettres manuscrites du XVI^e siècle a fait l'objet d'une réflexion³² : l'hypothèse était que « la numérisation en TEI des manuscrits devait faciliter leur diffusion et rendre le travail d'analyse historique moins fastidieux et plus scientifique » (p. 130). Travailler sur des documents anciens nécessite de construire de nouvelles métadonnées adaptées, ce qui rend nécessaire la collaboration décrite. A l'INRIA, cette richesse des métadonnées appliquées au document ancien est le point de départ d'un développement des techniques de comparaison et de recherche des

³¹ VISCOMI, Joseph. Digital facsimiles : reading the William Blake archives. *Computers and the Humanities*, 2002, 36, p. 27-48. Disponible sur Internet <http://ipsapp009.lwonline.com/content/search/4589/22/4/fulltext.pdf> > (consulté le 15.05.2002). Voir aussi en 3.8.7.

³² FEKETE, Jean-Daniel, DUFOURNAUD, Nicole. Analyse historique de sources manuscrites : application de TEI à un corpus de lettres de rémission du XVI^e siècle. *Document numérique*, 1999, 3, 1-2, p. 117-134.

métadonnées³³ : si l'auteur s'est contenté de l'exemple du chercheur isolé, il « semble que les solutions [...] explorées puissent être mises en oeuvre dans un cadre coopératif, par exemple pour combiner les travaux de différents chercheurs et vérifier la cohérence de cette articulation »³⁴.

Outre ces projets de postes de travail, avortés, aboutis ou encore en cours, et les sites de bibliothèques qui ont été le cadre de l'expression des besoins des chercheurs, ces derniers abordent le thème de la numérisation de documents anciens dans les articles qu'ils publient.

2.2. La voix des chercheurs

Des chercheurs prennent donc position dans la recherche sur la numérisation de documents anciens, se rendant compte que si les questions informatiques ne relèvent pas d'eux, les fonctionnalités et types de numérisation les concernent.

Deux établissements de recherche spécialisés en textes anciens se distinguent par leurs sites Internet : l'Institut de recherche en histoire des textes (IRHT), organisme français, et le Consortium pour ordinateurs en sciences humaines, association de collègues et d'universités canadiennes, qui publie sur l'Internet la revue *Early modern literature studies*.

2.2.1 L'IRHT

L'IRHT propose sur son site Internet une revue électronique semestrielle intitulée *Le médiéviste et l'ordinateur*. A l'automne 2001 est paru un numéro spécial issu d'un colloque : *La numérisation des manuscrits médiévaux : actes de la journée d'étude, Paris, 13 octobre 2000*³⁵. Les articles, écrits par des chercheurs spécialistes de l'histoire des textes, présentaient le point de vue de l'utilisateur à propos des documents anciens numérisés.

³³ ROLE, François. Représentation et exploitation de métadonnées complexes : le cas des documents anciens. *Document numérique*, 1999, 3, 1-2, p. 135-150.

³⁴ Ibid., p. 149.

³⁵ *Le Médiéviste et l'ordinateur* [en ligne]. Paris : Institut de recherche en histoire des textes. Semestriel. Disponible sur Internet <<http://irht.cnrs-orleans.fr/medieviste.htm>> (consulté le 05.06.2002)

Dès 1999, un article intitulé « Conclusions et propositions pour que l'étude des textes puisse aussi se faire sur l'Internet »³⁶ fait le point sur les besoins des chercheurs en matière de diffusion électronique des textes : les textes sources surtout sont demandés, en mode texte et en mode image, mais sans nécessairement un appareil critique élaboré. Egalement présenté comme une « exigence de base : que le texte soit accompagné de toutes les indications nécessaires à une utilisation scientifique, depuis le nom et les références des auteurs des différentes versions jusqu'aux règles de transcription appliquées, en passant par une description de la source ou des sources ». La question du droit de propriété intellectuelle des chercheurs est également posée : le rôle des métadonnées en ce domaine semble très important. En conclusion, « le texte (et bientôt le livre) électronique ne s'oppose pas, par conséquent, au texte papier. Chaque forme -- on dirait volontiers chaque format -- du texte correspond à des besoins ou à des activités spécifiques et complémentaires. Le texte papier ne craint pas les coupures de courant, les pannes de système ; il reste d'une commodité inégalée pour la lecture, le voyage, l'annotation rapide. En revanche, on ne peut l'explorer dans ses caractères, dans ses mots, dans ses structures syntaxiques, phraséologiques, statistiques... ; on ne peut le communiquer sans en perdre l'utilisation, on ne peut le coder, le baliser, l'enrichir indéfiniment et collectivement de valeur ajoutée par l'analyse et la critique. »³⁷

L'intérêt de la double numérisation est encore plus important pour le paléographe : « Les besoins des paléographes sont à mi-chemin entre ceux des philologues, travaillant sur des textes que la machine sait interpréter elle-même, et ceux des historiens de l'art, qui accèdent directement aux images »³⁸. Une double numérisation texte et image, avec des liens entre les deux formats, constitue une réponse possible aux besoins spécifiques des paléographes.

³⁶ PELLEN, René. Conclusions et propositions pour que l'étude des textes puisse se faire aussi sur Internet. *Le médiéviste et l'ordinateur*, hiver 1999, n° 38. Disponible sur Internet <<http://irht.cnrs-orleans.fr/meto/mo3816.htm>> (consultée le 15.05.2002)

³⁷ Ibid.

³⁸ SMITH, Marc. Numérisation et paléographie. *Le Médiéviste et l'ordinateur*, n° 40, automne 2001, *La numérisation des manuscrits médiévaux : actes de la journée d'étude, Paris, 13 octobre 2000*. Disponible sur Internet <<http://irht.cnrs-orleans.fr/meto/mo40-03.htm>> (consulté le 15.05.2002)

Toujours sur la double numérisation, le désaccord qui peut exister entre chercheurs et établissements de conservation est évoqué, les premiers préférant du mode texte susceptible d'être travaillé, les seconds le mode image, dans un but de communication des originaux.

Les outils de visualisation et de travail sur le texte : grossissement, multifenêtrage pour la comparaison, outils de transcription, outils d'édition classiques des traitements de textes, sont également plébiscités³⁹.

Pour conclure sur les attentes des chercheurs formulées par l'IRHT, les documents numérisés doivent pour être utilisables et donc utilisés répondre à leurs besoins : « une ultime précision pour dire que je n'ai retenu que les entreprises qui proposaient des images utilisables pour du « vrai » travail, images où l'on peut lire le texte et essayer d'interpréter les tracés »⁴⁰. Le critère de lisibilité est fondamental, de même que les informations permettant l'identification du document, et les outils de manipulation restent très importants, dans la mesure où ils donnent à la version numérique sa valeur ajoutée.

2.2.2 Le COSH

Le numéro spécial n° 4 de la revue *Early modern literature studies*⁴¹, du Consortium pour ordinateurs en sciences humaines, présente les comptes-rendus des communications données lors de conférences tenues pendant le congrès 1998 de Social Sciences and Humanities de l'Université d'Ottawa, au Canada. Ce congrès était conjointement organisé par le Consortium pour ordinateurs en sciences humaines et la Canadian Society for Renaissance Studies. Il s'agit ici d'une société savante exclusivement tournée vers l'étude de la Renaissance, à la différence de l'IRHT. Les opinions exprimées diffèrent d'ailleurs pour certaines de celles exprimées à l'IRHT : le texte électronique doit être moins un fac-similé de

³⁹ PORTET, Pierre. La numérisation des manuscrits scientifiques et techniques médiévaux sur le web : notes de lecture. *Le Médiéviste et l'ordinateur*, n° 40, automne 2001, *La numérisation des manuscrits médiévaux : actes de la journée d'étude, Paris, 13 octobre 2000*. Disponible sur Internet <<http://irht.cnrs-orleans.fr/meto/mo40-08.htm>> (consulté le 20.05.2002)

⁴⁰ Ibid.

⁴¹ *Early modern literature studies* [en ligne]. Sheffield : University of British Columbia. Triannuel. Disponible sur Internet <<http://www.shu.ac.uk/emls/emlsmast.html#edgp>> (consulté le 05.06.2002)

l'édition imprimée ou manuscrite originale qu'un nouveau type d'édition, qui entraîne un nouveau rapport au texte et permet de le manipuler.

A partir de l'exemple des *Emblèmes* d'Alciat, l'évolution du processus de lecture liée au passage du texte imprimé au texte électronique est étudiée⁴² : quelles représentations mentales nouvelles se forment dans cette situation encore nouvelle de la lecture sur écran ? Si la lecture proprement dite s'est perdue, le texte devient mouvant, et c'est là que réside l'intérêt de la numérisation des textes. Les fac-similés ne présentent que peu d'intérêt par rapport à un document texte où sont possibles la recherche en texte intégral, l'annotation, l'indexation de termes. Par opposition avec les nombreuses éditions sur papier des *Emblèmes*, l'édition électronique permet une lecture non suivie, l'établissement de renvois actifs par liens hypertextes, toutes possibilités adaptées au caractère particulier des *Emblèmes* : l'hypertexte est plébiscité.

Autre préoccupation de l'EMLS, celle de prendre en compte les besoins des spécialistes.

La conception de la « Virtual Furness Shakespeare Library », développée à l'Université de Pennsylvanie, s'est faite selon l'idée que l'évolution des études et plus généralement de l'enseignement et de la culture de la Renaissance donne aux éditions contemporaines, copiées sur celles du XIXe siècle et appauvries, une importance considérable⁴³. Il devient nécessaire de retourner à des éditions plus riches, d'où le projet qui en découle « The English Renaissance in Context : Teaching Shakespeare with the World Wide Web », mené par le département d'Anglais et le département des collections rares à Van Pelt, qui dépend de l'Université de Pennsylvanie, projet destiné à la recherche comme à l'enseignement. Les besoins des différents groupes d'utilisateurs sont décrits : « Scholars for example, need access to the widest possible range of early modern and later materials. Researchers do not want choices made for them. They will

⁴²FELTHAM, Mark, and BARKER, William. The Web and the Book: The Memorial Electronic Edition of Andrea Alciato's *Book of Emblems*. *Early Modern Literary Studies*, janvier 2000, 5, 3, n° 4. Disponible sur Internet <<http://purl.oclc.org/emls/05-3/fbemblem.html>> (consulté le 10.04.2002)

⁴³BUSHNELL, Rebecca. Reinventing Rare Books : The "Virtual Furness Shakespeare Library" at the University of Pennsylvania. *Early Modern Literary Studies*, janvier 2000, 5, 3, n° 4. Disponible sur Internet <<http://purl.oclc.org/emls/05-3/bushfurn.html>> (consulté le 10.05.2002)

profit from the use of searchable texts as well as facsimile materials. They do not need many background materials but they do need bibliographic information, and they would be interested in multiple editions of works »⁴⁴.

De la même façon, le projet « Internet Shakespeare Editions », a commencé par une enquête auprès de chercheurs : 49 ont répondu au questionnaire en avril 1997, permettant de déterminer quelle utilisation est souhaitée ou envisagée pour le corpus numérique final, et quel est le public susceptible de s'y intéresser. Les questions portaient sur l'intérêt d'une telle offre pour l'enseignement, le mode de consultation probable et les outils liés, le type de numérisation souhaitée, et enfin l'apparat critique qu'il serait bon d'inclure dans le projet. Pour ce qui est du mode de numérisation, la majorité souhaite avoir accès à des reproductions des originaux, contrairement aux conclusions tirées de l'étude sur les *Emblèmes*. L'auteur innove en se plaçant dès 1997 et tout à fait volontairement d'un point de vue marketing, analysant la demande pour élaborer une offre, en récusant toutefois toute utilisation commerciale des résultats de la recherche.

La voix des chercheurs se fait donc entendre à propos de la numérisation des documents anciens. Les idées présentées dans les articles sont assez diverses, voire opposées, notamment à propos du mode de numérisation, de la présence d'une transcription, de l'usage même de ces nouveaux outils, ce qui montre le lien étroit et à double sens qui existe entre un besoin et une fonctionnalité logicielle. Pour conclure sur le mode de numérisation, la préférence pour le mode texte ou le mode image est vraisemblablement conditionnée par plusieurs éléments :

- la discipline du chercheur, et le type de documents concernés : les linguistes et littéraires s'intéressent davantage au contenu, et les fonctions de recherche dans le texte leur sont utiles ; à l'inverse, les historiens du livre ou les paléographes ont eux besoins d'une reproduction la plus fidèle possible du document original, pour éviter tout risque d'interprétation.

⁴⁴ Ibid.

- la maîtrise de l'outil informatique : utiliser des logiciels aux fonctionnalités très développées n'est pas chose aisée pour tous les chercheurs en sciences humaines.
- le fait d'avoir ou non accès facilement à des documents originaux : si c'est le cas, une reproduction en fac-simile n'offre que peu d'intérêt.

Le projet DEBORA se situe dans ce mouvement d'analyse a priori des besoins des chercheurs : une enquête a été menée par l'équipe de recherche auprès de 62 chercheurs spécialistes du livre du XVI^e siècle. Le questionnaire abordait l'usage de documentation électronique, les pratiques actuelles des chercheurs, et les fonctionnalités qu'ils souhaitaient voir implémentées : plus de la moitié des chercheurs interrogés affirment avoir besoin de fonctionnalités adaptées au travail sur le livre ancien, parmi lesquelles l'annotation (12,9 % de chercheurs intéressés), le surlignage (11,9 %), la possibilité de calculer des statistiques sur le texte (10,9 %), et surtout celle de rechercher des occurrences (27,7 %). La recherche des occurrences, plébiscitée ici, demande une numérisation en mode texte, ou une transcription du texte des images. Il est vrai que la plupart des chercheurs interrogés pour cette enquête sont des littéraires, lexicographes ou lexicologues, ce qui va dans le sens de l'analyse qui vient d'être menée.

Les sites et logiciels destinés aux chercheurs ont ainsi inclus dans leur phase de conception une réflexion développée sur les besoins de leur public, et les fonctions à développer pour les satisfaire. A l'inverse, les usagers peuvent être pris en compte après une première distribution du produit, et influencer non la conception mais l'évolution de l'interface : on ne parlera plus alors de besoins, mais de pratiques.

3. Prendre en compte les retours des usagers pour faire évoluer une interface

Cette seconde méthode pour prendre en compte les usagers dans l'élaboration d'une bibliothèque numérique est davantage le fait de sites relevant d'établissements bibliothèques. Cette méthode consiste en effet à faire évoluer son interface en fonction des retours des usagers : expression des besoins comme précédemment, pratiques observées. Les postes de travail décrits plus haut sont pour certains disponibles sur Internet, mais ils ne proposent qu'un petit corpus, un seul texte parfois, et s'adressent donc presque uniquement à des spécialistes. Les établissements qui numérisent se posent eux dans une perspective d'accès plus large à leurs collections, et constituent des corpus numériques d'une importance parfois considérable. En contrepartie de cette concentration sur les questions de numérisation, les interfaces élaborées proposent moins de fonctions. Afin de poursuivre sur les corpus, les documents numérisés dans ce cadre sont beaucoup moins homogènes que ceux des projets précédemment présentés : le volume même des données présentes, et la volonté de toucher un plus vaste public en sont autant de causes. Les documents anciens sont toutefois bien présents : le mode de diffusion numérique s'intègre dans une politique de conservation, et ces documents ne sont plus soumis au droit d'auteur.

Dans le cadre de ces sites, la réflexion sur les attentes et pratiques vient dans un deuxième temps, quand le site est pour ainsi dire lancé. A l'interface initiale minimale vont être apportées des modifications, modifications souvent liées à l'usage observé du site : le fait que les ressources numérisées soient disponibles sur l'Internet a permis la constitution de groupes d'utilisateurs.

Deux attitudes se distinguent : celles des bibliothèques qui font évoluer leur interface en fonction de retours des usagers, et celles qui conduisent de vastes enquêtes d'usages, destinées à élaborer une politique réfléchie en ce domaine.

3.1. Les sites ayant évolué avec le feed-back

Cette première attitude consiste à faire évoluer son site en fonction des retours éventuels des usagers. Assez pragmatique et moins volontaire que la seconde (les

enquêtes d'usages), cette attitude permet toutefois d'intégrer un certain nombre de pratiques, et de mieux satisfaire ses usagers.

Les travaux menés par la Bibliothèque nationale de France en matière de consultation d'ouvrages numérisés sont intéressants, pour l'évolution de la politique appliquée : les résultats de l'enquête lancée sur *Gallica* n'ont jamais été publiés. Si quelques informations sur les pratiques observées peuvent être glanées dans différents articles, c'est essentiellement de la démarche des responsables de la bibliothèque numérique de la BnF dont il sera ici question.

Dans le programme de construction de la BnF est incluse la numérisation d'ouvrages anciens du patrimoine national, pour une mise à disposition sur des postes dédiés dans les murs de l'établissement :

« Sur les sites de la BnF, deux types de terminaux permettront de consulter les textes et les images numérisées : des PLS (postes de lecture simple) et des PLAO (Postes de Lecture Assistée par Ordinateur). Les PLS permettront de " feuilleter " ces documents. Ils seront installés dans la zone " grand public " de la bibliothèque. Les PLAO seront de véritables stations de travail sur PC équipées de toute une panoplie d'outils pour charger et héberger les textes sélectionnés, pour naviguer en mode image dans des textes qu'il sera loisible de repérer, surligner, souligner, annoter, comparer grâce au multifenêtrage et dont il sera possible d'extraire des sous-corpus. Le lecteur pourra également passer en mode texte pour récupérer ces textes et les travailler avec un simple logiciel de traitement de texte. Ces PLAO seront mis à la disposition des chercheurs. »⁴⁵

La BnF fait le pari que la lecture sur écran va se développer⁴⁶. L'arrivée d'Internet va rapidement modifier la donne : avec le lancement du programme *Gallica*, la bibliothèque numérique en ligne, les PLS destinés au grand public ne se justifient plus, et sont supprimés. Les fonctions disponibles sur l'interface *Gallica* sont celles des PLS : feuilletage et grossissement. Les PLAO font eux l'objet d'une enquête par observations. La réflexion des responsables du service de la

⁴⁵ MAZHOUD, O., PASCUAL, E., VIRBEL, Jacques. Annotating as a document management tool. Disponible sur Internet <<http://www.hit.uib.no/allc/mazhoud.pdf>> (consulté le 20.05.2002)

bibliothèque numérique de la BnF a aujourd'hui évolué : les PLAO (appelés ensuite PABN pour poste d'accès à la bibliothèque numérique), conçus pour une architecture informatique en réseau local, vont céder la place à un Intranet dont l'interface usager sera sur le plan graphique très proche de celle de *Gallica*. Des fonctionnalités supplémentaires seront tout de même conservées : possibilité de prendre des notes sur un bloc-notes électronique, plus haute définition d'image. Cette évolution, qui a aussi pour but de faciliter la gestion technique du réseau, n'a été précédée d'aucune étude spécifique auprès des usagers.

Une enquête par questionnaire avait bien été lancée sur *Gallica*, mais les résultats n'ont jamais été publiés : les questions portaient principalement sur le niveau de satisfaction des internautes et le contenu, et ne permettaient pas d'analyser leurs pratiques de façon précise. Une seconde enquête est aujourd'hui en préparation, qui portera tant sur l'interface Internet que sur les postes internes : un panel d'utilisateurs va être constitué et interrogé, pour déterminer quels sont les usages et établir l'orientation à donner au travail sur les interfaces et fonctionnalités. L'enquête sur les usages suit donc de près le lancement de l'Intranet : la constitution d'un groupe d'utilisateurs ne peut que suivre le lancement d'un nouveau service, et une enquête sur les pratiques ne peut porter que sur un groupe d'utilisateurs un tant soit peu anciens. Il ne s'agit dans le cas contraire que de tests en laboratoire.

Soulignons pour conclure que ce qui peut passer pour un recul sur les fonctionnalités proposées au lecteur s'inscrit dans une réflexion globale sur les missions de la BnF : en tant que bibliothèque nationale et établissement public, elle estime que la mission première consiste à mettre à disposition du public les documents patrimoniaux. La réflexion sur les usages, si importante et intéressante soit-elle, est une deuxième étape. Par ailleurs, le modèle contemporain de développement des bibliothèques numériques est l'Internet : chaque utilisateur consulte les fichiers à distance, et dispose de ses propres logiciels sur son ordinateur personnel. Développer un logiciel de lecture assistée n'est pas une priorité actuelle pour la BnF.

⁴⁶ WAGNEUR, Jean-Didier. La bibliothèque électronique de la BnF ou Quand le lecteur a le droit d'écrire sur les livres. *Revue de la BnF*, juin 1999, 2, p. 78-81.

Quant à l'équipe du projet Runeberg, composée d'étudiants de l'Université de Linköping, en Suède, elle a dès 1993 mis gratuitement sur l'Internet des textes de la littérature nordique, en mode texte⁴⁷. Divers problèmes sont rencontrés au fil des lectures par les usagers : erreurs de saisie dans les fichiers, productivité réduite pour cause de relecture attentive, difficulté à identifier une édition. Le passage à une numérisation en mode image est alors ressentie comme une évolution nécessaire pour cette bibliothèque numérique, qui veut faire connaître au monde la littérature nordique depuis ses origines. La production de fac-similé est mise en place, avec utilisation de l'interface développée par *Making of America*, la bibliothèque numérique américaine. L'équipe mène parallèlement une réflexion sur les coûts et l'intérêt d'une double numérisation en mode texte et en mode image. Soulignons que le passage d'un mode de numérisation à l'autre, voire à une double numérisation comme c'est le cas dans *DEBORA* a été décidé grâce aux retours des usagers autant que pour des questions de production. Le site n'a par la suite fait l'objet d'aucune enquête sur les pratiques.

3.2. Les études sur les usages : méthodes employées et équipes de recherche

La seconde attitude des bibliothèques numériques, très volontaire, consiste à utiliser des méthodes d'enquête sociologiques pour faire correspondre une politique aux pratiques des utilisateurs.

L'entrée de l'utilisateur réel dans le paysage des bibliothèques numériques est récente, et relève en bonne part de la recherche anglo-saxonne. Les démarches mises en oeuvre se rapprochent des réflexions qui ont pu être menées ces dernières années sur l'évaluation des bibliothèques traditionnelles, voire sur une méthode marketing, où l'analyse de la demande intervient dans la définition de l'offre : les innovations technologiques gagnent en valeur si elles servent l'utilisateur et font l'objet d'une utilisation régulière.

