

Chaîne de numérisation et bibliothèque virtuelle : de l'acquisition des images à la consultation sur le web

Aurélie Blanc

Sous la direction de Véronique Lacroix
Ingénieur d'études à la bibliothèque de la Maison de l'Orient et
de la Méditerranée (CNRS – Université Lumière Lyon 2).

Tuteur de stage : Franck Lebourgeois
Maître de conférences à l'INSA de Lyon

Résumé :

La Maison de l'Orient et de la Méditerranée a engagé un programme de numérisation d'ouvrages libres de droits couvrant les domaines de recherche de ses équipes. Elle souhaite diffuser ces ouvrages auprès d'un public élargi grâce à la mise en place d'une bibliothèque virtuelle consultable sur internet. Le présent rapport expose les différents aspects de la mission qui m'a été confiée dans le cadre de ce projet : pratiquer la chaîne de numérisation, réfléchir aux modalités d'exploitation du fonds numérisé et assurer la réalisation technique d'une solution permettant sa diffusion en ligne.

Mots clés :

Numérisation de documents ; bibliothèque numérique ; bibliothèque virtuelle ; bibliothèque électronique ; chaîne de numérisation ; système de gestion de bases d'images ; site web

Abstract :

The *Maison de l'Orient et de la Méditerranée* has started a digitization programme of works free from royalties dealing with the research fields of its teams. It would like to broadcast these works towards a widened audience thanks to a virtual library available on the internet. This report presents the various aspects of the mission I was entrusted with : using the digitizing system, considering how to make use of the digitized books and ensuring the realization of a technical solution allowing their online broadcasting.

Keywords :

Document digitization ; digital library ; virtual library ; electronic library ; digitizing system ; stock image management system ; web site

Remerciements

Tous mes remerciements à Véronique Lacroix pour son accueil, ses conseils et la confiance qu'elle m'a témoignée en me chargeant de cette mission de stage.

Je tiens également à remercier Marjorie Burghart, Pierre Philippe, Miled Rousset et Denis Duperray pour le temps qu'ils ont bien voulu me consacrer.

Mes remerciements s'adressent aussi à tous les membres de la bibliothèque et de la plateforme de numérisation, pour leur sympathie et la chaleur de leur accueil.

Sommaire

INTRODUCTION	5
I. CONTEXTE ET PRÉSENTATION DU PROJET	6
1. La Maison de l’Orient et de la Méditerranée	6
2. Le projet de numérisation	8
3. La mission de stage	10
II. MAÎTRISER LA CHAÎNE DE NUMÉRISATION	12
1. Description matérielle de la plateforme	12
1.1. Les stations d’acquisition et les serveurs.....	13
1.2. Les logiciels.....	14
1.3. Les outils de calibrage	20
1.4. Le système de gravure	21
1.5. Le réseau informatique	22
2. Pratiquer la chaîne de numérisation	22
2.1. Acquisition des images.....	22
2.2. Traitement des images	25
2.3. Nommage des fichiers	26
2.4. Intégration des images	28
III. CONCEVOIR LE SITE DE LA BIBLIOTHÈQUE NUMÉRIQUE	32
1. Analyses préalables	32
1.1. Veille comparative	33
1.2. Analyse de l’existant	49
2. Conception et réalisation du site	56
2.1. Concevoir et réaliser l’interface graphique.....	57
2.2. Concevoir, alimenter et interroger la base de données.....	75
CONCLUSION	82
BIBLIOGRAPHIE.....	83

Introduction

L'utilisation des ressources numériques est devenue une réalité quotidienne pour de nombreux chercheurs, y compris en sciences humaines et sociales. Toutefois, ces pratiques s'inscrivent encore pour l'essentiel dans le cadre d'une consultation qui se limite aux ressources secondaires, comme les OPACs ou les bases de données bibliographiques. Pour les chercheurs en sciences humaines, l'étape suivante réside dans la possibilité de consulter à distance les sources primaires elles-mêmes.

De nombreux projets institutionnels ont d'ailleurs été engagés dans ce sens. En France, le projet de numérisation baptisé Gallica, lancé par la bibliothèque nationale, a montré que la diffusion sur internet d'un corpus d'ouvrages de caractère patrimonial, numérisés en mode image, pouvait rencontrer un large public. Le temps est donc venu pour les bibliothèques de taille plus modeste d'entreprendre à leur tour des programmes de numérisation et de diffusion électronique de leurs fonds.

La Maison de l'Orient et de la Méditerranée Jean Pouilloux a fait le choix de s'inscrire dans cette politique en lançant un programme de numérisation d'ouvrages couvrant les domaines de recherche de ses équipes et issus des différents fonds qui constituent sa bibliothèque. En présentant les divers aspects de la mission de stage qui m'a été confiée à la Maison de l'Orient, du 1^{er} juin au 30 septembre 2004, le présent rapport rend compte de ce projet de bibliothèque numérique, de l'acquisition des images à la consultation sur le web.

I. Contexte et présentation du projet

1. La Maison de l'Orient et de la Méditerranée

La Maison de l'Orient et de la Méditerranée, créée en 1975 sous la direction de Jean Pouilloux, est une fédération de recherche qui dépend du CNRS et de l'Université Lumière Lyon 2. Membre du réseau national des Maisons des Sciences de l'Homme, elle fédère actuellement sept unités de recherche¹ et rassemble près de 350 personnes autour d'une grande variété de domaines relatifs à l'étude de la Méditerranée, du Proche et du Moyen-Orient, sous tous leurs aspects, passés ou présents. Cette fédération a pour objectif de regrouper et d'héberger, dans un site unique, des unités de recherche se consacrant, principalement dans le domaine des sciences humaines et sociales, à l'étude des civilisations et des aires géographiques du monde méditerranéen. Elle est chargée de susciter des programmes scientifiques fédérateurs, de gérer les moyens destinés à ces thèmes de recherche, de favoriser la formation à la recherche et de développer la diffusion des connaissances scientifiques.

Avec ses quelque 50 000 volumes de monographies et plus de 1 000 revues spécialisées, la bibliothèque de la Maison de l'Orient se situe au coeur même de cet organisme de recherche. Celle-ci offre des fonds de référence pour les recherches portant sur la préhistoire et la géophysique méditerranéennes, l'Égypte pharaonique, l'Orient ancien et le monde égéen, le monde gréco-romain (histoire, archéologie, épigraphie, littérature, papyrologie), le monde arabe (classique et contemporain), ainsi qu'un fonds en développement sur le monde byzantin. Ces

¹ Les sept unités de recherche de la Maison de l'Orient et de la Méditerranée sont les suivantes : Archéorient (environnements et sociétés de l'Orient ancien), Hisoma (Histoire et sources des mondes antiques), Gremmo (groupe de recherche et d'études sur la Méditerranée et le Moyen-Orient), Archéométrie et archéologie (origine, datation et technologie des matériaux), Romanitas (jeune équipe), bureau lyonnais de l'IRAA (institut de recherche sur l'architecture antique), Textes latins scientifiques et techniques.

collections continuent d'être enrichies et s'accroissent d'environ 1200 ouvrages par an.

La bibliothèque actuelle, qui est associée au SCD de l'Université Lumière Lyon 2, a été ouverte en 1995 dans un bâtiment expressément construit pour l'abriter. Cette création a permis de rassembler les fonds documentaires des différentes équipes de la fédération et de les rendre accessibles à un plus large public. L'organisation des collections tient compte des spécificités des fonds préexistants dont la bibliothèque est issue. Ces fonds sont au nombre de six :

- « Victor Loret », égyptologie (EGY)
- « Salomon Reinach », archéologie et histoire du monde classique (ACL & HCL)
- « Bibliothèque d'archéologie orientale », Proche-Orient ancien et monde égéen (AOR)
- « Fernand Courby », histoire et littérature grecques classiques, épigraphie grecque, archéologie et épigraphie de Chypre (TXT, HCL & AOR)
- « Antoine Bon », monde arabe médiéval, période byzantine, histoire de l'art musulman, monde arabe contemporain (BAB)
- « Francis Hours », préhistoire et géomorphologie (PHG)

A ces fonds s'ajoute un dépôt d'ouvrages de l'association « Sources chrétiennes » ainsi qu'un fonds ancien qui comprend environ 1 300 volumes. Ces ouvrages dits rares, anciens ou précieux (RAP) datent du 16^{ème} au 20^{ème} siècle et concernent les différents domaines de recherche de la Maison de l'Orient et de la Méditerranée. La répartition de l'ensemble de ces fonds entre les cinq étages de la bibliothèque obéit à des notions d'aires de civilisation et d'aires géographiques.

Le personnel de la bibliothèque compte un total de treize membres. Parmi eux se trouvent sept ingénieurs d'études et assistants ingénieur : Olivier Dubois, coordinateur de la bibliothèque et responsable du fonds EGY, Franck Capisano, qui s'occupe des références générales ainsi que des fonds BAB et PHG, Véronique

Lacroix, chargée du fonds ACL et AOL, Catherine Gautin, responsable des cotes HCL et TXT, Véronique Faure qui s'occupe des fonds AOR et RAP, Bénédicte Kuntziger-Planche, chargée des périodiques et de l'informatique documentaire, et enfin Marjorie Burghart, qui en plus d'être chargée du site web de la Maison de l'Orient est également affectée à l'informatique documentaire de la bibliothèque.

La Maison de l'Orient dispose par ailleurs d'une plateforme technologique CNRS qui bénéficie de moyens humains et informatiques importants. Celle-ci s'organise autour de trois pôles liés à la documentation :

- le pôle « Banques d'images et données textuelles ». Animé par Patrick Desfarges, ce pôle concentre ses efforts sur la numérisation de documents, la constitution de banques d'images et la mise en œuvre de systèmes d'informations.
- le pôle « Têtes de réseaux documentaires ». La mission de ce pôle, à la tête duquel se trouve Bruno Helly, assisté de Miled Rousset, est de travailler à l'élaboration de catalogues collectifs documentaires tels que le réseau Frantiq ou le réseau PMC (Premier Millénaire Chrétien).
- le pôle « Document numérique ». Ce pôle, dirigé par Andrea Iacovella, a pour objectif de mettre à disposition des laboratoires un pôle d'expertise sur le document numérique scientifique, de développer des partenariats dans des projets impliquant le document numérique et de promouvoir des initiatives nationales ou internationales autour du document numérique scientifique.

2. Le projet de numérisation

La bibliothèque de la Maison de l'Orient et de la Méditerranée a initié en 2002 un projet de numérisation d'ouvrages libres de droit couvrant les thèmes de

recherche de ses équipes. A ce jour, les fonds numérisés appartiennent essentiellement aux domaines de l'égyptologie et de la civilisation gréco-latine, qu'il s'agisse d'histoire, d'archéologie ou de textes classiques. Quelques ouvrages issus du fonds des livres rares, anciens ou précieux ont également été numérisés. Cependant, la plupart des ouvrages concernés datent de la fin du 19^{ème} ou du tout début du 20^{ème} siècle. La numérisation de ces fonds répond à la volonté de diffuser les ressources de la bibliothèque auprès d'un public élargi et de permettre une consultation à distance qui soit souple et respectueuse des originaux. Le lancement de ce projet poursuivait en effet quatre objectifs principaux :

- Améliorer la conservation des documents : la numérisation permet de créer un support de substitution dont la qualité permet de satisfaire la plus grande partie des chercheurs. Les originaux qui sont fragiles ou précieux ne sont plus ainsi consultés que de façon exceptionnelle.
- Elargir la diffusion des collections : la mise en ligne des ouvrages numérisés permet de les diffuser plus largement et autorise une consultation à distance qui, par sa souplesse, peut procurer aux chercheurs un gain de temps substantiel.
- Développer un nouvel outil pédagogique : la constitution d'une banque d'images peut servir de support pour travailler avec les étudiants dans le cadre de séminaires ou pour monter des actions de communication en direction de tout autre public.
- Valoriser l'établissement : la numérisation est un moyen d'assurer la promotion de l'établissement à travers la mise en avant de ses collections et des travaux de ses chercheurs.

La sélection effectuée obéit à la volonté de privilégier les ouvrages les plus fréquemment consultés ainsi que les ouvrages précieux, endommagés ou répondant à une thématique générale cohérente et clairement définie. La date de publication

des documents entre également en compte : si les ouvrages qui ne sont pas encore tombés dans le domaine public ne sont pas à exclure absolument, priorité a été donnée aux ouvrages libres de droit². L'acquisition des droits de propriété intellectuelle afférents aux usages liés à la numérisation est effectivement exclue des aides financières de l'Etat.

Les bibliothécaires en charge des différentes collections proposent à Véronique Lacroix, coordinatrice du projet de numérisation, des listes d'ouvrages qu'ils estiment dignes d'être numérisés. Faute de temps et de moyens humains, la bibliothèque n'a malheureusement pas pu faire évoluer ce projet de façon significative, depuis la date de son lancement. Quelques ouvrages ont pu être numérisés mais les aspects qui concernent la gestion de la base d'images, l'indexation et la diffusion des ouvrages numérisés en étaient encore, lors de mon arrivée en stage, à l'état de prototype.

3. La mission de stage

La mission de stage qui m'a été proposée lors de mon premier contact avec Véronique Lacroix consistait à se familiariser avec la chaîne de numérisation, à réfléchir aux modalités d'exploitation des pages numérisées et à assurer la réalisation technique des moyens permettant leur diffusion en ligne. La proposition de stage a initialement été formulée de la façon suivante :

« Dans le cadre de votre stage de quatre mois, nous vous proposons d'étudier les modalités de création, d'alimentation et de gestion d'une base d'images en ligne. Les différentes étapes pourront se décliner de la façon suivante : sélection et localisation des ouvrages à numériser, finalisation des procédures de numérisation, numérisation à des fins de familiarisation sur les outils et les formats, intégration des images dans une base de données gérée par le

² D'une manière générale, les œuvres sont protégées durant la vie de l'auteur et pendant une période de 70 ans à compter du 1^{er} janvier de l'année qui suit la mort de l'auteur. Toutefois, de nombreuses règles particulières dérogent à cette règle des 70 ans *post mortem auctoris*. Il existe par exemple des prorogations pour années de guerre qui peuvent allonger significativement la durée de protection d'une œuvre. Si l'on tient compte de ces prorogations, on peut ainsi obtenir un cumul de 85 années de protection *post mortem auctoris*. Pour cette raison, la plupart des ouvrages numérisés sont antérieurs au 20^{ème} siècle.

progiciel Transvision, création de liens sur les sommaires des documents vers les pages intégrées dans la base d'images, établissement de liens avec le catalogue de la bibliothèque ou création d'une base SQL pour l'indexation des documents, définition des modalités de publication en ligne du fonds numérisé de la bibliothèque.

Notre informaticienne en charge du site web de la Maison de l'Orient a fait une première proposition de bibliothèque numérique qui nécessite d'être finalisée. Nous souhaiterions également proposer une rubrique de liens vers d'autres sites de bibliothèques numériques en ligne. Enfin, dans le cadre général de notre politique de numérisation, nous souhaiterions prendre contact avec les autres institutions « concurrentes » afin de mettre en place une concertation en vue d'éviter d'éventuels doublons.

Dans le cadre de votre stage vous serez principalement encadrée par trois personnes : Véronique Lacroix, bibliothécaire coordinatrice du projet de numérisation au sein de la bibliothèque ; Pierre Philippe, ingénieur du Service Images pour les questions liées à la chaîne de numérisation ; Marjorie Burghart, informaticienne responsable du site web de la Maison de l'Orient, pour la diffusion en ligne. »

Dans les faits, ma mission de stage s'est trouvée partagée entre deux grands volets, distincts et complémentaires. Une première partie, en amont, concernait la chaîne de numérisation proprement dite : numérisation des ouvrages en mode image sur différents scanners, gestion et traitement des fichiers issus de la numérisation puis intégration des images traitées dans le progiciel Transvision. Une seconde partie, en aval, consistait à réfléchir aux modalités de diffusion du fonds numérisé de la bibliothèque puis à réaliser une solution concrète permettant la consultation des pages numérisées.

En pratique, l'essentiel de ma mission de stage a donc consisté à concevoir puis à réaliser, dans son intégralité, le site web de la bibliothèque numérique. La réalisation de ce site web a nécessité, entre autres choses, de procéder à l'adaptation et à l'interrogation d'une base de données MySQL dans laquelle se trouvent renseignés, pour chaque ouvrage, un certain nombre de champs choisis en fonction de leur pertinence par rapport aux usages prévisibles de la bibliothèque numérique.

II. Maîtriser la chaîne de numérisation

Le service Image de la Maison de l'Orient et de la Méditerranée dispose d'une plateforme technologique qui permet de numériser et d'exploiter des documents variés tels que des ouvrages, des photographies, des dessins ou des cartes, que ces derniers soient en couleur ou en noir et blanc. Cette plateforme technologique, qui est aujourd'hui mutualisée au sein du réseau des Maisons des Sciences de l'Homme, a été utilisée dans le cadre de la production de plusieurs bibliothèques numériques d'envergure. On citera notamment le projet DEBORA³, le projet Cefael⁴, la bibliothèque Pierre Alexandre⁵, ou encore, plus récemment, le programme PRESCRIP(T)S⁶, projet dans lequel le matériel et le personnel de la plateforme sont aujourd'hui pleinement engagés. Autant de travaux de recherche et de développement qui ont valu à la Maison de l'Orient de se voir attribuer par le CNRS le label de Centre de compétences thématiques « Banques d'images et données textuelles ».

1. Description matérielle de la plateforme

La plateforme de numérisation de la Maison de l'Orient se compose de plusieurs types de ressources⁷ :

³ DEBORA (Digital accEss to BOoks of the RenAissance) est un projet de numérisation coordonné par l'ENSSIB qui vise à développer des outils permettant l'accès à des collections datant du 16^{ème} siècle.

⁴ Le projet Cefael, acronyme de « Collections de l'École française d'Athènes en ligne », est une bibliothèque numérique regroupant l'intégralité des ouvrages publiés par l'École française d'Athènes depuis 1877, ce qui en fait le premier portail de publications électroniques sur les études grecques, avec plus de 500 volumes numérisés, soit 250 000 pages.

⁵ La bibliothèque Pierre Alexandre est une bibliothèque numérique qui rassemble des fonds de linguistique et d'ethnolinguistique africaines. L'édition scientifique de ce corpus tiré des collections du chercheur Pierre Alexandre est le fruit de la collaboration de l'Université Lumière Lyon 2 et du CNRS dans le cadre du programme régional BRAIN.

⁶ PRESCRIP(T)S est un portail de revues de sciences humaines et sociales pour la recherche et l'information du public scientifique.

⁷ Cf. PHILIPPE, Pierre. Description de la chaîne de numérisation de la Maison de l'Orient méditerranéen – Jean Pouilloux. In Journées d'études des 29 - 30 mai 2002 organisées par l'École française d'Athènes, la Maison de l'Orient méditerranéen et l'Université Lumière Lyon 2. *Modèles opératoires de production et de diffusion des collections scientifiques dans les bibliothèques numériques*. Athènes : Ecole française d'Athènes, 2002. p. 19 - 32.

- Des stations d'acquisition
- Des serveurs
- Des logiciels
- Des outils de calibrage
- Un système de gravure
- Un réseau informatique

1.1. Les stations d'acquisition et les serveurs

Le panel des périphériques d'acquisition présents sur la plateforme technologique de la MOM a été constitué de façon à permettre la numérisation d'un large éventail de documents. Deux chaînes de production coexistent :

Une chaîne consacrée à la numérisation des ouvrages

Utilisée essentiellement pour la numérisation en noir et blanc, cette chaîne regroupe :

- Un scanner pour les livres précieux (Digibook 5600)
- Un scanner de microfilms
- Un lecteur reproducteur de microfilms
- Un appareil photo numérique haute définition utilisé pour la reprographie

Une chaîne dédiée à la reproduction de documents de toute forme

Permettant la numérisation de documents scientifiques en couleur et de formats variés, cette chaîne regroupe :

- Un scanner à plat format A3
- Un scanner à plat format A4

- Un scanner de diapositives et de formats moyens avec chargeur permettant une automatisation de la numérisation
- Un scanner de diapositives avec chargeur et films en bandes et en rouleaux
- Un scanner format A3 recto/verso bitmap dédié à la reproduction de publications
- Un scanner format A3 recto/verso couleur
- Un scanner format A0 pour cartes et plans
- Un appareil photo numérique

Ces périphériques d'acquisition sont connectés à des stations configurées pour supporter les flux de numérisation et les traitements associés.

