



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



Université Claude Bernard – Lyon 1

DESS en INFORMATIQUE DOCUMENTAIRE

Rapport de stage

Contribution au développement d'un serveur de thèses électroniques

Carole CLERC

Effectué sous la direction de Jean-Michel Mermet
INSA de Lyon

1999

Contribution au développement d'un serveur de thèses électroniques

Carole Clerc

Résumé :

La diffusion des thèses sur Internet devient le fer de lance de la promotion de la recherche française aux yeux de la communauté scientifique internationale. De nombreuses idées émergent un peu partout dans le monde. Il faut maintenant les rassembler pour construire ensemble un catalogue de thèses en ligne unique. Ce projet ambitieux impose de changer les méthodes de travail, des étudiants et des bibliothèques, et de réactualiser certains textes gouvernementaux.

Mots clefs :

Thèse électronique, bibliothèque numérique, SGML, XML, PDF, feuille de style, métadonnée.

Abstract :

For the international scientific community, the diffusion of the dissertations on the internet is becoming the main purpose of the french research promotion. It's now time to gather these in order to build a unique on ligne dissertation directory. This ambitious project imposes changes on the students' and the libraries' working methods, and it also needs a change in governmental texts.

Keywords :

Electronic thesis, digital library, SGML, XML, PDF, stylesheet, metadata.

Table des matières

1. INTRODUCTION	4
2. LE PROJET CITHER ACTUEL.....	5
2.1 LES ANNEES 1996-1998, LE DEMARRAGE	5
2.2 UN PROJET QUI REPOSE SUR 2 AXES	6
2.2.1 <i>Le BackOffice</i>	6
2.2.1.1 Récupération des documents	6
2.2.1.2 Le traitement	6
2.2.1.3 L'archivage.....	8
2.2.2 <i>Le FrontOffice</i>	9
2.3 LES BESOINS	10
2.3.1 <i>Les missions remplies</i>	10
2.3.2 <i>Les nouveaux besoins</i>	11
3. LES DEVELOPPEMENTS EN COURS.....	13
3.1 LES FEUILLES DE STYLES	13
3.1.1 <i>Institutions impliquées dans le groupe de travail</i>	14
3.1.2 <i>Les objectifs à atteindre</i>	14
3.1.3 <i>Les projets similaires</i>	14
3.1.3.1 UMI	15
3.1.3.2 Virginia Tech (ETD-ML : Electronic Theses and Dissertations Markup Language) :	15
3.1.3.3 L'université de Laval.....	18
3.1.3.4 Les PUM et l'université Lyon2	20
3.1.3.5 La France.....	21
3.1.4 <i>La 1^{ère} réunion du groupe de travail</i>	21
3.1.5 <i>La 2^{ème} réunion du groupe de travail</i>	22
3.1.6 <i>L'enquête auprès des directeurs de thèses</i>	23
3.2 LA PROMOTION DU PROJET, UNE COMMUNICATION TOUS AZIMUTS.....	23
3.3 L'EXTENSION DU PROJET A D'AUTRES ETABLISSEMENTS	24
4. XML (EXTENDED MARKUP LANGUAGE)	25
4.1 LES AVANTAGES D'XML	25
4.2 XML ET CITHER.....	27
4.2.1 <i>Analyse de l'existant des outils XML</i>	27
4.2.1.1 Doc'Zilla.....	27
4.2.1.2 Hybrick.....	28
4.3 PRODUIRE DU XML A PARTIR DU RTF, L'EXPERIENCE DE LYON2 – LES PUM	28
5. BILAN FINANCIER.....	30
6. BILAN ET FUTUR DES PROJETS DE THESES ELECTRONIQUES	30
6.1 ASPECT RESSOURCES HUMAINES.....	30
6.2 ASPECTS TECHNOLOGIQUES.....	31
6.2.1 <i>Les métadonnées</i>	31
6.2.2 <i>Le multimédia</i>	32
6.2.3 <i>XML</i>	32
6.2.4 <i>Les prochains développements de CITHER</i>	33
7. ET EN DEHORS DE CITHER ?	33
8. GLOSSAIRE.....	35
9. REFERENCES BIBLIOGRAPHIQUES	37
10. BIBLIOGRAPHIE.....	40