47 ARONSSON, Lars. Project Runeberg's electronic facsimile editions of Nordic literature. Disponible sur Internet <<http://www5.hk-r.se/elpub99/ap-nsf/08cbc2f88424ad99c12566ff002>> (consulté le 13.05.2002)

3.2.1 Les méthodes appliquées

Les deux approches principales des usages des bibliothèques électroniques sont quantitative et qualitative : la première utilise les données statistiques disponibles grâce aux serveurs des bibliothèques numériques, la seconde privilégie les méthodes d'enquête sociologique, que ce soit les observations, les enquêtes par questionnaire, les entretiens, ou les tests avec un panel d'utilisateurs. De fait, même si l'on a choisi de ne pas occulter les enquêtes quantitatives afin de ne pas risquer de donner une image faussée du paysage de la recherche sur le sujet, seules les études qualitatives qui relèvent de la sociologie donnent des résultats intéressants et exploitables, tant il est difficile de déterminer les pratiques des usagers à partir de données statistiques de téléchargement. Encore que l'on peut se demander dans de nombreux cas si cette distinction est pertinente, dans la mesure où l'utilisation croisée de plusieurs approches est choisie.

Le tableau présenté page suivante permet de visualiser les méthodes utilisées pour ces enquêtes.

Etat de la littérature. Quelle place est accordée à l'utilisateur dans la recherche sur les bibliothèques numériques ?

	Analyse de statistiques du serveur	Questionnaire	Entretien	Observation	Observation de groupe
ABU	+				
CISTI		+			
Alexandria digital library		+		+	+
e-prints archives			+		
<i>Material mastery...</i>			+		
CIHM		+	+		
ECO 1	+	+			+
ECO 2		+			
F. W. T. Collection	+		+	+	+
<i>ILEJ</i>		+			
TLG / Perseus/ Chicago Homer					
Perseus 1999	+	+			
Perseus 2001	+				+
MOA (HCI group)	+	+			+
Frick Art Ref. (HCI group)					+
SCOT (HIC group)	+	+	+	+	+
Xlibris		+	+	+	+

3.2.2 Les équipes de recherche

Les équipes de recherche qui ont élaboré et analysé des enquêtes sur les usages des bibliothèques électroniques sont présentées ici.

D'après l'analyse qui précède, la recherche sur les bibliothèques numériques s'oriente désormais vers l'étude des usages et pratiques. Ceci se retrouve notamment pour ce qui est des bibliothèques de sciences et techniques : les bases de périodiques en texte intégral, les bases de données factuelles, les serveurs de *e-prints* sont concernés. Cette documentation est sans doute loin de *DEBORA* et des ouvrages anciens numérisés, mais elle illustre une tendance de la recherche actuelle, et peut avoir des points communs pour ce qui est de la méthode employée ou de l'analyse cognitive de la lecture sur écran. Cinq articles ont été retenus pour dresser un tableau de ces études d'usages, et serviront d'introduction à cette partie. Ces articles ont tous été édités dans le *Journal of the american society for information*, qui a consacré un numéro à ces questions en février 2000.

Ces enquêtes se font par le biais de questionnaires⁴⁸ ou de tests auprès de groupes d'utilisateurs volontaires⁴⁹. Un auteur insiste sur l'intérêt d'analyser les statistiques du serveur, dans la mesure où l'utilisateur ne parle que peu de ses problèmes, et ne parle jamais de ce qu'il ne sait pas ou ne comprend pas. Dans trois des cinq articles retenus ici, les résultats des enquêtes ont amené à modifier l'interface de consultation, pour mieux satisfaire les besoins des chercheurs. L'article écrit par l'équipe de recherche en sciences sociales de l'Université de l'Illinois précise que « how users meet infrastructure is a key practical, methodological challenge for digital library design » (p. 394). Les fonctionnalités utilisées ou souhaitées sont décrites : elles concernent surtout la recherche d'information et le suivi des publications.

La documentation scientifique est trop différente de celle des chercheurs en sciences humaines pour qu'une comparaison poussée ait du sens, mais ces articles

⁴⁸ BELZILE, Sylvie. NRC clients' perceptions of electronic resources and use of CISTI's virtual library. *Internet reference services quarterly*, vol. 4, (4), 1999, p. 5-16.

⁴⁹ HILL, Linda L., et al. Alexandria Digital Library : user evaluations studies and system design. *Journal of the american society for information*, février 2000, n° 51, vol. 3, p. 246-259.

aident à situer la recherche sur les usages et attentes des usagers pour les bibliothèques électroniques.

Les travaux de Robert R. Downs, de l'Université d'Etat de Montclair, aux Etats-Unis, portent eux sur l'étude des usages des documents anciens originaux avant de s'intéresser aux usages de la bibliothèque numérisée : perspective originale par rapport aux autres travaux décrits ici. Dans le cadre de la *Frederick Winslow Taylor Collection*, une première enquête par observations et entretiens a permis de déterminer un modèle conceptuel de recherche documentaire dans le processus de recherche⁵⁰. Les résultats de cette première enquête ont permis de développer un prototype de consultation d'images de documents numérisés, avec des outils adaptés aux pratiques des chercheurs. Une seconde enquête, par questionnaires et entretiens téléphoniques, a été soumise aux chercheurs utilisant ce logiciel : disponible sur l'Internet, il offrait 38 titres de la *Frederick Winslow Taylor Collection*. Les résultats permettent de conclure qu'à des degrés variables, les outils implémentés dans le logiciel répondaient aux besoins des chercheurs, et correspondaient au modèle conceptuel de recherche cité plus haut. Des préconisations ont pu être formulées pour l'évolution du logiciel. Les résultats permettent également d'affirmer que les chercheurs sont demandeurs de formations et d'aide apportées par des acteurs humains des bibliothèques. L'auteur de l'article précise pour conclure que ses prochains travaux pourraient porter sur les relations de collaboration existant entre les utilisateurs.

Depuis 1995, le Human-computer interaction group (HCI group) de l'Université Cornell, aux Etats-Unis, dirigé notamment par Robert Roeger et Geri Gay, travaille sur l'évaluation des bibliothèques numériques qui offrent des reproductions de collections de bibliothèques ou de musées. Différents axes de recherche sont développés⁵¹ : les métadonnées et question des droits (« backstage »), la

⁵⁰ DOWNS, Robert R., FRIEDMAN, Edward A. Digital library support for scholarly research. *Information processing and management*, 35, 1999, p. 281-291.

⁵¹ JONES, Michael L. W., GAY, Geri K., RIEGER, Robert H. Project soup : comparing evaluations of digital collection efforts. *D-Lib Magazine*, novembre 1999, vol. 5, n° 11. Disponible sur Internet <www.dlib.org/dlib/november99/11jones.html> (consulté le 7.05.2002)

maintenance et l'accès aux documents, la problématique des usages. Les responsables de cette équipe de recherche soulignent qu'il est nécessaire de toujours placer l'utilisateur au centre des bibliothèques numériques, mais également au centre de la recherche portant sur ces bibliothèques, pour assurer leur avenir et leur utilité. La politique de l'offre, souvent adoptée dans les bibliothèques traditionnelles, n'est pas adaptée aux nouveaux modèles de bibliothèques numériques ; de même, si les collections traditionnelles et les usages qui en découlaient peuvent apparaître figés, ce n'est plus le cas pour les bibliothèques numériques : constitution des collections et usages de ces collections par le bibliothécaire comme par l'utilisateur s'interpénètrent et s'influencent mutuellement. La question des méthodes de recherche utilisées pour mieux connaître les usages des bibliothèques numériques est également posée⁵² : la méthode « SCOT », pour « Social construction of technology », rassemble l'observation des pratiques de l'utilisateur, la collecte des opinions par les méthodes sociologiques classiques, l'analyse des chemins de navigation suivis grâce à des logiciels spécifiques, l'analyse des statistiques de connexion, et enfin la mise en place d'outils permettant d'obtenir un retour (« feedback ») des utilisateurs. Cette méthode a été utilisée lors des cinq suivis de projets réalisés par le Human-computer interaction group, et a permis de dresser une synthèse sur les besoins et attentes des utilisateurs.

Un travail de longue haleine est mené par Joan M. Cherry et Wendy M. Duff sur l'usage des documents numérisés proposés sur le site canadien *Early Canadiana Online*. Le « Canadian Institut for Historical Microreproductions » (CIHM) a pour mission de fournir aux chercheurs les reproductions en fac-similé de documents anciens du patrimoine national. Une enquête sur les usages des utilisateurs a été lancée en 1998, parallèlement à un premier projet de numérisation⁵³ : les réponses des utilisateurs liaient étroitement ces deux projets initialement distincts, et les

⁵² RIEGER, Robert, GAY, Geri. Tools and techniques in evaluating digital imaging projects. *RLG Diginews*, vol.3, n°3. Disponible sur Internet <<http://www.rlg.org/preserv/diginews/diginews3-3.htm>> (consulté le 7.05.2002)

⁵³ BURNINGHAM, B. Attitudes of the Canadian Research Community toward Creating and Accessing Digitized Facsimile Collections of Historical Documents. *Computers and the Humanities*, 33, 1999, p. 409-419. disponible sur Internet <<http://ipsapp008.lwwonline.com/content/search/4589/15/4/fulltext.pdf>> (consulté le 22.05.2002)

responsables ont donc publié des résultats portant sur les attentes des chercheurs en matière de numérisation des collections. La conclusion principale est que, si les attentes en cette matière sont grandes, de nombreuses réserves sont émises quant à l'utilisation qui pourra en être faite : la reproduction électronique n'est considérée par les chercheurs que comme un pis-aller, faute de pouvoir consulter le document original. Une deuxième étude comparative sur l'usage et la satisfaction est lancée au printemps 1999⁵⁴ : par questionnaires, analyses des connexions et entretiens avec des usagers volontaires, elle permet de déterminer quels usages correspondent à quels types de documents (documents originaux, microfiches, version numérisée disponible à la date de l'enquête). Les résultats doivent notamment servir à améliorer la version électronique.

Enfin, une troisième enquête⁵⁵ portant uniquement sur le site a pour but de déterminer si usages et usagers ont évolué depuis la précédente enquête, et quels sont les usages actuels pour l'enseignement et la recherche. Les auteurs soulignent combien connaître les usages des bibliothèques numériques est important, pour développer des outils adaptés aux besoins des chercheurs.

L'Internet Library of Early Journals (ILEJ) est une bibliothèque électronique issue de la collaboration de plusieurs universités anglaises, et consacrée à des périodiques anglais des XVIII^e et XIX^e siècles, numérisés en mode image, avec parfois le texte issu d'une reconnaissance de caractères. Les métadonnées bibliographiques suivent deux DTD (dont la TEI), dans des documents en SGML. Une évaluation de l'interface a eu lieu auprès de 377 utilisateurs volontaires, la plupart venant de l'université⁵⁶. Si le taux de satisfaction est globalement très élevé

⁵⁴ CHERRY, Joan M., DUFF, Wendy M. Studying digital Library users over time: a follow-up survey of Early Canadiana Online. *Information Research*, 7, 2. Disponible sur Internet <<http://InformationR.net/ir/paper123.html>> (consulté le 5.05.2001)

⁵⁵ DUFF, Wendy M., CHERRY, Joan M. Use of historical documents in a digital world : comparisons with original materials and microfiche. *Information research*, vol. 6, n° 1, octobre 2001. Disponible sur Internet <<http://www.shef.ac.uk/~is/publications/infres/paper86.html>> (consulté le 04.05.2002)

⁵⁶ GARTNER, Richard. The Internet Library of Early Journals : an electronic library of primary sources on the Internet. Disponible sur Internet <[http://www5.bth.se/elpub99/ap_nsf/08c6c2f88424ad99c12566ff002a0c10/6f7b5fad7be3e268c12566ff00361328/\\$FILE/194-203.pdf](http://www5.bth.se/elpub99/ap_nsf/08c6c2f88424ad99c12566ff002a0c10/6f7b5fad7be3e268c12566ff00361328/$FILE/194-203.pdf)> (consulté le 20.05.2002)

pour la plupart des 11 critères retenus, certains points sont à améliorer : la vitesse de chargement, la qualité et la faible quantité des documents textes produits par reconnaissance de caractères. La qualité de l'image rend tout à fait possible la lecture sur écran : seule la moitié des utilisateurs disent imprimer les documents qu'ils veulent lire.

Perseus Digital Library enfin a fait l'objet d'une évaluation⁵⁷, pour déterminer quels problèmes rencontrait l'utilisateur, et donc quelles améliorations étaient à apporter.

En 1999, un rapport sur le projet Perseus, signé par Gary Marchionini entre autres, comprend la synthèse d'une enquête d'usages menée par questionnaire mis en ligne en 1998.

Plus récemment, en 2001, des données recueillies dans le cadre de travaux dirigés d'étudiants en grec ont été analysées : observations, entretiens, statistiques du serveur. Les améliorations attendues par les utilisateurs sont : augmenter le nombre de liens hypertextes, en prenant garde de ne pas perdre l'utilisateur ; développer des outils qui aident l'étudiant à se forger une représentation exacte du savoir ; donner des outils pour créer et gérer ses propres documents ; aider à développer l'enseignement et le travail collaboratif

3.2.3 Synthèse sur les pratiques observées

Décrire pour chaque groupe de recherche les pratiques observées dans l'utilisation des bibliothèques numériques aurait occasionné beaucoup de redites, sans donner d'aperçu global de la situation. Aussi a-t-on préféré les présenter à part. Seuls les résultats des enquêtes portant sur des bibliothèques numériques de sciences humaines sont rapportés ici : si les enquêtes sur les bibliothèques de sciences et techniques sont intéressantes pour les méthodes, les pratiques de leurs usagers ne relèvent pas de notre sujet.

⁵⁷ JEFFREY A. et al. Knowledge management in the Perseus digital library. *Ariadne*, septembre 2000, n° 25. Disponible sur Internet <http://www.ariadne.ac.uk/issue25/rydberg-cox/intro.html> (consulté le 3.05.2002)

Le premier objet d'étude de ces enquêtes sur l'usage des bibliothèques numériques est le profil de l'utilisateur.

Son statut est assez homogène, et se partage entre les enseignants-chercheurs, dans le cadre de leurs travaux, et les particuliers qui viennent consulter les sites par curiosité. La seconde enquête menée sur *Perseus* fait apparaître 42 % d'usage dans le cadre professionnel : cette proportion est de 75 % pour *ILEJ*, qui propose des fonds de périodiques très spécialisés. La part des professionnels a fortement augmenté entre les deux enquêtes d'*ECO* : 13 % de chercheurs en 1998, 45 % lors de la dernière enquête. Environ la moitié des utilisateurs de ces bibliothèques numériques sont donc des particuliers. L'enquête menée par *Perseus* en 1999 permet de constater que la moitié des connexions se font depuis le domicile, 17 % depuis le bureau, et 3 % seulement depuis une bibliothèque : les bibliothèques numériques sont intégrées à la vie quotidienne de leurs utilisateurs, et justifient leur existence par les possibilités d'accès à distance qu'elles offrent.

Les différentes disciplines représentées par les utilisateurs reflètent cette situation : histoire et littérature sont largement dominantes, mais la généalogie est également très présente, et semble être la préoccupation des particuliers qui viennent d'être cités. La deuxième enquête d'*ECO* par exemple a touché 33 % d'historiens pour 29 % de généalogistes.

La fréquence d'utilisation est le dernier élément de cette typologie des utilisateurs : la première enquête d'*ECO* comptait 70 % de nouveaux utilisateurs, proportion qui passe à 50 % lors de la seconde enquête ; 3 % font de la bibliothèque un usage quotidien, 18 % un usage hebdomadaire et 29 % un usage mensuel. *Perseus*, qui propose des fonds plus spécialisés, fait l'objet d'une utilisation encore plus régulière : 27,5 % d'usage hebdomadaire, le même pourcentage pour un usage plus fréquent, 24,3 % pour un usage mensuel. On peut distinguer à la lueur de ces chiffres deux types d'utilisateurs : des professionnels pour qui la bibliothèque numérique est un outil de travail régulier, des particuliers qui s'y promènent ou viennent chercher un renseignement ponctuel.

Dernier détail : la majorité des utilisateurs habitent le pays d'origine de la bibliothèque numérique, et le faible pourcentage d'utilisateurs étrangers est américain pour les bibliothèques européennes ou canadiennes, et canadien ou

européen pour les bibliothèques américaines. L'usage de l'anglais est une constante.

L'utilisation des fonctionnalités de recherche et de visualisation est également considérée.

Les différents modes de recherche font l'objet d'une utilisation intensive : 57 % des utilisateurs de *Perseus* déclarent s'en servir, de même que ceux de la *Frederick Winslow Taylor Collection*. Ces pourcentages montent à 60 et 74 % dans les deux enquêtes d'*ECO*. Il s'agit pour cette dernière bibliothèque de la recherche en texte intégral. D'autres modes d'accès à l'information sont également utilisés : parmi elles, les liens hypertextes (*Perseus*, *Frederick Winslow Taylor Collection*), les index (*Frederick Winslow Taylor Collection*, *ECO*), les tables des matières (*ECO* : 27 % en 1999 et 52 % en 2001).

Cette grande utilisation de la recherche est à l'origine d'une lecture partielle des documents : les utilisateurs de la *Frederick Winslow Taylor Collection* affirment que lire les passages entourant les termes trouvés et mis en surbrillance suffit à satisfaire leurs besoins. De la même manière, près d'un tiers de ceux d'*ECO* disent ne lire qu'une partie des documents. Les résultats présentés ne permettent malheureusement pas de conclure s'il s'agit là d'un comportement commun aux professionnels comme aux particuliers, ou propre à l'un de ces deux groupes : on aurait tendance à l'attribuer aux seconds, mais il pourrait s'agir d'une interprétation abusive. Les étudiants utilisant *Perseus* se plaignent d'ailleurs de cette fragmentation qui les égare. Notons que la lecture intégrale sur écran est le fait de 3 % des personnes ayant répondu au questionnaire d'*ECO* en 2001.

Autre façon de lire les textes proposés : les imprimer. Dans son enquête sur les usages de la *Frederick Winslow Taylor Collection*, R. Downs constate que les bibliographies qui accompagnent les textes ne sont généralement lues qu'imprimées : la lecture sur écran est là encore partielle, et l'usage du papier reste très ancré dans les comportements. Pour ce qui est d'*ECO*, 30 % des documents consultés sont imprimés. Si l'on ajoute ce chiffre au tiers représentant la lecture partielle sur écran et aux 3 % affirmant lire intégralement en ligne, 60 % environ

des documents consultés sont lus d'une façon ou d'une autre. Mais il n'est pas sûr que ce qui est imprimé soit lu...

Les outils de travail sur le contenu des documents font l'objet d'une utilisation beaucoup moins marquée : les utilisateurs de Perseus (dont la moitié rappelons-le sont des étudiants) affirment ne pas se servir des fonctions d'annotations et de création de documents en raison de leur complexité. De même, des fonctions d'édition comme celle de 'copier/coller' augmentent chez eux la sensation de fragmentation et de perte citée à propos de la lecture. Pour *ECO*, 19 % des utilisateurs introduisent des fragments de la bibliothèque numérique (textes ou images) dans leurs propres travaux (articles, thèses, cours).

Les enquêtes utilisées pour cette synthèse posaient, outre des questions sur les usages, des questions sur la satisfaction des usagers. Ces taux sont très élevés (75 % pour *ILEJ*, 90 % pour *ECO* en 2001), et ne donnent pas d'indications précieuses sur les usages. Une question liée est pourtant intéressante : 3,2 % seulement des personnes interrogées à propos d'*ILEJ* affirment que, s'ils en avaient eu la possibilité, ils auraient préféré consulter un fac-similé imprimé. Lors de la première enquête sur *ECO*, les utilisateurs devaient dire si le document original leur aurait été utile après consultation de la bibliothèque numérique : la moitié disent que non, 39 % occasionnellement, et 11 % souvent. C'est dire si les bibliothèques numériques, au dire même des utilisateurs, remplissent une mission de diffusion des collections rares et précieuses.

Les commentaires qui accompagnent les questions de satisfaction permettent aussi de cerner les difficultés des utilisateurs : les erreurs liées à la transcription automatique ne sont pas perçues comme un problème lourd pour les utilisateurs de la Frederick Winslow Taylor Collection. Par contre, ceux d'*ECO* se plaignent de la lenteur des impressions. Ils insistent pour que la collection proposée en ligne soit encore augmentée, et suggèrent de faire des tables des matières des documents actifs, avec des liens hypertextes vers les chapitres.

Pour ce qui est du travail collaboratif enfin, aucune pratique n'est observée et rapportée dans les résultats d'enquêtes. Les responsables de la bibliothèque numérique Perseus envisagent de développer de tels outils, notamment pour

l'enseignement. Actuellement, une seule base de données alimentée par plusieurs utilisateurs semble active dans le monde de la recherche, mais elle est présentée comme encore expérimentale : mise en place à l'IRHT, elle doit permettre de construire une base de connaissance sur les grands humanistes et les relations qu'ils entretenaient. Structurée avec le logiciel Access de Microsoft, « la base de données offre donc à ce jour des modules de consultation et d'entrée de données. Elle est nourrie par l'ensemble des chercheurs intéressés au projet et sert de lien entre plusieurs équipes rassemblées dans un réseau de recherche dont le but est l'étude des modalités de la transmission des textes par les humanistes. Ce réseau compte, outre les sections de l'IRHT concernées, douze équipes ». Cette base de connaissances collaborative est construite sur le modèle des bases de connaissance développées dans les entreprises privées, et elle dépasse le cadre d'une petite équipe de recherche locale.

4. Conclusion

A la lumière des différents articles utilisés ici, il est possible de situer la place de l'utilisateur dans la recherche sur les bibliothèques numériques : il est présent, mais de deux façons différentes.

Une première méthode de travail, souvent appliquée aux sites et postes de travail destinés aux spécialistes, en collaboration avec les chercheurs eux-mêmes, place l'étude des attentes et besoins de l'utilisateur en amont de la conception des interfaces. Cette approche, pour intéressante qu'elle soit, reste assez théorique, et comporte le risque d'aboutir à des interfaces très développées dont les fonctionnalités ne seront que peu utilisées. Les corpus concernés par ces numérisations sont relativement restreints, et obéissent à des politiques documentaires très spécifiques.

A l'inverse, une approche plus pragmatique est le fait des établissements de conservation, soit des bibliothèques traditionnelles, dont la mission et la préoccupation est de diffuser le plus largement possible leurs collections. Les interfaces sont alors plus simples, et les études d'usages suivent, souvent de plusieurs années, le lancement sur l'Internet : les retours des lecteurs et résultats d'enquêtes sont alors la base du travail mené sur les interfaces de consultation.

La synthèse des résultats d'enquêtes proposée ici ne fait d'ailleurs apparaître que des usages de fonctionnalités relativement simples : recherche, feuilletage, impression. Il faut souligner que ces résultats ne concernent pas uniquement les chercheurs, mais que la mise sur l'Internet de reproductions de documents anciens accroît le nombre de lecteurs, et diversifie leurs origines. Il ne faudrait donc pas conclure ici que les fonctions développées sur des postes de travail spécifiques, comme les annotations par exemple, ne sont pas utilisées. Seules des enquêtes d'usages plus précises, et locales, pourraient nous renseigner sur ce point, mais les postes destinés aux chercheurs n'ont pas fait l'objet d'enquêtes a posteriori : un champ d'étude reste ouvert.