La plateforme est dotée d'un serveur de bases de données, nommé Sun Entreprise 3000, qui héberge principalement le système de gestion de bases d'images (SGBI), nommé Transvision. La plateforme est également dotée d'un serveur de fichiers centralisant toutes les données de production.

1.2. Les logiciels

1.2.1. Le SGBI Transvision

Transvision est un système de gestion de bases d'images qui a été développé à la Maison de l'Orient, avant d'être exploité commercialement dans le secteur privé. Le progiciel Transvision permet notamment de :

- Stocker des images enrichies d'informations textuelles de façon organisée dans une base de données
- Mettre à la disposition des utilisateurs des outils de recherche efficaces afin d'exploiter au mieux leurs bases d'images

- Mettre à la disposition des producteurs les principaux outils de retouche d'images ainsi que la possibilité d'utiliser un plug-in du logiciel d'acquisition.
- Rendre le producteur autonome en lui donnant la possibilité d'approvisionner lui-même sa banque d'images et de définir les droits d'accès à ses données, sans faire appel à l'administrateur de la base.

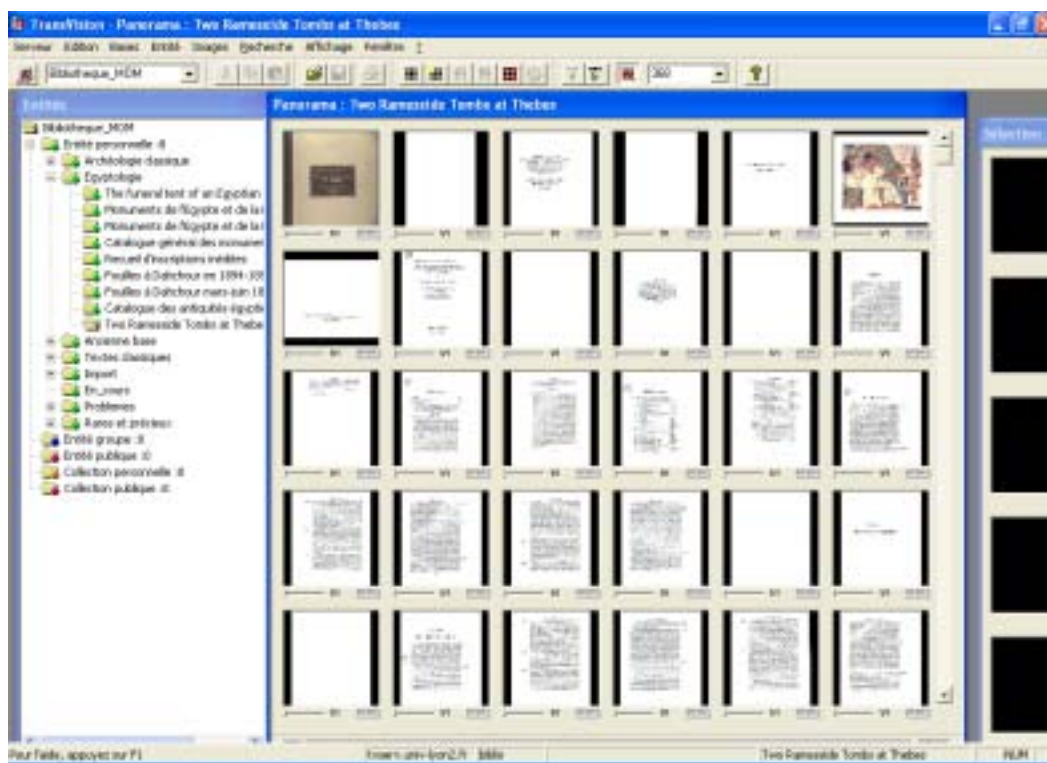
Le progiciel Transvision organise les images en les rassemblant dans des unités de base nommées « objets-images ». Un « objet-image » est constitué :

- d'une image principale, directement visualisée dans le panorama
- d'une ou plusieurs images secondaires qui viennent compléter la représentation de l'image principale, en offrant d'autres vues de l'objet figuré.
- de champs textuels, tels que le titre ou l'auteur, qui permettent d'indexer l'objet-image. Ces derniers sont définis par le producteur de la base en collaboration avec l'administrateur de la base.

Les objets-images s'organisent à leur tour au sein d'une arborescence hiérarchisée d'entités et de sous-entités permettant leur organisation logique. Le principe de cette arborescence reproduit l'organisation classique des fichiers en dossiers et sous-dossiers.

Dans le cas de la numérisation d'ouvrages, chaque page numérisée constitue à elle seule, avec les champs textuels qui la caractérisent, un « objet-image ». Toutes les pages, soit tous les objets-images, qui correspondent à un tome donné se trouvent rassemblées au sein d'une sous-entité qui porte le nom de l'ouvrage considéré. De même, tous les ouvrages, c'est-à-dire toutes les sous-entités, relevant d'une même thématique sont rassemblés au sein d'une entité englobante qui porte le nom de cette thématique. C'est ainsi, par exemple, que toutes les sous-

entités correspondant à des ouvrages d'égyptologie ont été rangées dans une entité « Egyptologie ».



Fenêtre du SGBI Transvision

Transvision permet par ailleurs un double mode de recherche : recherche visuelle sous la forme de panoramas d'images organisées hiérarchiquement et recherche en plein texte grâce à l'indexation automatique des champs de description des objets au fur et à mesure de leur saisie. Cette recherche peut-être mono ou multicritères sur un ou plusieurs champs et sur une ou plusieurs entités que l'on peut sélectionner.

Chaque producteur peut, à chaque niveau de l'arborescence qu'il a constituée, autoriser la consultation et / ou la mise à jour de ses données pour lui-même ou pour un groupe d'utilisateurs. Il peut également autoriser la consultation de tout ou partie de ses données au grand public en vue d'une diffusion sur le web. Il peut encore filtrer l'affichage des champs textuels.

Le progiciel Transvision est composé de plusieurs fonctions « serveur » :

- serveur de base images
- serveur d'indexation
- serveur http intégré
- serveur de listes
- client de production sur les systèmes Windows et MacOS.

Ces bases sont consultables à travers une interface web. De plus, une passerelle Z3950 en permet l'interrogation à partir de clients compatibles.

1.2.2. Le robot Transvision

Le robot Transvision est un produit associé au SGBI Transvision qui a été développé en interne, à la Maison de l'Orient. Il s'agit d'une interface graphique qui permet de paramétrer un enchaînement de modules qui répondent à des besoins de traitement standard, pour tel ou tel projet. Le robot Transvision comprend trois types de modules :

Les modules d'acquisition : ces modules permettent de définir la source des images à traiter.

- Acquisition à partir d'une entité Transvision⁸
- Acquisition à partir d'un répertoire d'entrée⁹
- Acquisition à partir d'un spool d'acquisition¹⁰

Les modules de traitement : ces modules permettent d'appliquer de façon automatisée des traitements aux images issues de la source préalablement définie.

Les traitements disponibles sont les suivants :

⁸ Ce module permet de se connecter à un serveur, de choisir une base puis une entité à partir de laquelle on va lire des fichiers image.

⁹ Ce module permet de sélectionner le répertoire dans lequel on récupère les fichiers image qu'on souhaite intégrer dans le SGBI.

¹⁰ Un spool d'acquisition est un répertoire dans lequel le Robot vient vérifier de manière régulière si aucune nouvelle image n'a été placée. Lorsque le robot détecte une nouvelle image, il la lit et la transfère vers les modules de traitement.

- Conversion du mode et du format des images
- Changement de la résolution et de la taille des images
- Module de repagination permettant d'établir une correspondance entre la numérotation automatique des fichiers image (pagination logique) et le foliotage réel de l'ouvrage (pagination physique) et de définir des zones de pagination spécifiques (pages préliminaires, pages finales, table des matières...)
- Séparation des vues correspondant à une double page pour constituer deux vues d'une page.
- Indexation textuelle des images
- Modification de valeurs dans les champs d'indexation
- Module permettant le renommage, l'indexation et la reconstitution de la structure d'un ouvrage (chapitres, articles, etc.) dans une base Transvision à partir d'un lot de fichiers images
- Découpe de zones dans l'image
- Recadrage d'histogramme¹¹
- Rotations
- Accentuation¹²
- Détourage Kodak¹³
- Possibilité de sauvegarder des images résultant de traitements intermédiaires

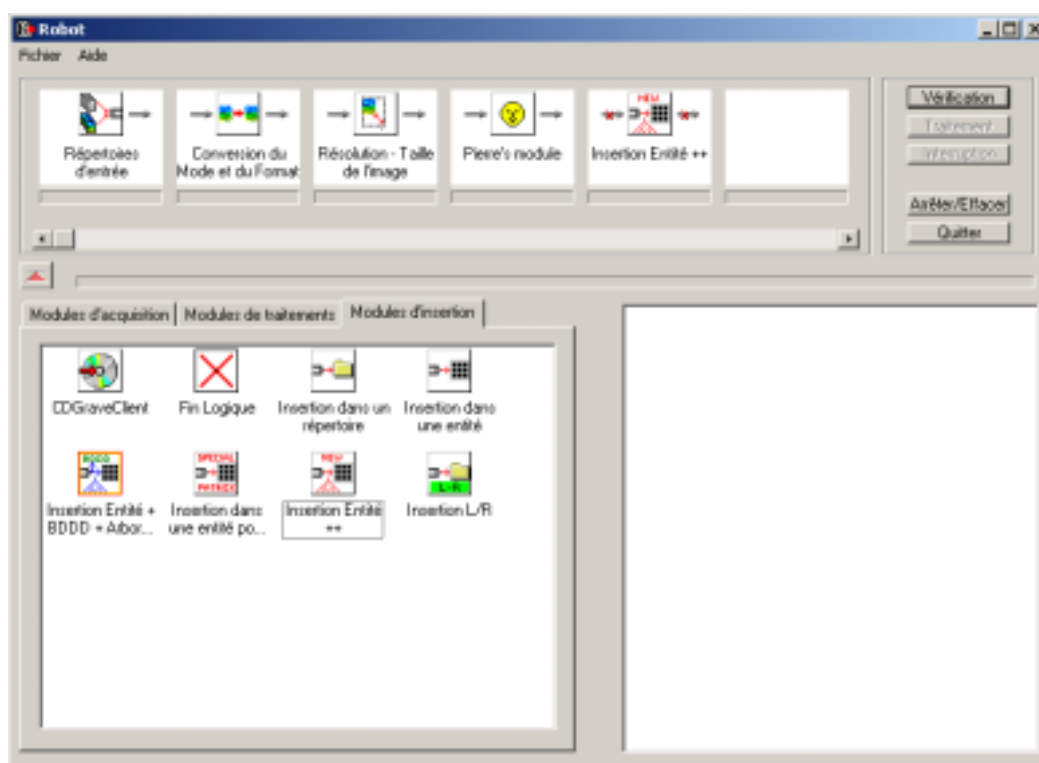
Les modules d'insertion : ces modules permettent de définir la destination des fichiers issus du traitement.

- Insertion dans une entité Transvision
- Insertion dans une entité Transvision avec séparation des doubles pages
- Insertion dans un répertoire
- Insertion dans une entité Transvision avec distinction des pages de droite et des pages de gauche

¹¹ Ce module permet de recadrer l'histogramme d'une image en niveaux de gris ou en couleurs.

¹² Ce module permet de renforcer les contours des objets contenus dans une image.

¹³ Ce module permet d'éliminer le bord noir que présentent les images au format Kodak.



Fenêtre du Robot Transvision

1.2.3. Book Restorer

Book Restorer est un logiciel développé par la société i2s qui permet de retoucher des images résultant de la numérisation d'ouvrages. Il permet de lancer sur la totalité d'un document des scripts de traitement que l'on a préalablement paramétrés sur un échantillon d'images représentatives du document. L'utilisation de ce logiciel est particulièrement indiquée lorsqu'on envisage la numérisation comme un préalable à la réimpression de l'ouvrage. La version 3.0 propose les traitements suivants :

Correction d'éclairage	Redimensionnement
Redressement de pages	Détourage
Cadrage	Suppression des taches
Détection de zones permettant	Détection des bords de page avec

d'isoler automatiquement des éléments graphiques telles que les lettrines ou les illustrations	possibilité de séparer automatiquement les vues de double page
Correction de courbure de ligne	Binarisation
Transformation (rotation, symétrie)	Conversion des couleurs (binaire, niveaux de gris, RVB)
Teinte, saturation, luminosité	Négatif
Étalement des niveaux	Courbes colorimétriques
Luminosité / contraste	Filtres (flou, accentuation, moyenneur, épaissement, amincissement)

1.2.4. Cameleo

Le logiciel Cameleo, développé par la société Caldera, est un outil de type *workflow* qui permet le pilotage de périphériques et l'enchaînement de traitements sur les images dans un système Unix. Ce logiciel est particulièrement adapté pour gérer des scanners générant de gros volumes de données. Il est muni d'une base image qui est utile pour le contrôle et le catalogage des archives générées sur CD-ROM.

1.2.5. Automate de gravure

Associé à une tour de gravure, l'automate de gravure utilisé à la Maison de l'Orient est une interface développée en interne qui permet de graver sur CD-ROM des projets entiers de numérisation. Il assure le découpage des volumes, la labellisation de chacun d'entre eux et l'impression de ces labels sur les supports.

1.3. Les outils de calibrage

Les outils de calibrage permettent d'assurer une reproduction fiable et cohérente des couleurs originales du document du début à la fin de la chaîne de

numérisation. En effet, chaque périphérique d'acquisition, d'impression ou de visualisation reproduit différemment les couleurs de l'image originale, et ce en fonction de facteurs variés : caractéristiques optiques, caractéristiques électroniques, comportement des drivers, consommables utilisés (toner, papier, etc.), usure de l'appareil. La technologie ICC (International Color Consortium), employée via le moteur Colorsync, permet de caractériser et de calibrer chaque composant. Elle consiste en la création de tables de conversion établies entre des valeurs colorimétriques de référence et des valeurs mesurées. Les outils utilisés sur la plateforme de la MOM sont les suivants :

- Des colorimètres permettant le calibrage des écrans
- Des spectrophotomètres pour la caractérisation des imprimantes
- Des chartes de référence qui sont numérisées ou imprimées et dont le résultat est mesuré
- Un logiciel générateur de profils (Gretag ProfileMaker)

Notons que ces caractérisations ne sont applicables que si l'on parvient à une bonne stabilité matérielle et logicielle sur l'ensemble de la chaîne.

1.4. Le système de gravure

Les « bruts de scan », c'est-à-dire les fichiers directement issus de la numérisation, sont systématiquement archivés, et ce afin de se ménager la possibilité de toujours pouvoir revenir aux fichiers originaux, vierges de tout traitement et de toute réduction de taille. Dans un premier temps, ces fichiers sont sauvegardés sur bande magnétique. Ils sont ensuite archivés sur CD-ROM, support qui présente l'avantage d'être économique et largement diffusable. Une tour robotisée munie d'un graveur, d'une imprimante thermique, d'un bras et d'un logiciel développé en interne permet de graver et de labelliser ces données.

1.5. Le réseau informatique

Le choix d'une configuration en réseau permet de mutualiser les ressources matérielles entre plusieurs postes de travail. La mise en réseau des scanners de production implique cependant de pouvoir assurer le transit d'importants volumes de données. Si l'on souhaite traiter les documents en continu, au fur et à mesure de leur numérisation, il est nécessaire de disposer d'un réseau capable de transporter des flux entre les stations d'acquisition, les stations de calcul et les serveurs. La plateforme est actuellement montée sur un réseau à 1 Gbits/s.

2. Pratiquer la chaîne de numérisation

Tous les éléments de la plateforme ne sont pas mis à contribution dans le projet de numérisation qui nous intéresse, loin de là. La mise en œuvre de ce dernier, comparée à celle d'un projet tel que Cefael ou Prescrip(t)s, reste en effet relativement artisanale. Les paragraphes qui suivent décrivent les différentes opérations effectivement réalisées dans le cadre d'un projet de numérisation dont le volume de production ne nécessite pas une automatisation complète de la chaîne de numérisation.

2.1. Acquisition des images

Le projet de numérisation dans lequel s'inscrit mon stage nécessite principalement l'utilisation de deux scanners, sur lesquels j'ai été amenée à travailler : le scanner Digibook 5600, pour la numérisation des pages de texte et des planches en noir et blanc, et le scanner à plat format A3 Fujifilm Finescan 2750, pour la numérisation des planches en couleur. Dans tous les cas, il s'agit d'une numérisation en mode image : le but est d'offrir une représentation du livre en tant qu'« objet archéologique ». Selon le support numérisé, on est cependant amené à choisir entre trois modes de numérisation :

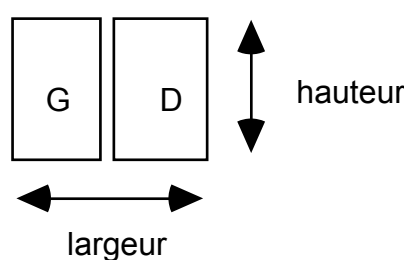
- Mode binaire, que l'on réserve aux pages de texte sans illustration
- Mode niveaux de gris, pour les planches monochromes
- Mode RVB, pour les planches en couleurs

Le mode binaire, pour la numérisation des pages de texte, permet d'obtenir des fichiers image qui sont moins lourds et plus lisibles que les fichiers en niveaux de gris.

Le choix de la résolution est fonction de la qualité désirée et des usages prévus. Une haute résolution sera par exemple nécessaire si l'on envisage une réimpression des images. Si l'on choisit de numériser en mode binaire, on peut forcer la résolution car le poids du fichier résultat restera relativement faible. En revanche, on aura tendance à baisser la résolution pour la numérisation en niveaux de gris ou en couleurs. D'une façon générale, la résolution choisie pour la numérisation tourne autour de 300 dpi.

Sur le scanner Digibook 5600, qui permet de numériser des ouvrages anciens ou précieux, le protocole de numérisation est le suivant :

- Choisir une abréviation servant à désigner l'ouvrage numérisé, définir puis sélectionner le répertoire d'enregistrement des images. Les pages numérisées sont stockées et nommées automatiquement par Digibook dans l'ordre de la numérisation et par ordre croissant.
- Mesurer la hauteur et la largeur de l'ouvrage en veillant à le positionner correctement sur les plateaux.



- Définir le format d'entrée en renseignant les champs suivants :

nom	edfou
largeur	430
hauteur	330
résolution	300
nombre de pages ¹⁴	2

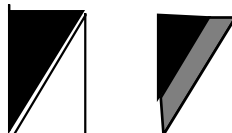
- Définir le format de sortie en renseignant de nouveau les champs de hauteur et de largeur. Attention, dans le cas où l'on choisit de séparer les doubles pages, il convient de diviser par 2 la largeur définie dans le format d'entrée :

largeur	215
hauteur	330

- Faire le point blanc. Il s'agit d'effectuer des corrections d'éclairage en réglant la luminosité sur une feuille blanche. Il est souvent nécessaire de faire le point blanc plusieurs fois dans la journée, en fonction des variations de la luminosité ambiante.
- Faire la mise au point. Il s'agit de lancer une mise au point automatique qui se règle en fonction d'une feuille pourvue de lignes noires horizontales. Il est nécessaire de faire cette mise au point plusieurs fois, notamment pour les ouvrages épais, quand on constate que l'image obtenue commence à devenir floue.

¹⁴ Le nombre de pages correspond au nombre de fichiers générés pour chaque acquisition. Si on entre le chiffre 1, chaque image correspondra à une double page, ce qui peut s'avérer utile s'il s'agit de numériser des planches qui s'étendent sur deux pages. Si on entre le chiffre 2, chaque acquisition générera deux images correspondant à chacune des deux pages. C'est le cas le plus fréquent.

- Sélectionner le mode binaire ou le mode niveau de gris, selon la nature du support numérisé.



- Régler le contraste et le nettoyage de fond¹⁵
- Lancer la prévisualisation en appuyant sur la pédale ou sur la barre d'espace
- Positionner les cadres de sélection qui délimitent la prise de vue finale et gommer les éventuelles imperfections
- Enregistrer l'image. Les images issues de la numérisation sont enregistrées en local, dans le répertoire sélectionné. C'est seulement dans un second temps qu'on les copie, après vérification, sur le réseau informatique. Le format d'enregistrement est le format TIFF.