1. Introduction

Les INSA (Instituts Nationaux des Sciences Appliquées) font partie des écoles formant le plus grand nombre d'ingénieurs. De nombreuses recherches sont en cours dans leurs laboratoires. L'INSA de Lyon est implanté sur le campus de la Doua à Villeurbanne. Sur ce campus se trouvent également l'UCBL sciences (Université Claude Bernard de Lyon), les écoles d'ingénieurs ISTIL (Institut des Sciences et Technique de l'Ingénieur de Lyon), CPE (Chimie Physique Electronique), l'ENSSIB (Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques), l'ISFA (Institut de Sciences Financières et d'Assurances), l'IN2P3 (Institut National de Physique Nucléaire et de Physique des Particules), l'IUT A (Institut Universitaire de Technologie) et des unités du CNRS (Centre National de Recherche Scientifique). L'INSA possède sa propre unité documentaire scientifique et technique. Doc'INSA assure à la fois les fonctions d'une bibliothèque universitaire classique, spécialisée dans le domaine des sciences et techniques de l'ingénieur et les fonctions d'un centre de documentation proposant un large choix de services et impliqué dans plusieurs projets préparatifs les bibliothèques de demain.

Avec une telle implantation, le public potentiel de Doc'INSA est vaste. Doc'INSA est accessible notamment :

- aux élèves ingénieurs de l'INSA et aux étudiants de l'UCBL,
- aux enseignants et chercheurs de l'INSA et de tous les établissements universitaires de la région,
- aux ingénieurs des entreprises (souvent PME-PMI).

Pour satisfaire ce public exigeant, de nombreux services ont été mis en place :

- la lecture sur place,
- le prêt à domicile,
- le prêt entre bibliothèques (PEB),
- l'accès en intranet à une trentaine de bases de données scientifiques, techniques et économiques,
- la veille technologique pour les entreprises et les laboratoires,
- des formations à la recherche de l'information.

De plus, Doc'INSA est le dépositaire officiel des mémoires de thèses produits dans les laboratoires de l'INSA, soit environ 130 documents par an. Il était fort dommage de laisser tant de littérature grise accessible uniquement sous forme de microfiches. La croissance exponentielle d'Internet et l'apparition de nouvelles technologies permettent maintenant de diffuser sans problème ces travaux de recherche en leur apportant une dimension supplémentaire : le multimédia. C'est pourquoi Doc'INSA s'est lancé dans une politique de « bibliothèque numérique » avec le projet CITHER (Consultation en texte Intégral des THèses En Réseau) sous la responsabilité technique de Jean-Michel Mermet.

2. Le projet CITHER actuel

2.1 *Les années 1996-1998, le démarrage*

Il est intéressant de remarquer que tous les projets de thèses électroniques dans le monde débutent ces années là. En effet, à cette époque, le réseau Internet se démocratise. Le public s'intéresse de plus en plus à ce nouveau média. Les pages personnelles en HTML (facilement accessibles par les utilisateurs) commencent à apparaître en masse. Les grandes institutions de divers pays s'implantent alors sur le Web, et chacun essaie de promouvoir son savoir-faire. Il ne faut pas oublier que le World Wide Web a initialement été créé au CERN par des chercheurs et pour des chercheurs en Physique. Ceci explique pourquoi les divers travaux de recherche sont présents en nombre sur "la toile" dès le début. Les chercheurs s'en servent pour publier des prépublications, pour se faire connaître, faire connaître leurs travaux et leurs laboratoires à la communauté scientifique internationale. Il était donc naturel que les thèses apparaissent sur ce nouveau média sous forme de documents électroniques, permettant ainsi à leurs auteurs d'inclure des photographies haute définition, des séquences vidéo ou de son, etc. A Lyon, le projet de thèses électroniques voit le jour à l'INSA fin 1997 sous le nom de projet CITHER¹. Il débute avec le PFE (Projet de Fin d'Etude) « Chaîne d'Édition Numérique (CEN) » [HUNE-98] d'un élève ingénieur INSA en 5^{ème} année d'informatique, Marc Etienne Huneau et se poursuit en 1998 avec le stage de DESSID (Diplôme d'Etudes Supérieures Spécialisées en Informatique Documentaire) de Jean-Michel Mermet, intitulé « Coordination et mise en place d'un serveur de thèses en texte intégral à l'INSA de Lyon. Conception du Front Office » [MERM-98]. Voyons plus en détail en quoi consiste CITHER.