Partie 3 : Tableaux statistiques et analyse des sites de bibliothèques numériques

1. Analyse des sites visités

Les différentes parties des 32 tableaux analytiques des sites sélectionnés présentés en annexe 3 (p. XIII-XLVII) sont reprises pour en dégager des conclusions, afin de comprendre comment *DEBORA* se situe dans les perspectives des bibliothèques numériques d'aujourd'hui.

1.1. Les éditeurs et la date de création des sites

Editeur	Bibliothèque	12
	Université	29
	Autre institution publique, ou entreprise privée	8
	Particulier(s)	3
Date du site	1990-1995	6
	1996-1999	13
	2000-2002	5
	Non précisée	8

La date de création relevée sur les sites indique une majorité de bibliothèques numériques créées ou mises en ligne entre 1996 et 1999.

Les sites peuvent être créés par plusieurs institutions, ce qui explique le total de 46 éditeurs pour 32 sites. Les universités sont les plus nombreuses car elles travaillent plus volontiers en association que les bibliothèques, par exemple. Celles-ci semblent se cantonner à la numérisation de leurs fonds propres pour les valoriser, tandis que les bibliothèques numériques constituées par les universités

présenteraient plus de fonctionnalités de recherche à destination de leurs chercheurs.

Les entreprises privées ayant contribué à la numérisation de livres apparaissent parfois sur les sites au même titre que les bibliothèques, et ont donc été considérées dans ce cas-là comme co-éditeurs. Toutefois, les éditeurs privés de véritables bibliothèques numériques sont rares. Deux sites privés observés ont présenté de véritables originalités et témoignent d'une réflexion très approfondie (dont *Athena e-textes* et *Studiolum*). Il s'agit plutôt de documents numérisés en mode texte et constituant une collection virtuelle.

Athena e-textes, développé par Pierre Perroud, cherche surtout à donner un maximum de textes et de liens sur un auteur ou un thème. Il ne s'intéresse pas au travail sur écran, à l'annotation, etc, mais veut surtout diffuser des œuvres en ligne, afin qu'elles soient imprimées et divulguées.

Les perspectives de *Studiolum* et de son créateur Tamàs Sajo sont différentes : il s'agit de permettre vraiment un travail sur le texte et sur l'écran, grâce à de nombreuses fonctionnalités de traitement de texte, d'annotations, d'affichage, etc. Une description plus complète est présentée en 3.8.6.

1.2. Typologie des sites

Type	Inscription préalable	2
	Accès restreint ou payant	4
	Fonds anciens numérisés	22

L'inscription préalable et gratuite, constatée sur les sites d'*ILEJ* et du *Dunhuang Project*, sert notamment à étudier le public par le biais de l'enquête qui y est associée, car ces sites concernent des collections spécialisées relativement ciblées. L'accès aux sites peut être restreint, au bénéfice des membres de l'université du Michigan par exemple (*EEBO/TCP*), ou payant, lorsque l'un des éditeurs est une entreprise privée (Proquest pour *EEBO*). Une partie du site reste accessible à tous pour servir de vitrine.

Nombre de données (pages ou folios)	1-100	12
	100-1000	7
	1000 et +	9
	Non précisé	4

Le nombre de documents annoncé peut être trompeur, selon que l'on comptabilise les livres entiers numérisés, ou simplement les pages ou les documents d'archives. D'autre part, le nombre de documents annoncés à terme n'a parfois aucun rapport avec la réalité : *EEBO/TCP* annonce ainsi 25 000 documents alors qu'il n'en propose que huit en consultation.

Les bibliothèques numériques importantes sont logiquement les plus anciennes et celles constituées par les bibliothèques nationales.

La Bibliothèque du Congrès regroupe dans *American Memory* plus de 100 collections thématiques correspondant à 7 millions de documents. Ce total extraordinaire (s'il ne résulte pas d'une projection) n'a pu être atteint que grâce à la politique de coopération menée par la Bibliothèque du Congrès en direction des bibliothèques américaines, dans le but de numériser les collections les plus intéressantes pour l'histoire des Etats-Unis. L'exemple ici retenu est celui des archives d'histoire religieuse (*Early religious petitions*) de la bibliothèque de l'Etat de Virginie (423 documents). D'autre part, les initiatives indépendantes de la Bibliothèque du Congrès sont fédérées par celle-ci au travers de son portail, qui recense ainsi au sein de ses collections le *Making of America* de l'université de Cornell (267 livres et 100 000 articles de journaux), alors que le site de celle-ci ne renvoie pas à celui de la Bibliothèque du Congrès. Tous ces sites sont ainsi disponibles via un seul portail d'accès, qui classe ces collections de bibliothèques numériques au sein d'une seule grande bibliothèque virtuelle, en proposant des fonctionnalités de recherche pour orienter le lecteur.

Mode de numérisation	Image	17
	Image et texte	11
	Texte	4
Couleurs	Oui	19
	Non	9

La numérisation en mode image et en couleurs concerne surtout les fonds anciens et les ouvrages illustrés. L'inconvénient de ce mode de numérisation est qu'il produit des images très lourdes, ce qui explique le choix du noir et blanc pour les collections imprimées qui s'y prêtent, c'est-à-dire les collections imprimées postérieures au XIXe siècle. Ce dernier mode de numérisation reste le moyen le moins coûteux et le plus rapide de numériser les fonds d'une bibliothèque importante, par exemple les revues savantes du XIXe siècle pour la BIUM de Paris, dans *Medic@*.

Le simple mode texte reste limité aux sites créés par des particuliers ou de petites bibliothèques, et à ceux dédiés à l'étude de la littérature ou de la philologie. Les particuliers sont parfois des universitaires (*Eliohs*), qui saisissent eux-mêmes les textes pour une question de coût, et ne proposent pas de fonctionnalités de recherche particulière. Dans cette logique, la bibliothèque de Lisieux a décidé de numériser ses fonds en faisant saisir les livres par des bénévoles. Le site d'*Eliohs* propose ainsi ses propres textes, ainsi que des liens vers des titres d'autres bibliothèques électroniques, son but étant de proposer une sélection cohérente d'ouvrages sur l'histoire des idées en Europe ; les titres sont classés par siècle et par ordre alphabétique, et c'est le seul mode d'accès au document possible. De nombreuses bibliothèques virtuelles dédiées à un domaine d'étude particulier utilisent au contraire le mode texte pour les fonctionnalités de recherche d'occurrences qu'il permet : l'exemple retenu ici est celui de *Studiolum*, qui sera développé plus bas, mais on peut citer également le *projet ARTFL*⁵⁸ qui s'appuie sur un corpus de textes déjà numérisés par le CNRS.

⁵⁸ Voir la liste des sites visités et non retenus en annexe 2.

La conjonction de l'image et du texte est rendue possible grâce à la reconnaissance optique de caractères (OCR), qui s'applique plus facilement aux ouvrages imprimés avec des caractères typographiques normalisés (surtout à partir du XIXe siècle). Le mode texte ainsi obtenu permet la recherche des mots du texte, ce qui rejoint le souhait général des chercheurs. Ceci explique qu'on retrouve cette configuration dans les bibliothèques numériques des universités ou destinées en priorité aux chercheurs. Des tentatives de transcription automatique de textes imprimés modernes ou mêmes de manuscrits médiévaux existent, mais elles ne sont pas encore disponibles sur les sites visités. De même, la bibliothèque *Menasseh Ben Israel*, qui propose des manuscrits imprimés en hébreu, projette de mettre au point un système d'OCR spécialement adapté.

On peut néanmoins citer le site *ILEJ*, qui propose la recherche en plein texte sur les images des pages de journaux du XIXe siècle, mais aussi du XVIIIe siècle, ce qui est peu fréquent. Mieux encore, le site *The William Blake Archives* est original en ce qu'il propose une numérisation en mode image et une transcription d'excellente qualité, avec des fonctionnalités de recherches et des liens entre l'image et la transcription. Ce type de travail n'est possible que sur un corpus bien défini et limité.

Enrichie	Oui	21
	Non	11

La majorité des bibliothèques retenues ici ajoute des informations aux notices bibliographiques des livres numérisés : du point de vue scientifique, les enrichissements les plus intéressants sont les liens vers une bibliographie, ou mieux, vers des articles sur l'ouvrage ou le sujet, comme dans le site *Menasseh Ben Israel*. Celui-ci propose également des documents d'archives concernant l'imprimeur des livres du site, qui est également l'auteur de plusieurs d'entre eux. On peut citer *The William Blake Archives*, *Athena e-textes*, *Galileo Galilei* ou bien encore les sites de livres d'emblèmes dont l'objectif est de développer un véritable appareillage critique autour des œuvres numérisées.

Niveau de description des métadonnées	1	2
	2	12
	3	10
	4	6

Le niveau de description des métadonnées est très divers. Le niveau 4, le plus faible, ne comporte souvent que le nom de l'auteur, le titre de l'œuvre et parfois, la date d'édition et l'éditeur. C'est le cas des sites de livres d'emblèmes, du site personnel *Athena e-textes*.

Les bibliothèques ou les universités présentent en général des notices catalographiques et un appareillage critique de niveau 2 ou 3. Il existe alors une volonté de décrire, de développer une notice claire et précise mais il n'y a aucun effort de réflexion sur les métadonnées.

Rares sont les sites qui suivent les règles de la TEI ou même participent à son élaboration. L'équipe du *Digital Scriptorium* mène une réflexion particulièrement développée sur la TEI des manuscrits. Elle a conçu, en septembre 2001 un « Data dictionary, version 5 »⁵⁹ de grande qualité. Le but est de collecter et organiser toutes les informations fournies sur les manuscrits par tous les partenaires du *Digital Scriptorium* (à Columbia et à Berkeley).

Comme le projet *DEBORA*, *Digital Scriptorium* organise les métadonnées sur 4 niveaux :

- Niveau 1 : l'identification du manuscrit (la ville, l'institution de conservation ou la bibliothèque, la cote, la provenance, le nombre de folios, la reliure, l'état matériel du document, la bibliographie, le mode de reproduction, etc)
- Niveau 2 : les informations « physiques » ou « part » (le nombres de parties, le support, le filigrane, les dimensions, les renseignements géographiques du manuscrit, les datations, l'alphabet, l'écriture, le nombre et le nom des scribes, les décorations et artistes, etc)

⁵⁹ Consultable sur <http://sunsite.berkeley.edu/Scriptorium/datadic5.html>

- Niveau 3 : les informations bibliographiques concernant le texte (auteur(s), titre, langue utilisée, historique du texte, incipit, explicit, manques éventuels, etc)
- Niveau 4 : les information sur l'image (titre de l'image, indexation iconographique iconclass, notes à destination du photographe)

Cette distinction par niveau de métadonnées est assez comparable à celle de *DEBORA*, du moins pour les 3 premiers niveaux rappelés ici : niveau 1 : le manuscrit dans la collection ; niveau 2 : la structure du livre ; niveau 3 : la page. Le projet *DEBORA* prévoit un quatrième niveau qui comprendrait les informations intégrées dans la base soit par reconnaissance automatique des caractères ainsi que les données collaboratives.

Il faut signaler enfin l'extraordinaire richesse des notices présentées par le site *The William Blake Archives*. Chaque exemplaire (« copy ») d'une même œuvre est décrite, afin de reconstituer l'historique du document (provenance, passages dans les ventes aux enchères, prix de vente), et d'en donner sa description bibliophilique et bibliothéconomique.

Format	Png	2
	Tiff	5
	Jpeg	20
	Gif	14
Logiciels	Adobe acrobat reader	5
	Photoshop	1
	Logiciels spécifiques : Mapedit, Folio Views, Ebrary, Imaging, Inote	5

Dans la mesure où les sites de bibliothèques numériques sont consultables via l'Internet, les formats de numérisation d'images restent traditionnels : Jpeg dans la majorité, Gif, Tiff et en dernier lieu Png. Png (Portable Network Graphics) a été créé pour améliorer Gif et Tiff. Il présente des avantages notamment la fonctionnalité d'image « high resolution ».

Les pages numérisées de livres anciens peuvent faire l'objet d'un traitement de l'image banal, grâce à un logiciel de graphisme comme Adobe Photoshop, proposé par la Bibliothèque du Congrès. Il devient alors possible d'inscrire des notes dans le corps du texte, comme pour *DEBORA*. Un autre logiciel, Imaging de Kodak, présente des fonctionnalités graphiques analogues à Photoshop, mais aussi des possibilités d'annotation spécifiques pour les livres numériques, (*Menasseh Ben Israel*). Cependant, les notes ainsi produites restent strictement personnelles et ne peuvent être indexées, à la différence de *DEBORA*.

Certaines bibliothèques numériques utilisent des logiciels qui doivent être téléchargés lors de la consultation : Adobe Acrobat Reader (*Christian Classics Ethereal Library, Gallica, Canadiana, EEBO, EEBO/TCP*), Adobe Photoshop (*American Memory*), Mapedit (*The English Emblem Book Project*), Folio Views (*Studiolum*), Ebrary (*Internet Public Library*), Imaging de Kodak (*Menasseh Ben Israel*). Ces logiciels, qui font pour certains l'objet d'une exploitation commerciale, apportent des possibilités de travail sur l'image et de recherche sur le texte ou sur l'image.

1.3. Les fonctionnalités proposées

1.3.1 Outils de visualisation

Vignettage	15
Mosaïque	2
Affich. annotations	9
Fonds	1
Luminosité	2
Grossissement	24
Autres (liens internes)	Main : 2 transcription : 3 rotation de l'image : 3

En général, seuls le vignettage et le grossissement sont possibles.

Le vignettage permet la navigation au sein du livre ; il est généralement l'alternative à la présentation de la structure du livre sous forme de liens vers les pages, sauf pour *DEBORA* qui laisse le choix à son lecteur entre les deux.

On remarque que l’affichage des annotations (ou métadonnées) est assez souvent permis, dans un cadre à gauche du document numérisé, au début du texte ou encore dans le texte lui-même grâce à la technique de l’hypertexte.

Les principales fonctionnalités des logiciels de lecture sont le grossissement (Adobe, Ebrary), la rotation de l’image (Adobe), l’affichage de la table des matières (Adobe, Ebrary), l’impression payante (Ebrary). Le grossissement est possible de différentes manières selon les logiciels : loupe, icône, pourcentage, simple clic sur l’image.

Mapedit, implementé par *The English Emblem book project*, permet de relier une image (un emblème) et le texte correspondant. On peut donc combiner les affichages : texte seul, image seul, texte et image.

Quant à Folio Views (*Studiolum*) et à Imaging (*Menasseh Ben Israel*), ils seront décrits plus tard.

1.3.2 Recherche

Titre	22	
Plein texte	15	
Illustrations	9	
Autres	Auteur	10
	Sujet	12
	Editeur	1
	Cote	4
	Occurrences	2
	Limitations des recherches	4

Les fonctionnalités de recherche sur le texte restent classiques : titre, auteur, plus rarement sujet. Cependant, les bibliothèques numériques semblent très attachées à développer la recherche en texte intégral, puisque 14 d’entre elles la proposent. Certains sites permettent des recherches particulièrement développées : c’est le cas du *Digital Scriptorium* (cote, reliure, provenance, pays, ville, titre, incipit, scribe, décoration, artiste, date, etc.), du *William Blake Archives*, de *Studiolum* (« Query », « Advanced Query », « Query in two fields »). Seuls les sites

présentant un corpus défini et limité peuvent prétendre à créer des outils de recherche performants.

La recherche sur les images est très variée. La plus simple consiste à présenter les images en vignettes et à cliquer sur celle qui intéresse. La plus élaborée est celle permise par des systèmes d'indexation iconographique du type Iconclass. Le site *William Blake Archives* indexe toutes ses images selon une classification très précise, puis présente une description des images dans laquelle sont repris les mots clés. Cela permet donc une recherche sur l'image par mots-clés, dont les résultats sont satisfaisants.

Au total, on constate que peu de sites proposent des fonctionnalités développées. La priorité des bibliothèques numériques reste avant tout la numérisation de nouveaux documents, l'augmentation du corpus et non le travail sur le texte ou l'amélioration des outils de visualisation.

1.3.3 Travail personnel sur le document

Corpus personnel	4
Fichier personnel	5
Impression	27 Limitée ou payante : 2
Autres	

Le travail personnel est très peu pris en compte : les préoccupations des bibliothèques numériques sont surtout de régler les problèmes de conservation des documents anciens et de communiquer au grand public des ouvrages difficilement consultables. Aussi, elles permettent dans leur grande majorité l'impression des pages et le résultat est généralement de bonne qualité. C'est particulièrement vrai pour les bibliothèques qui utilisent le logiciel Acrobat reader, comme *Gallica*.

Quelques exceptions importantes : *Studiolum*, *William Blake Archives* et *Menasseh Ben Israel* (bien que le logiciel Imaging ne soit utilisable que pour un seul livre pour l'instant). C'est la raison pour laquelle ces trois sites seront étudiés à part.

1.4. Travail collaboratif et documents associés

Messagerie	15
Newsgroup	4
Banque de données	1
Autres	

Les fonctionnalités de travail collaboratif restent très classiques et limitées à la messagerie et parfois à des newsgroups. Des rubriques, du type « users voice », ou « users comment » sont parfois proposées. Les webmasters attendent plus des commentaires critiques ou des suggestions sur la base de données en général et non des corrections sur les documents, les métadonnées, etc.

Le seul site qui permette l'enrichissement collaboratif de sa base de données est le *Dunhuang Project* de la British Library. Il s'adresse à des spécialistes des manuscrits chinois, et la collaboration est réservée aux chercheurs accrédités. Ceux-ci doivent d'abord prendre contact par messagerie, pour faire parvenir leurs commentaires et surtout leurs corrections concernant les problèmes de translittération du chinois à l'anglais (le site est bilingue). Si leur niveau scientifique est jugé valable par le responsable du site (le conservateur du département des collections orientales), le chercheur se voit proposer une accréditation pour pouvoir intervenir sur la base directement. D'autre part, l'inscription, obligatoire et gratuite pour pouvoir consulter la base, s'assortit d'une enquête sur le profil universitaire des usagers. Ceci permet aux administrateurs du site de sélectionner les chercheurs en fonction de leurs domaines de recherches, et de les faire bénéficier d'un dossier personnel pour travailler sur écran pendant leur visite du site, avec la possibilité de garder en mémoire leurs précédentes recherches.

L'équipe du *Digital Scriptorium* mène également une réflexion intéressante sur le travail collaboratif dans le cadre de la rédaction de la TEI. *Digital Scriptorium*, dans son « Data dictionary » présenté plus haut, propose, à chacun des 4 niveaux de description des métadonnées (« Manuscrit », « Part », « Text » et « Image »),

trois champs particulièrement intéressants dans une perspective de travail collaboratif.

Le premier champ « Acknowledgments » permet d'inscrire et de remercier les personnes qui ont contribué au travail critique autour du manuscrit. Ce champ veut promouvoir « le dialogue entre les chercheurs qui sont spécialistes dans un domaine précis et les bibliothécaires qui sont responsables auprès du public de la propagation des connaissances des manuscrits »⁶⁰. Ce champ permet également de dater l'information contenue dans la base de données.

Le second champ « Source », définit les sources des informations contenues dans la base de données. Ce champ n'est pas destiné à l'utilisateur final mais à une « in-house verification » :

Enfin le champ « Inputter » contient le nom de la personne responsable de la sélection et de l'organisation de l'information sur le manuscrit. Cette personne peut être différente en fonction du manuscrit bien sûr mais également des différents niveaux des métadonnées.

Le travail collaboratif induit par ces trois champs ne passe pas par la constitution d'une base de données collaborative, comme l'imagine l'équipe de *DEBORA*. Il s'agit d'une collaboration qui passe d'abord par les canaux traditionnels : le dialogue, la messagerie, le travail entre chercheurs. Ensuite elle s'inscrit dans les métadonnées de façon explicite afin de promouvoir une connaissance des personnes, une facilité de correction et de datation de l'information. Un médiateur existe, c'est celui dont le nom est inscrit dans le champ « Inputter » ; cependant son travail peut être divisé entre plusieurs personnes selon les manuscrits. Cela évite la surcharge de travail.

Enquêtes	2
Liens	19
Webmaster	23
Autres	2 (lettres en ligne), 1 : résultat de la recherche sur les OCR

⁶⁰ Disponible sur Internet <http://sunsite.berkeley.edu/Scriptorium/datadic5.html> (consulté le 12.04.2002)

Les deux enquêtes rencontrées portent sur le profil du public ; les sites d'*ILEJ* et du *Dunhuang Project* demandent de remplir un formulaire d'inscription comme dans une bibliothèque de recherche traditionnelle. Ces sites concernent en effet des collections spécialisées s'adressant en priorité aux chercheurs, et en second lieu aux étudiants.

Aucun site ne présente d'enquête sur les usages ; toutefois, l'institution responsable d'une bibliothèque numérique peut le faire localement, comme c'est le cas de l'*ILEJ* ou encore de *Gallica*.

2. Description approfondie de Menasseh Ben Israel, Studiolum et The William Blake Archives

Trois sites ont été sélectionnés parce qu'ils envisagent une bibliothèque numérique d'une façon proche de celle de *DEBORA* : ils développent ou utilisent des logiciels de lecture et de travail sur le texte ou l'image, avec des fonctionnalités d'annotations et de recherche notamment. *Menasseh Ben Israel*, qui propose des documents en mode image, utilise un logiciel de travail de l'image tandis que *Studiolum*, qui propose du texte, s'est concentré sur la recherche. Le *William Blake Archives* a choisi lui d'allier le mode image et le mode texte avec une transcription de grande qualité.

2.1. Menasseh Ben Israel

A l'Université d'Amsterdam la Bibliotheca Rosenthaliana a acquis une collection de livres imprimés par *Menasseh Ben Israel* au XVII^e siècle. La collection comprend 70 éditions sur les 75 existantes, ainsi que six lettres autographes de l'imprimeur, également numérisées. La numérisation d'ouvrages de commentaires sur cet imprimeur et son oeuvre est en projet, et une bibliographie est actuellement disponible.

Les ouvrages de la collection sont numérisés en mode image, en format GIF, sans autre fonctionnalité que le feuilletage. Ces images sont en noir et blanc, et doivent être téléchargées une par une. Le noir et blanc a dû être choisi pour diminuer les

temps de chargement : des vues miniatures en couleurs sont disponibles pour les pages de titre. Deux niveaux de description bibliographique coexistent : une description brève incluant titre, auteur, éditeur, date d'édition et cote au minimum, et une description longue incluant des données d'exemplaires, telles que les éventuelles dédicaces.