2.2. Traitement des images

Pour la retouche d'images, le logiciel du scanner Digibook propose quelques fonctionnalités, mais c'est avec Book Restorer que sont effectuées les éventuelles corrections. Le logiciel Book Restorer est essentiellement utilisé pour les corrections suivantes : redressement de page, redressement de lignes, correction de courbure de page, gestion des marges d'impression. Toutefois, pour des raisons de temps, toutes les images, loin de là, ne font pas l'objet de retouche.

¹⁵ Seuil au-delà duquel tout point noir est remplacé par un point blanc, et inversement.

En revanche, toutes les images subissent au moins deux traitements : le passage au format JPEG et la réduction de taille, deux étapes préalables à leur mise en ligne sur le web. Le format TIFF est en effet peu adapté aux transferts rapides, contrairement au format JPEG, qui est le plus répandu sur le web. Les fichiers originaux en format TIFF, c'est-à-dire les « bruts de scan », sont toutefois conservés et archivés. Lorsqu'il s'agit de pages de texte, numérisées en mode binaire, on applique un troisième traitement : la conversion en niveaux de gris. Cette conversion est nécessaire car le mode binaire n'est pas implémenté pour le format JPEG.

Ces trois derniers traitements s'effectuent de façon automatisée sur l'ensemble des images. Généralement, ils sont appliqués lors de l'insertion des images dans le progiciel Transvision, grâce aux modules de traitement du Robot Transvision. Ils peuvent également s'effectuer sur Photoshop, grâce à l'élaboration d'un script dans lequel on spécifie le répertoire source, la taille souhaitée pour les images, une éventuelle conversion en niveaux de gris, l'enregistrement en format JPEG, et le répertoire destination.

2.3. Nommage des fichiers

Le nommage des fichiers issus de la numérisation est une opération lourde de conséquences qui doit obéir à un plan général cohérent. La structure de nommage doit être élaborée en envisageant tous les cas de figure pour l'ensemble des collections. Autant que faire ce peut, il convient de tenir compte, pour le nommage des volumes, répertoires et fichiers, de la norme ISO 9660¹⁶.

Le nom du fichier commence généralement par l'identifiant de l'organisme producteur. Pour les bibliothèques, il s'agit de la lettre « B » suivie du numéro RBCCN¹⁷ de la bibliothèque concernée. La structure de nommage adoptée pour

¹⁶ ISO 9660 : 1988, p.10, section 7.5.1

¹⁷ RBCCN = Répertoire des Bibliothèques du Catalogue Collectif National des publications en série. Le numéro RBCCN de la bibliothèque de la Maison de l'Orient est le suivant : 693872201. Ce numéro est délivré par les agences régionales du CCN-PS.

notre projet de numérisation est le suivant : « identifiant organisme producteur + identifiant ouvrage _ identifiant image + identifiant fichier ». En voici le détail :

Identifiant	Nbre de caractères	Exemple
Identifiant organisme producteur = lettre B (pour bibliothèque) suivie du n°RBCCN de la MOM	10	B693872201
Identifiant ouvrage = code barre du document (num-inv ou cote du document s'il n'a pas de code barre)	6	087975
Rang de reliure (si n ouvrages reliés ensembles) : 0 (ouvrage unique), 1 (premier ouvrage relié avec 0), 2 (second ouvrage relié avec 0 et 1), et ce jusqu'à 9	1	0
Séparateur : caractère underscore	1	_
Partie de l'ouvrage : d (pages préliminaires), t (page de texte), p (pages de planches), f (pages finales), a (matériel d'accompagnement)	1	d
Type de vue codé sur une lettre : p pour pleine page, z pour double page, d pour détail suivi d'une lettre indiquant le rang de détail ("da" pour l'image du premier détail, "db" pour celle du second, etc.)	1 ou 2	p
Résolution codée en 2 (2000 x 3000 pixels) ou 4 (4000 x 6000 pixels) ou h (> 4000 x 6000 pixels)	1	2
Type de fichier codé sur une lettre : v pour vignette (128x192 pixels), t pour tiers-écran et p pour plein écran (512x768 pixels)	1	v
Identifiant image (page) = le n° de page (4 caractères) suivi d'une lettre codant le type de pagination	4	0062
point extension : .gif ou .jpg, etc.		.jpg
Total : 27 caractères + 4 caractères d'extension (point compris) = 31 caractères maximum (30 minimum)		

Exemples de nom de fichier : B6938722010879750_d0062p2v.jpg (30 caractères)

B6938722010879750_0065ada2v.jpg (31 caractères)

2.4. Intégration des images

Après avoir éventuellement retouché, rassemblé et correctement ordonné les fichiers correspondant aux pages de texte, aux planches monochromes et aux planches de couleurs, on procède à leur intégration dans le SGBI Transvision. Celle-ci se fait à l'aide du Robot Transvision. Grâce au module « Acquisition à partir d'un répertoire d'entrée », on sélectionne le répertoire source, où sont stockés les fichiers images à intégrer.

On applique ensuite les modules de traitement. On sélectionne systématiquement les modules « Conversion du mode et du format des images » et « Changement de la résolution et de la taille des images », afin de passer au format JPEG et de réduire la taille des images, si cela n'a pas déjà été fait sur Photoshop.

On applique ensuite le module de traitement baptisé « Pierre's module ». Ce module permet de renommer et de repaginer automatiquement les fichiers images grâce à l'insertion d'un fichier tabulé. On passe ainsi d'une numérotation logique, générée automatiquement par le numériseur, à une pagination physique, qui correspond au foliotage réel de l'ouvrage.

Pagination logique (sortie numériseur)	Pagination physique (structuration réelle de l'ouvrage)
0001	Page préliminaire 1
0002	Page préliminaire 2
0003	Première section de texte : page 1
0004	Première section de texte : page 2
...	...
0156	Deuxième section de texte : page 1
0157	Deuxième section de texte : page 2
...	...
0312	Table des matières : page 1
0313	Table des matières : page 2
0314	Page finale 1
0315	Page finale 2

Exemple de passage d'une pagination logique à une pagination physique

« Pierre's module » crée ensuite une arborescence dans Transvision et procède à l'indexation des pages numérisées. Les informations permettant de réaliser ces opérations sont spécifiées dans le fichier tabulé que l'on a préalablement rédigé puis inséré dans le Robot Transvision.

Ce fichier tabulé se présente sous la forme suivante :

1	3	B693872201050683_P_####%_0000	1	Elite des Monuments céramographiques T.4 Texte
		" Ch. Lenormant et J. De Witte		Elite des Monuments céramographiques Tome quatrième, Texte
		Paris : Leleux 1861 "		
4	14	B693872201050683_T1_####%_0000	1	Elite des Monuments céramographiques T.4 Texte
		" Ch. Lenormant et J. De Witte		Elite des Monuments céramographiques Tome quatrième,
		Texte Paris : Leleux 1861 "		
15	298	B693872201050683_T2_####%_0000	1	Elite des Monuments céramographiques T.4 Texte
		" Ch. Lenormant et J. De Witte		Elite des Monuments céramographiques Tome quatrième,
		Texte Paris : Leleux 1861 "		

Exemple de fichier tabulé

En voici la syntaxe :

```

page_logique_début_1      page_logique_fin_1      nouveau_nom_image_1      page_physique_début_1
chemin_entité_enregistrement_1 " contenu_champs_1_1  contenu_champs_2_1
contenu_champs_3_1      contenu_champs_4_1 "
page_logique_début_2      page_logique_fin_2      nouveau_nom_image_2      page_physique_début_2
chemin_entité_enregistrement_2 " contenu_champs_1_2  contenu_champs_2_2
contenu_champs_3_2      contenu_champs_4_2 "
...
page_logique_début_m      page_logique_fin_m      nouveau_nom_image_m      page_physique_début_m
chemin_entité_enregistrement_m " contenu_champs_1_m  contenu_champs_2_m
contenu_champs_3_m      contenu_champs_4_m "

```

Syntaxe des fichiers tabulés

Explications :

Les deux premiers chiffres (ici en bleu) correspondent aux numéros des première et dernière pages d'une section de l'ouvrage, dans sa numérotation logique. On leur fait correspondre le numéro de la première page physique correspondante (ici en noir).

Dans la structure du nom de l'image (ici en vert) peuvent se trouver différents préfixes, qui correspondent aux différentes sections de l'ouvrage : C pour couverture, P pour page préliminaire, T1 pour section de texte n° 1, T2 pour section de texte n° 2 (etc.), F pour pages finales, D pour dépliant. La chaîne de caractères #####% indique que la numérotation se fait par incrémentation automatique sur 4 chiffres.

Le chemin de l'entité d'enregistrement (ici en violet) correspond à l'arborescence que l'on souhaite créer dans Transvision. Généralement, il s'agit de l'abréviation du titre de l'ouvrage. Celle-ci servira à désigner la nouvelle sous-entité propre à ce volume.

Le texte qui se trouve entre guillemets (ici en rouge) correspond au contenu des champs d'indexation. Les champs qu'il est possible de remplir, pour la base qui nous concerne, sont les suivants :

```

"auteur      titre_ouvrage      sous-titre      éditeur      chrono      ISBN
texte      notes      titre_article      auteur_article"

```

Dans notre exemple, seuls 4 champs sont remplis : auteur, titre_ouvrage, éditeur, chrono (= date de publication).

La rédaction de ce fichier tabulé compte sans doute parmi les étapes les plus délicates de la chaîne de numérisation. Elle réclame une certaine attention. La consultation de l'ouvrage dans sa forme matérielle permet de la faciliter mais celle-ci n'est pas toujours possible car certains livres numérisés ne sont pas propriété de la bibliothèque. Ils ont donc été empruntés puis rendus après numérisation. Il faut dans ce cas se fonder sur la consultation des fichiers image pour retrouver la structuration logique de l'ouvrage et de sa pagination, ce qui est moins commode.

Enfin, après avoir appliqué les différents modules de traitement, on choisit un module d'insertion, généralement « Insertion dans une entité Transvision », qui permet de spécifier l'entité dans laquelle on souhaite enregistrer les fichiers traités. Une fois les fichiers images « transvisés », on peut consulter les collections d'ouvrages numérisés grâce à l'interface offerte par le progiciel Transvision. L'étape suivante consiste alors à réaliser le site internet à partir duquel les internautes pourront consulter ces ouvrages.

III. Concevoir le site de la bibliothèque numérique

Le second versant de mon stage, de loin le plus important, a consisté à réfléchir aux modalités de publication en ligne des fonds numérisés puis à réaliser le site internet et la base de données nécessaires à cette publication. Une première ébauche, assez rudimentaire, avait déjà été réalisée par la webmestre en charge du site internet de la Maison de l'Orient. Celle-ci souhaitait que je propose une interface qui soit plus aboutie et qui tienne compte des évolutions prévisibles du contenu de la bibliothèque numérique. Elle ne me donna que deux consignes restrictives : respecter la charte graphique du site internet de la Maison de l'Orient et opter pour une interface qui soit simple, lisible, dépouillée. Ma responsable de stage ajouta qu'elle souhaitait voir figurer une liste de liens pointant vers d'autres bibliothèques numériques en relation avec les thèmes de recherche propres à la Maison de l'Orient.

1. Analyses préalables

Une première étape consistait à faire une analyse critique du prototype existant. Une telle analyse ne pouvait se faire qu'à la lumière d'un travail de recherche visant à recenser les sites d'un certain nombre de bibliothèques numériques représentatives de ce qui peut se faire en la matière. Il s'agissait donc de visiter et d'étudier l'interface de sites internet donnant accès à des pages d'ouvrages numérisées en mode image. Ce travail de veille poursuivait en fait deux objectifs distincts : un premier objectif était d'ordre technique et formel. Le but était d'étudier les modalités de présentation et de navigation dont les webmestres usaient pour donner accès aux pages d'ouvrages numérisées, quel que soit le contenu des documents mis en ligne. Un second objectif était d'ordre intellectuel. Il s'agissait de répertorier les bibliothèques numériques dont le contenu recoupait

les thèmes de recherche propres à la Maison de l'Orient, et ce afin de proposer une sélection de liens susceptible d'intéresser les chercheurs travaillant dans ces domaines.

1.1. Veille comparative

Le travail de veille réalisé en début de stage nous a permis de découvrir et de recenser les sites de nombreuses bibliothèques numériques. Parmi les outils qui permettent de s'orienter dans ce domaine, on citera le site du Ministère de la culture et de la communication, qui propose un catalogue des fonds culturels numérisés¹⁸. Ce catalogue permet d'identifier les établissements français qui procèdent à la numérisation de leurs collections ainsi que la nature des fonds concernés. Le site Culture.fr, édité lui aussi par le Ministère de la culture, propose quant à lui deux rubriques consacrées aux bibliothèques numériques¹⁹. Dans le cadre de son site Educnet²⁰, consacré aux nouvelles technologies au service de l'éducation et de la formation, le ministère de l'Éducation nationale propose lui aussi des liens intéressants. Enfin, l'IFLA et l'UNESCO proposent également des outils d'orientation²¹.

La visite de bibliothèques numériques permet de se faire une idée de la variété des modalités possibles pour la publication en ligne d'un fonds numérisé. Elle

¹⁸ Le catalogue des fonds culturels numérisés est accessible à l'adresse suivante : <http://www.culture.fr/culture/mrt/numerisation/fr/f_02.htm>. Ce catalogue décrit les programmes numériques des bibliothèques, mais aussi des archives départementales, des musées, des services de l'inventaire et de l'archéologie, ainsi que des établissements publics dépendant du ministère. En cours de constitution, il décrit actuellement 429 projets dont 124 provenant des bibliothèques.

¹⁹ Le site Culture.fr propose une rubrique intitulée « Numérisation du patrimoine culturel » consultable à l'adresse suivante : <<http://www.culture.fr/EtudesEtRecherchesCulturelles/c389>>. Il propose également une rubrique intitulée « Bibliothèques numériques » accessible à l'adresse suivante : <<http://www.culture.fr/BibliothequesMediatheques/c391>>

²⁰ Dans la rubrique « Dossiers et guides » du site Educnet se trouve une sous-rubrique consacrée aux bibliothèques numériques. Elle est accessible à l'adresse suivante : <<http://www.educnet.education.fr/dossier/rechercher/biblionum1.htm>>

²¹ Sur le site de l'Unesco se trouve un répertoire de projets de numérisation intitulé « UNESCO-IFLA Directory of Digitized Collections ». Il est accessible à l'adresse suivante : <<http://www.unesco.org/webworld/digicol/>>. Le site de l'IFLA (International Federation of Library Associations and Institutions) propose lui aussi une rubrique dédiée aux projets de numérisation. Celle-ci s'intitule « Digital Libraries Resources and Projects » et se consulte à l'adresse suivante : <<http://www.ifla.org/II/diglib.htm>>.

exerce le regard critique et forme le jugement sur ce qu'il convient d'éviter ou d'adopter. Pour la grande majorité des bibliothèques numériques visitées, nous avons dû nous contenter d'un rapide examen. Nous avons ensuite sélectionné une dizaine de sites²² que nous avons soumis à une analyse plus approfondie. Les parties suivantes rapportent, dans ses grandes lignes, l'analyse comparative de deux d'entre eux : la bibliothèque Pierre Alexandre et le projet Cefael, deux bibliothèques numériques réalisées avec le concours de la Maison de l'Orient.

1.1.1. Page d'accueil

La bibliothèque Pierre Alexandre²³ est une bibliothèque numérique qui rassemble des fonds de linguistique et d'ethnolinguistique africaines. Sa page d'accueil se présente de la façon suivante :



²² Ancient Near East and the Mediterranean World, disponible sur : <<http://www.lib.uchicago.edu/e/dl/proj/neh2/>>
 Archim, disponible sur : <<http://www.culture.gouv.fr/documentation/archim/accueil.html>>
 Base Enluminures, disponible sur : <<http://www.enluminures.culture.fr/>>
 Bibliothèque Pierre Alexandre, disponible sur : <<http://sophia.univ-lyon2.fr:8090/pierre.alexandre/>>
 Collections de l'École française d'Athènes en ligne, disponible sur : <<http://cefael.efa.gr/>>
 Collections medic@ de la BIUM, disponible sur : <<http://www.bium.univ-paris5.fr/histmed/medica.htm>>
 Conservatoire numérique des arts et métiers, disponible sur : <<http://cnum.cnam.fr/>>
 Dossiers numérisés de l'INHA, disponible sur : <<http://www.inha.fr/bibliotheque/exposition.html>>
 Gallica, disponible sur : <<http://gallica.bnf.fr/>>.

L'interface offerte par la page d'accueil de la bibliothèque Pierre Alexandre répond bien, selon nous, à sa fonction d'accueil :

- Titre et sous-titre bien visibles, qui identifient clairement l'espace virtuel que constitue la bibliothèque
- Présentation aérée et bien structurée, à la fois sobre et attrayante
- Un fond blanc qui assure une très bonne lisibilité
- Illustrations iconiques permettant d'identifier la nature du site (logos, photo de Pierre Alexandre)
- Informations textuelles synthétiques visant à présenter brièvement le projet et à faciliter la navigation de l'internaute
- Barre de navigation qui permet de s'orienter facilement dans le site
- Déclinaison de couleurs harmonieuses pour signaler les différentes rubriques
- Présence d'illustrations fonctionnant comme des points de repère : le logo de la bibliothèque et la photo de Pierre Alexandre, présents à la même place sur toutes les pages, permettant de signer l'identité de la bibliothèque et d'assurer l'homogénéité graphique du site.

En revanche, on peut noter que la longueur des lignes, pour le texte de présentation, est trop importante. On sait en effet que la lecture à l'écran s'accommode mal des lignes de plus de quinze mots : plus la ligne est longue, plus l'œil a du mal à retrouver le début de la ligne suivante.

Le projet Cefael²⁴, acronyme de « Collections de l'École française d'Athènes en ligne », est une bibliothèque numérique regroupant l'intégralité des ouvrages publiés par l'École française d'Athènes depuis 1877. La page d'accueil de Cefael répond elle aussi très bien aux fonctions attendues d'une page d'accueil :

²³ <<http://sophia.univ-lyon2.fr:8090/pierre.alexandre/>> (consulté le 15/06/2004)

²⁴ Disponible sur : <<http://cefael.efa.gr/>> (consulté le 15.06.2004)

- Informations textuelles succinctes qui permettent d’appréhender immédiatement le contenu du site, d’en comprendre la portée et les enjeux.
- Illustration iconique, semblable à un tampon, qui signe l’identité de la bibliothèque numérique
- Menu de navigation, en haut de la page, qui permet d’accéder aux trois principales rubriques du site : Présentation / Parcours / Recherche
- Recherche d’un design harmonieux
- Mention de la date de création et de dernière modification



On peut toutefois émettre quelques réserves sur les points suivants :

- Le menu qui permet de pénétrer dans le site est d’une taille étonnamment réduite. De plus, les icônes accolées aux intitulés des rubriques sont d’une couleur bleu clair qui ressort mal sur le fond bleu du site. La déclinaison des différents bleus, si elle est esthétique, nuit aux exigences de contraste visuel. La présentation des rubriques, tant par sa taille que par ses couleurs, manque donc de lisibilité.
- Absence de titre lisible et explicite au profit d’un sigle dont la signification n’est forcément pas transparente pour le profane. Au premier abord, il n’est

pas évident de comprendre qu'il s'agit de la bibliothèque numérique de l'Efa.

- Déséquilibre visuel entre une grande zone bleue très dépouillée et une bande noire étroite dont la densité textuelle n'est pas très lisible.

1.1.2. Navigation

Dans la bibliothèque Pierre Alexandre, la présence d'une barre de navigation fixe en haut de la page permet à l'internaute de ne jamais s'égarer. Où qu'il se situe dans l'arborescence du site, il a toujours la possibilité de revenir à la page d'accueil ainsi que dans n'importe quelle autre des rubriques du site.



La dénomination des boutons de navigation est assez claire. On notera que les rubriques apportant des éléments informatifs sur le projet sont clairement dissociées, par leur emplacement et leur couleur, de celles qui permettent de consulter le contenu de la bibliothèque en elle-même, ce qui facilite l'orientation de l'internaute.