2.2 *Un projet qui repose sur 2 axes*

L'objectif principal de CITHER est alors d'offrir gratuitement par Internet l'accès à un maximum de thèses soutenues à l'INSA de Lyon depuis janvier 1997, de promouvoir la recherche et les laboratoires de l'Institut.

2.2.1 Le BackOffice

Le BackOffice correspond à toutes les opérations qui sont effectuées en dehors de la vue du public et des auteurs. Il comprend la récupération des documents sous forme électronique, leur traitement et leur archivage.

2.2.1.1 *Récupération des documents*

Actuellement, une présentation du projet CITHER est faite aux futurs docteurs lors de leur visite obligatoire à Doc'INSA pour la vérification des références bibliographiques. Lors de ce premier contact, une démonstration du serveur leur est faite. Les objectifs du projet et les bénéfices que peut en retirer le doctorant sont présentés. La procédure de remise de la thèse sous forme électronique est détaillée [CONS-98]. Une demande d'autorisation de diffusion à retourner est également fournie. Cette autorisation est un véritable contrat rédigé par maître Alain Marter², qui engage la responsabilité de l'INSA auprès de l'auteur de la thèse [CONT-99]. Enfin, dans la période s'étalant de la soutenance à 3 mois après la soutenance, le doctorant est invité à fournir à Doc'INSA une version électronique de sa thèse, corrigée si le jury l'a demandé.

2.2.1.2 *Le traitement*

La conversion de la thèse en document publiable sur le réseau peut commencer quand l'opérateur dispose :

- du contrat de diffusion cosigné par l'auteur et par l'INSA,
- d'un exemplaire papier de la thèse,

¹ <URL : <http://csidoc.insa-lyon.fr/these/index.html>>

- de la forme électronique de la thèse,
- du formulaire d'enregistrement de la thèse.

et qu'il a vérifié que la thèse est cataloguée à Doc'INSA.

Les choix des formats sources et de publication ont été décidés suite à des études menées par Marc-Etienne Huneau et Jean-Michel Mermet au cours de leur stage. Le résultat de cette enquête a permis de mettre en évidence que deux logiciels étaient principalement utilisés à l'INSA de Lyon pour la rédaction des mémoires de thèses : LaTeX ou TeX (7% des mémoires rendus) et Word (92% des mémoires rendus). Les fichiers arrivant à Doc'INSA sont donc des fichiers Word ou des fichiers PostScript (format d'impression des documents LaTeX ou TeX). Le fichier issu de la chaîne de traitement informatique, quant à lui, est un fichier PDF enrichi. C'est à dire qu'il contient des liens permettant d'obtenir toute l'interactivité que les documents électroniques peuvent offrir :

- une table des matières active offrant l'accès à n'importe quelle partie du document en permanence,
- des renvois actifs aux notes de bas de page, à la bibliographie etc,
- des liens hypertextes actifs et valides.