Un livre intitulé *Minhagim* est également téléchargeable avec le logiciel Imaging, développé de 1995 à 1999 par Kodak. Ce logiciel propose de nombreuses fonctionnalités de travail sur l'image.

Visualisation :

- Interface graphique avec barre de menu, barre de tâche, boutons de fonctions
- possibilité d'afficher en vignettage, ou avec deux cadres : un reprenant les vignettes, et la page en cours à droite
- zoom réglable
- rotation de l'image
- impression
- choix entre différents niveaux de numérisation : noir et blanc, niveaux de gris, couleurs. Cette fonctionnalité n'était pas active, la page affichée étant en noir et blanc.

Travail sur l'image :

- dessin : ligne à main levée, marqueur, ligne droite et encadrement vide ou plein, avec choix de la couleur utilisée
- annotations : note sur l'image, post-it sur l'image, lien vers fichier texte externe : annotation externe, fusion des annotations dans la page, ce qui permet l'exportation du fichier avec les annotations vers un logiciel qui ne peut pas les gérer séparément
- cachets : peut servir d'annotation ; différents modèles sont proposés : 'approuvé', 'brouillon', 'reçu', 'refusé', la date étant automatiquement insérée ; il est également possible de créer ses propres cachets

Il n'y a aucune fonctionnalité de recherche sur la page, ni d'indexation des annotations, ni d'indications sur la structure du livre ou de feuilletage, que le chargement image par image rend impossible.

Ce logiciel présente l'intérêt d'une grande liberté d'annotations sur l'image, et l'envoi des images annotées, ce qui semble être une piste vers le travail collaboratif.

2.2. Studiolum

Le but de *Studiolum* est de recréer l'espace intime et propice à la recherche qu'était la bibliothèque d'un humaniste : des œuvres importantes de la Renaissance sont numérisées et diffusées sur CD-Rom. Seuls quelques ouvrages sont disponibles via l'Internet à titre d'exemple. *Studiolum* numérise en mode texte des livres d'emblèmes, des bestiaires, des herbiers, des traités d'art et d'antiquité, des géographies, des collections de proverbes, des monnaies anciennes, des dictionnaires, des miroirs de princes, des lexiques, etc. Dans le cas d'œuvres en latin et en grec, sont intégrées les translations en anglais, français, italien ou allemand éditées aux XVI^e et XVII^e siècles.

Les nouvelles versions du CD-Rom *Studiolum* sont publiées tous les 6 mois.

Visualisation :

- Logiciel Folio Views 4.2 développé par OpenMarket de 1992 à 1999 pour Windows.
- Interface graphique avec barre de menu, barre de tâche, boutons de fonctions, zone d'affichage avec 7 onglets : all, search, browse, document, contents, hitlist, object.
 - o All : affiche 4 zones d'affichage simultanées : liste des occurrences, texte, titre du chapitre courant, sommaire des titres des chapitres permettant la navigation
 - o Search : 3 zones : titre du chapitre courant, texte avec terme recherché en surbrillance, liste des lignes comportant des occurrences en surbrillance, classées par chapitre. D'une session sur l'autre, les recherches restent en mémoire.

- o Browse : 3 zones d'affichage simultanées : texte, titre du chapitre courant, sommaire des titres des chapitres permettant la navigation
- o Document : 2 zones : texte et titre courant
- o Contents : 1 zone : sommaire des titres de chapitres en liens hypertextes
- o Hitlist : 1 zone : liste des lignes comportant des occurrences en surbrillance, classées par chapitres
- o Object : affichage séparé des images
- possibilité d'afficher en mosaïque horizontale ou verticale, ou en vignettage, plusieurs fichiers.
- Raccourci vers un titre sur le bureau
- zoom réglable
- boutons de défilement
- impression

Fonctionnalités de recherche :

- Dans le menu « search », plusieurs possibilités de recherche :
 - o le bouton « query » : recherche en texte intégral
 - o le bouton « advanced query » : recherche par choix du terme à rechercher dans la liste des mots indexés, possibilité de préciser la partie du texte à explorer
 - o « query in one field » ou « query in two fields » permettent de lancer des recherches sur un ou plusieurs champs : le texte intégral, les différents de descriptions (notes, index, errata, titres de chapitre, titre de livre, texte grec, texte latin, n° chapitre, n° page, etc), qui sont différents selon les documents ouverts. La saisie des termes à rechercher est contrôlée, ou choix dans l'index. Possibilité de rechercher un mot pour le corriger : le terme trouvé est automatiquement barré, le curseur est positionné à côté.

Les termes recherchés trouvés dans le texte sont mis en surbrillance, un bouton de défilement permet de passer d'occurrence en occurrence. Une fenêtre « hit reference » s'ouvre en bas de l'écran, et affiche les lignes comportant le terme recherché.

- Un menu déroulant placé en bas de l'écran rappelle les recherches déjà effectuées. Le bouton « clear query » permet d'effacer les surbrillances. Un historique des recherches effectuées est disponible.

Fichier caché (ou personnel) :

- Même interface que celle de consultation, mais toutes les fonctionnalités de travail et d'annotations sont actives.
- possibilité de modifier le texte lui-même, la mise en page (supprimer les illustrations, ajouter des sauts de paragraphe), et d'enregistrer les modifications.
- fonctionnalités de traitement de texte (dont style)
- Post-it, surligneur (couleur, police), marque-page (cible dans le texte)
- modification de la structure du livre (ajout de titres de chapitre)
- Possibilité d'établir des liens hypertextes : au sein du document, vers des pages web externes, vers des « objets » (image, tableau), vers un programme informatique, vers des notes (fenêtre masquée, que l'on a nommée, avec le logiciel Popup), vers une recherche prédéfinie
- Insertion de tableau, icônes de mise en page spécifiques (fractionner/fusionner)

Le fichier caché s'adapte aux nouvelles versions de *Studiolum*.

Les fonctionnalités de recherche, d'annotation et de création de liens de *Studiolum* sont très développées. Comme elles ne concernent pas des images mais du texte, elles permettent un travail collaboratif, comme l'explique l'auteur du site :

« Les fichiers cachés sont spécialement utiles si plusieurs personnes travaillent avec le même CD-ROM, ou pour rassembler différents commentaires de plusieurs fichiers cachés. »

Par exemple, l'auteur recommande de créer des fichiers cachés de commentaires (« feedback shadow files ») pour les envoyer régulièrement à *Studiolum*.

2.3. The William Blake Archives

Ce site est édité par des chercheurs de trois universités américaines : Rochester, Californie et Caroline du Nord ; il est notamment sponsorisé par la Bibliothèque du Congrès et l'Institute for Advanced Technology in the Humanities, de l'Université de Virginie. Il propose en accès libre sur l'Internet un grand nombre d'œuvres de William Blake, ainsi qu'un appareillage critique sur l'artiste : son but est de créer un environnement hypermédia d'une grande qualité tant scientifique que technique. La particularité de ce site est de lier les images et le texte, pour que l'utilisateur puisse passer de l'un à l'autre. Aux images numérisées des ouvrages sont donc liées des transcriptions des textes.

Visualisation :

- Accès par un logiciel de navigation Internet
- Interface graphique avec image ou transcription au centre, et diverses fonctionnalités :
 - Feuilletage
 - Affichage des données d'exemplaire
 - Informations techniques sur le fichier
 - Zoom
 - Affichage des différents exemplaires du même texte, pour comparaison (exemplaires codés par des lettres)
 - Raccourci vers les index
 - Menu déroulant proposant des liens vers :
 - o l'image haute résolution
 - o une description détaillée de l'image
 - o la transcription du texte de la page
 - o les notes ajoutées par l'éditeur du site

Possibilité de recherche :

Via l'icône « search », plusieurs possibilités : recherche de texte ou recherche d'image.

- La recherche d'image :
 - o Restriction aux ouvrages ou à tous les documents
 - o Recherche boudéenne (and, or)
 - o Recherche par catégorie : 'figure, types, characters, postures, attributes, animal, vegetation, structure, objet'. Ces catégories sont détaillées dans des listes de termes indexés.
- La recherche dans le texte :
 - o Restriction aux ouvrages ou à tous les documents
 - o Recherche boudéenne (and, or, keyword, phrase, proximité)
 - o Trois types de recherche : dans les transcriptions, dans les mots des titres, dans les descriptions des images. Ce dernier type de recherche permet donc indirectement de rechercher par mots clés dans les images.
 - o les termes recherchés trouvés dans le texte apparaissent en couleur
- une icône permet de supprimer la dernière recherche

Travail sur le texte et l'image :

- Fonctions d'éditations (possibilité de copier/coller dans un traitement de texte)
- Impression (qualité réduite)
- Logiciel d'annotation d'image Inote, qui permet l'affichage des notes des éditeurs, la saisie et l'enregistrement de notes sur un fichier local. Ce logiciel est développé depuis 1998 par l'Université de Virginie. Certaines fonctionnalités n'étaient pas accessibles lors des consultations du site.⁶¹

⁶¹ Voir le site du constructeur disponible sur Internet <<http://www.iath.virginia.edu/inote/>> (consulté le 01.06.2002)

3. Synthèse de l'analyse des sites

3.1. Classement des sites visités

Le tableau suivant classe les bibliothèques numériques selon les deux critères évoqués dans la méthodologie de la recherche : le mode de numérisation (mode image ou mode texte) et le statut de la collection (collection existante ou collection virtuelle).

Il faut souligner que les sites qui proposent une numérisation en mode texte peuvent proposer des illustrations, mais ces dernières demeurent marginales, comme dans *Studiolum*. De même, les sites qui proposent des documents numérisés en mode texte et en mode image ont dû procéder à deux numérisations successives la plupart du temps, une par scanner et une par saisie manuelle, comme pour le site *The William Blake Archives* ; une véritable transcription automatique grâce aux OCR ne fonctionne que sur le site d'*ILEJ*, pour des journaux des XVIII^e et XIX^e siècles. Dans tous les cas, lorsque les deux modes images et texte sont de même importance pour le site, on a considéré le mode image comme plus important, car dans le projet *DEBORA*, la transcription découle des images.

	Mode Image	Mode Texte
Collection existante	<i>American Memory, Gallica, Alciato, Dunhuang Project, EEBO/TCP, Early Manuscripts at Oxford University, Early Virginia Religious Petitions, Emblem Project Utrecht, The English Emblem Book project, The Universal Library-Antique books, Galileo Galilei, HUMI Project, Making of America, Medic@, Menasseh Ben Israel, Biblioteca digital Dioscórides, Mateo, Mnemosyne</i>	<i>Lisieux</i> ⁶²
Collection virtuelle	<i>Digital Scriptorium, Early English Books Online, Electronic text center, Internet Public Library, Christian Classics Ethereal library, Early canadiana on-line, Projet Charrette, ILEJ, William Blake Archives</i>	<i>Perseus, Athena e-textes, Eliohs, Studiolum</i>

⁶² Voir la liste des sites visités et non retenus en annexe 2.

Le déséquilibre constaté dans le tableau s'explique en partie par le choix initial du corpus des bibliothèques numériques : on a retenu en priorité celles traitant du livre ancien, pour pouvoir les comparer avec *DEBORA*. Seule la bibliothèque numérique de Lisieux a utilisé la numérisation en mode texte pour une collection existante, mais n'a pas été retenue pour cette raison dans les tableaux d'analyse. De manière générale, on a écarté les bibliothèques numériques en mode texte, sauf lorsqu'elles concernaient le livre ancien, ou qu'elles proposaient des fonctionnalités de recherche intéressantes.

3.2. Conclusions sur l'analyse des sites

Les bibliothèques numériques proposant des fonctionnalités de recherche et de visualisation développées concernent autant, dans le tableau ci-dessus, les collections existantes (*Dunhuang Project, Menasseh Ben Israel*) que des collections virtuelles (*Digital Scriptorium, William Blake Archives*). Dans tous les cas, il s'agit de collections spécialisées destinées à un public ciblé de chercheurs.

Ces bibliothèques numériques ont été développées par des universités en liaison avec des chercheurs (*Digital Scriptorium, Menasseh Ben Israel, William Blake Archives, Perseus*). *Studiolum* est un cas à part car il a été créé par un particulier, Tamás Sajo, mais celui-ci est également professeur à l'Université Centrale de Budapest (Hongrie). Ces bibliothèques ont donc été créées par et pour des chercheurs. Par conséquent, il faut imaginer que les attentes de ces derniers ont été prises en compte dans la conception des logiciels, par des enquêtes préalables sur leurs besoins ou collaboration lors de la conception, comme indiqué dans la deuxième partie. Aucun des sites visités ne présentait une enquête d'usage ou du moins des résultats. On se situe donc plus dans le cas d'une politique de l'offre, mais d'une offre très ciblée, conçue pour un public connu à priori.

Ce public de chercheurs est donc la principale cible des bibliothèques numériques comparables à *DEBORA*, mais il faut noter qu'ils sont également impliqués dans leur gestion, notamment pour l'évaluation scientifique des commentaires qui peuvent être envoyés au responsable.

L'effort de développement des fonctionnalités de recherche et de visualisation porte essentiellement sur le travail sur écran et sur la recherche dans le texte

numérique. L'effort en direction de l'indexation des ouvrages peut être parfois très poussé.

Les sites conçus et développés par de grandes bibliothèques nationales sont de deux sortes : soit des portails améliorés, soit des bibliothèques numériques ne proposant que très peu de fonctionnalités. Le public visé est alors le grand public, moins exigeant que les chercheurs. Pour ces bibliothèques, le travail sur écran n'est pas une priorité. Elles offrent en compensation la possibilité d'imprimer les pages des livres, voire de télécharger un livre complet. Dans le meilleur des cas, l'effort porte sur la qualité de la numérisation et de l'impression. Le format PDF, proposé par Adobe Acrobat Reader permet cette qualité d'impression. Il n'est donc pas étonnant qu'il soit proposé par *Gallica* et par *Early Canadiana Online*.

Les fonctionnalités proposées sont donc étroitement liées au public destinataire de la bibliothèque numérique, et ne sont développées que lorsque le site vise un public de chercheurs bien identifié, correspondant à un domaine d'études bien circonscrit.

L'étude de ces fonctionnalités, au travers de l'exemple des trois sites les plus intéressants, permet de dégager des conclusions sur les usages qu'ils induisent chez les chercheurs concernés. Ces usages induits sont à comparer avec ceux que serait susceptible de générer le logiciel *DEBORA*.

La recherche sur le texte lui-même est une des possibilités les plus développées dans les sites concernant une collection spécialisée. Lorsqu'il s'agit de fonctionnalités de recherche sur des documents anciens, la numérisation en mode texte est encore incontournable. *Studiolum*, à partir d'un texte saisi manuellement, propose ainsi les possibilités de recherche les plus complètes, de même que *The William Blake Archives*, qui propose pourtant les pages des livres en mode image également. Dans ce cas précis, toute transcription automatique analogue à *DEBORA* est rendue impossible par le fait que Blake gravait lui-même son texte et ses illustrations en lithogravure. La bibliothèque numérique de *Menasseh Ben Israel*, exclusivement numérisée en mode image, a un projet de reconnaissance par OCR de caractères hébreux du XVII^e siècle, qui n'a pas encore abouti.

Les annotations elles-mêmes sont peu présentes sur les sites en mode image. L'exemple de *Menasseh Ben Israel* montre les limites d'annotations inscrites directement sur la page grâce à un logiciel de traitement de l'image : pour pouvoir les échanger avec un tiers, il faut que celui-ci possède lui-même le logiciel, mais surtout, ce mode d'annotations ne permet pas leur indexation ni leur enregistrement dans un fichier texte, et donc de recherche sur leur contenu.

Ces différentes possibilités induisent de nouvelles pratiques de travail sur écran et d'échange d'informations pour les chercheurs.

Studiolum propose lui la possibilité de se créer un fichier personnel (« fichier caché » : *shadow file*), comme dans *DEBORA*. Ce fichier permet non seulement d'annoter, mais de modifier le texte lui-même, tout comme l'administrateur de la base : il fait l'objet d'un enregistrement distinct, et l'utilisateur final dispose du droit d'écriture. Le fait que *Studiolum* soit disponible sur CD-Rom (seul un site de démonstration est en ligne) suppose des échanges d'informations locaux. Les usages de travail collaboratif ne changent pas dans ce cas de nature mais d'échelle. Quant à la constitution d'un corpus personnel, cela suppose la possibilité de télécharger un livre en entier et non de le copier page à page ; seul *Gallica* le propose.

Ainsi, les fonctionnalités les plus abouties en matière de recherche et de collaboration, et donc les plus proches du logiciel *DEBORA*, restent liées à la numérisation en mode texte, alors que la majorité des projets de numérisation utilisent le mode image. Les problèmes de transcription automatique des typographies anciennes (antérieures au XIXe siècle), et *a fortiori* des écritures manuscrites médiévales, n'ont pas encore permis de résoudre l'opposition entre la numérisation en mode texte et en mode image. Le projet *DEBORA*, grâce à son système de transcription, se place ainsi à l'avant-garde des bibliothèques numériques dans ce domaine.

Conclusion

Les bibliothèques numériques disponibles aujourd'hui sont très diverses dans leurs objectifs initiaux comme dans leurs réalisations. Le mode d'intégration des besoins ou pratiques des chercheurs est intimement lié au statut de l'établissement porteur du projet, aux collections concernées mais surtout au public visé. C'est majoritairement en fonction du public que la typologie des bibliothèques numériques s'établit.

Les bibliothèques, dont une des missions actuelles est de toucher le grand public, voient dans la numérisation un moyen d'élargir l'accès à leurs collections patrimoniales, pour lesquelles les exigences de conservation interdisent une communication directe trop fréquente. Le mode image est alors préféré, car il faut fournir au lecteur un fac-similé le plus proche possible du document original. Le grand public trouve ici de quoi satisfaire sa curiosité, si la collection numérique présente les documents comme une exposition virtuelle. Ces collections numériques sont souvent importantes en quantité, et proposées avec des interfaces de consultation simples : la préoccupation est de permettre un plus large accès au document, tandis que la réflexion sur les usages vient dans un second temps. Cette réflexion s'organise dans la ligne de la recherche récente sur l'évaluation des services offerts par les bibliothèques, et inclut des enquêtes empruntant à la sociologie ses méthodes d'observation. Les pratiques et le profil du public sont de cette manière un peu mieux connus, même si les résultats d'enquêtes publiés sont peu nombreux, et parfois difficilement interprétables sans les données brutes.

La plupart des sites de bibliothèques numériques observés dans cette étude relèvent de cette situation : les corpus mis à disposition du public sont importants, mais les fonctions de l'interface sont limitées à la visualisation, c'est à dire au feuilletage, parfois au multifenêtrage, ou à la recherche sur des notices bibliographiques peu détaillées. La collaboration active permise par l'informatique n'est pas

ouvertement envisagée ; ainsi, les outils proposés relèvent davantage de la communication que de la collaboration.

Le second type de bibliothèque numérique observé concerne des projets de numérisation presque exclusivement destinés aux chercheurs, qui ont été conçus par des équipes de recherche mêlant informaticiens et chercheurs. Leurs interfaces proposent des fonctionnalités de consultation plus élaborées que les sites des bibliothèques dont il vient d'être question : recherche, feuilletage, annotations, constitution de corpus. La structuration des documents et l'élaboration de métadonnées permettant d'identifier très précisément le document affiché font également l'objet de recherches. Par contre, un volume moindre de documents est proposé à l'utilisateur, et les corpus sont souvent constitués autour de thèmes très spécialisés.

Les équipes de recherche qui ont participé à l'élaboration de ces interfaces comportaient souvent, on l'a dit, des spécialistes des textes, qui sont les utilisateurs finaux de la collection : leurs attentes et leurs besoins ont été pris en compte dès l'origine, mais aucune enquête postérieure sur les pratiques n'est disponible. La numérisation des collections se fait soit en mode texte soit en mode image, car les besoins sont différents selon les disciplines : si la qualité de l'image est suffisamment bonne pour garantir la justesse de leur lecture et leur fournir les informations attachées au support, les historiens du livre, pour qui la mise en page est fondamentale, ou encore les historiens de l'art, les paléographes et historiens du texte sont partisans d'une numérisation en mode image. Les linguistes et les philologues, eux, travaillent avec le mode texte, qui leur permet de relever des occurrences de mots, d'établir des statistiques, de rechercher dans le texte intégral. Pour nombre de chercheurs, l'idéal est encore une double numérisation, qui allierait l'image de l'original à une transcription codée en ASCII. La démarche ici adoptée consiste à élaborer une interface en fonction des besoins exprimés, mais aucune étude ne permet actuellement de savoir quelle appropriation de ces interfaces a été faite par la suite par les chercheurs.

La collaboration qui est induite par les fonctionnalités de ces bibliothèques « de recherche » reste classique. Le partage de fichiers et l'écriture à plusieurs mains

sur le même document semble en effet plus simple à développer pour des fichiers texte que pour des fichiers image : dans le cas d'une bibliothèque numérique proposant des images de documents comme *DEBORA*, seules les notes sont concernées par les éventuelles fonctions collaboratives. L'image est fixe, tandis que le texte est aisément modifiable : la lecture active dont il a été question dans l'état de l'art ne concerne que le texte. Ainsi, *Studiolum* (mode texte) inclut des possibilités de travail collaboratif sur le texte lui-même, par le biais d'un fichier caché : il devient possible pour les utilisateurs organisés de créer ensemble un nouveau document, sur le modèle des bases de connaissances collaboratives issues du knowledge management, comme celles citées à propos de l'IRHT ou de *Dunhuang Project*. Pour des sites proposant de l'image, comme *Menasseh Ben Israel*, l'envoi des annotations est possible après intégration des informations saisies par l'utilisateur dans le document lui-même : enregistrées comme des images, ces notes ne sont plus modifiables par le destinataire. Il ne peut s'agir ici que d'une collaboration de niveau 1, pour reprendre les niveaux définis dans le cadre du projet *DEBORA* : la simple communication d'information. Ne perdons toutefois pas de vue que les informations dont il s'agit ici sont des documents de travail, ce qui est peut-être encore peu communiqué entre chercheurs, et pourrait effectivement ouvrir de nouveaux espaces de discussions.

DEBORA en tant qu'interface de consultation se situe dans ce mouvement : élaborée avec des informaticiens, des spécialistes de l'information, et en collaboration avec des historiens du livre, il propose notamment une double numérisation et des possibilités de recherche d'images destinées aux chercheurs. Les besoins de ces derniers ont été étudiés, mais il n'a pas été possible jusqu'ici de lancer une enquête sur les pratiques : *DEBORA* n'a pas été mis à la disposition du public.

Le type de numérisation, le corpus numérisé, le public visé et la prise en compte des usagers sont donc liés dans le processus d'élaboration initial.