Deux types de consultation des ressources sont proposés :

- L'exploration par la rubrique « parcours » permet de consulter la bibliothèque de façon thématique : consultation des ressources dans un parcours par pays / par matière et spécialité / par lieu et langue / par langue. Ce type de navigation thématique est idéal pour se faire une idée du contenu de la bibliothèque dans son ensemble. Ajoutons que la consultation par pays se fait selon un mode iconique, grâce à la présence d'une carte géographique découpée en autant de zones sensibles qu'il y a de pays concernés.
- La rubrique « recherche » permet d'effectuer une recherche documentaire précise en fonction des nombreux champs indexés (auteur, titre, table des

matières...). Deux niveaux de recherche sont proposés : recherche simple et recherche experte.

La multiplicité des parcours possibles permet d'optimiser l'exploitation des richesses documentaires de la bibliothèque numérique. On notera en particulier le souci de varier les canaux de perception sollicités (recherche sur le mode textuel ou spatial, sur carte géographique)

La bibliothèque Cefael présente elle aussi une barre de navigation fixe placée en haut de la page et présente sur l'ensemble du site :



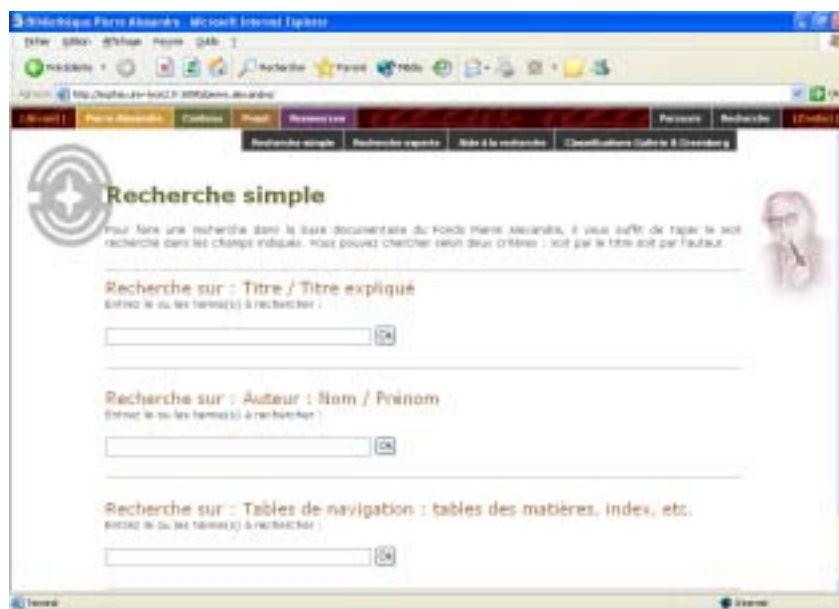
On remarque que cette barre de navigation respecte le principe du marquage redondant, qui est un principe général en ergonomie²⁵. Elle associe effectivement, à chaque fois, désignations textuelle et iconique (l'icône en forme de maison et le mot *Accueil*, par exemple). On remarque par ailleurs que le système d'onglets fonctionne comme un indicateur de positionnement : d'emblée, l'internaute sait ici qu'il se trouve dans la rubrique parcours. Ce type d'indicateur participe de la bonne ergonomie du site. Enfin, on note la présence d'un historique de navigation sous la forme d'un rappel de l'arborescence : Parcours < Site < Portail. Si la présence d'historiques de ce type permet normalement à l'internaute de mieux se repérer, celui-ci est malheureusement dans le mauvais sens, ce qui risque d'entraîner une confusion cognitive.

1.1.3. Interfaces de recherche

La bibliothèque Pierre Alexandre propose deux interfaces de recherche : recherche simple et recherche experte. On notera, sur ces deux interfaces, la

²⁵ cf. CARO Stéphane, BETRANCOURT Mireille. Ergonomie des documents numériques. *Techniques de l'ingénieur, traité Informatique*, février 2001, p. 5.

présence d'un lien intitulé « Aide à la recherche », qui permet à l'internaute de les utiliser au mieux.

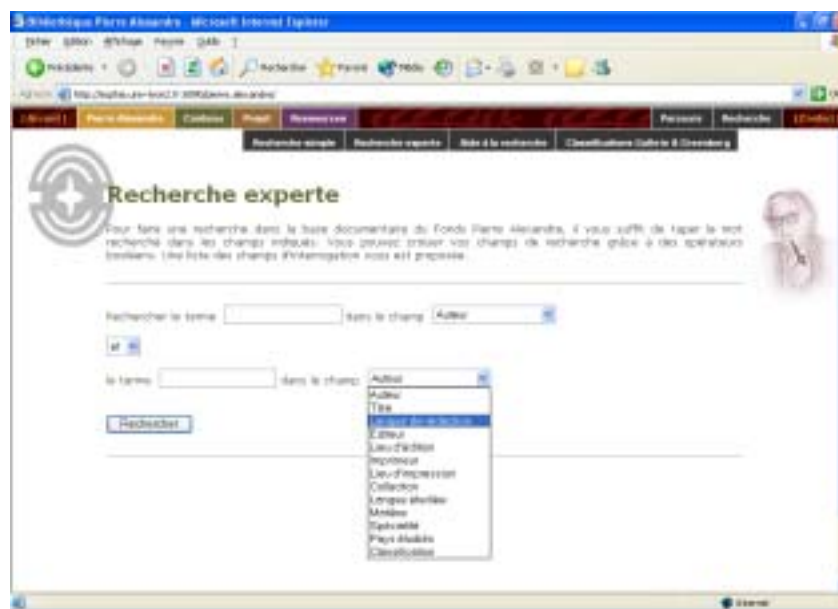


L'interface de recherche simple, agréablement présentée, nous semble toutefois trop limitée. Elle ne permet en effet qu'une interrogation monocritère de la base : soit par mot du titre soit par auteur, sans possibilité de combiner ces deux critères essentiels. L'internaute ne sera-t-il pas tenté d'utiliser directement et exclusivement la seconde interface de recherche ? Par ailleurs, la signification de la barre oblique (/) ne nous semble pas transparente, au premier coup d'œil. Elle indique en fait que l'interrogation ne doit porter que sur l'un des deux critères (Titre OU Titre expliqué, Nom OU Prénom). Une brève explication textuelle n'aurait-elle pas mieux valu ?

Cette interface propose également une recherche sur les mots des tables et index, qui ont été saisis. Cette possibilité de recherche est très intéressante mais a-t-elle sa place dans une interface de recherche dite simple, présentée au même niveau que l'auteur ou le titre de l'ouvrage ?

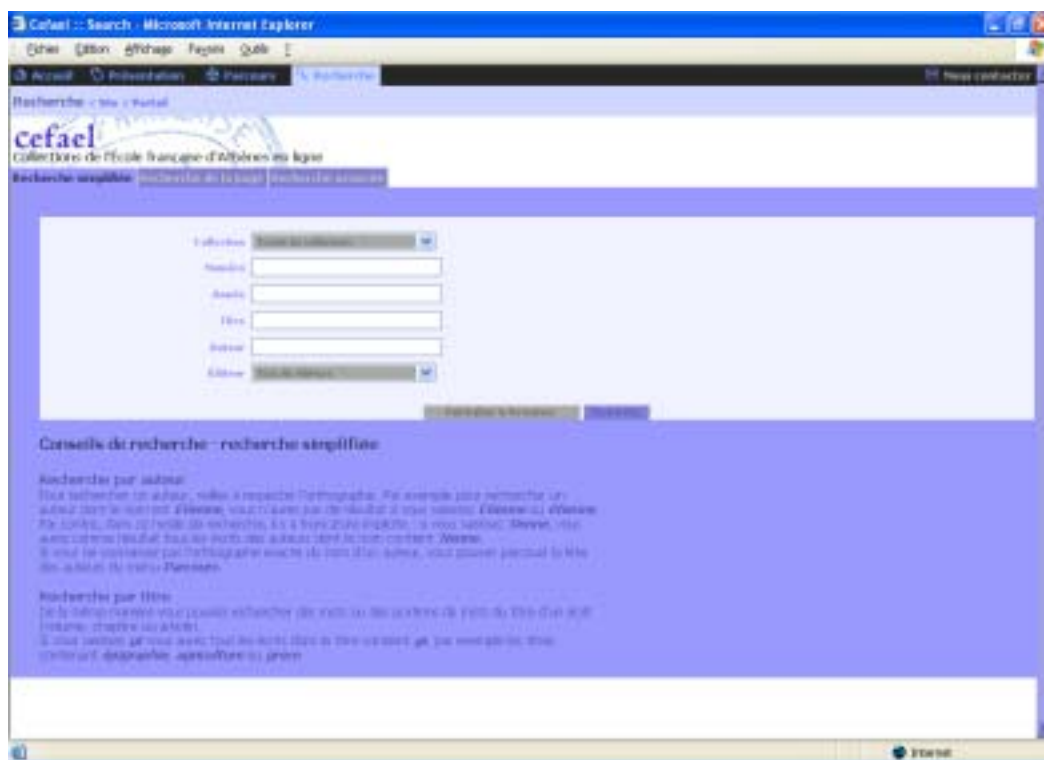
En revanche, l'interface de la recherche experte nous paraît à la fois facile d'utilisation et suffisamment puissante, même si elle ne permet de combiner que deux critères de recherche. La possibilité de combiner ces critères selon deux

opérateurs (ET / OU) offre une grande souplesse d'interrogation. La liste des champs interrogeables est par ailleurs très complète.



Dans les deux interfaces, on note que la casse des lettres est indifférente : que l'internaute saisisse « Ebert » ou « ebert », par exemple, il obtiendra les mêmes résultats.

La bibliothèque Cefael offre quant à elle trois interfaces de recherche : « recherche simplifiée », « recherche de la page », « recherche avancée ». Contrairement à la recherche simple de la bibliothèque Pierre Alexandre, la recherche simplifiée de Cefael permet la combinaison de plusieurs critères de recherche. L'opérateur par défaut entre les différents champs est le « ET ». Cette interface est à la fois souple, puissante et facile d'utilisation. En revanche, la casse des lettres n'est pas indifférente, ce qui présente l'inconvénient de générer des silences indésirables : si l'on tape auteur = « moretti » au lieu de « Moretti », par exemple, aucun résultat n'apparaît. Heureusement, une note explicative, consultable sans qu'il soit nécessaire de cliquer sur un lien, précise qu'il faut strictement respecter la casse et l'accentuation, ce qui reste malgré tout assez contraignant.

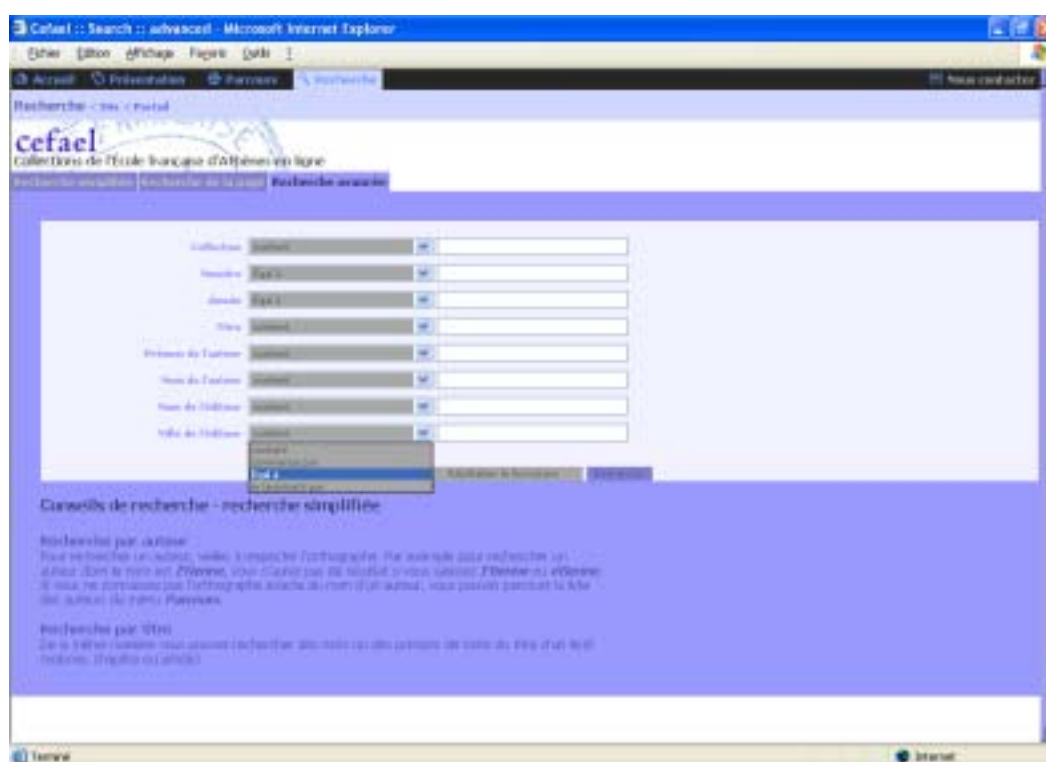


Une seconde interface a été conçue pour permettre la recherche d'une page donnée. Cette possibilité n'est pas inintéressante, même si on peut se demander dans quelle mesure l'utilisateur aura l'occasion d'y avoir recours. Rares sont les cas où l'on connaît précisément la collection, le numéro, le volume et la page qu'on recherche...



Enfin, une troisième interface permet à l'utilisateur d'effectuer une recherche affinée. La différence avec l'interface de recherche simplifiée réside dans le nombre de critères combinables, qui est ici plus important, mais surtout

dans les modalités de recherche proposées pour chaque critère : une liste déroulante permet pour chacun d'entre eux de préciser si l'on veut l'expression exacte (option « égal à »), une troncature droite illimitée (option « commence par »), une troncature gauche illimitée (option « se termine par »), une troncature droite et gauche illimitée (option « contient »). Pour les champs numériques (année, numéro), on se voit proposer la possibilité de définir une borne inférieure (option « de »), une borne supérieure (option « jusqu'à »), ou une expression exacte (option « égal à »). Autant d'options qui garantissent une excellente souplesse dans la recherche. En revanche, l'utilisation des opérateurs booléens « ou / sauf » n'est nullement prévue.



1.1.4. Résultats de la recherche

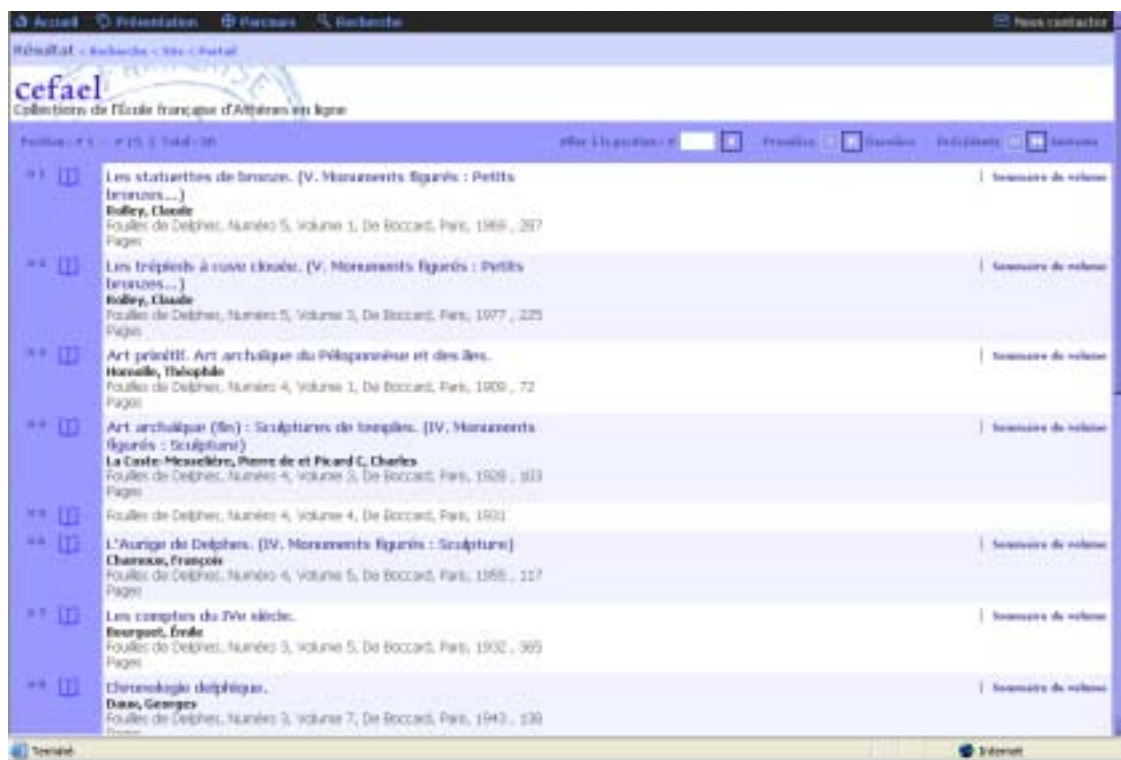
Dans la bibliothèque Pierre Alexandre, une fois sa requête composée et validée, l'internaute arrive à une page de résultats qui se présente de la façon suivante :

Ref.	Titre	Auteur(s)	
fpa_0027	Manuel de linguistique bantoue	R.P. Van Bubb, G.	consulter
fpa_0030	Introduction à l'étude des langues bantoues du Congo belge	Burssens, Ansaet	consulter
fpa_0059	Les préfixes nominaux dans les parlers peul, haoussa et bantous	Homburger, L.	consulter
fpa_0109	Temas de linguística banta / Não serão de um mesmo tronco as línguas bantas e as indo-europeias? III). Les langues bantoues et les langues indo-européennes n'auraient-elles pas une origine commune ?	De Sá Nogueira, Rodrigo	consulter
fpa_0109	Temas de linguística banta / Não serão de um mesmo tronco as línguas bantas e as indo-europeias? III). Les langues bantoues et les langues indo-européennes n'auraient-elles pas une origine commune ?	De Sá Nogueira, Rodrigo	consulter
fpa_0171	Premières applications de la géographie linguistique aux langues bantoues	De Boeck, L.-B.	consulter
fpa_0187	Le groupe sud-ouest des langues bantoues	Homburger, L.	consulter

On note que l'équation de recherche est rappelée en haut du tableau, ainsi que le nombre de réponses trouvées, ce qui s'avère fort utile. Il est à noter que l'interface de Cefael, si elle mentionne également le nombre total de résultats, ne prévoit pas le rappel de l'équation de recherche. L'alternance de couleurs pour les lignes du tableau, également adoptée dans Cefael, permet une meilleure lisibilité des résultats. En revanche, on peut regretter le fait qu'il faille cliquer sur le lien « Consulter », à droite, pour visualiser l'un des résultats. Il semble plus intuitif de cliquer directement sur l'intitulé des résultats, comme cela se fait dans l'interface de Cefael.

Dans la bibliothèque Pierre Alexandre, tous les résultats, aussi nombreux soient-ils, s'affichent sur la même et unique page, alors que sur Cefael ils s'affichent 15 par 15, ce qui permet de limiter l'utilisation de l'ascenseur vertical. Cette dernière option nécessite en revanche l'utilisation de boutons pour naviguer

entre les pages de résultats ainsi que la présence d'un indicateur précisant la numérotation des résultats effectivement visualisés sur la page courante.



Dès cette étape, l'internaute peut, dans Cefael, consulter le sommaire du volume grâce à un lien présent sur la droite. En cliquant sur le titre de l'ouvrage ou sur l'icône en forme de livre, à gauche, il peut également accéder directement à la consultation « monopage », ce qui est impossible à ce stade dans la bibliothèque Pierre Alexandre. On peut se demander si le double mode d'accès, dans Cefael, à la consultation monopage (icône et texte du titre) est un bien (redondance qui permet de multiplier les incitations à la consultation) ou un mal (redondance qui risque de troubler l'internaute, susceptible de se demander si ces deux liens conduisent ou non à la même page).

Enfin on remarque, dans Cefael, que la notice de l'ouvrage est présente dans le corps même de la liste des résultats, sous une forme minimaliste. Au contraire, la bibliothèque Pierre Alexandre se contente de mentionner les trois critères nécessaires à l'identification de l'ouvrage (référence, titre, auteur(s)). Ce

n'est que dans un second temps qu'on peut consulter la notice de l'ouvrage, qui est beaucoup plus détaillée que dans Cefael.