Ainsi, on peut distinguer deux politiques différentes influant sur l'élaboration des bibliothèques numériques : la politique de l'offre, traditionnelle dans le domaine culturel, et une politique de la demande. Cette distinction reprend celle

précédemment établi dans l'état de l'art, entre analyse préalable des besoins et étude postérieure des pratiques. Elle permet également de situer certaines interfaces de sites visités sur l'Internet. *DEBORA* est plus difficile à situer, car le projet ambitionne à la fois de s'adapter aux besoins et d'accompagner l'émergence de nouveaux usages. Il semble que son positionnement, par moment ambigu, gêne aujourd'hui son développement, et empêche notamment sa diffusion. On peut prendre ici pour exemple la question du financement : les bibliothèques proposant sur Internet leurs fonds anciens numérisés le font gratuitement, et le coût en est supporté par la collectivité pour autant que ces établissements soient publics, tandis que les outils logiciels développés pour les chercheurs sont souvent d'accès restreints, sur inscription, celle-ci pouvant être payante. Les plus développés d'entre eux sont distribués sur CD-Rom pour la version complète, qui regroupe le logiciel et le corpus numérisé (comme c'est le cas pour *Studiolum*, *Perseus*, ou *Menasseh Ben Israel*) : une installation en réseau est alors possible, ce qui ouvre la porte à un mode d'organisation collaborative, mais un investissement initial doit être consenti. Il est possible d'imaginer une diffusion commerciale pour un outil élaboré en étroite collaboration avec ses utilisateurs finaux, mais la gratuité de l'accès aux informations disponibles sur l'Internet est déjà très ancrée dans les usages.

Quelles sont, dans ces conditions, les perspectives pour l'étude des usages de *DEBORA* ?

Développé en priorité pour les chercheurs, le projet *DEBORA* devra assumer ce choix dans sa première phase de diffusion, qui semble aujourd'hui fondamentale. Il sera ensuite possible d'élargir l'accès aux collections de *DEBORA* au grand public. Lancer le logiciel, ne serait-ce qu'auprès d'un public restreint de chercheurs, permettrait en effet son expérimentation par des utilisateurs, voire la constitution de groupes d'utilisateurs, ce qui rendrait possible une enquête sur les pratiques liées à *DEBORA*, et sur l'émergence éventuelle de pratiques collaboratives. Cette nécessaire expérimentation en situation réelle permettrait à *DEBORA* d'allier les deux types d'enquêtes, celles concernant les attentes exprimées et celles sur les pratiques observées, pour comprendre comment s'articulent ces deux réalités. Il

semble en effet, d'après les comptes-rendus d'enquêtes étudiés, qu'il existe entre elles un écart, et celui-ci n'est pas toujours pris en compte. La mesure de cet écart correspondra en réalité au changement des pratiques dû au travail sur écran, aux « résonances » du travail informatique.

Mais, pour convaincre un public de chercheurs de travailler à moyen ou long terme avec *DEBORA*, un deuxième choix doit intervenir, à propos des collections à numériser. En l'état actuel, les fonds numérisés viennent de deux bibliothèques, ce qui fait de *DEBORA* une bibliothèque numérisée présentant une collection virtuelle. Si la coopération avec la bibliothèque de Coïmbra se renforce, il faudrait construire un corpus susceptible d'intéresser un public de chercheurs précis, en numérisant selon une thématique bien identifiable, à la manière des collections thématiques d'*American Memory* (qui sont en fait des collections spécialisées de différentes bibliothèques). Construire le corpus de *DEBORA* sur le seul critère de la date d'impression des livres peut en effet réduire sa visibilité auprès des chercheurs. Les bibliothécaires doivent donc s'investir sur ce point, pour qu'à côté de la réalisation informatique, *DEBORA* soit perçu comme une collection, c'est à dire un ensemble de documents ayant une signification propre. Ce travail à mener sur la politique de numérisation est liée à la question du travail collaboratif : l'éventuelle base de données communes suppose la désignation d'un responsable pour sa gestion. Trois modèles de structuration d'un réseau *DEBORA* sont envisageables, liés au choix du schéma de développement du corpus numérisé : une gestion par bibliothèque, où chaque établissement alimenterait le réseau après avoir joué le rôle de modérateur, comme *Digital Scriptorium* le fait déjà pour les métadonnées bibliographiques ; une gestion globale, qui rendrait nécessaire l'apparition d'un centre *DEBORA*, dans un projet jusque-là organisé sur le mode coopératif ; une gestion dans le cadre d'un réseau local, dans laquelle les concepteurs et diffuseurs de *DEBORA* n'interviendraient pas. Les deux dernières possibilités évoquées ont pour principal inconvénient de restreindre progressivement l'ampleur des pratiques collaboratives, que *DEBORA* ambitionne justement de développer chez les chercheurs.

Les usages des bibliothèques numériques restent aujourd'hui un champ de réflexion ouvert pour la recherche : des études pionnières ont été menées, mais beaucoup reste à étudier. En s'engageant plus avant dans cette voie, comme cela avait été prévu au début du projet, et en tenant compte des réalisations et réflexions actuelles d'autres groupes de recherche comme celui du PNER⁶³, *DEBORA* peut trouver une nouvelle dynamique pour son développement propre, et participer à l'émergence de nouvelles pratiques de recherche, voire de nouveaux usages. La priorité actuelle pour nourrir cette dynamique est la diffusion de *DEBORA* auprès de son public-cible, les chercheurs.

⁶³ Voir en annexe 2.

Bibliographie

Publications du groupe de recherche **DEBORA**

DEBORA, WP2. *Rapport final d'évaluation de l'usage des outils et interface DEBORA.* 2001.

HEMBISE, C., et al. *Etat de l'art des recherches sur les pratiques et sur les usagers des bibliothèques virtuelles*, 1999. Disponible sur Internet <<http://debor.enssib.fr/del21.doc>> (consulté le 12.03.2002)

NICHOLS, David M., et al. *DEBORA : developing an interface to support collaboration in digital library.* Disponible sur Internet <<http://www.comp.lanec.ac.uk/computing/research/eseq/projects/ariadne/docs/ecdl2000.html>>, (consulté le 23.01.02).

PEMBERTON, Duncan et NICHOLS, David. *DEBORA : Digital access to books of the Renaissance.* Disponible sur Internet <<http://www.exploit-lib.org/issue3/debora/>>, (consulté le 19.04.02).

Articles

ARONSSON, Lars. Project Runeberg's electronic facsimile editions of Nordic literature. Disponible sur Internet <<http://www5.hk-r.se/elpub99/ap-nsf/08cbc2f88424ad99c12566ff002>> (consulté le 13.05.2002)

BELZILE, Sylvie. NRC clients' perceptions of electronic resources and use of CISTI's virtual library. *Internet reference services quarterly*, vol. 4, (4), 1999, p. 5-16.

BOZZI, Andrea. Pour un système de philologie numérique. *Document numérique*, 1999, 3, 1-2, p. 93-101.

BURNINGHAM, B. Attitudes of the Canadian Research Community toward Creating and Accessing Digitized Facsimile Collections of Historical Documents. *Computers and the Humanities*, 33, 1999, p. 409-419. disponible sur Internet <<http://ipsapp008.lwwonline.com/content/search/4589/15/4/fulltext.pdf>> (consulté le 22.05.2002)

BUSHNELL, Rebecca. Reinventing Rare Books : The "Virtual Furness Shakespeare Library" at the University of Pennsylvania. *Early Modern Literary Studies*, janvier 2000,

5, 3, n° 4. Disponible sur Internet <<http://purl.oclc.org/emls/05-3/bushfurn.html>> (consulté le 10.05.2002)

CALABRETTO, Sylvie, PINON, Jena-Marie, BOZZI, Andrea. BAMBI : système de gestion de manuscrits anciens pour historiens. *Document numérique*, 1999, 2, 3-4, p. 31-50.

CHERRY, Joan M., DUFF, Wendy M. Studying digital Library users over time: a follow-up survey of Early Canadiana Online. *Information Research*, 7, 2. Disponible sur Internet <<http://InformationR.net/ir/paper123.html>> (consulté le 5.05.2001)

COVI, Lisa M. Material mastery : situating digital library use in university research practices. *Information processing and management*, n° 35, 1999, p. 293-316.

CUBAUD, Pierre, GIRARD, Didier. ABU : une bibliothèque numérique et son public. *Document numérique*, 2, (3-4), 1999, p.13-30.

DOWNS, Robert R., FRIEDMAN, Edward A. Digital library support for scholarly research. *Information processing and management*, 35, 1999, p. 281-291.

DUFF, Wendy M., CHERRY, Joan M. Use of historical documents in a digital world : comparisons with original materials and microfiche. *Information research*, vol. 6, n° 1, octobre 2001. Disponible sur Internet <<http://www.shef.ac.uk/~is/publications/infres/paper86.html> > (consulté le 04.05.2002)

FEKETE, Jean-Daniel, DUFOURNAUD, Nicole. Analyse historique de sources manuscrites : application de TEI à un corpus de lettres de rémission du XVI^e siècle. *Document numérique*, 1999, 3, 1-2, p. 117-134.

FELTHAM, Mark, and BARKER, William. The Web and the Book: The Memorial Electronic Edition of Andrea Alciato's *Book of Emblems*. *Early Modern Literary Studies*, janvier 2000, 5, 3, n° 4. Disponible sur Internet <<http://purl.oclc.org/emls/05-3/fbemblem.html>> (consulté le 10.04.2002)

GARTNER, Richard. The Internet Library of Early Journals : an electronic library of primary sources on the Internet. Disponible sur Internet <[http://www5.bth.se/elpub99/ap.nsf/08c6c2f88424ad99c12566ff002a0c10/6f7b5fad7be3e268c12566ff00361328/\\$FILE/194-203.pdf](http://www5.bth.se/elpub99/ap.nsf/08c6c2f88424ad99c12566ff002a0c10/6f7b5fad7be3e268c12566ff00361328/$FILE/194-203.pdf)> (consulté le 20.05.2002)

GOLOVCINSKY, Gene, PRICE, Morgan N., SCHILIT, Bill N. From reading to retrieval : freeform ink annotations as queries. Disponible sur Internet <<http://www.fxpal.com/PapersAndAbstracts/papers/gol99a.pdf>> (consulté le 15.05.2002)

GOLOVCHINSKY, Gene, MARSHALL, Catherine C. Hypertext Interaction Revisited. Disponible sur Internet <<http://www.fxpal.com/PapersAndAbstracts/papers/gol2000a.pdf>> (consulté le 20.05.2002)

HILL, Linda L., et al. Alexandria Digital Library : user evaluations studies and system design. *Journal of the american society for information*, février 2000, n° 51, vol. 3, p. 246-259.

HU, Shuen-Cheng, FURUTA, Richard, URBINA, Eduardo. An electronic version of *Don Quixote* for humanities scholars. *Document numérique*, 1999, vol. 3, n° spécial 1-2, p. 75-91.

RYDBERG-COX, Jeffrey A., et al. Knowledge management in the Perseus digital library. *Ariadne*, septembre 2000, n° 25. Disponible sur Internet <http://www.ariadne.ac.uk/issue25/rydberg-cox/intro.html> (consulté le 3.05.2002).

JONES, Michael L. W., GAY, Geri K., RIEGER, Robert H. Project soup : comparing evaluations of digital collection efforts. *D-Lib Magazine*, novembre 1999, vol. 5, n° 11. Disponible sur Internet www.dlib.org/dlib/november99/11jones.html (consulté le 7.05.2002)

KUNSTMANN, Pierre, MARTINEAU, France. Chrétien de Troyes sur le Web: présentation, analyse et interprétation des manuscrits d'*Yvain ou le Chevalier au Lion*. Editors of *CHWP* 2000. Disponible sur Internet <http://texttechnology.mcmaster.ca/kunstmannmartineau.html> (consulté le 15.04.2002)

LECOLINET, Eric, ROBERT, Laurent. Conception d'un poste d'édition et de lecture d'hypermédias littéraires. *Document numérique*, 1999, 3, 1-2, p. 114.

LECOLINET, Eric, ROBERT, Laurent, ROLE, François. Text-image coupling for editing literary sources. *Computers and the Humanities*, 2002, 36, p. 49-73. Disponible sur Internet http://ipsapp008.lwwonline.com/content/search/4589/22/1/full_text.pdf (consulté le 16.05.2002).

MARCHIONINI, Gary, FOX, Edward A. Progress toward digital libraries : augmentation through integration. *Information processing and management*, n° 35 p. 219-225, 1999.

MARCHIONINI, Gary. Evaluation Report on the Perseus Project : Publication Model 1998-1999. Disponible en ligne http://www.ils.unc.edu/~march/perseus/final_report98-99.pdf (consulté le 10.06.2002)

MARSHALL, Catherine C. Annotation : from paper books to the digital library. Disponible sur Internet <http://www.csdl.tamu.edu/~marshall/dl97.pdf> (consulté le 25.04.2002)

MARSHALL, Catherine C. Toward an ecology of hypertext annotation. Disponible sur Internet <http://www.csdl.tamu.edu/~marshall/ht98-final.pdf> (consulté le 25.04.2002)

MARSHALL, Catherine C. The future of annotation in a digital (paper) world. 1998. Article disponible en ligne [URL : <http://www.csdl.tamu.edu/~marshall/uiuc-paper-complete.pdf>], consulté le 27 avril 2002.

MARSHALL, Catherine C., et al. Introducing a digital library reading appliance into a reading group. In *Proceedings of ACM Digital Libraries 99*, ACM Press, pp. 77-84, 1999. Disponible sur Internet <http://www.csdl.tamu.edu/~marshall/dl99.pdf> (consulté le 22.05.2002)

MAZHOUD, O., PASCUAL, E., VIRBEL, Jacques. Annotating as a document management tool. Disponible sur Internet <<http://www.hit.uib.no/allc/mazhoud.pdf>> (consulté le 20.05.2002)

MUELLER, Martin. Electronic Homer. *Ariadne*, n° 25, septembre 2000. Disponible sur Internet <http://www.ariadne.ac.uk/issue25/mueller/intro.html> (consulté le 4.05.2002).

PELLEN, René. Conclusions et propositions pour que l'étude des textes puisse se faire aussi sur Internet. *Le médiéviste et l'ordinateur*, hiver 1999, n° 38. Disponible sur Internet <<http://irht.cnrs-orleans.fr/meto/mo3816.htm>> (consulté le 15.05.2002)

PELLEN, René. Le « Projet Charrette » à Poitiers. *Le Médiéviste et l'ordinateur*, n° 40, automne 2001, *La numérisation des manuscrits médiévaux : actes de la journée d'étude, Paris, 13 octobre 2000*. Disponible sur Internet <<http://irht.cnrs-orleans.fr/meto/mo40-07.htm>> (consulté le 20.05.2002)

PETERSON BISHOP, Ann, et al. Digital libraries : situating use in changing information infrastructure. *Journal of the american society for information*, n° 51, vol. 4, p. 394-413, mars 2000.

PINFIELD, Stephen. How do physicists use an e-print archive? : implications for institutionnal e-print services. *D-Lib Magazine*, vol. 7, n° 12, décembre 2001. Disponible sur Internet <<http://www.dlib.org/dlib/december01/pinfield/12pinfield.html>> (consulté le 5.05.2002)

PORTET, Pierre. La numérisation des manuscrits scientifiques et techniques médiévaux sur l'Internet : notes de lecture. *Le Médiéviste et l'ordinateur*, n° 40, automne 2001, *La numérisation des manuscrits médiévaux : actes de la journée d'étude, Paris, 13 octobre 2000*. Disponible sur Internet <<http://irht.cnrs-orleans.fr/meto/mo40-08.htm>> (consulté le 20.05.2002)

RIEGER, Robert, GAY, Geri. Tools and techniques in evaluating digital imaging projects. *RLG Diginews*, vol.3, n°3. Disponible sur Internet <<http://www.rlg.org/preserv/diginews/diginews3-3.htm.l>> (consulté le 7.05.2002)

ROBERT, Laurent, LECOLIVET, Eric. Digital Annotation and Exploration Techniques for Handling Image-Based Hypermedia. Disponible sur Internet <<http://perso.enst.fr/~lrobert/paper/interact01.pdf>> (consulté le 16 mai 2002).

ROWLAND, Fytton. Two large-scale surveys of electronic publication in the United-Kingdom. Disponible sur Internet <[http://www5.bth.se/elpub99/ap.nsf/08c6c2f88424ad99c12566ff002a0c10/d6f40b8c378ae03c12566ff003749a1/\\$FILE/131-136.pdf](http://www5.bth.se/elpub99/ap.nsf/08c6c2f88424ad99c12566ff002a0c10/d6f40b8c378ae03c12566ff003749a1/$FILE/131-136.pdf)> (consulté le 5.05.2002).

ROLE, François. Représentation et exploitation de métadonnées complexes : le cas des documents anciens. *Document numérique*, 1999, 3, 1-2, p. 135-150.

SHU CHING Yang. An interpretative and situated approach to an evaluation of Perseus digital libraries. *Journal of the american society for Information, Science and Technology*, vol. 52, 14, 2001, p. 1210-1223.

SMITH, Marc. Numérisation et paléographie. *Le Médiéviste et l'ordinateur*, n° 40, automne 2001, *La numérisation des manuscrits médiévaux : actes de la journée d'étude, Paris, 13 octobre 2000*. Disponible sur Internet <<http://irht.cnrs-orleans.fr/meto/mo40-03.htm>> (consulté le 15.05.2002)

VENTABERT, Gusnard de. Représentation et exploitation électroniques de documents anciens. *Document numérique*, 1999, 3, 1-2, p. 57 à 73.

VISCOMI, Joseph. Digital facsimiles : reading the William Blake archives. *Computers and the Humanities*, 2002, 36, p. 27-48. Disponible sur Internet <<http://ipsapp009.lwwonline.com/content/search/4589/22/4/fulltext.pdf>> (consulté le 15.05.2002). Voir aussi en 3.8.7.

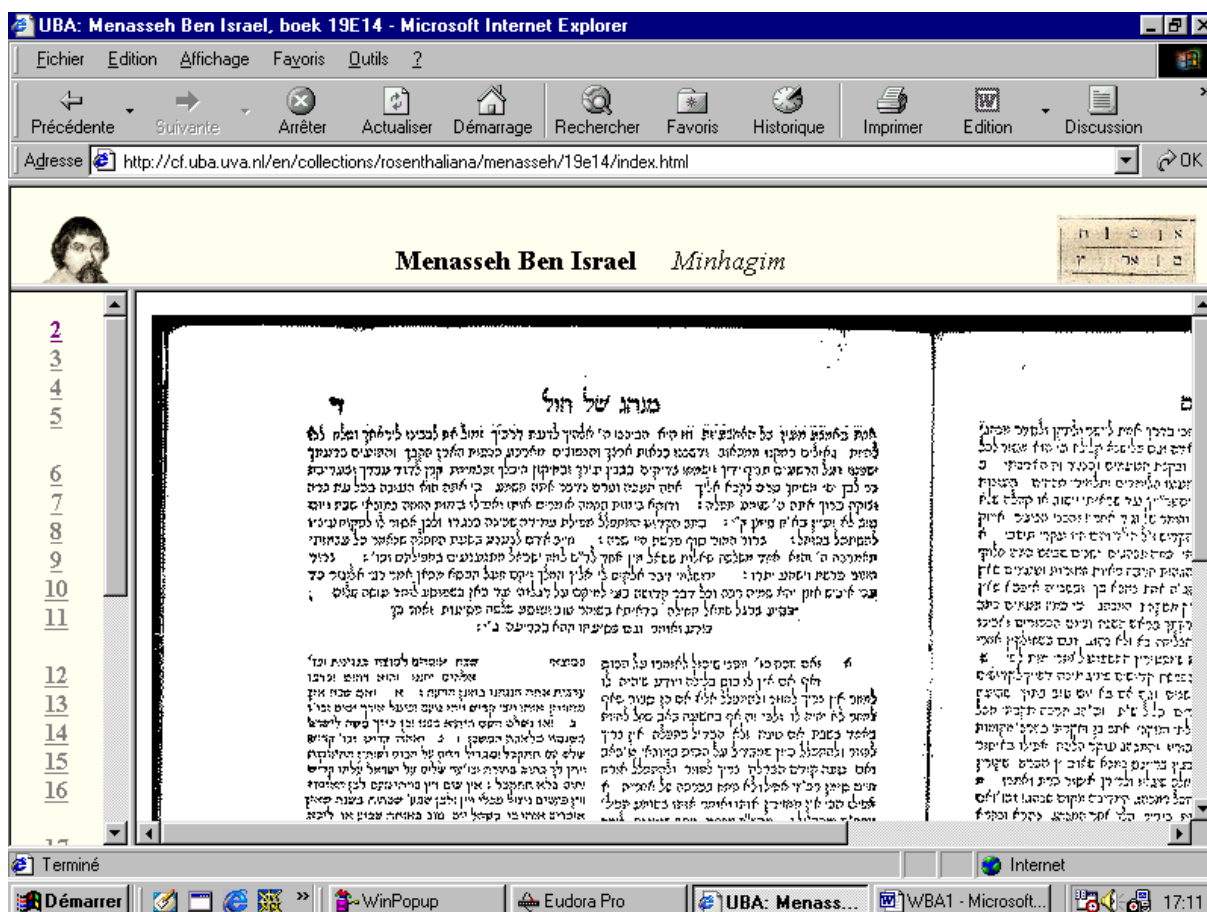
WAGNEUR, Jean-Didier. La bibliothèque électronique de la BnF ou Quand le lecteur a le droit d'écrire sur les livres. *Revue de la BnF*, juin 1999, 2, p. 78-81.