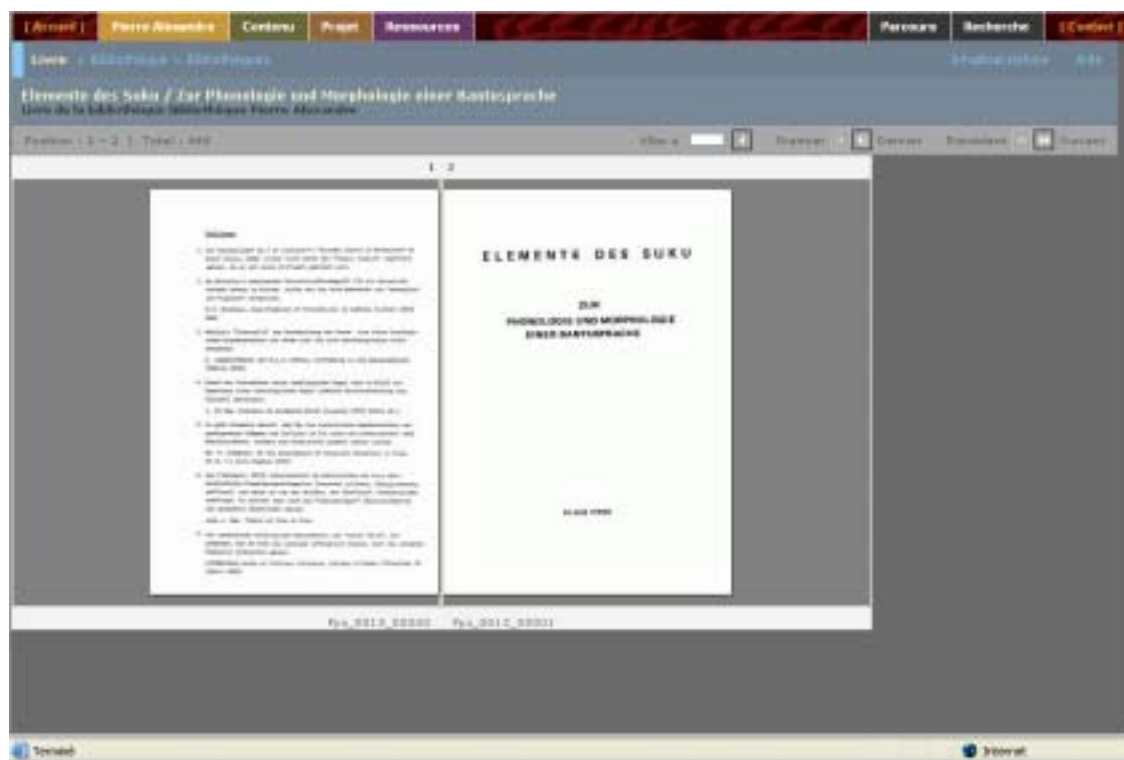
1.1.5. Consulter l'ouvrage

Après avoir cliqué, dans la bibliothèque Pierre Alexandre, sur le lien correspondant à l'ouvrage qu'il désire consulter, l'internaute obtient la notice détaillée de l'ouvrage et se voit alors proposer, dans un menu en haut à droite, deux modes de consultation distincts : « consulter l'ouvrage » et « consulter les tables ». Cette dernière option constitue une étape intermédiaire qui permet de visualiser la structure de l'ouvrage et d'accéder, dans un troisième temps, aux pages de tel ou tel chapitre. Autant d'étapes qui diffèrent la consultation des pages numérisées. En cliquant sur le lien « consulter », dans la liste des résultats, l'internaute pouvait s'attendre à consulter directement l'ouvrage. En fait, il doit encore cliquer sur « Consulter l'ouvrage », pour accéder, enfin, aux pages numérisées. A moins qu'il ne choisisse d'abord de consulter la table, ce qui ajoute encore une étape supplémentaire... Le cheminement pour arriver jusqu'aux pages numérisées est ainsi beaucoup plus tortueux que dans Cefael.

The screenshot shows a web interface for a library. At the top, there are navigation tabs: Accueil, Pierre Alexandre, Connexion, Aide, Recherche, and Un menu. Below these are three buttons: Consulter l'ouvrage, Consulter les tables, and Aller à la fin. The main content area is titled 'Notice détaillée' with the reference 'fpa_0171'. A circular logo is on the left, and a small portrait of a man is on the right. The notice is presented as a table with the following data:

Titre	Premières applications de la géographie linguistique aux langues bantoues		
Titre abrégé			
Auteur(s)	De Boeck, L.-B.		
Langue(s) étudiée(s)			
Matière	Linguistique		
Spécialité			
Classification	Linguistique		
Éditeur	Institut Royal Colonial belge	Langues	
Lieu d'édition	Bruelles	Lang. d'expression	
N° d'édition		année d'édition	1942
Format	26 x 22 cm		
Page(s)	220 p.		

L'interface de consultation des pages numérisées se présente alors de la façon suivante :



On note d'abord qu'il ne s'agit toujours pas de la consultation « monopage ». Il faudra pour cela attendre encore une étape. On note ensuite une rupture nette dans la gamme de tons utilisés, ce qui rompt l'harmonie visuelle du site dans son ensemble. On passe brutalement d'une déclinaison de tons chauds à une page qui décline des couleurs froides, ce qui donne le sentiment que cette interface n'a pas été conçue pour ce site en particulier.

En revanche, la navigation est bien pensée : visualisation des deux pages en regard, comme lorsqu'on lit un ouvrage physique, possibilité d'accéder directement à telle page, à la page suivante ou précédente, ou encore à la dernière ou première page. On note par ailleurs que le nombre total de pages numérisées de l'ouvrage est mentionné, tout comme la position, au sein de cet ensemble, des pages visualisées. L'abstraction des documents électroniques rend la présence de ces indicateurs de volume et de positionnement indispensable. Le lecteur est ainsi capable de se situer au sein de l'entité que constitue l'ouvrage. On note enfin que

le titre de l'ouvrage est rappelé en haut de la fenêtre de façon à n'être jamais perdu de vue.



L'interface monopage de Cefael, dont le graphisme et les couleurs s'harmonisent avec le reste du site, propose les mêmes indicateurs de volume et de positionnement, permettant au lecteur de se repérer. Là encore, le titre de l'ouvrage est rappelé, mais il est ici accompagné des autres champs de la notice, ce qui permet de lever toute ambiguïté. La navigation « inter-pages » est là aussi bien pensée : possibilité d'accéder directement à telle page, grâce à une liste déroulante, à la page suivante ou précédente, ou encore à la dernière ou première page. Il peut également accéder au sommaire de l'ouvrage, ce qui n'est plus possible, à ce stade, dans la bibliothèque Pierre Alexandre.

En revanche, la consultation est directement monopage. L'absence de consultation intermédiaire permet d'économiser des clics pour accéder à la lecture des articles, mais l'internaute perd en recul ce qu'il gagne en temps. La seule consultation monopage ne permet pas au lecteur de percevoir l'ouvrage à distance,

dans sa dimension matérielle. Cette impression est accentuée par le fait que l’affichage par défaut se fait à une grande échelle : l’internaute a le sentiment d’avoir le nez collé sur la page. Il ne perçoit ni la globalité ni la matérialité de la page consultée. Il peut toutefois, dans un second temps, réduire l’échelle de consultation grâce à une liste déroulante.

Une fois parvenu à l’écran de consultation des pages numérisées, l’internaute doit encore, dans la bibliothèque Pierre Alexandre, cliquer sur l’une des deux pages pour lire véritablement l’ouvrage. S’affiche alors l’écran de consultation rapprochée :

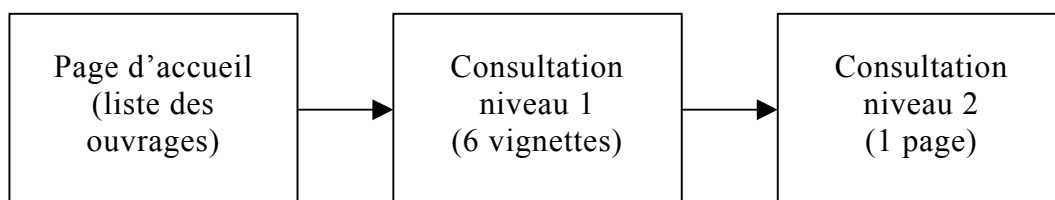


La référence qui s’affiche en haut de la fenêtre, à la place du titre de l’ouvrage, n’offre que peu d’intérêt pour le profane. On remarque par ailleurs que si l’internaute peut augmenter ou réduire l’échelle de la page, il ne peut en revanche naviguer de page en page, ce qui constitue selon nous un grave défaut ergonomique : l’internaute est contraint, pour passer à la page suivante, de revenir à l’étape précédente en cliquant sur « Livre ». Ce dernier lien, qui s’inscrit dans un historique de navigation inversé et assez peu parlant, est d’ailleurs difficile à trouver. Enfin parvenu au niveau où il peut appréhender le contenu textuel de l’ouvrage, l’internaute ne peut malheureusement plus le feuilleter... ce qui ne

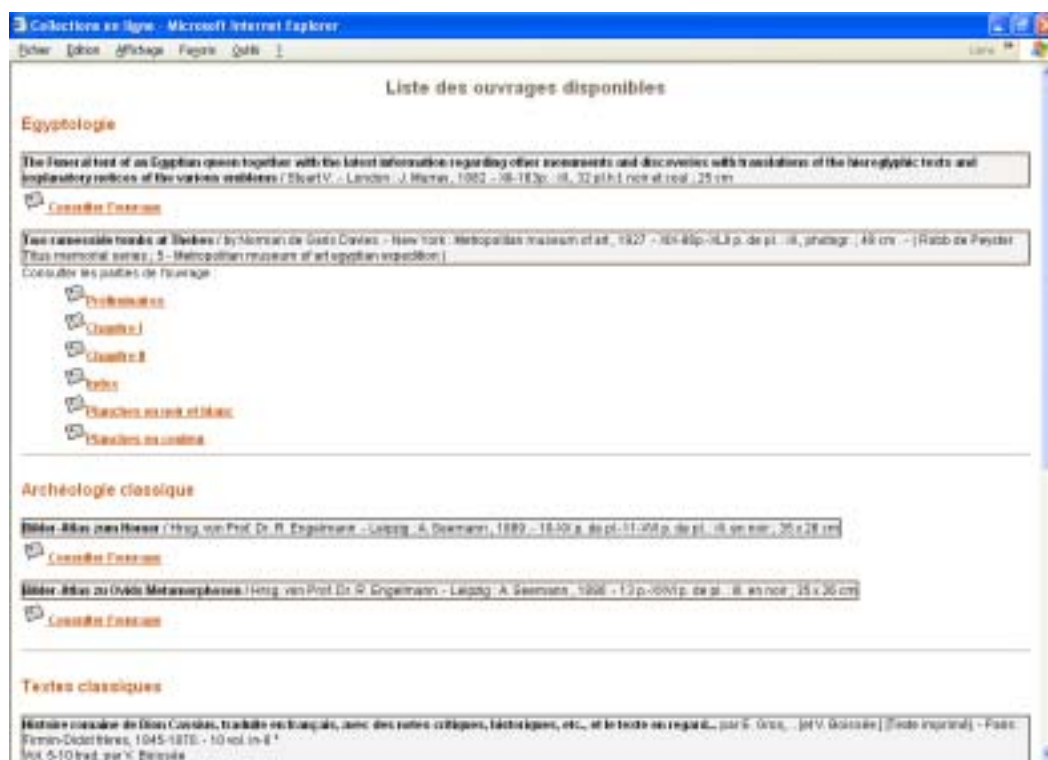
laisse pas d'être problématique. C'est toute la barre de navigation de l'étape précédente qu'il conviendrait de rappeler dans cette ultime étape de consultation.

1.2. Analyse de l'existant

Le travail préalable de veille comparative nous a permis de procéder à l'analyse de l'existant avec un regard plus averti. Le prototype réalisé par l'informaticienne, qui a permis de tester un premier dispositif technique de consultation des ouvrages, présente une architecture rudimentaire qui se décline sur trois niveaux : l'écran d'accueil, qui liste les ouvrages numérisés, un écran de « consultation globale », qui présente les pages numérisées par groupes de six vignettes, et un écran de « consultation rapprochée », qui permet de consulter l'ouvrage page par page.



1.2.1. Page d'accueil



La première page de la bibliothèque numérique se présente comme une page relativement austère. Dépouillée de toute illustration, elle n'est que peu attrayante. D'une façon générale, on note que son apparence ne reflète pas son statut de « page d'accueil ». Cette première page se présente plus comme une page de sous-rubrique que comme la page d'accueil de l'entité distincte que constituera la bibliothèque numérique au sein du tout englobant que représente le site de la Maison de l'Orient. Il conviendrait au contraire de marquer nettement l'entrée de l'internaute dans l'espace virtuel autonome que représente la bibliothèque numérique. Dans cette perspective, la création d'une page dont le design réponde effectivement à une fonction d'accueil semble s'imposer. Même s'il convient de rester sobre pour garantir une bonne lisibilité, une interface plus travaillée et plus originale serait appréciable. Cette page d'accueil étant amenée à remplir un rôle d'écran référent au sein de la bibliothèque numérique, elle devra pouvoir être facilement distinguée, par son graphisme et sa structuration, des autres pages de cette même bibliothèque.

Il est à remarquer qu'aucun texte de présentation ne vient informer l'internaute sur la nature des pages qu'il s'apprête à visiter. Or il est essentiel que celui-ci puisse se faire une idée précise et rapide non seulement du contenu de la bibliothèque numérique mais aussi et surtout de la nature du projet mis en œuvre et de la démarche qui a présidé à son élaboration. Par ailleurs, on note que le logo de la Maison de l'Orient ne figure pas sur cette première page, ce qui nuit à l'authentification de la bibliothèque numérique. Dans la mesure où l'internaute est susceptible d'arriver directement sur la page d'accueil de la bibliothèque numérique, sans passer par les autres pages du site de la Maison de l'Orient, il est essentiel de faire en sorte que la marque de fabrique « MOM » soit immédiatement perceptible.

En ce qui concerne la liste des ouvrages, plusieurs remarques s'imposent : les notices d'ouvrages, étroitement encadrées, ne sont pas très lisibles. La police dans laquelle sont écrits les titres d'ouvrage ne se lit pas très bien et l'encadré sur fond gris, qui serre le texte de trop près, n'est pas très heureux. Par ailleurs, cette liste gagnerait à être présentée d'une façon moins linéaire : la succession strictement verticale des titres d'ouvrages produit une impression de catalogue potentiellement rébarbatif.

On peut en outre se demander si le format choisi pour l'affichage des notices (format ISBD) est le plus approprié pour la liste générale des ouvrages numérisés. Cette liste constitue en effet une exposition d'ensemble dont le but est de renseigner globalement l'internaute sur la nature des fonds numérisés. Est-il nécessaire d'être aussi précis, à ce stade de la consultation, pour un public qui n'a pas forcément l'œil exercé à la lecture des formats bibliothéconomiques ? Ne vaudrait-il pas mieux adopter, dans un premier temps, une description des ouvrages qui soit plus synthétique et réserver la description technique à un second niveau de consultation ?

On constate que l'accès aux pages numérisées se fait grâce à une icône agrémentée de texte. Celle-ci se trouve nettement séparée de la désignation de l'ouvrage en lui-même. Sur un plan cognitif, il semble que la dissociation de la mention de l'ouvrage et de la zone cliquable pour accéder à sa consultation puisse être source de conflit. Il serait sans doute plus intuitif, et donc préférable, d'un point de vue ergonomique, de cliquer directement sur le titre de l'ouvrage choisi plutôt que de cliquer sur une icône accompagnée de son explication textuelle. Ce mode de navigation alourdit d'ailleurs inutilement la présentation.

L'examen de cette première page pose également la question de savoir si l'accès aux différentes sections des ouvrages doit pouvoir se faire dès le premier niveau de consultation. Lorsque les titres des ouvrages numérisés commenceront à se faire nombreux, cette option s'avérera impossible à maintenir. Il conviendrait donc de cliquer d'abord sur la désignation de l'ouvrage pour accéder à une nouvelle page qui présenterait, dans un second temps, le découpage de cet ouvrage en différentes sections.

On note enfin que ce prototype rudimentaire n'offre qu'un seul et unique mode d'accès aux ouvrages numérisés, rigide et contraignant. Il serait intéressant de varier les modes de navigation afin d'offrir à l'internaute un accès plus riche et plus souple aux différents ouvrages : navigation textuelle à travers une liste de titres, navigation iconique à partir d'une carte figurant différentes aires géographiques, navigation thématique, etc. En tout état de cause, le volume croissant des ouvrages numérisés rendra progressivement indispensable la conception d'une interface de recherche permettant une interrogation multicritère. Il conviendra de veiller à ce que les différents modes de navigation et la complexité de l'interface de recherche restent adaptés et proportionnés à la nature et à la quantité des fonds effectivement numérisés.

1.2.2. Consultation globale

Après avoir cliqué sur l'icône correspondant à l'ouvrage choisi, l'internaute accède à la page suivante :



L'idée d'insérer un niveau de consultation intermédiaire qui permet de visualiser plusieurs pages numérisées à la fois permet d'appréhender l'ouvrage dans une relative globalité. Le choix d'y faire figurer un ensemble de six pages correspond selon nous à un bon compromis : les vignettes sont suffisamment grandes pour que l'on puisse discerner la structuration des pages, et donc la nature de leur contenu, tout en restant assez nombreuses pour donner à l'internaute un recul minimal et une vue d'ensemble acceptable. Toutefois, ces vignettes gagneraient à être davantage mises en valeur : les pages blanches sur fond blanc ressortent très mal. Il conviendrait d'adopter un fond plus soutenu afin de rehausser les pages des ouvrages numérisés. Ces derniers constituent la richesse et le cœur même de la bibliothèque numérique. Peut-être faudrait-il aussi davantage expliciter l'ordre de succession des vignettes au sein de la page web.

Les boutons de navigation ne s'accompagnent d'aucun commentaire. Il conviendrait d'en expliciter la fonction par une indication textuelle du type « passer aux six vignettes suivantes ». Lorsqu'il clique sur la double flèche

orientée vers la droite, l'internaute est effectivement susceptible de se demander si cette action va simplement décaler d'un cran la succession des vignettes ou s'il va directement passer aux six vignettes suivantes. Dans le cadre d'une bibliothèque numérique, il est capital de veiller à ce que l'internaute ne s'égaré pas dans la succession des pages. Il est d'ailleurs à noter que cette consultation par « sauts », de six vignettes en six vignettes, n'est peut-être pas la plus ergonomique : elle brise la linéarité de la lecture. Une consultation par « glissement », de page en page, telle qu'on la trouve par exemple sur le site de l'INHA²⁶, est sans doute plus intuitive, mais elle est également plus difficile à réaliser...

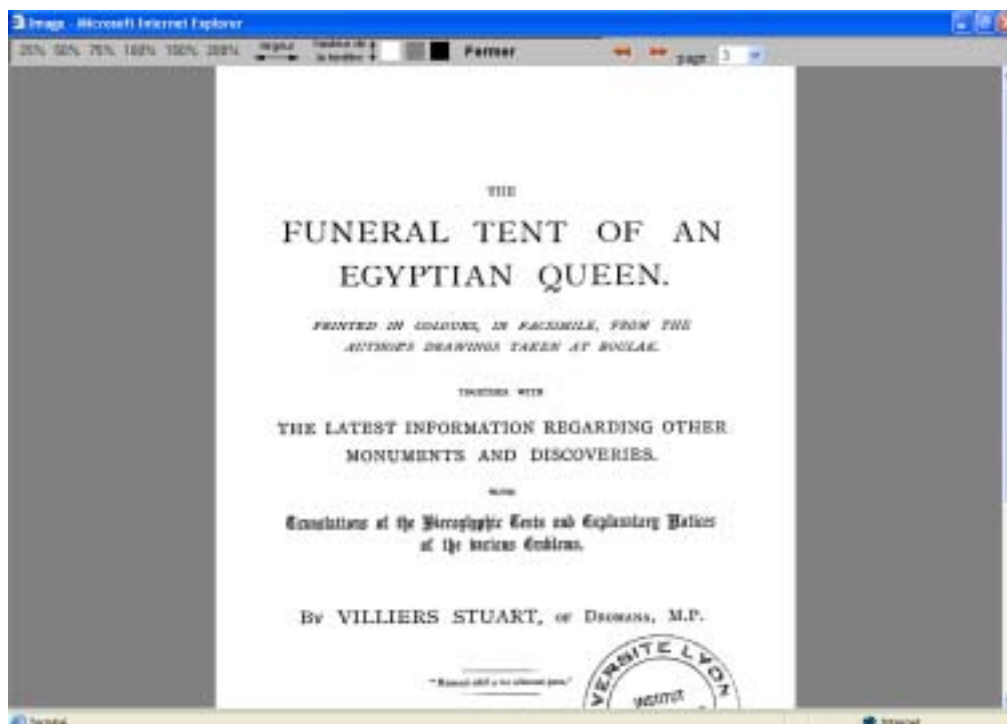
La présence de boutons de navigation permettant d'accéder directement aux six dernières ainsi qu'aux six premières pages de l'ouvrage serait plus qu'appréciable. D'emblée, une telle navigation permettrait d'appréhender l'ouvrage dans sa globalité. Contrairement à l'édition électronique en mode texte, une bibliothèque numérique qui propose des pages numérisées en mode image s'attache à reproduire le livre dans son aspect matériel et concret. Dans cette perspective, le mode de lecture proposé doit tenter de se rapprocher, dans la mesure du possible, des habitudes de lecture traditionnelles. Or le plus souvent, lorsqu'un lecteur se saisit d'un ouvrage, il l'appréhende dans sa totalité en accédant directement aux premières et dernières pages de l'objet que constitue le livre.

Compte tenu du contexte, le bouton de navigation nommé « sommaire » est susceptible d'être mal interprété : il peut suggérer un lien vers le sommaire de l'ouvrage visualisé alors qu'il redirige l'internaute vers la liste des ouvrages numérisés, c'est-à-dire vers la page dont il vient. Un bouton nommé « Retour à la liste des ouvrages » serait moins ambigu.

²⁶ Institut National d'Histoire de l'Art. *Dossiers d'images numérisées* [en ligne]. Disponible sur : <<http://www.inha.fr/bibliotheque/expo-presentation.html>> (consulté le 13.06.2004).

1.2.3. Consultation rapprochée

En cliquant sur l'une des vignettes, l'internaute accède à une page du type suivant :



Par défaut, lorsque l'internaute clique sur une vignette, la page numérisée s'affiche sur toute la largeur de la page web nouvellement ouverte. Pour permettre à l'internaute d'appréhender le folio dans sa globalité, il serait selon nous préférable que l'affichage par défaut se fasse tel qu'il qu'on le voit ci-dessus, c'est-à-dire tel qu'il devient lorsqu'on clique sur « hauteur de la fenêtre ». Il faut que l'internaute garde le sentiment de consulter un « objet livre » constitué d'« objets pages » clairement identifiables et délimités. C'est là que réside l'intérêt d'une bibliothèque numérique en mode image par rapport à l'édition électronique : elle présente le livre comme un objet archéologique. Il convient donc de faire en sorte que la page se détache sur un fond dont la couleur permette de la mettre en valeur et de bien discerner ses limites. On remarque que trois couleurs de fond sont proposées : blanc, gris et noir. La possibilité d'afficher la page sur fond blanc nous semble superflue car elle présente l'inconvénient de noyer la page de l'ouvrage numérisé dans la page web qui la présente.

A ce niveau aussi, la présence de boutons de navigation permettant d'accéder directement à la dernière ainsi qu'à la première page de l'ouvrage serait appréciable. Par ailleurs, l'accès direct à une page donnée via la saisie ou la sélection d'un numéro de page est un mode de navigation souple et naturel qu'il convient de conserver et de mettre en valeur. Toutefois, il faudrait préciser s'il s'agit de la numérotation automatique des pages numérisées ou s'il s'agit du foliotage réel de l'ouvrage numérisé. Sans cette précision, ce mode de navigation risque de rester ambigu pour l'internaute.

La présentation étalée, selon un axe horizontal, des différentes échelles possibles pour l'affichage de la page n'est peut-être pas la meilleure qu'on puisse imaginer : elle prend beaucoup de place et n'est explicitée par aucune indication. Une liste déroulante, annoncée par un terme explicatif tel qu'« échelle », conviendrait sans doute mieux. On constate par ailleurs que l'emplacement du bouton « Fermer » est pour le moins inattendu. L'internaute qui souhaite fermer une fenêtre a spontanément tendance à diriger la souris vers le coin supérieur droit de l'écran. Parce que ce bouton se distingue, en raison de sa fonction, des autres options proposées, il conviendrait de l'isoler des autres boutons en le plaçant en haut et à droite de la barre de commandes.

2. Conception et réalisation du site

Après avoir tiré des enseignements de la veille comparative et procédé à l'analyse de l'existant, il restait à proposer et à réaliser une solution pour la consultation des ouvrages numérisés. Pour ce faire, nous disposions des outils suivants : Macromédia Dreamweaver MX, Adobe Photoshop 6.0, Adobe Illustrator 10, Easy PHP 1.7.

La conception d'un site de ce type présente deux volets distincts et complémentaires : un premier aspect concerne la réalisation de l'interface

graphique tandis qu'un second volet porte sur la conception, l'alimentation et l'interrogation d'une base de données MySQL.

2.1. Concevoir et réaliser l'interface graphique

La conception de l'interface de la bibliothèque numérique devait obéir à une consigne restrictive majeure : le strict respect de la charte graphique du site de la Maison de l'Orient et de la Méditerranée. La bibliothèque numérique est effectivement appelée à être intégrée comme une sous-rubrique du site général de la Maison de l'Orient. Il convenait donc de reprendre les différents éléments qui structurent le site de la MOM. Une seconde contrainte, d'ordre technique, s'est imposée d'elle même : force était de se limiter aux solutions informatiques que nous étions nous même capable de mettre en œuvre.

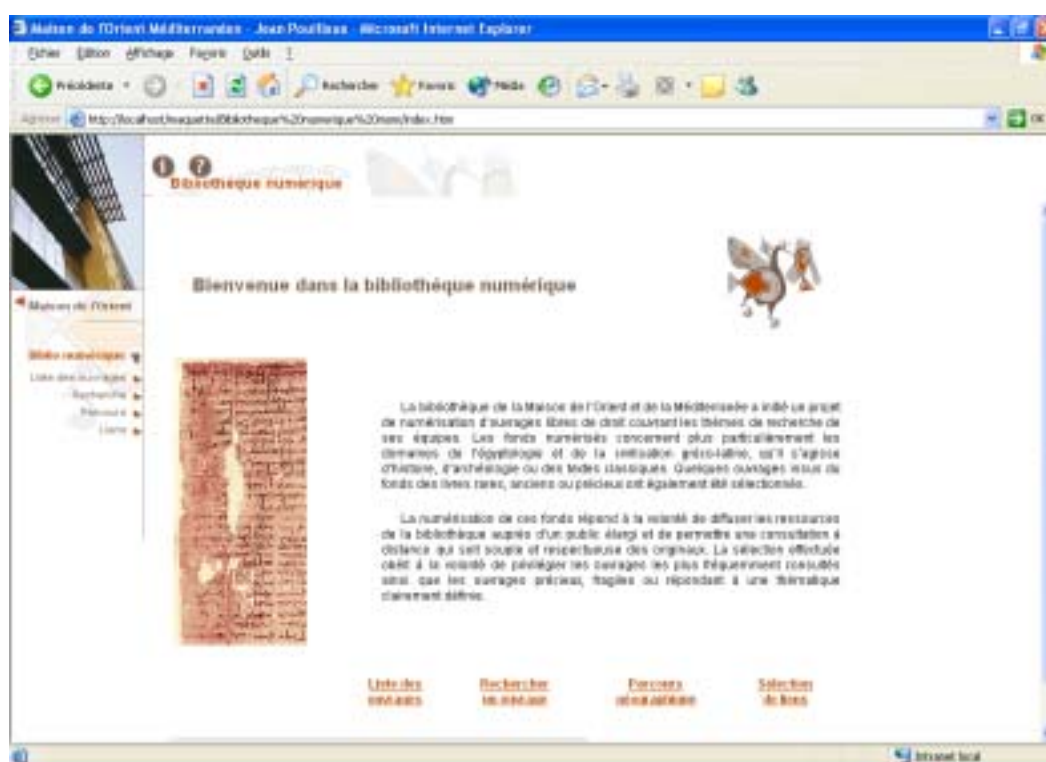
L'une des premières choses à faire consistait à récupérer la feuille de style du site de la MOM, nommée « wwwmom.css ». Il convenait également de récupérer les frames qui structurent la présentation des pages du site : le frame supérieur, dans lequel l'intitulé de la rubrique se trouve mentionné ainsi que le frame de gauche, qui propose deux menus de navigation : un menu général²⁷, qui permet de naviguer parmi les 9 principales rubriques du site, et un menu local qui présente les sous-rubriques de la rubrique dans laquelle on se trouve.

2.1.1. Page d'accueil

La conception de la page d'accueil obéit à des impératifs de clarté et de dépouillement. Elle comporte un titre qui marque clairement l'entrée de l'internaute dans l'espace que constitue la bibliothèque numérique. La présence du logo de la Maison de l'Orient identifie d'emblée le producteur de la bibliothèque numérique. La page offre également un texte de présentation qui répond aux exigences de concision propres à l'écriture pour écran. On a veillé à ce que la

longueur des lignes de texte ne soit pas excessive, afin d'assurer un plus grand confort et de ne pas ralentir la lecture. Les couleurs de police, comme celle du fond, étaient imposées par la feuille de style. Elles garantissent une bonne lisibilité.

La navigation peut se faire de deux façons différentes : soit par le menu local de gauche, que l'on retrouve sur toutes les pages du site, soit par la liste de liens présente dans le corps même de la page d'accueil. Il nous a en effet paru nécessaire de doubler la navigation sur la première page de la bibliothèque numérique dans la mesure où le menu local du site de la Maison de l'Orient, qui est un menu déroulant relativement discret, n'est pas immédiatement perceptible. Les quatre rubriques que nous avons jugé bon de créer sont les suivantes : liste des ouvrages, rechercher un ouvrage, parcours géographique, sélection de liens.



²⁷ Ce dernier apparaît à la place de la photographie de la Maison de l'Orient lorsqu'on passe la souris sur l'intitulé « Maison de l'Orient ».

2.1.2. Liste des ouvrages

Cette rubrique permet à l'internaute de prendre connaissance de l'ensemble du fonds numérisé. Pour une meilleure lisibilité, nous avons fait le choix de ne retenir que les champs essentiels à une première identification des ouvrages : l'auteur²⁸, le titre²⁹ et l'année de publication. Afin de limiter l'utilisation de l'ascenseur vertical, les résultats ne s'affichent que 8 par 8. Au bas du tableau, des icônes permettent de passer à la page suivante, à la page précédente, ainsi qu'à la première ou à la dernière page de la liste. Ces icônes, imposées par la charte graphique, ont une signification identique sur l'ensemble du site de la Maison de l'Orient, répondant en cela aux critères de qualité ergonomiques³⁰ définis par l'INRIA.

Auteur	Titre de l'ouvrage	Année
Champollion Le Jeune	Monuments de l'Égypte et de la Nubie - tablettes descriptives continues aux manuscrits autographes rédigés sur les lieux. Tome 1	1844
Champollion Le Jeune	Monuments de l'Égypte et de la Nubie - tablettes descriptives continues aux manuscrits autographes rédigés sur les lieux. Tome 2	1844
Dumont, Albert Chaplain, Jules	Les Céramiques de la Grèce précoce. Tome premier : résumé de la peinture des vases grecs depuis les céphales jusqu'au ve siècle avant Jésus-Christ suite d'un choix de vases peints trouvés en Grèce	1888
Dumont, Albert Chaplain, Jules	Les Céramiques de la Grèce précoce - vases peints et brèves copies. Première partie - vases peints	1881
Dumont, Albert Chaplain, Jules	Les Céramiques de la Grèce précoce. Tome second : réalignes archéologiques suivis d'un choix de terres cuites, bronzes et métaux trouvés en Grèce	1888
Engelstein, Richard	Bilder atlas zu Ovids Metamorphosen	1988
Engelstein, Richard	Bilder atlas zum Homer	1988
Geis, Dieter, Hermann de	Two recessed tombs at Thebes	1907

²⁸ Il peut s'agir de plusieurs auteurs.

²⁹ Le titre de l'ouvrage comprend en fait le titre, le sous-titre et la tomaisn, trois éléments nécessaires et suffisants pour distinguer les ouvrages d'une même série.

³⁰ Il s'agit ici du critère « Homogénéité/Cohérence ». cf. SCAPIN D. L., BASTIEN J.M.C. *Inspection d'interfaces et critères ergonomiques*. Rocquencourt : INRIA., 14 p. Rapport de recherche n° 2901 : INRIA : 1996.

Pour garantir une certaine souplesse, nous avons fait en sorte que l'internaute puisse trier les ouvrages par ordre alphabétique d'auteur, par ordre alphabétique du titre ou par année de publication. Il suffit pour cela de cliquer sur les intitulés de colonne. Afin de rendre l'interface vivante et interactive, nous avons également fait en sorte que le passage de la souris sur les différentes lignes du tableau s'accompagne d'un changement de couleur selon un mode alterné : selon qu'on passe la souris sur le titre du premier ou du second ouvrage, le fond de la ligne correspondante devient jaune ou gris clair, et ainsi de suite. Nous avons également choisi de faire apparaître le nombre total de résultats, afin que l'internaute puisse évaluer le volume de la bibliothèque numérique, ainsi que la numérotation des résultats effectivement affichés sur la page, afin qu'il puisse toujours se repérer.

2.1.3. Rechercher un ouvrage

L'internaute peut interroger la base de données selon six critères différents : recherche par auteur, par titre, par année de publication, par éditeur, par mot sujet et par lieu. Compte tenu du volume de ces fonds, il nous paru superflu de concevoir deux interfaces de recherche, permettant de choisir entre une « recherche simple » et une « recherche experte ». L'unique interface permettant de rechercher un ouvrage se présente de la façon suivante :



Un lien intitulé « Aide à la recherche » vient expliciter les modalités d'interrogation, qui sont les suivantes : entre les différents champs, l'opérateur booléen implicite est le « ET ». Le champ « Nom de lieu » correspond à une indexation du contenu des ouvrages par site géographique. Les champs « Mot sujet 1 » et « Mot sujet 2 » correspondent à une indexation des ouvrages par mot clé. Le doublement du champ « mot sujet » permet de combiner deux mots clés différents.

Quel que soit le champ interrogé, à l'exception de l'année de publication, la troncature illimitée, à droite comme à gauche, fonctionne automatiquement. Par exemple, si l'on cherche un ouvrage de Champollion Le Jeune, on obtiendra le même résultat que l'on saisisse « Champollion Le Jeune », « Champollion », « champ » ou même « pollion ». De même, il n'est pas nécessaire de saisir l'intégralité des titres d'ouvrage : la saisie de quelques mots, à condition qu'ils se suivent effectivement dans le titre recherché, d'un seul mot ou même d'une simple syllabe suffit à obtenir les résultats voulus.

La casse des lettres est indifférente : on peut saisir en minuscules comme en majuscules. De même, l'accentuation n'est pas prise en compte. Si on cherche un

ouvrage d'André Thévet, par exemple, on obtiendra donc le même le résultat que l'on saisisse « Thévet », « thévet », « Thevet » ou « thevet ».

L'interrogation des champs « auteur », « mot sujet » et « nom de lieu » peut se faire de deux façons distinctes : soit par une saisie libre dans les champs correspondants soit par l'utilisation d'index accessibles grâce aux icônes en forme de points d'interrogation. L'internaute clique sur le point d'interrogation correspondant au champ qu'il souhaite interroger. Dans la fenêtre d'index nouvellement ouverte, il sélectionne un terme en cliquant dessus. Automatiquement, la fenêtre d'index se referme et le terme choisi s'inscrit dans le champ correspondant.

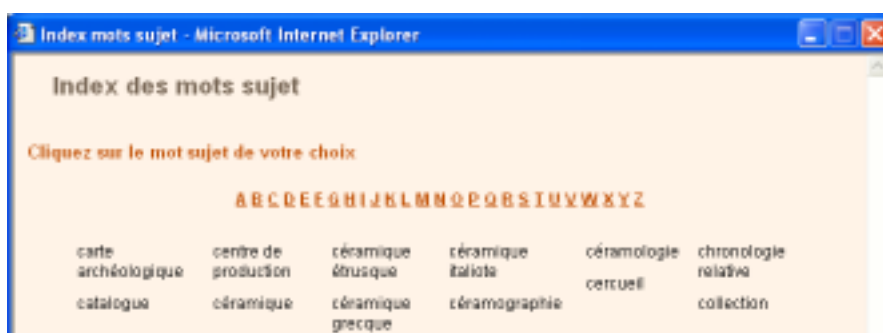
Les points d'interrogation renvoient à des index du type suivant :



Ces index permettent de faciliter l'interrogation de la base de données dans la mesure où ils indiquent à l'internaute l'ensemble des auteurs, mots sujet et noms de lieu pour lesquels il existe effectivement un ou plusieurs résultats. Pour l'index des mots sujet, appelé à contenir un grand nombre de termes, nous avons ajouté un « alphabet interactif » qui permet de sélectionner et d'afficher les mots sujet en fonction de leur initiale. Il se présente de la façon suivante :



Lorsqu'on clique sur la lettre C, par exemple, seuls les mots sujet commençant par C s'affichent :



Il n'est pas exclu d'étendre le principe de cet alphabet interactif aux autres index. Cela dépendra de l'accroissement de leur volume, difficilement prévisible à ce jour.

Après avoir composé son équation de recherche et cliqué sur l'icône « Rechercher », l'internaute accède à l'une des deux pages suivantes, suivant que des résultats correspondent ou non à sa requête :

3 Maires de l'Orient Méditerranéen - Jean-Paul Bass - Microsoft Internet Explorer

Recherche numérique

Résultats

Auteur = W. J. de Lézervat, Ch.
8 réponses(s) / Résultats 1 à 8.

Auteur	Titre de l'ouvrage	Pages
W. J. de Lézervat, Ch.	Ette des monuments éternographiques : matériaux pour l'histoire des régions et des moeurs de l'Italie. Tome 1 - plaques	1044
W. J. de Lézervat, Ch.	Ette des monuments éternographiques : matériaux pour l'histoire des régions et des moeurs de l'Italie. Tome 1 - texte	1044
W. J. de Lézervat, Ch.	Ette des monuments éternographiques : matériaux pour l'histoire des régions et des moeurs de l'Italie. Tome 2 - plaques	1067
W. J. de Lézervat, Ch.	Ette des monuments éternographiques : matériaux pour l'histoire des régions et des moeurs de l'Italie. Tome 2 - texte	1067
W. J. de Lézervat, Ch.	Ette des monuments éternographiques : matériaux pour l'histoire des régions et des moeurs de l'Italie. Tome 3 - plaques	1088
W. J. de Lézervat, Ch.	Ette des monuments éternographiques : matériaux pour l'histoire des régions et des moeurs de l'Italie. Tome 3 - texte	1088
W. J. de Lézervat, Ch.	Ette des monuments éternographiques : matériaux pour l'histoire des régions et des moeurs de l'Italie. Tome 4 - plaques	1081
W. J. de Lézervat, Ch.	Ette des monuments éternographiques : matériaux pour l'histoire des régions et des moeurs de l'Italie. Tome 4 - texte	1081

[< << >> >]

Nouvelle recherche Modifier la recherche

3 Maires de l'Orient Méditerranéen - Jean-Paul Bass - Microsoft Internet Explorer

Recherche numérique

Résultats

Auteur = W. J. de Lézervat, Ch.
0 réponse(s) / Résultats 1 à 0.

Aucun résultat ne correspond à votre requête.

Nouvelle recherche Modifier la recherche

Un script PHP permet de rappeler à l'internaute l'équation de recherche qu'il a entrée. Deux autres scripts permettent d'afficher le nombre de résultats trouvés et le numéro des résultats effectivement affichés dans la page. A chaque fois, l'internaute peut retourner à l'interface de recherche selon deux modalités distinctes : le bouton « Nouvelle recherche » réinitialise l'interface de recherche tandis que le bouton « Modifier la recherche » permet de conserver les critères de recherche déjà entrés. Comme dans la liste des ouvrages, l'internaute peut trier les résultats par ordre alphabétique d'auteur, par ordre alphabétique du titre ou par

date de publication. Le passage de la souris sur l'intitulé des différents résultats s'accompagne là encore d'un changement de couleur selon un mode alterné.