Table des annexes

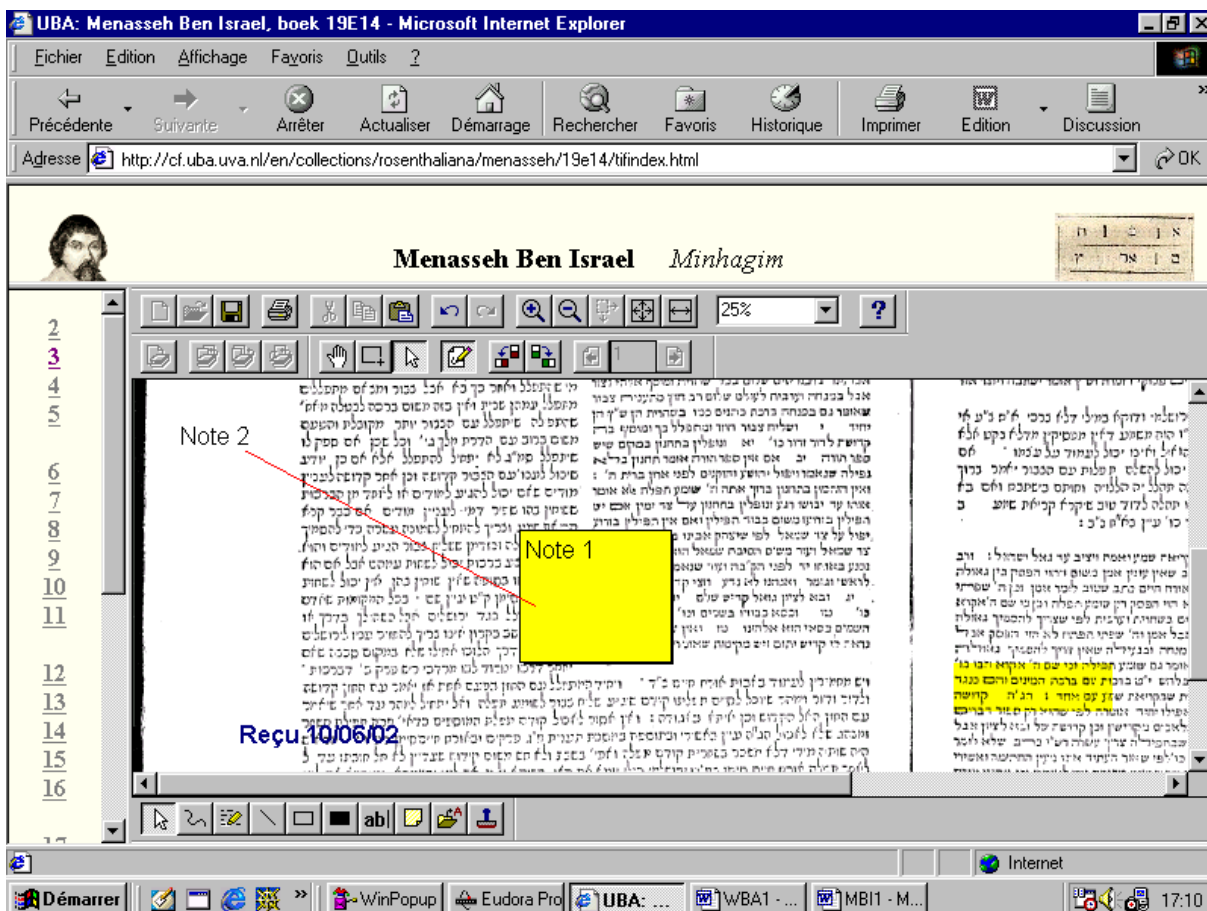
ANNEXE 1 : COPIES D'ÉCRANS	I
ANNEXE 1.1 : <i>MENASSEH BEN ISRAEL</i>	I
ANNEXE 1.2 : <i>STUDIOLUM</i>	III
ANNEXE 1.3 : <i>THE WILLIAM BLAKE ARCHIVES</i>	VI
ANNEXE 2 : LISTE DES SITES VISITÉS ET NON RETENUS.....	VIII
ANNEXE 3 : TABLEAUX DES BIBLIOTHÈQUES NUMÉRIQUES	XIII

Annexe 1 : Copies d'écrans

Annexe 1.1 : Menasseh Ben Israel



Ecran 1 : affichage simple (sans Imaging) d'un ouvrage hébreu de la bibliothèque numérique de Menasseh Ben Israel



Ecran 2 : affichage avec le logiciel Imaging d'un ouvrage en hébreu de la bibliothèque numérique de *Menasseh Ben Israel*, avec des annotations sur l'image.

Annexe 1.2 : *Studiolum*

The screenshot shows the 'Folio Views - [studiolum]' application window. The interface includes a menu bar (File, Edit, View, Insert, Search, Layout, Tools, Table, Window, Help), a toolbar with navigation and editing icons, and a search field containing '[Field 1a text: caudam.]'. Below the toolbar is a navigation pane with icons for search, browse, document, contents, hitlist, and object. The main area is divided into two panes. The left pane displays a list of 29 items, with item [1] 'INSIGNIA DVCATVS MEDIOLA' selected. The right pane shows the detailed view of item [1], titled 'INSIGNIA DVCATVS MEDIOLAN.'. It features a drawing of a tree with a snake coiled around its trunk. Below the drawing is a block of Latin text:


[a2v] Exiliens infans simosi [simosi] è faucibus anguit,
 Est gentilijs nobile stemma tuus.
 Talia Fellaem gesisse nomismata regem,
 Vidimus, hisque suam concelebrare genus.
 Dum se Ammonē, satum matrem anguis imagine lusam,
 Diuini & sobolem seminis esse docet.
 Ore exit tradunt sic quosdam enitier angues,
 An quia sic Pallas de capite orta Iouis.

The bottom of the window shows a status bar with 'Record: 28 / 336', 'Hit: 0 / 0', and 'Query:'. The taskbar at the very bottom displays the Windows taskbar with icons for 'Démarrer', 'logiciels', 'Folio Vi...', 'Document...', 'Sans titre...', 'Adobe Ph...', and system tray icons including 'Bureau' and the time '22:02'.

Ecran 1 : affichage de l'interface de lecture *Studiolum*

- [2] FOEDERA ITALORVM. [a2v-a3]
- [3] IN SILENTIVM. [a3r]
- [4] ETIAM FEROCISSIMOS DOMA
- [5] GRATIAM REFERENDAM [a3v-
- [6] CONCORDIA. [a4r]
- [7] POTENTISSIMVS affectus amc
- [8] NON WLGANDA CONSILIA [a
- [9] IN VICTORIAM DOLO PARTAM
- [10] REVERENTIAM IN MATRIMOI
- [11] IN AVAROS, VEL QVIBVS MEI
- [12] AMICITIA ETIAM POST MORT
- [13] ILLICITVM NON SPERANDVM
- [14] NEC VERBO NEC FACTO qu
- [15] DESIDIAM ABIICIENDAM [a7v
- [16] PAVPERTATEM SVMMIS inge
- [17] IN OCCASIONEM [a8r-a8v]
- [18] OCNI EFFIGIES DE HIS QVI r
- [19] VIRTVTI FORTVNA COMES. [
- [20] EX PACE VBERTAS. [b1v]
- [21] IN EOS QVI SVPRA VIRES ql
- [22] PRINCEPS SVBDITORVM inc

**INSIGNIA DVCVTVS ME-
DIOLAN.**



[a2v] *Exiliens infans sinnosi [sinuosi] è faucibus anguit,*
Est gentilijs nobile stemma tuis.
Talia Pellaeum gessisse nomisma regem,
Vidimus, hisque suum concelebrare genus.
Dum ea Ammonae satum matrem anguis imactus lusam

ANDIDO LECTORI S. P. [a1v]
 ...fuisse. Vtilissimum itaque hobis uisum **est**, si notulis quibusdam obiter, rudioribus...
 INSIGNIA DVCVTVS MEDIOLAN. [a2r-a2v]
 ...**Est** gentilijs nobile stemma tuis....
 FOEDERA ITALORVM. [a2v-a3r]
 ...Difficile **est** nisi docto homini tot tendere...
 ...Ruptaue (quod facile **est**) perit omnis gratie conchae,...

Ecran 2 : affichage des résultats de recherche d'occurrences dans *Studiolum* : les termes trouvés sont en surbrillance dans le texte, une liste des occurrences est donnée dans la fenêtre inférieure

verse

Times New Roman 11

Contents HitList Object

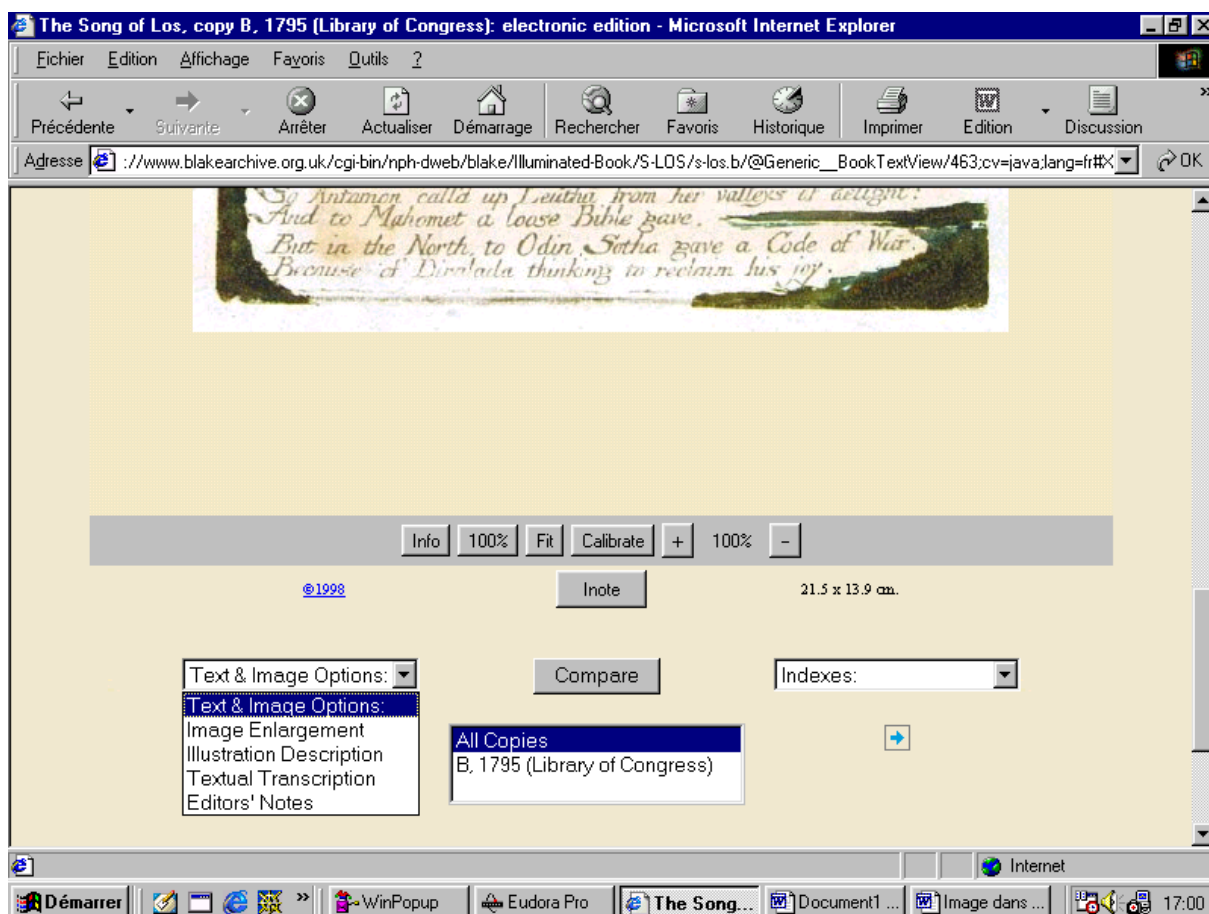
46] EX BELLO PAX. [c3v]

EX BELLO PAX.

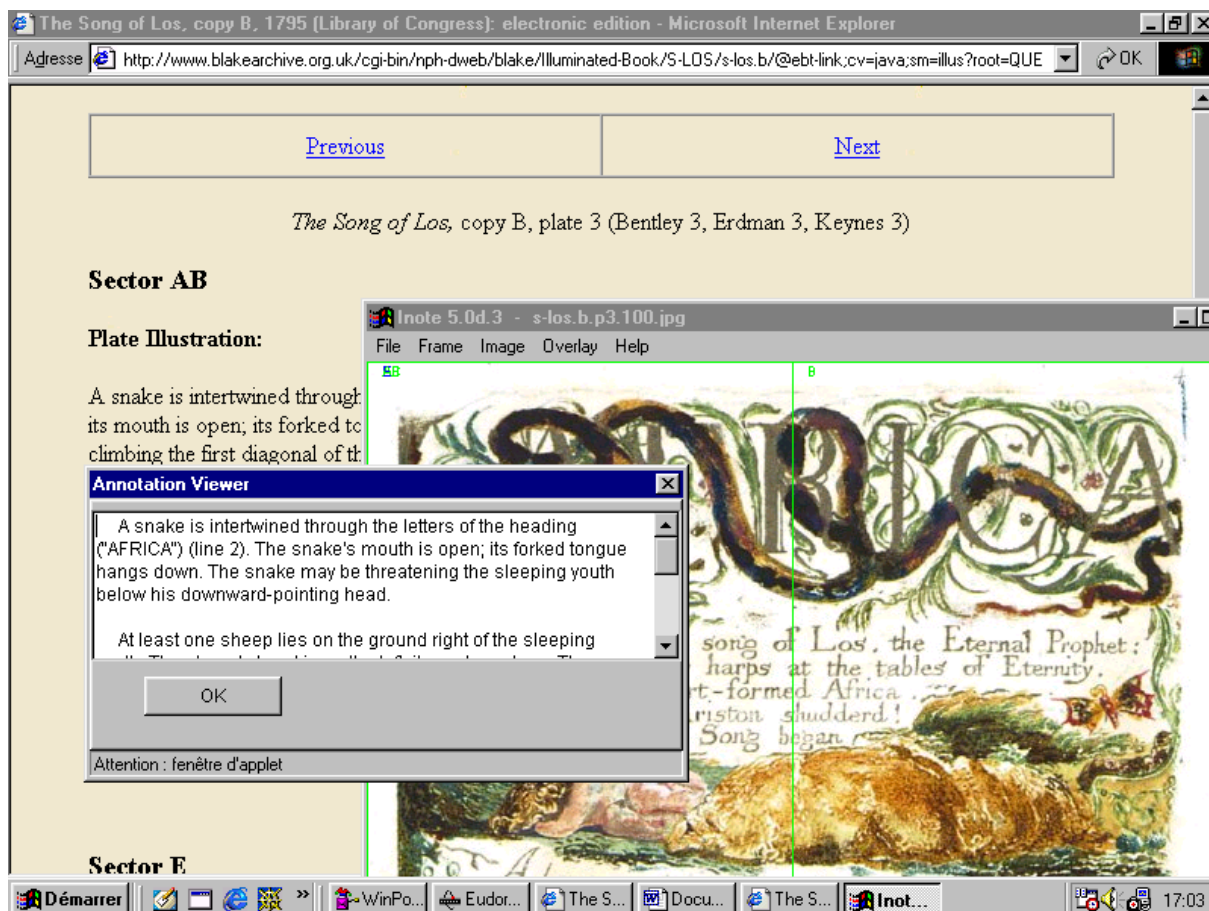
En galea intrepidus quam miles gesserat, & quae
Saepius hostili sparsa texte modifié, le vers suivant est un lien hypertexte:
Parta pace apibus tenuis concessit in usum.
Alueoli atque fauos grataque melle gerit.
Arma procul iaceant, fas sit tamen sumere bellum,
Quin aliter pacis non potes arte frui: gras, souligné, droit

Ecran 3 : affichage du fichier « shadow » de *Studiolum*, avec des modifications sur le texte.

Annexe 1.3 : The William Blake Archives



Ecran 1 : affichage de l'interface de consultation des images, avec les liens vers les autres possibilités d'affichage



Ecran 2 : affichage de trois fenêtres : description des images, logiciel Inote d'annotation des images (image cible et annotations déjà enregistrées)

Annexe 2 : Liste des sites visités et non retenus

Cette liste présente les 31 sites qui ont été visités mais n'ont pas été sélectionnés dans les tableaux. Une ligne, après l'adresse URL, explique pourquoi ils n'ont pas été retenus. Un commentaire plus long indique les aspects intéressants du site pour les études d'usages, ou par rapport à *DEBORA*.

Ces sites ont tous été visités entre le mois de mars et le mois d'avril 2002. La date de consultation ne sera donc pas répétée pour chacun d'entre eux.

Programme de numérisation pour l'enseignement et la recherche (PNER) [**en ligne**]. Disponible sur Internet <<http://www.pner.org>>

Ce projet de portail de la Fondation Maison des Sciences de l'Homme (1999-2001), à destination de l'enseignement et de la recherche, ne présente pour l'instant aucune collection en ligne, mais de nombreux articles de groupes de recherche, et notamment sur les usages.

Au cours de l'année 2001, le PNER a mis en place un « Programme d'Actions Concertées de Recherche » dans les domaines de la numérisation et de la production numérique.

Le pôle usages, en particulier, a retenu quatre thèmes d'études : des études sur les contenus numériques, des études sur les usages et les pratiques dans le domaine de l'enseignement scolaire, de l'enseignement supérieur et de la recherche

The Erasmus Text project **[en ligne]**. Disponible sur Internet

<<http://smith2.sewanee.edu/Erasmus>>

Mode texte

Projet Libris **[en ligne]**. Disponible sur Internet <<http://libris.univ-lille3.fr>>

Base de données de cartes postales anciennes du nord de la France.

Bibliotheca universalis **[en ligne]**. Disponible sur Internet

<<http://www.konbib.nl/gabriel/bibliotheca-universalis/digit.htm>>

Projet de réseau constitué des bibliothèques nationales qui numérisent leurs collections. Il manque encore une plateforme technique commune.

Memoria **[en ligne]**. Disponible sur Internet

<<http://www.unesco.org/webworld/mdm>>

Répertoire de ressources sur les bibliothèques numériques.

Elise Project (Electronic Library Image Service for Europe) **[en ligne]**. Disponible sur Internet <<http://severn.dmu.ac.uk/elise>>

Projet.

Fathom : the source for online learning **[en ligne]**. Disponible sur Internet

<<http://fathom.com>>

Cours payants, accompagnant une bibliothèque numérique servant à l'enseignement.

ARTFL Project **[en ligne]**. Disponible sur Internet

<<http://humanities.uchicago.edu/ARTFL/ARTFL.htm>>

Mode texte, plusieurs collections thématiques virtuelles.

Projet Gutenberg **[en ligne]**. Disponible sur Internet <<http://www.gutenberg.net>>

Mode texte, collection virtuelle.

Repertorio informatizzato dell'antica letteratura catalana (RIALC) **[en ligne]**.

Disponible sur Internet <<http://www.riald.unina.it>>

Mode texte, collection virtuelle. Recherche par index ou par incipit des poèmes (XIV^e et XV^e siècles, Catalogne et Occitanie).

Project Runeberg [**en ligne**]. Disponible sur Internet <<http://www.lysator.liu.se/runeberg>>
Projet.

Institut de recherche et d'histoire des textes (IRHT) [**en ligne**]. Disponible sur Internet <<http://irht.cnrs-orleans.fr>>

Base de données sur le livre médiéval, accessible sur place mais pas en ligne.

Conservatoire numérique des arts et métiers (CNUM) [**en ligne**]. Disponible sur Internet <<http://cnum.cnam.fr>>

Mode image, collection existante, mais qui traite des livres scientifiques du XVIII^e siècle, et sans possibilité de recherche sur les textes eux-mêmes.

Les collections de la bibliothèque municipale de Lyon [**en ligne**]. Disponible sur Internet <<http://www.bm-lyon.fr>>

Base de données sur les collections spécialisées, mais pas d'ouvrages numérisés.

La bibliothèque électronique de Lisieux [**en ligne**]. Disponible sur Internet <<http://www.bmlisieux.com>>

Mode texte, alors que la numérisation concerne une collection existante, celle de la bibliothèque municipale de Lisieux. Celle-ci a fait le choix de faire appel à des bénévoles pour saisir le texte d'ouvrages d'intérêt local.

Renascence Editions [**en ligne**]. Disponible sur Internet : <<http://darkwing.uoregon.edu/~rbear/ren.htm>>

Mode texte

Dante [**en ligne**]. Disponible sur Internet : <<http://www.nd.edu/~italnet/Dante>>

Exposition de quelques feuillets, aucune fonctionnalité

Bibliotheca Augustana [**en ligne**]. Disponible sur Internet :

<http://www.fh-augsburg.de/~harsch/augusta.html>>

Mode texte

Manuscrits de la Bibliothèque inter-universitaire de Graz **[en ligne]**. Disponible sur Internet <<http://www-ub.kfunigraz.ac.at/sosa/katalog>>

En allemand

Les Très Riches heures de Metz **[en ligne]**. Disponible sur Internet :

<http://www.mairie-metz.fr:8080/METZ/MED/EXPOS/MED_Psautier.html>

Pas d'accès

Bibliothèque inter-universitaire de Montpellier **[en ligne]**. Disponible sur Internet :

<<http://boree.cnusc.fr:8100/textes/manus-med.html>>

Site espagnol **[en ligne]**. Disponible sur Internet

<<http://emilia.dc.fi.udc.es/CICYT96/index.html>>

Site intéressant qui présente les images numérisées sous forme de livre, avec les métadonnées auparavant. Malheureusement image de piètre qualité et qui ne s'affiche pas en entier. Le système n'est pas au point.

Emblem books **[en ligne]**. Disponible sur Internet :

<<http://www.netnik.com/emblemata/>>

Petit site, avec une mauvaise numérisation d'images et du texte en mode texte.

Liberty Library of Constitutional Classics **[en ligne]**. Disponible sur Internet :

<<http://www.constitution.org/liberlib.htm>>

Mode texte.

La Prose du Transsibérien de Blaise Cendrars **[en ligne]**. Disponible sur Internet :

<<http://www.franceweb.fr/poesie/transib.htm>>

Très beau site en mode texte avec des photographies.

100 frontispices de livres de médecine du XVI^e S. au début du XIX^e S. [en ligne].

Disponible sur Internet :

<<http://www.bium.univ-paris5.fr/expo/>>

Exposition virtuelle : 663 images correspondant à la reproduction (en 4 tailles différentes) des frontispices de 100 ouvrages sélectionnés dans les collections de la bibliothèque, plus 263 images diverses prises à l'intérieur des ouvrages, ou agrandissements d'éléments intéressants de certains frontispices ; 27 animations pour les pages inter-titres et les génériques.

Hypterotomachia Poliphili [en ligne]. Disponible sur Internet :

<<http://mitpress.mit.edu/e-books/HP/>>

1 livre ancien (1499) numérisé par MIT Press et Design Knowledge Systems group (Université de Delft), Liane Lefaivre

Berkeley Digital Library Sunsite [en ligne]. Disponible sur Internet :

<<http://sunsite.berkeley.edu/>>

Portail qui donne accès à de nombreuses collections (papyrus, manuscrits, photos, livres), à des sites d'aide à la numérisation, des sites de recherche, répertoire de liens.

The Life of King Edward the Confessor [en ligne]. Disponible sur Internet :

<<http://www.lib.cam.ac.uk/MSS/Ee.3.59/>>

Un manuscrit enluminé du XIII^e siècle numérisé par Cambridge University, Department of Manuscripts. Très bonne résolution.

La bibliothèque virtuelle [en ligne]. Disponible sur Internet :

<<http://www.bm-troyes.fr/pages/bib-virtuelle/>>

Exposition virtuelle d'un livre de chœur du XV^e siècle.