2.1.4. Parcours géographique

La rubrique « Parcours géographique » propose un mode de navigation de nature iconique. Elle permet de présenter les ouvrages en fonction de la civilisation dont ils traitent et d'offrir ainsi à l'internaute une visualisation des aires géographiques concernées par le programme de numérisation. Pour l'instant, trois grandes aires sont représentées : l'Egypte, le monde gréco-romain et les civilisations du Proche-Orient.



Après avoir cliqué sur l'une de ces civilisations, l'internaute accède à une liste de résultats sur laquelle figure un lien intitulé « Retour parcours », ce qui permet à l'internaute de revenir à la page dont il vient. La liste de résultats se présente ainsi :

Auteur	Civilisation	Somme
Dumont, Albert (1788-1870), Jules	Céramiques de la Grèce propre. Terre cuite - vases de la peinture des vases grecs depuis les origines jusqu'à la période alexandrine. Clés de lecture d'un choix de vases peints trouvés en Grèce	1000
Dumont, Albert Chastel, Jules	Céramiques de la Grèce propre. Vases peints et terres cuites. Première partie - vases peints	1001
Dumont, Albert Chastel, Jules	Céramiques de la Grèce propre. Terre cuite - reliefs archaïques jusqu'à l'époque de l'art hellénistique et romain. Gravés en relief	1002
Engelmann, Richard	Atlas des Civilisations	1003
Engelmann, Richard	Atlas des civilisations	1004
Denon, Louis (1747 - 1819)	Histoire-général Tome 2	1005
Denon, Louis (1747 - 1819)	Histoire-général Tome 1	1006
Denon, Louis (1747 - 1819)	Histoire-général Tome 1	1007

Cette répartition par grandes civilisations correspond en fait aux fonds dont sont issus les ouvrages numérisés : les ouvrages cotés en EGY³¹ trouvent naturellement leur place dans la zone cliquable intitulée « Egypte », ceux qui sont cotés en HCL, ACL, TXT ressortissent de la zone intitulée « Monde gréco-romain », tandis que les ouvrages cotés en BAB appartiennent à la zone intitulée « Proche-Orient ». Le choix de la carte géographique s'est trouvé limité par la question des droits d'auteur : il convenait de se procurer une illustration pour laquelle la Maison de l'Orient était titulaire des droits. Pour ce faire, nous nous sommes adressée au service cartographie de la Maison de l'Orient. Ce dernier nous a fourni une carte que nous avons retouchée sur Photoshop afin d'harmoniser les couleurs avec celles de la charte graphique et d'ajouter l'intitulé des aires géographiques.

Le problème de cette représentation réside dans le fait que le découpage géographique ne coïncide pas toujours avec la notion de civilisation : le monde gréco-romain, par exemple, s'est étendu bien au-delà de la péninsule grecque et de la botte italienne, y compris en Egypte. Fallait-il pour autant s'interdire toute

³¹ La signification des cotes est la suivante : EGY pour « égyptologie », HCL pour « Mondes grec et romain : histoire et civilisations », ACL pour « Mondes grec et romain : archéologie », TXT pour les textes grecs et latins, BAB pour « Monde arabo-musulman, monde byzantin et Orient chrétien ».

représentation géographique des fonds numérisés ? La représentation cartographique se comprend ici sur le mode symbolique : les ouvrages qui appartiennent à la civilisation gréco-latine sont accessibles à partir d'une zone cliquable qui est symboliquement centrée sur l'Italie et la Grèce. Il nous a semblé que l'internaute était en mesure de saisir la dimension symbolique de cette représentation. Il n'est toutefois pas exclu d'ajouter une note explicative, ou même de supprimer cette représentation géographique au profit d'un classement thématique qui serait malheureusement dépourvu de support visuel... Un classement chronologique serait également possible mais compte tenu de la nature des fonds numérisés, cette solution nous a paru moyennement pertinente.

2.1.5. Consulter un ouvrage

Après avoir sélectionné un ouvrage en cliquant sur l'un des résultats, l'internaute accède à une page du type suivant :

Notice	Sommaire	Consultation
Auteurs :	Albert Dumont Jules Chaplain	
Titre :	Céramiques de la Grèce propre. Tome second : mélanges archéologiques suivis d'un choix de terres cuites, bronzes et marbres trouvés en Grèce	
Date de publication :	1890	
Éditeur :	Paris : F. Didot	
Description physique :	266 p. (111) de pl., ill. en noir et en coul. ; 24 cm	
Langue(s) :	Français	
Indicateur par sujet :	Céramique grecque, peinture grec, peinture grecque, terre cuite, Égoutte de terre cuite, métal, marbre, bronze, bronze	
Indicateur par lieu :	Messalie, Grande Grèce, Grèce insulaire, Grèce, Monde gréco-romain	
Indicateur chronologique :	Ép. archaïque, Ép. classique	

Grâce à un système d'onglets, cette page permet tout à la fois de visualiser la notice de l'ouvrage, de prendre connaissance de son sommaire et de consulter

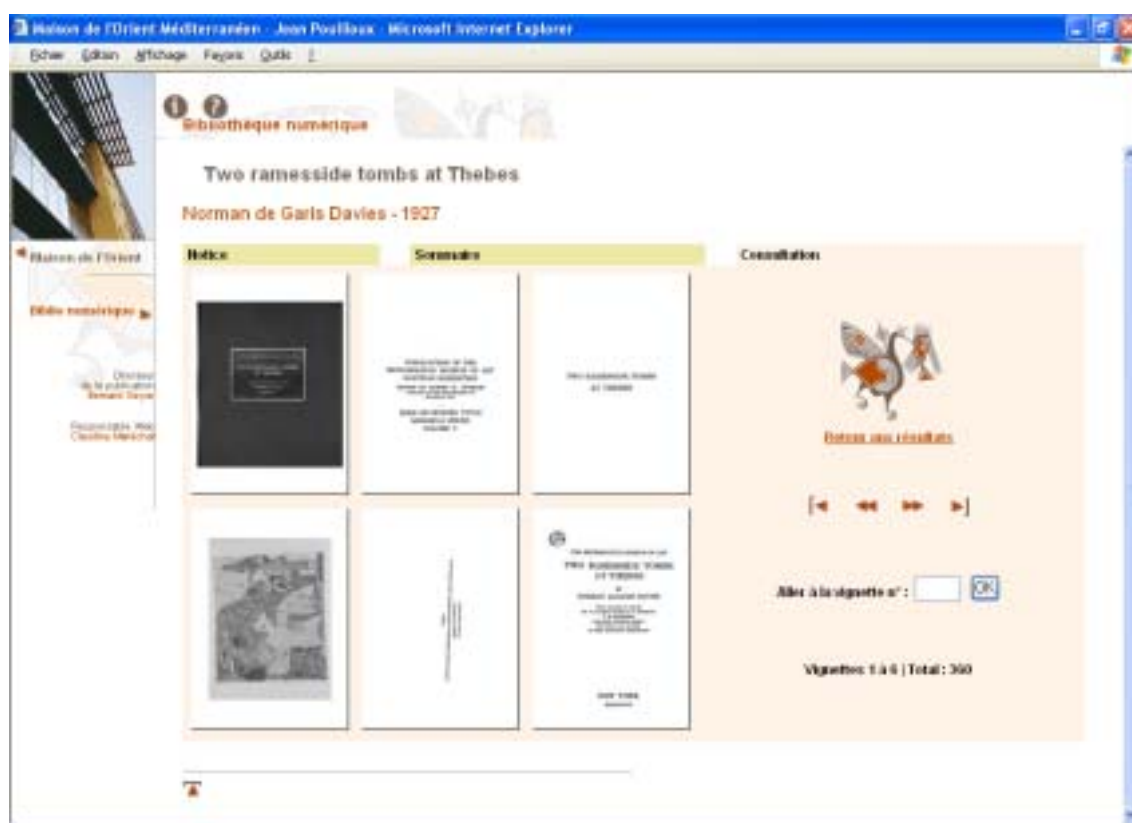
ses pages numérisées. A chaque fois, le titre de l'ouvrage est rappelé en gros caractères en haut de la page, suivi du nom de l'auteur et de la date de publication. Quel que soit l'onglet consulté, l'internaute sait ainsi de quel ouvrage il s'agit. De même, quel que soit l'onglet consulté, un lien permet toujours de retourner aux résultats de la recherche.

En plus de l'auteur, du titre et de la date de publication, l'internaute peut prendre connaissance, dans l'onglet « Notice », du lieu d'édition et du nom de l'éditeur (champ « édition »), de la collation de l'ouvrage (champ « description physique »), des langues dans lesquelles est écrit l'ouvrage (champ « langue(s) »), ainsi que des mots sujet, des noms de lieu et des époques chronologiques qui caractérisent le contenu de l'ouvrage (champs « indexation par sujet » et « indexation par lieu » et « indexation chronologique »). A l'exception des livres extérieurs à la bibliothèque ou issus du fonds ancien, le contenu de ces champs est tiré du catalogage effectué sur Loris, le logiciel documentaire de la bibliothèque.

L'onglet « Sommaire » permet de visualiser la structure de l'ouvrage numérisé. Ce sommaire est constitué de liens hypertexte qui permettent d'accéder directement à la première page de chaque partie dont le titre est mentionné. L'indentation des titres vers la droite est fonction de leur niveau hiérarchique, défini dans la base de données.



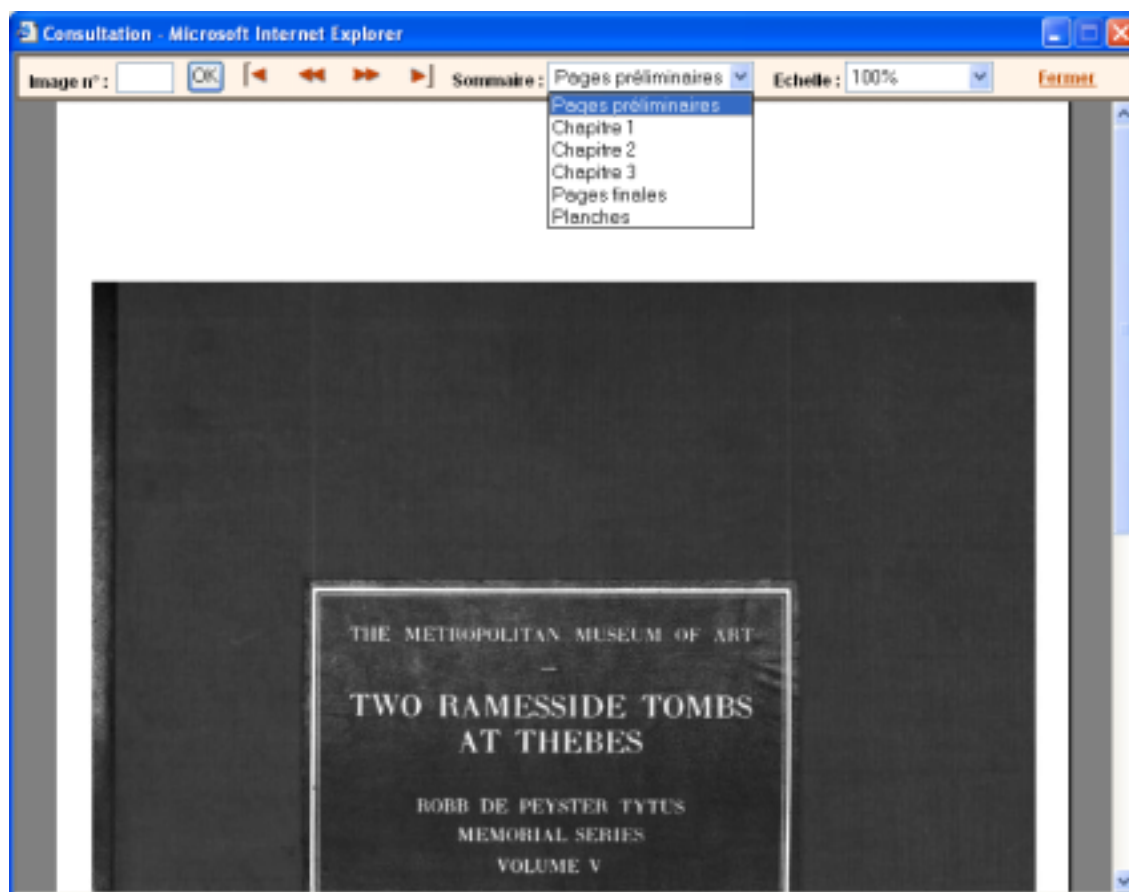
Pour l'onglet « Consultation », nous avons choisi de reprendre le système de visualisation par groupe de 6 vignettes qui existait dans le prototype, jugeant que ce niveau de consultation intermédiaire correspondait à un bon compromis. Nous avons pris soin, en revanche, de faire apparaître ces vignettes sur un fond coloré qui permette de les mettre en valeur :



Lorsqu'on passe la souris sur les icônes de navigation, leur signification est explicitée. L'internaute peut ainsi « Revenir aux 6 premières vignettes », « Revenir aux 6 vignettes précédentes », « Aller vers les 6 vignettes suivantes » ou « Aller vers les 6 dernières vignettes ». Dans le souci de lever toute ambiguïté sur la signification des icônes, ce principe d'explicitation³² est d'ailleurs généralisé à l'ensemble des icônes de la bibliothèque numérique. L'internaute peut également accéder directement à la vignette qu'il désire en saisissant son numéro puis en cliquant sur OK. Il s'agit ici de la pagination logique, c'est-à-dire de la numérotation des images, et non de la pagination physique de l'ouvrage. Nous avons d'ailleurs pris soin d'employer le terme « vignette », et non page, pour éviter la confusion entre pagination logique et pagination physique. Enfin, un indicateur de positionnement vient préciser le numéro des vignettes effectivement visualisées dans l'onglet et mentionne le nombre total d'images existantes pour l'ouvrage concerné.

Pour visualiser un folio en pleine page, l'internaute clique sur l'une des vignettes visualisées. La page de consultation rapprochée se présente alors sous la forme suivante :

³² Il s'agit simplement, dans le code HTML, de l'option « alt = " " », ajoutée dans les balises d'images.



C'est seulement à ce niveau que l'internaute peut véritablement lire l'ouvrage. Plusieurs modes de navigation s'offrent alors à lui : il peut choisir d'accéder directement à telle ou telle image, en saisissant son numéro. Là encore, nous avons préféré le terme « image » au mot « page », pour bien signifier qu'il ne s'agit pas de la pagination physique. L'utilisateur peut également, dans l'ordre des icônes, revenir à la première image ou à l'image précédente, ou encore accéder à l'image suivante ou à la dernière image. Une liste déroulante, qui rappelle la structuration de l'ouvrage consulté, permet en outre d'accéder directement à la première page de telle ou telle partie de l'ouvrage.

Une autre liste déroulante permet par ailleurs de régler l'échelle pour l'affichage de la page. Cette échelle comprend 8 degrés : 200 %, 150 %, 100 %, 75 %, 50 %, 25 %, largeur page, hauteur page. Par défaut, nous avons souhaité faire en sorte que la page s'affiche en « hauteur page », ce qui permet à l'internaute d'appréhender le folio dans sa globalité, sur un fond gris foncé qui le

met en valeur. Nous avons par ailleurs veillé à ce que la barre du haut s'inscrive dans une frame afin qu'elle reste toujours visible pour l'internaute, même lorsque l'échelle de consultation nécessite l'utilisation de l'ascenseur vertical. Enfin, un lien placé en haut à droite permet de fermer la fenêtre.

2.1.6. Sélection de liens

Dès le début de notre stage, Véronique Lacroix avait manifesté la volonté de voir figurer sur le site une sélection de liens vers d'autres bibliothèques numériques en rapport avec les thèmes de recherche de la Maison de l'Orient. Les sites proposant des ressources intéressantes dans les domaines de l'Antiquité, de l'archéologie ou de l'histoire du bassin méditerranéen se comptent par centaines. Il existe d'ailleurs des pages répertoriant et classant les sites de ce type³³, qui sont d'ailleurs de natures très diverses³⁴. La difficulté, dans un premier temps, était donc de ne pas se perdre dans le labyrinthe incroyablement touffu des liens et renvois entre les différents sites web. Une seconde difficulté consistait à définir ce qu'il convenait d'entendre par « bibliothèque numérique », car cette expression peut recouvrir des réalités très différentes selon qu'on l'envisage dans un sens strict ou dans un sens large.

Dans l'une de ses rubriques, le site de la bibliothèque de l'Université Paris VIII³⁵ propose par exemple un classement de bibliothèques numériques en prenant soin de préciser : « Nous entendons par *bibliothèque numérique* un site mettant à disposition une collection de données numérisées ». En fait, il s'agit d'une liste de ressources électroniques recouvrant des réalités techniques parmi les plus diverses : de la numérisation en mode image à l'édition électronique en mode texte, en passant par des fichiers PDF ou des fichiers TXT... Dans cette

³³ A titre d'exemple, on peut mentionner : IOANNIS GEORGANAS. Mediterranean Archaeology Resources [en ligne]. Disponible sur : <http://www.geocities.com/i_georganas/main.html> (consulté le 12.08.2004).
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE. Liste de sites consacrés à l'Antiquité gréco-romaine [en ligne]. Disponible sur : <http://callimac.vjf.cnrs.fr/Ant_class.html> (consulté le 14.07.2004)

³⁴ Certains de ces sites sont des bases de données fournissant des ressources secondaires tandis que d'autres offrent un accès aux ressources primaires.

perspective, des sites tels que celui du « Perseus Project »³⁶ ou de l'« Epigraphische Datenbank Heidelberg »³⁷ sont effectivement comptés au rang des bibliothèques numériques.

Après une brève discussion avec Véronique Lacroix, qui permit d'éclaircir les choses, je compris que par *bibliothèque numérique* elle entendait en fait « site permettant d'accéder à des documents numérisés en mode image », c'est-à-dire des projets qui sont techniquement comparables à celui engagé par la Maison de l'Orient. Or les sites de ce type sont infiniment moins nombreux que ceux qui proposent de l'édition électronique en mode texte, surtout si on se limite aux sites qui offrent un contenu en rapport avec les domaines de recherche de la Maison de l'Orient. Fallait-il donc effectuer une sélection se fondant prioritairement sur des critères formels ou sur l'intérêt scientifique du contenu ? Fallait-il écarter des ressources aussi intéressantes, dans le domaine de l'Antiquité, que la Perseus Digital Library, l'Epigraphische Datenbank Heidelberg ou la Bibliotheca Classica Selecta³⁸ parce qu'elles ne répondaient pas à ces critères formels ? Estimant que les chercheurs n'auraient aucun mal à trouver des liens vers les ressources de ce type, signalées sur de nombreux sites, Véronique Lacroix préférait se limiter strictement aux sites offrant des ressources numérisées en mode image. Notre sélection, qui privilégie donc l'aspect formel, se limite pour l'instant à 6 bibliothèques numériques³⁹. La page de liens se présente de la façon suivante :

³⁵ BIBLIOTHEQUE UNIVERSITE PARIS 8. Bibliothèques numériques [en ligne]. Disponible sur : <<http://www-bu.univ-paris8.fr/Ref/BibTextes.html>>. (consulté le 10.08.2004)

³⁶ TUFTS UNIVERSITY. The perseus digital library [en ligne]. Disponible sur : <<http://www.perseus.tufts.edu/>>. (consulté le 26.07.2004)

³⁷ HEIDELBERG UNIVERSITÄT. Epigraphische Datenbank Heidelberg [en ligne]. Disponible sur : <<http://www.uni-heidelberg.de/institute/sonst/adw/edh/indexe.html>>. (consulté le 14.08.2004)

³⁸ UNIVERSITE CATHOLIQUE DE LOUVAIN. Bibliotheca Classica Selecta. [en ligne]. Disponible sur : <<http://bcs.fltr.ucl.ac.be/>> (consulté le 13.06.2004)

³⁹ Il s'agit de : Cefael (publications de l'Ecole française d'Athènes), Casalini Digital Division (publications de l'Ecole française de Rome), Collections Medica (corpus des médecins de l'Antiquité édité par la BIUM de Paris V), bibliothèque numérique du pôle international de la préhistoire, Ancient Near East and the Mediterranean World (Université de Chicago), Cuneiform Digital Library Initiative (Université de Californie et Max Planck Institute for the History of Science).