Annexe 3 : Tableaux des bibliothèques numériques

Identification du site

Nom du site	Projet <i>DEBORA</i>
Editeur	INSA Lyon / ENSSIB
URL	http://rfv6.insa-lyon.fr/debora/client.htm
Date du site	1999-2002

Base de données

Type	Livres imprimés du XVI ^e siècle (France, Portugal)
Nombre de données	5 livres disponibles en ligne
Mode de numérisation	Image et texte
Couleurs	Non, mais différents niveaux de gris restituent l'effet de transparence de la page
Mise à jour	3 octobre 2001 : dernière version (01b10) du logiciel
Enrichie	Oui, métadonnées
Niveau de description des métadonnées	2
Format, logiciel de lecture	Format AdHoc, logiciel <i>DEBORA</i>

Outils de visualisation

Vignettage	Oui
Mosaïque	Oui
Affich. annotations	Oui
Fonds	Oui : l'effet de transparence de la page peut être atténué grâce à la luminosité
Luminosité	Oui
Grossissement	Oui
Autres	

Recherche

Titre	Oui
Plein texte	Oui
Illustrations	Oui
Autres	

Travail personnel sur le document

Corpus personnel	Oui
Fichier personnel	Oui
Impression	Non (en projet)
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	Oui
Autres	

Documents associés

Enquêtes	
Liens	
Webmaster	trinh@rfv.insa-lyon.fr

Identification du site

Nom du site	<i>Alciato</i>
Editeur	Memorial Université de Newfoundland, William Barker , Mark Feltham , et Jean Guthrie.
URL	http://www.mun.ca/alciato/
Date du site	1995

Base de données

Type	<i>Le Livre des Emblèmes</i> d'Alciat, en trois versions : latin, anglais, anglais/latin.
Nombre de données	
Mode de numérisation	Mode texte (textes), mode image (illustrations des emblèmes)
Couleurs	Noir et blanc pour les illustrations
Mise à jour	16 mars 2002
Enrichie	Commentaires, bibliographie
Niveau de description des métadonnées	Niveau 4. Auteur, œuvre et date
Format, logiciel de lecture	Gif

Outils de visualisation

Vignettage	Oui, en cliquant sur « thumbnails »
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Oui, en cliquant une fois sur l'image
Autres	On peut afficher en regard la version latine et la version anglaise

Recherche sur le document

Titre	Liste des titres
Plein texte	
Illustrations	Oui, en cliquant sur une vignette, on retrouve le texte correspondant
Autres	

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	alciato@mun.ca
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Oui, avec d'autres sites s'intéressant aux emblèmes
Webmaster	alciato@plato.ucs.mun.ca

Identification du site

Intitulé	<i>American Memory</i>
Editeur	Bibliothèque du Congrès en coopération avec de nombreuses autres bibliothèques américaines
URL	http://memory.loc.gov
Date du site	1998 (projet 1996-1998)

Base de données

Type	Portail pour l'accès vers d'autres sites de bibliothèques numériques, et collections numérisées de la Bibl. du Congrès
Nombre de données	7 millions de documents provenant de 100 collections historiques
Mode de numérisation	Image, parfois texte
Couleurs	Oui (photoshop)
Mise à jour	17 avril 2002
Enrichie	Oui : transcription du texte (avec version SGML en option, pour naviguer dans le texte) ; pages pédagogiques pour les professeurs
Niveau de description des métadonnées	Niveau 2 ou 3, suivant le type des collections (description « sommaire » pour les archives : titre, auteur, date, mots-clés)
Format, logiciel de lecture	Jpeg, tiff, mpeg, mp3, quicktime, Realmedia, SGML, MrSID, WAV : pas de problème de lecture en particulier, sauf pour la vidéo ; photoshop en option

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	
Fonds	Oui (photoshop)
Luminosité	Oui (photoshop)
Grossissement	Oui ; Résolution + haute : Tiff (10 020 K) ; vitesse chargement 43,4 K/s, càd 3 min, sous photoshop ; Résolution normale : Jpeg (408 K)
Autres	

Recherche

Titre	Oui, par titre de collection
Plein texte	Non
Illustrations	Non
Autres	Liste alphabétique des collections, recherche d'une collection par thème, époque, lieu, département de la Bibliothèque du Congrès (par exemple « Rare Book and Special Collections Division ») À l'intérieur d'une collection : recherche par mots clés ; index pour les sujets, titres, auteurs (parfois genres, lieux)

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Autres bibliothèques participant au projet
Webmaster	American Memory Help Desk

Identification du site

Nom du site	<i>Athena e-textes</i>
Editeur	Pierre Perroud
URL	http://un2sg4.unige.ch/athena/html/
Date du site	Copyrights 1986, 1994, 2000

Base de données

Type	Textes littéraires classiques (français, suisse) ; textes scientifiques ; images et textes en histoire de l'art ; base de données de renseignements minéralogiques ; héberge la base de données sur Maupassant éditée par Thierry Selva.
Nombre de données	Nombreuses car répertoire de textes numérisés sur d'autres sites (par ex. BnF)
Mode de numérisation	Mode texte pour les textes et mode image pour les illustrations qui sont placées en frontispice des textes
Couleurs	Oui, pour les illustrations
Mise à jour	2002
Enrichie	Bibliographies, commentaires
Niveau de description des métadonnées	Niveau 4. Juste auteur, œuvre et année d'édition
Format, logiciel de lecture	Jpeg et gif

Outils de visualisation

Vignettage	Oui
Mosaïque	
Affich. annotations	Oui
Fonds	
Luminosité	
Grossissement	Oui, possibilité d'agrandir les vignettes
Autres	

Recherche sur le document

Titre	Oui
Plein texte	Oui, par mots ou expression dans le texte
Illustrations	
Autres	

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	Oui
Newsgroup	Liens vers des newsgroups
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Bibliothèques virtuelles, institutions, sites consacrés aux auteurs, aux thèmes...
Webmaster	Pierre.Perroud@terre.unige.ch

Identification du site

Nom du site	<i>Biblioteca digital Dioscórides</i>
Editeur	Biblioteca Universidad Complutense (Madrid)
URL	http://www.ucm.es/BUCM/diosc/00.htm
Date du site	

Base de données

Type	Incunables, manuscrits du XVe au XIXe en médecine
Nombre de données	1400 livres (500.000 images)
Mode de numérisation	Mode image
Couleurs	Noir et blanc
Mise à jour	2002
Enrichie	
Niveau de description des métadonnées	Niveau 2. Notice Marc, liens vers d'autres bibliothèques (requête Z39 50), liens vers les notices d'autorité auteur, matière
Format, logiciel de lecture	Jpeg et Tif

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	
Autres	

Recherche sur le document

Titre	Oui
Plein texte	
Illustrations	
Autres	Sujet, auteur, collection...

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Oui
Webmaster	dioscorides@ucm.es

Identification du site

Nom du site	(<i>Charrette</i>) Le projet <i>Charrette</i>
Editeur	Université de Princeton (USA) / Uni. de Poitiers
URL	http://www.princeton.edu/~lancelot/ site miroir : http://www.mshs.univ-poitiers.fr/cescm/lancelot/index.html
Date du site	1991 (1 ^{ère} démo avec images) ; 1994

Base de données

Type	Les huit traditions manuscrites du <i>Chevalier à la Charrette</i> de Chrétien de Troyes (12e s.) en mode image et texte, et le texte de l'édition critique Foulet-Uitti (mode texte)
Nombre de données	8 manuscrits
Mode de numérisation	Image et texte (par OCR, clés de transcription compliquées)
Couleurs	Oui
Mise à jour	18 septembre 2001 (Poitiers)
Enrichie	Oui : introduction littéraire, commentaires stylistiques
Niveau de description des métadonnées	Niveau 3 (cotes et localisations des manuscrits)
Format, logiciel de lecture	Jpeg (image), sgml (transcription)

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Non (pleine page)
Autres	La transcription n'est pas en regard du texte, mais les liens vers les 8 manuscrits sont disponibles dans la marge de la transcription

Recherche

Titre	
Plein texte	
Illustrations	
Autres	Par manuscrit, par n ^{os} de vers

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	
Webmaster	kduitti@princeton.edu ; Rene.Pellen@mshs.univ-poitiers.fr

Identification du site

Nom du site	<i>Christian Classics Ethereal library</i>
Editeur	Calvin College
URL	http://www.ccel.org/
Date du site	

Base de données

Type	Textes religieux réunis dans un but d'édification
Nombre de données	Centaines d'œuvres du ccel plus liens vers d'autres sites
Mode de numérisation	Mode texte et mode image
Couleurs	Noir et blanc
Mise à jour	Avril 2002
Enrichie	Oui, tableaux, bibliographies, notices biographiques
Niveau de description des métadonnées	Niveau 3. Auteur, titre, éditeur et date
Format, logiciel de lecture	Xml, html, pdf, rtf (implementation de ThML : theological markup language) ; Png pour le mode image

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Oui et fonctionnalité d'image « high resolution » png
Autres	Pour les textes en mode image, on peut afficher de différentes manières à côté la transcription

Recherche sur le document

Titre	Oui
Plein texte	Oui
Illustrations	
Autres	Par auteur

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	Fonctionnalité « progress » où on peut suggérer une correction à faire directement dans le langage source puis soumission au webmaster
Newsgroup	Suscribe Ccel-list à envoyer à majordomo@calvin.edu
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Oui
Webmaster	hplantin@calvin.edu

Identification du site

Nom du site	<i>Digital Scriptorium</i>
Editeur	U.C. Berkeley (Columbia University, San Francisco State University...)
URL	http://sunsite.berkeley.edu/Scriptorium/
Date du site	2000 (projet depuis 1997)

Base de données

Type	Projet commun à plusieurs universités américaines de base de données d'image (manuscrits du Moyen-Age et de la Renaissance)
Nombre de données	Environ 8500 images numérisées (2400 docs, non complets, environ une dizaine d'images numérisées par codex)
Mode de numérisation	Mode image
Couleurs	Oui
Mise à jour	16.04.01
Enrichie	Excellent 'Data dictionary', qui présente les métadonnées historiques, physiques et bibliographiques du document et de l'image (http://sunsite.berkeley.edu/Scriptorium/datadic5.html). Utilise l'indexation iconographique d'Iconclass.
Niveau de description des métadonnées	Niveau 1 : Scriptorium utilise la TEI (text encoding initiative,) pour coder les transcriptions des textes (http://www.tei-c.org), et participe à la création une DTD (document type definition) pour coder les métadonnées des manuscrits en XML
Format, logiciel de lecture	Jpeg, gif

Outils de visualisation

Vignettage	Oui
Mosaïque	
Affich. annotations	Oui, à côté de la vignette, on a le numéro de folio et une case « caption » (titre de l'image) et où sont notées les métadonnées de l'exemplaire (par exemple, l'état du folio, ou une description de l'image)
Fonds	
Luminosité	
Grossissement	3 tailles : petit, moyen, grand
Autres	

Recherche sur le document

Titre	Oui
Plein texte	Oui
Illustrations	Oui
Autres	Recherche sur de très nombreuses métadonnées : côte, reliure, provenance, pays, ville, titre, incipit, scribe, décoration, artiste, date, etc. Il existe un autre mode de recherche, avec le logiciel Luna Insight

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui, en 1 ex. et uniquement les petites images (pour les profs :20 ex.).
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	

Documents associés

Enquêtes	
Liens	
Webmaster	manager@sunsite.berkeley.edu

Identification du texte

Nom du site	<i>Digitale bibliotheek</i>
Editeur	Koninklijke bibliotheek (Pays-Bas)
URL	http://www.kb.nl/kb/resources/frameset_galerie-en.html
Date du site	

Base de données

Type	Présente des projets de numérisation et héberge des sites comme le « Medieval Illuminated Manuscripts Rijksmuseum Meermanno-Westreenianum » ou encore « A Hundred Highlights from the Koninklijke Bibliotheek »
Nombre de données	Plusieurs manuscrits enluminés
Mode de numérisation	Mode image
Couleurs	Oui
Mise à jour	
Enrichie	
Niveau de description des métadonnées	Niveau 2. Notice complète et historique du document
Format, logiciel de lecture	Jpeg

Outils de visualisation

Vignettage	Oui
Mosaïque	
Affich. annotations	Oui, des notes explicatives apparaissent avec les vignettes
Fonds	
Luminosité	
Grossissement	Oui, zoom in et out. Système de languette pour tirer l'image vers soi
Autres	

Recherche sur le document

Titre	Oui
Plein texte	
Illustrations	Oui, avec le système iconclass
Autres	Recherche par sujet, recherche « experte » (titre, auteur, mot-clés), recherche par l'indexation Iconclass

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	
Autres	

Travail collaboratif

Messagerie	Propose un guestbook où on peut noter ses impressions et où l'on a accès aux réflexions des autres
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	
Webmaster	

Identification du site

Nom du site	(<i>Dunhuang Project</i>) International <i>Dunhuang Project</i>
Editeur	British Library (Oriental and India Office Collections)
URL	http://idp.bl.uk
Date du site	1993 (projet) ; 2002

Base de données

Type	Manuscrits et documents imprimés des collections de la British Library, venant d'Asie Centrale (région de Dunhuang, route de la Soie) ; site bilingue anglais-chinois ; projet de base de données internationale sur Dunhuang ; accès libre à condition de remplir un formulaire (nom et mot de passe : guilleray / debora)
Nombre de données	15 000 images (28 000 titres référencés)
Mode de numérisation	Image
Couleurs	Oui
Mise à jour	04.01.2002
Enrichie	Oui : détails sur le doc., le texte, bibliographie, contexte historique et géographique de Dunhuang, illustrations ; À l'intention des chercheurs, conservateurs, étudiants
Niveau de description des métadonnées	Niveau 2 (cote du document, région d'origine, expédition qui l'a ramené, taille, support)
Format, logiciel de lecture	Jpeg

Outils de visualisation

Vignettage	Oui
Mosaïque	
Affich. annotations	Non : pas en parallèle avec le texte
Fonds	
Luminosité	
Grossissement	Oui : clic gauche
Autres	

Recherche

Titre	Oui
Plein texte	
Illustrations	
Autres	Sur les métadonnées, par mots-clés : langue, écriture/ site / sujet / forme et taille du document (en cours) ; par catalogue et par cote

Travail personnel sur le document

Corpus personnel	Déchargement sur autorisation
Fichier personnel	Pour les chercheurs : « Your details », avec un mot de passe. Ce dossier permet de garder en mémoire les recherches des sessions précédentes.
Impression	Oui
Autres	

Travail collaboratif

Messagerie	Oui : à la fin du site, suggestions, commentaires, après avoir inscrit son e-mail et son adresse
Newsgroup	Non
Banque de données	Oui, 2 possibilités : - ajouter au catalogue un commentaire (sur les métadonnées), après s'être identifié. La date est notée automatiquement. - faire inclure ses propres travaux (menu "help" : "Adding your own research") au niveau des champs concernés, à condition d'être accrédité. Cette possibilité s'adresse sans doute aux chercheurs chinois (le site est en effet bilingue), susceptibles de faire des commentaires sur les erreurs dans les données, voire de les corriger.
Autres	

Documents associés

Enquêtes	Sur les chercheurs, par le biais de l'inscription pour le fichier personnel : université, discipline, thème de recherche, langues pratiquées, visites en Asie centrale, liens pertinents pour le site
Liens	Oui (par ex. vers des sites de transcription en caractères chinois)
Webmaster	Susan Whitfield ; Colin Chinnery (pour ajouter ses propres travaux sur le site)
Autres	Lettre d'information en ligne (1994-2001)

Identification du site

Nom du site	<i>Early Canadiana Online</i> (Notre Mémoire en ligne)
Editeur	Institut canadien de microreproductions historiques (ICMH), BN du Canada, BU de Laval, BU de Toronto
URL	http://www.canadiana.org/eco/francais/collections.html
Date du site	1998-2001

Base de données

Type	Journaux et imprimés depuis l'arrivée des premiers pionniers au Canada (XVI-XIXe). Une partie de la base n'est accessible qu'avec un login
Nombre de données	1.037.000 pages du patrimoine imprimé canadien
Mode de numérisation	Mode image
Couleurs	Noir et blanc
Mise à jour	Mai 2002
Enrichie	
Niveau de description des métadonnées	Niveau 2. En plus, identification par n°ICMH : notes sur l'exemplaire dans « cible » : notes techniques (numérisation) et bibliographiques, présence d'illustrations en couleur, état matériel du document : qualité de la lecture (transparence, décoloration), taux de réduction de la numérisation...
Format, logiciel de lecture	2 possibilités de lecture : une première image gif et on peut charger ensuite une version Pdf avec les fonctionnalités d'acrobat reader

Outils de visualisation

Vignettage	
Mosaïque	Oui (acrobat)
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Oui, loupe (acrobat)
Autres	Fonctionnalités d'acrobat : Main, visualisation en cascade, rotation de l'image, possibilité de déplacement avec des vues miniatures...

Recherche sur le document

Titre	Oui
Plein texte	Oui
Illustrations	
Autres	Sujet, éditeur, auteur, bibliographie, affichage des index (browse : bouquinage), limitations possibles (anglais, français...)

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	
Webmaster	cihmcmh@nlc-bnc.ca

Identification du site

Nom du site	<i>Early English Books Online (EEBO)</i>
Editeur	University of Michigan , Oxford University , Council on Library and Information Resources (CLIR), ProQuest Information and Learning
URL	http://wwwlib.umi.com/eebo
Date du site	

Base de données

Type	Textes en anglais (1475-1700) : Pollard & Redgrave's Short-Title Catalogue (1475-1640), Wing's Short-Title Catalogue (1641-1700), and the Thomason Tracts (1640-1661). numérisés par « ProQuest Information and Learning » accès réservé, pas de possibilité d'inscription ; site gratuit de démonstration
Nombre de données	125 000 titres prévus !
Mode de numérisation	Image
Couleurs	Non : noir et blanc
Mise à jour	
Enrichie	
Niveau de description des métadonnées	Niveau 2 (notice MARC + signature, notes sur l'exemplaire, le colophon)
Format, logiciel de lecture	Adobe acrobat en option

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Oui
Autres	Déplacement par n° de page

Recherche

Titre	
Plein texte	
Illustrations	Oui, notamment les marques d'imprimeur et les colophons
Autres	

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui : mauvaise qualité de l'image
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	
Webmaster	

Identification du site

Nom du site	<i>Early English Books Online Text Creation Partnership (EEBO/TCP)</i>
Editeur	University of Michigan
URL	http://www.lib.umich.edu/eebo/
Date du site	(2000)

Base de données

Type	Livres imprimés en anglais (1475-1700) du catalogue de la bibliothèque de l'université du Michigan. But : proposer la transcription des textes numérisés en mode image
Nombre de données	8 livres sur le site de démonstration du prototype (à terme, dans 5 ans, 25 000 volumes sur les 125 000 de EEBO)
Mode de numérisation	Image / texte avec OCR (balises SGML)
Couleurs	Non
Mise à jour	
Enrichie	
Niveau de description des métadonnées	Niveau 2
Format, logiciel de lecture	Adobe est fourni

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	
Autres	

Recherche

Titre	Oui
Plein texte	Oui
Illustrations	
Autres	Booléenne, de proximité, de citations, avec index de mots

Travail personnel sur le document

Corpus personnel	
Fichier personnel	? accès réservé aux membres de l'université ou accrédités
Impression	Oui (mauvaise qualité)
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	
Webmaster	Hillary Nunn (hunn@umich.edu) / dlps-info@umich.edu /

Identification du site

Nom du site	<i>Early Manuscripts at Oxford University</i>
Editeur	Université d'Oxford
URL	http://image.ox.ac.uk/
Date du site	2000 (projet 1995-2000)

Base de données

Type	Manuscrits celtiques et médiévaux (9 ^e -16 ^e siècles) (11 projets de bibl. numériques)
Nombre de données	80 titres complets
Mode de numérisation	Image
Couleurs	Oui
Mise à jour	
Enrichie	Non
Niveau de description des métadonnées	Niveau 3 (cotes et localisations des manuscrits)
Format, logiciel de lecture	Gif (vignettes), jpeg (grossissements)

Outils de visualisation

Vignettage	Oui
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Non. Très bonne résolution, mais format trop grand, impossible à réduire
Autres	

Recherche

Titre	Non
Plein texte	
Illustrations	
Autres	Classement par bibliothèque et par cote, mais pas de recherche possible

Travail personnel sur le document

Corpus personnel	
Fichier personnel	Non : copie image par image
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	
Webmaster	Matthew J. Dovey ; Katherine Ferguson .

Identification du site

Intitulé	<i>Early Virginia Religious Petitions</i>
Editeur	Bibliothèque du Congrès / Bibliothèque de Virginie
URL	http://memory.loc.gov/ammem/repehtml/repehome.html
Date du site	(Projet 1996-1998 : American Memory)

Base de données

Type	Archives (1774-1802), en Virginie, et cartes du 18 ^e s.
Nombre de données	423 documents
Mode de numérisation	Image / texte par OCR sans corrections
Couleurs	Oui
Mise à jour	Septembre 1999
Enrichie	Oui : analyse de l'acte
Niveau de description des métadonnées	Niveau 3 (analyse diplomatique : date, lieu, analyse de l'acte)
Format, logiciel de lecture	Jpeg, tiff, mpeg, mp3, quicktime, Realmedia, SGML, MrSID, WAV : pas de problème de lecture en particulier

Outils de visualisation

Vignettage	Oui
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Oui / option d'image haute résolution
Autres	

Recherche

Titre	Non
Plein texte	Oui, pour pallier l'absence de renseignements bibliographiques par la recherche sur des expressions du texte
Illustrations	
Autres	Par mot-clé, par lieu, par tranche chronologique

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Fait partie de <i>American Memory</i>
Webmaster	American Memory Help Desk

Identification du site

Nom du site	<i>Electronic text center</i>
Editeur	Université de Virginie
URL	http://etext.lib.virginia.edu/
Date du site	Projet de 1992

Base de données

Type	Héberge de nombreuses bases de données dont la plupart ne sont consultables qu'en passant par un proxy serveur. En ligne : <i>Early American Fiction Project</i> , d'autres sur CD Rom. Caractères chinois, russe, hébreux, etc.
Nombre de données	70000 textes littéraires, 350000 images (illustrations, couvertures, manuscrits, images de journaux, de livres anciens...)
Mode de numérisation	Mode texte et mode image
Couleurs	Oui
Mise à jour	
Enrichie	Commentaires, notes, bibliographies sur les oeuvres
Niveau de description des métadonnées	Niveau 3. Notices biographiques, bibliographies sur chaque œuvre. Numérisation de la tranche et de la reliure des documents
Format, logiciel de lecture	Jpeg, tiff

Outils de visualisation

Vignettage	Oui
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Oui, une fois
Autres	

Recherche sur le document

Titre	Oui
Plein texte	Oui
Illustrations	Oui, on obtient une liste des thèmes illustrés, et si on clique, on peut visualiser plusieurs images de manuscrits différents traitant du même thème
Autres	Fonctionnalité de « browse » par page

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	Rubrique 'contact us' avec possibilité de faire des commentaires
Newsgroup	
Banque de données	
Autres	Grâce à sa politique de formation des usagers, de nombreuses bases de données ont été créées par des usagers (voir rubrique « user projects »)

Documents associés

Enquêtes	
Liens	Oui
Webmaster	etext@virginia.edu

Identification du site

Nom du site	<i>ELIOHS : Electronic Library Of HiStoriography</i>
Editeur	Rolando Minuti (Florence) et Guido Abbattista (Turin)
URL	http://www.unifi.it/riviste/eliohs/
Date du site	Mars 1996

Base de données

Type	Thème : histoire des idées en Europe du XVI ^e au XIX ^e siècles ; Portail vers d'autres sites, et numérisation de ses propres livres ; catalogue : du 16 ^e au 20 ^e s. Pour le 16 ^e s., 3 liens vers les livres qui sont sur d'autres sites.
Nombre de données	Non indiqué
Mode de numérisation	Texte
Couleurs	Non
Mise à jour	Février 2002
Enrichie	Oui : répertoires de bibliothèques numériques, notices biographiques avec cartes géographique des villes européennes
Niveau de description des métadonnées	Niveau 4 : limitée à l'auteur, au titre, à la date de publication
Format, logiciel de lecture	Html pour les textes édités par <i>ElioHS</i>

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	
Autres	

Recherche

Titre	
Plein texte	
Illustrations	
Autres	Catalogue par ordre alphabétique d'auteurs, mais pas de recherche par nom d'auteur

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Oui, car le principe du site est d'être un portail vers d'autres bibliothèques numériques
Webmaster	

Identification du site

Nom du site	<i>Emblem Project Utrecht</i>
Editeur	Université d'Utrecht, département de littérature et langue hollandaise
URL	http://www.let.uu.nl/emblems/html/index.html
Date du site	2002

Base de données

Type	Livres d'emblèmes hollandais du 17 ^e siècle
Nombre de données	3 livres numérisés (en projet : 25 livres)
Mode de numérisation	Mode image
Couleurs	Oui
Mise à jour	Février 2002, fréquents ajouts (transcriptions, améliorations de lecture...)
Enrichie	Commentaires, index, bibliographies
Niveau de description des métadonnées	Niveau 3. Adaptation personnelle de la TEI
Format, logiciel de lecture	Jpeg

Outils de visualisation

Vignettage	Oui
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Oui
Autres	

Recherche sur le document

Titre	Oui
Plein texte	Oui
Illustrations	Oui, recherche par motif pictural
Autres	Recherche par sources, par sujets, liste des occurrences avec liens dans le texte

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	Rubrique 'feedback' : suggestions et commentaires
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Oui, dans une bibliographie
Webmaster	Els.Stronks@let.uu.nl

Identification du site

Nom du site	<i>The English Emblem Book project</i>
Editeur	Pennstate University Libraries' Electronic text center
URL	http://emblem.libraries.psu.edu/
Date du site	

Base de données

Type	Livres d'emblèmes anglais XVI-XVII ^e
Nombre de données	9 livres numérisés
Mode de numérisation	Mode image
Couleurs	Oui
Mise à jour	Avril 2002
Enrichie	Chronologie, bibliographie
Niveau de description des métadonnées	Niveau 4. Auteur, titre, éditeur et année d'édition
Format, logiciel de lecture	Jpeg, utilisation du logiciel Mapedit qui permet de cliquer sur une partie du texte ou sur une image et cela renvoie à l'image ou au texte correspondant.

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	
Autres	

Recherche sur le document

Titre	
Plein texte	
Illustrations	
Autres	Recherche de sites web (applet ou mini programme java)

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	Serveur de discussion sur inscription
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Liens avec 4 bases de données produites par le même éditeur : « English poetry », « English Verse drama », « African American poetry » et « American poetry »
Webmaster	mxc52@psu.edu

Identification du site

Nom du site	<i>Galileo Galilei</i>
Editeur	Biblioteca Nazionale Centrale, Florence ; Istituto e Museo di Storia della Scienza, Florence ; Max Planck Institute for the History of Science, Berlin
URL	http://www.mpiwg-berlin.mpg.de/Galileo_Prototype/INDEX.HTM
Date du site	1999

Base de données

Type	Reproduction du manuscrit <i>Opere</i> de Galilée
Nombre de données	Fac-similé d'environ 300 pages
Mode de numérisation	Mode image
Couleurs	Oui
Mise à jour	Mars 2002
Enrichie	Bibliographie, historique du manuscrit
Niveau de description des métadonnées	Niveau 1. Tout un travail de description du manuscrit
Format, logiciel de lecture	Jpeg et gif

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	Oui, annotations
Fonds	
Luminosité	
Grossissement	Oui, ; avec le niveau « high resolution »
Autres	Trois niveaux de visualisation de l'image « overview level », « working level » et « high resolution ». Transcriptions

Recherche sur le document

Titre	
Plein texte	Oui, index des mots, des dessins, des calculs
Illustrations	
Autres	

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	
Webmaster	

Identification du site

Nom du site	<i>Gallica</i>
Editeur	Bibliothèque nationale de France
URL	http://gallica.bnf.fr/
Date du site	Octobre 1997

Base de données

Type	Textes, images et sons, du Moyen Age au XXe
Nombre de données	80 000 documents numérisés, libres de droit.
Mode de numérisation	Mode image et mode texte (base de données Frantext de l'INALF)
Couleurs	Oui, pour certaines images
Mise à jour	Mars 2002
Enrichie	
Niveau de description des métadonnées	Niveau 2 : notice en format Marc
Format, logiciel	Format TIFF et format PDF (acrobat reader)

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	Table des matières à gauche
Fonds	
Luminosité	
Grossissement	Oui
Autres	main

Recherche sur le document

Titre	Oui
Plein texte	Oui, dans les œuvres en mode texte et en mode image
Illustrations	
Autres	Recherche sujet, auteur ; dossiers et parcours thématiques

Travail personnel sur le document

Corpus personnel	Oui
Fichier personnel	annotation, surlignage, pose de signets et bloc note possibles pour les PABN (postes d'accès à la bib. numérique), sur place à Tolbiac
Impression	Oui
Autres	

Travail collaboratif

Messagerie	gallica@bnf.fr
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	
Webmaster	

Identification du site

Nom du site	<i>The HUMI Project</i>
Editeur	British Library, Early Printed Collections (Londres) ; Keio University, (Tokyo) et NTT pour la numérisation
URL	http://prodigi.bl.uk/gutenbg/default.asp
Date du site	Novembre 2000

Base de données

Type	Deux bibles de Gutenberg : la numérisation veut permettre de les comparer plus facilement
Nombre de données	2 manuscrits : King's copy, Grenville copy
Mode de numérisation	Image
Couleurs	Oui
Mise à jour	11.2000
Enrichie	Oui : fonctionne comme un cd-rom, avec des encarts sur l'époque etc
Niveau de description des métadonnées	Niveau 2 : métadonnées très détaillées sur les deux exemplaires (reliure, rubrication, décoration, notes manuscrites)
Format, logiciel de lecture	Jpeg

Outils de visualisation

Vignettage	Oui
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Oui (clic gauche, puis retour écran précédent pour diminuer)
Autres	

Recherche

Titre	Oui : par ceux des livres de la Bible
Plein texte	Non
Illustrations	Non
Autres	

Travail personnel sur le document

Corpus personnel	
Fichier personnel	Non : copie page par page
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Depuis la British Library
Webmaster	Graham Jefcoate ; Kristian Jensen (Head of Incunabula)

Identification du site

Nom du site	<i>Internet Library of Early Journals (ILEJ)</i>
Editeur	Universités de Birmingham , Leeds , Manchester and Oxford
URL	http://www.bodley.ox.ac.uk/ilej/
Date du site	1997

Base de données

Type	Journaux du XVIII ^e et du XIX ^e siècles. Inscription préalable à la consultation
Nombre de données	
Mode de numérisation	Image
Couleurs	Non
Mise à jour	18 sept. 1998
Enrichie	
Niveau de description des métadonnées	Niveau 4 : titre de la revue, ses dates, et son contenu
Format, logiciel de lecture	Gif

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	
Autres	Liens vers les titres des revues, les n ^o s et les titres des articles, puis les pages

Recherche

Titre	Oui
Plein texte	Oui, par OCR
Illustrations	
Autres	Auteur

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	6 contacts
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	Oui, au moment de l'inscription (en annexe)
Liens	
Webmaster	6

Identification du site

Nom du site	<i>Internet Public Library</i>
Editeur	University of Michigan School of Information
URL	http://www.ipl.org/reading/books/
Date du site	1995

Base de données

Type	Textes en anglais, classiques ou récents (philosophie, histoire, médecine, arts, sciences, religion...) sur IPL et portail vers d'autres bases de données
Nombre de données	Plusieurs milliers de titres
Mode de numérisation	Mode image et mode texte
Couleurs	Noir et blanc
Mise à jour	
Enrichie	
Niveau de description des métadonnées	Niveau 3
Format, logiciel de lecture	Jpeg, Png; pour certains livres, installation de logiciel de lecture EBRARY

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Oui, avec Ebrary
Autres	Affichage ou non de la table des matières avec Ebrary

Recherche sur le document

Titre	Oui
Plein texte	
Illustrations	
Autres	Par auteur, par la Dewey

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui, payant sur Ebrary
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Oui
Webmaster	readroom@ipl.org

Identification du site

Nom du site	<i>Making of America (MOA)</i>
Editeur	Université de Cornell (USA)
URL	http://cdl.library.cornell.edu/moa
Date du site	1999-2000

Base de données

Type	Périodiques et livres américains (1815-1926), a priori libres de droits. Au sein des collections numériques de Cornell, 12 bases, dont Making of America (7 en noir et blanc, 5 en couleur) (ex : manuscrits de sagas islandaises, en couleur)
Nombre de données	267 livres et 100 000 articles de 955 journaux (905 750 pages)
Mode de numérisation	Image
Couleurs	Non
Mise à jour	2002
Enrichie	
Niveau de description des métadonnées	Niveau 2
Format, logiciel de lecture	Gif

Outils de visualisation

Vignettage	Non (liens classés par périodiques, par n°s, par articles)
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	
Grossissement	Oui
Autres	

Recherche

Titre	Oui
Plein texte	Oui : reconnaissance de termes par OCR ; critères de proximité, opérateurs booléens, dates
Illustrations	
Autres	

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui, page par page
Autres	

Travail collaboratif

Messagerie	Reference question? Comments?
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Non (pourtant, fait partie de <i>American Memory</i>)
Webmaster	

Identification du site

Nom du site	<i>Mateo</i>
Editeur	Université de Mannheim (Allemagne)
URL	http://www.uni-mannheim.de/mateo/
Date du site	1998

Base de données

Type	3 bases de données : <i>MARABU</i> (livres anciens et manuscrits de la bibliothèque), <i>CAMENA</i> (poésie latine des XVI, XVII et XVIII ^e siècles allemands), <i>DISCOTHECA RARA</i> (propose des CD-Roms de livres anciens numérisés)
Nombre de données	
Mode de numérisation	Mode image
Couleurs	Oui
Mise à jour	Avril 2002
Enrichie	Oui, transcriptions selon TEI (text encoding initiative), indexs, bibliographies en html
Niveau de description des métadonnées	Niveau 2. Notice complète et texte décrivant l'histoire du document
Format, logiciel de lecture	Jpeg, gif

Outils de visualisation

Vignettage	Oui
Mosaïque	
Affich. annotations	Affichage de l'introduction à côté de la vignette de la page de titre
Fonds	
Luminosité	
Grossissement	Oui, une fois en cliquant sur les vignettes, image de très bonne qualité
Autres	

Recherche sur le document

Titre	
Plein texte	
Illustrations	
Autres	

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	mateo@uni-mannheim.de
Newsgroup	
Banque de données	
Autres	Rubrique : « users's voice »

Documents associés

Enquêtes	
Liens	Oui
Webmaster	webmaster@rz.uni-mannheim.de

Identification du site

Nom du site	<i>Medic@</i>
Editeur	BIUM de l'université de Paris 5
URL	http://www.bium.univ-paris5.fr/histmed/medica.htm
Date du site	Octobre 2000

Base de données

Type	Rééditions de textes anciens en histoire de la médecine, collections de la BIUM : - l'ensemble des éditions princeps du XVI ^e siècle (1525 à 1595) des écrits d'Hippocrate et l'édition de référence d'Emile Littré (1839-1861). Même chose pour Galien. - textes fondateurs par disciplines médicales (ophtalmologie, odontologie, etc)
Nombre de données	105 volumes pour un total de près de 65.000 pages.
Mode de numérisation	Image
Couleurs	Oui
Mise à jour	Avril 2002
Enrichie	Oui, résumé des ouvrages dans les notices et présentation historique de l'édition ; possibilité de télécharger la police grecque accentuée « Greek Old Face ».
Niveau de description des métadonnées	Niveau 3 (titre, auteur, lieu éd., date, foliotation ou pagination particulière, cote)
Format, logiciel de lecture	Jpeg

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	
Fonds	Non : l'effet de transparence de la page demeure pour l'édition de 1525 d'Hippocrate.
Luminosité	
Grossissement	Non
Autres	Affichage du sommaire avec les liens vers les pages

Recherche

Titre	Non
Plein texte	
Illustrations	
Autres	Catalogue complet des ouvrages numérisés, ou liste limitée aux ouvrages « voisins » (du même auteur)

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	
Webmaster	ferreira@bium.univ-paris5.fr

Identification du site

Nom du site	<i>Menasseh Ben Israel</i>
Editeur	Bibliotheca Rosenthaliana, Amsterdam University
URL	http://www.uba.uva.nl/nl/collecties/rosenthaliana/menasseh/boeken.html
Date du site	1997

Base de données

Type	Collection d'ouvrages en hébreu imprimés par Menasseh au XVII ^e
Nombre de données	70 ouvrages
Mode de numérisation	Mode image
Couleurs	Noir et blanc
Mise à jour	2001
Enrichie	Oui : bibliographie, lettres autographes, liens vers des articles
Niveau de description des métadonnées	Niveau 2
Format, logiciel de lecture	Gif et jpeg, tif. Pour un des ouvrages présentés <i>Minhagim</i> , téléchargement de pages numérisées grâce au logiciel Imaging développé par Kodak pour Windows avec de nombreuses fonctionnalités

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	Oui
Fonds	
Luminosité	
Grossissement	Avec Imaging, zoom, possibilité d'afficher en miniature, ou page et miniature
Autres	

Recherche sur le document

Titre	Oui
Plein texte	
Illustrations	
Autres	

Travail personnel sur le document

Corpus personnel	Oui, on peut télécharger une image d'un livre (<i>Minhagim</i>) et travailler en traitement de texte, inscrire différentes notes...
Fichier personnel	Oui
Impression	Oui
Autres	

Travail collaboratif

Messagerie	http://www.uba.uva.nl/en/reactions/
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	
Webmaster	ootjers@uba.uva.nl
Autres	Recherche sur les OCR en hébreu, mais en hollandais donc pas compréhensible

Identification du site

Nom du site	<i>Mnemosyne</i>
Editeur	Editeur inconnu, mais travaille en coopération avec la Koninklijke Bibliotheek (La Haye), l'université d'Aberdeen
URL	http://www.mnemosyne.org/
Date du site	2000-2001

Base de données

Type	Banque de données d'images avec une indexation qui fonctionne avec le logiciel Iconclass. Le fonds vient pour une grande partie de la Koninklijke Bibliotheek (La Haye) : livres anciens, manuscrits, bestiaires, emblèmes...
Nombre de données	Une dizaine de livres
Mode de numérisation	Mode image
Couleurs	Oui
Mise à jour	
Enrichie	Lien avec le dictionnaire latin édité sur le site <i>Perseus</i> (voir plus haut), en cliquant sur <u>Lewis and Short</u>
Niveau de description des métadonnées	Niveau 3. Auteur, titre, différentes éditions qui ont servi à établir les transcriptions
Format, logiciel de lecture	Jpeg

Outils de visualisation

Vignettage	Oui pour certaines des images
Mosaïque	
Affich. annotations	Affichage de la transcription selon différents modes, soit automatiquement, soit en cliquant sur un lien
Fonds	
Luminosité	
Grossissement	Oui, mais cela dépend des bibliothèques qui ont numérisé
Autres	

Recherche sur le document

Titre	
Plein texte	
Illustrations	Logiciel Iconclass qui indexe les images par mot-clés sur le principe des répertoires iconographiques hiérarchisés
Autres	

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	<i>Perseus</i> , dictionnaire d'art et d'architecture du projet Getty
Webmaster	mnemosyne@planet.nl

Identification du site

Nom du site	<i>Perseus</i>
Editeur	Gregory Crane, Université de Tufts
URL	http://www.perseus.tufts.edu/
Date du site	Copyright 1997

Base de données

Type	Textes grecs, latins, Renaissance anglaise, Collection sur Londres,...
Nombre de données	important
Mode de numérisation	Mode texte plus illustrations et cartes
Couleurs	Noir et blanc
Mise à jour	fréquente
Enrichie	Index, dictionnaire, cartes et plans, analyse morphologique des mots..
Niveau de description des métadonnées	Niveau 3 : notice basique, avec juste les n° ISBN et OCLC
Formats, logiciels	Jpeg
Autres	Beaucoup d'outils de recherche, grammaire, etc

Outils de visualisation

Vignettage	
Mosaïque	
Affich. annotations	Oui
Fonds	
Luminosité	
Grossissement	Oui, pour les images, cartes et plans
Autres	

Recherche sur le document

Titre	Oui
Plein texte	Oui
Illustrations	Oui
Autres	Limites au type de doc., à la langue. Possibilité de recherche en navigant sur une carte géographique

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	Oui
Autres	

Travail collaboratif

Messagerie	
Newsgroup	perseusproject@yahogroups.com
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Oui
Webmaster	webmaster@perseus.tufts.edu

Identification du site

Nom du site	<i>Studiolum</i>
Editeur	Tamàs Sajo
URL	http://www.studiolum.com/frame.htm
Date du site	Dernière version du CD-Rom 2000-2001

Base de données

Type	Sur le CD-Rom à vendre : Bibliothèque de la Renaissance (XVI-XVIIe) virtuelle : livres d'emblèmes, bestiaires, herbiers, traités d'art, collections de proverbes, miroirs des princes, lexiques, dictionnaires... Sur le site, quelques livres exemples.
Nombre de données	
Mode de numérisation	Mode texte et intercallation d'images
Couleurs	Illustrations en noir et blanc
Mise à jour	
Enrichie	Commentaires sur les auteurs, les thèmes, histoire de l'édition et de l'exemplaire
Niveau de description des métadonnées	Niveau 1
Format, logiciel de lecture	Logiciel de lecture Folio views, numérisation des images GIF

Outils de visualisation

Vignettage	Oui, sur dossier perso on peut faire ce qu'on veut avec les images
Mosaïque	Oui
Affich. annotations	Oui, sur dossier perso
Fonds	
Luminosité	
Grossissement	Oui
Autres	On peut afficher la version <i>Studiolum</i> et son dossier perso à côté, soit à la verticale, soit à l'horizontale. Sur le dossier perso, on peut ajouter des fichiers, des images, des logiciels. En fait sur le dossier perso, on peut agir comme on veut et faire intervenir de nombreuses fonctionnalités pour modifier le texte ou les images

Recherche sur le document

Titre	Oui
Plein texte	Oui (index des mots, liste et calcul des occurrences d'un mot, navigation entre les occurrences, possibilités de limitations de cette recherche au latin, grec...)
Illustrations	
Autres	Index des mots, fréquence des occurrences

Travail personnel sur le document

Corpus personnel	
Fichier personnel	On peut créer autant de fichiers qu'on veut « shadow files » avec le texte identique mais modifiable à volonté, avec annotations...
Impression	Oui
Autres	

Travail collaboratif

Messagerie	studiolum@studiolum.com
Newsgroup	
Banque de données	
Autres	Lettre en ligne

Documents associés

Enquêtes	
Liens	Oui
Webmaster	

Identification du site

Nom du site	The Universal Library, Antique books
Editeur	Carnegie Mellon University
URL	http://www.antiquebooks.net/
Date du site	

Base de données

Type	Textes littéraires et images de botaniques
Nombre de données	Livres conservés sur différents serveurs, 5 livres sur le serveur de la bibliothèque de l'Université (Digital Library) ont des fonctionnalités développées
Mode de numérisation	Mode image
Couleurs	Oui
Mise à jour	
Enrichie	
Niveau de description des métadonnées	Niveau 4. Auteur, titre, année
Format, logiciel de lecture	Jpeg, gif

Outils de visualisation

Vignettage	Oui pour les illustrations
Mosaïque	
Affich. annotations	
Fonds	
Luminosité	Oui
Grossissement	Limité pour certains, d'autres ont une loupe
Autres	Main (page suivante), possibilité de rotation de l'image

Recherche sur le document

Titre	Oui
Plein texte	
Illustrations	
Autres	Recherche par mots-clés, auteurs, sujets

Travail personnel sur le document

Corpus personnel	
Fichier personnel	
Impression	
Autres	

Travail collaboratif

Messagerie	Oui, rubrique « Reader comments »
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Sur le site de l'université, répertoire de liens vers des bibliothèques numériques
Webmaster	bookmaster@antiquebooks.net

Identification du site

Nom du site	<i>The William Blake Archives</i>
Editeur	Universités de Rochester, Californie et Caroline du Nord ; sponsorisé par la Bibliothèque du Congrès
URL	http://www.blakearchive.org/main.html
Date du site	Copyright 1996-2002

Base de données

Type	Œuvres et lettres de William Blake et œuvres sur cet artiste
Nombre de données	19 livres illustrés, trentaine de manuscrits, lettres autographes, peintures, gravures
Mode de numérisation	Mode image et mode texte (transcription)
Couleurs	Oui
Mise à jour	Avril 2002
Enrichie	Biographie, glossaire, chronologie (en projet), bibliographie
Niveau de description des métadonnées	Niveau 1. Notice extrêmement complète qui reprend toutes les données d'exemplaires similaires à celles de la TEI
Format, logiciel de lecture	Tiff, logiciel Inote

Outils de visualisation

Vignettage	Oui pour les illustrations
Mosaïque	
Affich. annotations	Oui, grâce au logiciel Inote
Fonds	
Luminosité	
Grossissement	Oui
Autres	

Recherche sur le document

Titre	Oui
Plein texte	Oui, dans les transcriptions et sur les mots clés décrivant les images
Illustrations	Oui, indexation des images (similaire au système Iconclass)
Autres	

Travail personnel sur le document

Corpus personnel	Oui
Fichier personnel	Oui
Impression	Oui
Autres	Formulaire de demande de reproduction des images et des textes

Travail collaboratif

Messagerie	blake@jefferson.village.virginia.edu ; rubrique « User comments »
Newsgroup	
Banque de données	
Autres	

Documents associés

Enquêtes	
Liens	Oui
Webmaster	