A chaque fois, un texte de présentation vient préciser la nature du contenu de la bibliothèque numérique. Pour accéder au site de ces bibliothèques numériques, l'internaute peut soit cliquer sur le nom de la bibliothèque choisie, à gauche, soit sur son icône, à droite. Nous avons pris soin de demander l'autorisation, auprès des institutions concernées, de faire figurer ces illustrations sur notre site afin d'en faire des liens hypertexte. Trop contentes de se voir proposer une promotion gratuite, lesdites institutions n'ont opposé aucune résistance, au contraire. Nous avons par ailleurs veillé à ce que la cible du lien hypertexte s'affiche dans une nouvelle fenêtre, et non pas dans la fenêtre en cours⁴⁰, et ce afin d'éviter que l'internaute ne sorte de notre propre site. En bas de la page, figure un lien vers le courriel de la bibliothèque qui invite ce dernier à signaler l'existence d'autres bibliothèques numériques.

⁴⁰ En termes techniques, il s'agit en fait de choisir l'option « blank » (affichage dans une nouvelle fenêtre) plutôt que l'option « self » (affichage dans la fenêtre en cours), qui est l'option par défaut.

2.2. Concevoir, alimenter et interroger la base de données

Dans un second temps, pour rendre le site effectivement interactif et opérationnel, il nous a fallu procéder à l'élaboration, l'alimentation et l'interrogation, grâce à des scripts PHP, d'une base de données MySQL. A l'origine, il était question de concevoir une base de données spécifique, dans son intégralité. Finalement, l'informaticienne en charge du site web de la Maison de l'Orient a estimé qu'il serait plus simple de reprendre une base de données bibliographique conçue par l'une de ses stagiaires.

2.2.1. La base BibliOnline

La base de données réalisée par la stagiaire en question est une base initialement conçue pour les publications des chercheurs de la Maison de l'Orient. Baptisée BibliOnline, elle permet de renseigner de nombreux champs : auteur(s), avec précision éventuelle du rôle tenu par celui-ci (éditeur, traducteur, directeur, collaborateur), titre de l'ouvrage ou de l'article, titre de la revue ou de l'ouvrage collectif, support, ISBN, cote, etc. Cette base s'accompagne d'une interface web permettant, entre autres choses, de saisir les enregistrements, d'indexer les références bibliographiques d'après un thésaurus hiérarchisé et d'interroger le contenu de la base. Le formulaire de saisie et l'interface d'interrogation se présentent comme suit :

2.2.2. Adapter la base de données

Si le fait de reprendre une base de données existante nous a d'abord permis de gagner du temps, il s'est avéré que la structure de cette base de données, conçue pour un autre usage, ne répondait qu'imparfaitement aux besoins induits par les choix que nous avons faits lors de la conception de notre interface graphique. Nous avons donc été amenée à modifier cette base de données pour l'adapter aux

usages que nous souhaitons en faire. Si nous avons conservé, dans le souci de ne pas dénaturer la structure de cette base, les nombreuses fonctionnalités qui ne nous étaient pas immédiatement utiles pour notre projet, nous avons en revanche ajouté les tables et les champs dont nous avons besoin. A ce jour, la structure de la table est la suivante :

Table	Action	Enregistrements	Type	Taille
<input type="checkbox"/> auteurs	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	21	MyISAM	5,9 Ko
<input type="checkbox"/> bib_aut	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	58	MyISAM	7,8 Ko
<input type="checkbox"/> bib_chrono	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	43	MyISAM	1,4 Ko
<input type="checkbox"/> bib_coll	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	18	MyISAM	2,5 Ko
<input type="checkbox"/> bib_judec	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	211	MyISAM	8,3 Ko
<input type="checkbox"/> bib_lieu	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	112	MyISAM	2,0 Ko
<input type="checkbox"/> bib_lie	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	5	MyISAM	2,3 Ko
<input type="checkbox"/> bib_pro	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	2	MyISAM	2,1 Ko
<input type="checkbox"/> bib_requis	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	1	MyISAM	2,3 Ko
<input type="checkbox"/> collections	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	0	MyISAM	2,6 Ko
<input type="checkbox"/> indexation	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	92	MyISAM	4,8 Ko
<input type="checkbox"/> localisations	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	0	MyISAM	3,0 Ko
<input type="checkbox"/> productions	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	5	MyISAM	3,2 Ko
<input type="checkbox"/> reffib	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	38	MyISAM	12,0 Ko
<input type="checkbox"/> reffib_lie	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	0	MyISAM	2,1 Ko
<input type="checkbox"/> reffib	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	0	MyISAM	6,1 Ko
<input type="checkbox"/> t_chrono	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	17	MyISAM	2,4 Ko
<input type="checkbox"/> t_nom_lieu	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	40	MyISAM	1,8 Ko
<input type="checkbox"/> t_sommaires	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	88	MyISAM	37,8 Ko
<input type="checkbox"/> users	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	3	MyISAM	2,6 Ko
20 tables		Sommaire	1 216	113,5 Ko

Si l'indexation par sujet était possible grâce à la table « indexation », accompagnée de sa table intermédiaire « bib-index », qui permet de faire le lien entre les mots sujet et les références bibliographiques, l'indexation par nom de lieu et par époque chronologique n'était nullement prévue. Sur le même modèle, nous avons donc créé une table « t_nom_lieu » et une table « t_chrono », accompagnées de leurs tables intermédiaires respectives « bib_lieu » et « bib_chrono ». De même, aucun champ n'était prévu pour la mention de la langue des ouvrages. Dans la table « reffib » nous avons donc ajouté un champ intitulé « langue ». Surtout, rien n'était prévu pour la saisie des sommaires des ouvrages. Nous avons donc créé une table intitulée « t_sommaires ». La structure de cette dernière se présente de la façon suivante :

id_refbib	ordre	niveau	titre	page	url
1016	1	1	Pages préliminaires		
1016	2	1	Table		
1016	3	2	planches I - V	1	
1016	4	2	planches VI - 100	13	
1016	5	2	planche 101	25	
1016	6	2	planche 102	27	
1016	7	2	planches 101V - 100	28	
1016	8	2	planches 100B - 100VI	54	
1016	9	2	planches 100VI - 100VB	64	
1016	10	1	Planches	100	
1015	1	1	Pages préliminaires		

Le premier champ, intitulé « id_refbib » correspond à l'identifiant de la référence bibliographique concernée, c'est-à-dire le numéro de l'ouvrage dont on va décrire la structure. Le champ « ordre » correspond à l'ordre des parties de l'ouvrage dans leur succession linéaire, tandis que le champ « niveau » correspond au niveau hiérarchique de la partie concernée. Le champ « page » correspond au numéro de la première page de la partie de l'ouvrage concernée. Il s'agit de la pagination physique de l'ouvrage, c'est-à-dire de son foliotage réel, et non de la pagination logique issue du numériseur. Enfin, le champ « URL » permet d'introduire un lien hypertexte en direction de la page numérisée présente dans le SGBI Transvision.

La clé primaire de la table réside dans l'association des champs « idrefbib » et « ordre » : pour un ouvrage donné, il ne peut y avoir deux parties présentant le même numéro d'ordre. L'interrogation de la table « t_sommaire » permet d'afficher, dans l'onglet sommaire de la page de consultation, la structure d'un ouvrage donné et d'accéder ainsi directement à telle ou telle partie de l'ouvrage. L'écran suivant permet par exemple de décrire la structure de l'ouvrage 1008 :

№	№	№	Titre de la notice (chapitre)	Page
1007	18	1	Notes on the various chapters	139
1007	19	1	Appendix	143
1007	20	2	Meteorological observations	146
1007	21	2	Table of temperatures	156
1007	22	2	General itinerary for return journey from second l.	162
1008	1	1	Couverture	
1008	2	1	Pages préliminaires	
1008	3	1	Préface	ix
1008	4	1	Sommaire	xi
1008	5	1	Table des planches	xix
1008	6	1	Introduction	xx
1008	7	1	Texte	1
1008	8	2	Chapitre I. The tomb of Userhêt	1
1008	9	2	Chapitre II. The tomb of Apy	26
1008	10	1	Index	77
1008	11	1	Planches	87
1008	1	1	Pages préliminaires	
1008	2	1	Introduction	v

On peut représenter la structure de l'ouvrage 1008 de la façon suivante :

Couverture

Pages préliminaires

Préface

Sommaire

Table des planches

Introduction

Texte

Chapitre I. The tomb of Userhêt

Chapitre II. The tomb of Apy

Index

Planches

Par ailleurs, l'utilisation d'une base de données préexistante a nécessité quelques accommodements. Nous avons par exemple utilisé le champ « pagination » pour entrer l'ensemble de la description physique des ouvrages (pagination, illustrations, format). L'ensemble des changements effectués dans la base de données supposait en outre des modifications dans l'interface web que constitue le formulaire de saisie. Il faudra ainsi prévoir la saisie d'un champ langue, d'une indexation par lieu ainsi que de la structure des sommaires. En fin de compte, l'élaboration d'une base de données spécifique et adaptée n'aurait sans doute pas demandé beaucoup plus de travail.

2.2.3. Alimenter et interroger la base de données

Après avoir adapté la structure de la base de données, il convenait de l'alimenter. Nous avons donc renseigné les champs nécessaires pour tous les ouvrages ayant déjà été numérisés. La plupart des ouvrages avaient déjà été catalogués dans Loris, le logiciel documentaire dont se sert la bibliothèque. L'indexation par mot sujet et par nom de lieu, notamment, avait déjà été faite. Nous ne l'avons pas retouchée. Il a donc suffi de reprendre le contenu des champs qui nous intéressaient. En revanche, les ouvrages issus du fonds des livres rares, anciens ou précieux ne sont pas catalogués dans Loris, tout comme les quelques ouvrages qui ont été prêtés à la Maison de l'Orient le temps de leur numérisation. Pour ceux-là, il a fallu soit avoir recours à l'exemplaire physique soit consulter l'exemplaire numérique stocké dans Transvision.

En revanche, l'alimentation de la table « t_sommaires » a nécessité d'analyser la structure de l'ensemble des ouvrages ayant été « transvisés » puis de saisir leurs sommaires. La saisie des sommaires des ouvrages n'avait en effet jamais été faite. Cette opération s'est avérée relativement laborieuse, surtout lorsque l'impossibilité de consulter l'exemplaire physique nous a obligée à passer par une consultation des pages numérisées dans Transvision. La perception de la structure globale d'un ouvrage est en effet beaucoup plus aisée lorsqu'on peut l'appréhender sous sa forme matérielle. Par ailleurs, certains ouvrages anciens ne présentent ni sommaire ni table des matières. Il faut alors les feuilleter dans leur intégralité pour tenter de dégager leur structure logique. Dans certains cas, comme celui d'un ouvrage de Champollion Le Jeune⁴¹, qui reproduit des notes manuscrites, cela peut s'avérer particulièrement difficile. Pour des questions de temps, nous avons dû parfois nous en tenir à une structuration assez rudimentaire.

Une fois la base alimentée, il restait enfin à coupler l'interface graphique avec la base de données. Pour ce faire, nous avons eu recours à des scripts PHP. La

⁴¹CHAMPOLLION LE JEUNE, Monuments de l'Égypte et de la Nubie : notices descriptives conformes aux manuscrits autographes rédigés sur les lieux. Tome 2, Paris : F. Didot frères, 1844. 345 p.

programmation en PHP, que ce soit pour l'interrogation de la base de données ou pour certains aspects de l'interface graphique, comme l'affichage de l'équation de recherche ou encore la création du système des onglets, est sans doute ce qui nous a demandé le plus de temps et le plus d'efforts. Nous y sommes parvenue grâce à la consultation de sites tels que *PHP Facile*⁴² ou *PHPdébutant.org*⁴³, à la fréquentation de leurs forums ou à la demande de conseils auprès de personnes plus aguerries que nous dans ce genre d'exercices.

⁴² HEUTE, Thomas et HEUTE, Damien. PHP Facile [en ligne]. Disponible sur : <http://www.toutestfacile.com/phpinit.php?tef_site=php>. (consulté le 08.07.2004)

⁴³ PHPdébutant. PHPdébutant.org [en ligne]. Disponible sur : <<http://www.phpdebutant.org/>>. (consulté le 05.07.2004)

Conclusion

Le bilan que je tire de cette expérience professionnelle de quatre mois à la Maison de l'Orient et de la Méditerranée est très positif : la confiance que m'a témoignée ma responsable de stage et la liberté d'action qu'elle m'a laissée m'ont permis de mener de façon autonome un projet d'un grand intérêt, tant d'un point de vue technique que culturel. Ce stage, très complet, m'a donné l'occasion de mettre en pratique des compétences variées, acquises tout au long de mon année de formation : réflexion sur le concept de bibliothèque numérique ; conception et réalisation d'un site web ; création, alimentation et interrogation d'une base de données MySQL ; pratique des langages HTML et PHP ; compétences documentaires et bibliothéconomiques (indexation, formats bibliothéconomiques, interfaces de recherche, consultation de fonds anciens) ; connaissances théoriques et pratiques sur la chaîne de numérisation du patrimoine (scannage, traitement des images, gestion d'un SGBI).

Ces quatre mois de stage m'ont ainsi permis d'appréhender, dans leur globalité, les différents processus qui concourent à l'élaboration d'une bibliothèque virtuelle, de l'acquisition des images à la consultation sur le web. La mission qui m'a été confiée est aujourd'hui presque achevée et les fonds numérisés de la Maison de l'Orient devraient être prochainement consultables à partir de son site web. Mais au-delà des bibliothèques virtuelles et des aspects techniques qui leur sont propres, ce stage m'a aussi et surtout permis de me familiariser avec le fonctionnement, au quotidien, d'une bibliothèque de recherche, et ce dans un contexte qui m'est cher, celui des langues anciennes, de l'histoire et de l'archéologie.

Bibliographie

La présente bibliographie ne saurait être exhaustive. Les domaines de la numérisation et des bibliothèques virtuelles étaient trop vastes pour que l'on prétendît donner, dans le cadre de ce rapport, une bibliographie complète sur le sujet. On s'est donc ici limité à reproduire la liste des principaux ouvrages, articles et sites web personnellement consultés dans le cadre de notre stage.

Numérisation et recherche scientifique :

JOURNEES D'ETUDES DES 29 – 30 MAI 2002 ORGANISEES PAR L'ECOLE FRANÇAISE D'ATHENES, LA MAISON DE L'ORIENT MEDITERANEEN ET L'UNIVERSITE LUMIERE LYON 2. *Modèles opératoires de production et de diffusion des collections scientifiques dans les bibliothèques numériques, expérimentés à partir de la numérisation des publications de l'Ecole française d'Athènes*. Athènes : Ecole française d'Athènes, 2002. 93 p.

JOURNEE D'ETUDES ORGANISEE PAR L'ECOLE NORMALE SUPERIEURE DE LETTRES ET SCIENCES HUMAINES ET LE LABORATOIRE RECONNAISSANCE DE FORMES ET VISION DE L'INSA DANS LE CADRE DE L'INSTITUT DES SCIENCES DE DOCUMENT NUMERIQUE RHONE-ALPES LE 7 MARS 2002. *Valorisation des documents et numérisation des collections*. Villeurbanne : E.N.S.S.I.B., 2002. 56 p.

JOURNEE D'INFORMATION ET D'ETUDE ORGANISEE PAR LE LABORATOIRE RECONNAISSANCE DE FORMES ET VISION DE L'INSA DANS LE CADRE DE L'INSTITUT DES SCIENCES DU DOCUMENT NUMERIQUE RHONE-ALPES LE 25 JUIN 2001 A L'ENSSIB. *La Numérisation des collections*. Villeurbanne : E.N.S.S.I.B., 2001. 62 p.

HELLY, Bruno. Enjeux et défis de la numérisation rétrospective des documents et publications scientifiques en sciences humaines et sociales. *La lettre d'information SHS* [en ligne]. Mai 2004, n° 69, p. 19-24. Disponible sur : <<http://www.cnrs.fr/SHS/actions/lettre.php>> (consulté le 31.05.2004)

Numérisation et patrimoine

BEQUET, Gaëlle, CEDELLE, Laure. Numérisation et patrimoine documentaire. *Bulletin des Bibliothèques de France*, mars 2000, t. 45, n° 4, p. 67-72.

BURESI Charlette, CEDELLE-JOUBERT Laure (sous la dir. de). *Conduire un projet de numérisation*. Villeurbanne : E.N.S.S.I.B., Paris : Lavoisier, 2002. 326 p. (La boîte à outils).

DUPOIRIER, Gérard (sous la dir. de). *Les bibliothèques numériques*. Paris : Hermès, 1999. 299 p.

LUPOVICI, Catherine. De la bibliothèque classique à la bibliothèque numérique : continuité et rupture. *Documentaliste*, décembre 2000, volume 37, n° 5-6, p. 286-297.

Les bibliothèques numériques sur la toile

La liste suivante rassemble une sélection de sites web permettant de s'orienter dans le domaine des bibliothèques numériques.

MINISTERE DE LA CULTURE ET DE LA COMMUNICATION. Numérisation du patrimoine culturel [en ligne]. Disponible sur : <<http://www.culture.gouv.fr/culture/mrt/numerisation/index.html>>. (consulté le 3.06.2004)

MINISTERE DE LA CULTURE ET DE LA COMMUNICATION. Culture.fr, le portail de la culture / Bibliothèques numériques [en ligne]. Disponible sur : <<http://www.culture.fr/BibliothequesMediatheques/c391>> (consulté le 10.06.2004)

MINISTERE DE LA CULTURE ET DE LA COMMUNICATION. Culture.fr, le portail de la culture / Numérisation du patrimoine culturel [en ligne]. Disponible sur : <<http://www.culture.fr/EtudesEtRecherchesCulturelles/c389>>. (consulté le 7.06.2004)

MINISTERE DE L'EDUCATION NATIONALE, DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE. Educnet / Dossiers et guides / Bibliothèques numériques [en ligne]. Disponible sur : <<http://www.educnet.education.fr/dossier/rechercher/biblionum1.htm>> (consulté le 23.07.2004)

BIBLIOTHEQUE UNIVERSITE PARIS 8. Bibliothèques numériques [en ligne]. Disponible sur : <<http://www-bu.univ-paris8.fr/Ref/BibTextes.html>>. (consulté le 10.08.2004)

UNITED NATIONS EDUCATIONAL SCIENTIFIC AND CULTURAL ORGANIZATION. UNESCO-IFLA Directory of Digitized Collections [en ligne]. Disponible sur : <<http://www.unesco.org/webworld/digicol/>> (consulté le 16.07.2004)

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. Digital Libraries Resources and Projects [en ligne]. Disponible sur : <<http://www.ifla.org/II/diglib.htm>>. (consulté le 16.07.2004)

Aspects techniques

CACALY, Serge (sous la dir. de). *Gestion et diffusion d'un fonds d'images*. Paris : Nathan université, 2002. 127 p.

CARO Stéphane, BETRANCOURT Mireille. Ergonomie des documents numériques. *Techniques de l'ingénieur, traité Informatique*, février 2001, p. 1-14.

LE MOAL, Jean-Claude, HIDOINE, Bernard. *Bibliothèques numériques : cours INRIA, 9-13 octobre 2000*. Paris : ADBS Editions, 2000. 243 p. (Sciences de l'information. Série Etudes et techniques)

SCAPIN D. L., BASTIEN J.M.C. *Inspection d'interfaces et critères ergonomiques*. Rocquencourt : INRIA., 14 p. Rapport de recherche n° 2901 : INRIA : 1996.

HEUTE, Thomas, HEUTE, Damien. PHP Facile [en ligne]. Disponible sur : <http://www.toutestfacile.com/phpinit.php?tef_site=php>. (consulté le 08.07.2004)

PHP DEBUTANT. PHPdébutant.org [en ligne]. Disponible sur : <<http://www.phpdebutant.org/>>. (consulté le 05.07.2004)

MUNZ, Stefan. Selfhtml, faire soi-même les fichiers HTML [en ligne]. Disponible sur : <<http://fr.selfhtml.org/>> (consulté le 05.07.2004)