La conversion commence alors. Elle consiste en l'exécution du programme CEN. L'opérateur dispose d'un guide de conversion pour le guider, étape par étape. En voici les grandes lignes :

- Une macrocommande Word transforme le fichier source en fichier PostScript enrichi de liens à destination d'Acrobat Distiller.
- Les fichiers PostScript sont ensuite modifiés par le CEN qui ajoute des liens pointant sur les autres fichiers de sorte qu'une thèse peut être parcourue, sans se soucier de son éventuel découpage en plusieurs fichiers PDF [HUNE2-98].
- Les fichiers PostScript sont convertis par Acrobat Distiller en PDF.
- Les fichiers PDF sont retraités par Acrobat Exchange pour renseigner les champs titre, sujet, auteur... et les vues miniatures de pages sont créées.
- Une page HTML nommée « pont d'embarquement » est créée. Elle pointe vers les différents fichiers PDF de la thèse, et contient la référence bibliographique et les métadonnées. C'est sur cette page que pointe l'OPAC (Open Public Access Catalog) lors d'une consultation de thèses.

² Alain Marter est avocat, spécialisé en droit économique, propriété intellectuelle et droit pénal.

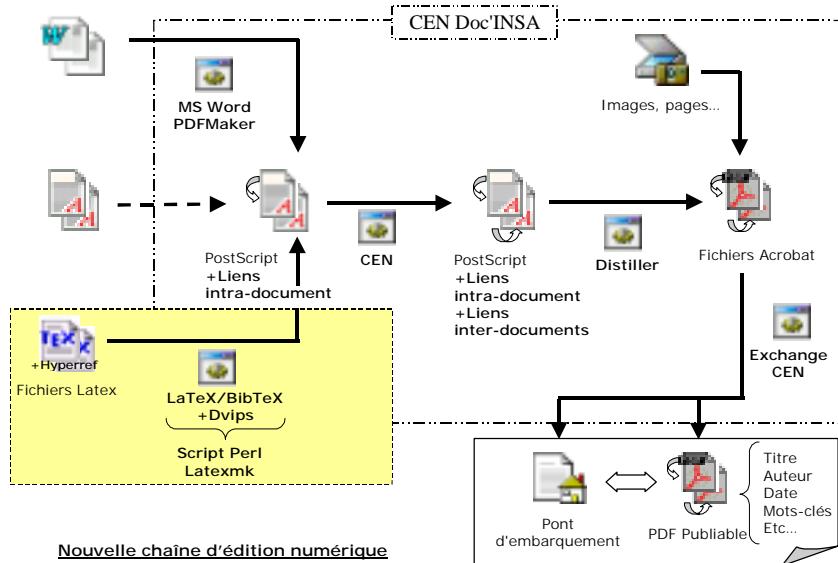


Fig1 : Chaîne d'Edition Numérique utilisée à Doc'INSA

La documentation du logiciel CEN est disponible en ligne :

- La « Chaîne d'édition numérique – Référence utilisateur » à
<URL : http://csidoc.insa-lyon.fr/these/doc/cen_ref_utilisateur.pdf>
- Le « Manuel utilisateur » à
<URL : http://csidoc.insa-lyon.fr/these/doc/manuel_utilisateur.pdf>

2.2.1.3 L'archivage

Une sauvegarde quotidienne des données présentes sur le serveur est effectuée sur bande DAT. Les fichiers PDF sont protégés contre la modification par mot de passe. Une attention particulière est portée à la pérennité des URL des ponts d'embarquement pointant sur les thèses. Enfin les fichiers sources originaux de la thèse ainsi que tous les fichiers pertinents créés lors de la conversion sont conservés sur CD-ROM. Une dizaine de thèses est gravée par CD-ROM et chaque CD-ROM est produit en 3 exemplaires gardés sous scellés en des lieux différents, de manière à se protéger au maximum contre une destruction des données. La durée de vie d'un CD-ROM est d'une quinzaine d'années environ. Doc'INSA prévoit donc un

nouvel enregistrement des données dans 10 ans, sur le support qui sera le plus adapté à cette époque.

2.2.2 Le FrontOffice

Le FrontOffice est la partie visible du projet, pour le lecteur et pour l'auteur. C'est le site CITHER localisé à <URL : <http://csidoc.insa-lyon.fr/these>>.

La création de ce site a fait l'objet d'une étude particulière par Jean-Michel Mermet quant à sa navigation, sa conception graphique et la conception structurelle de l'information. La métaphore marine a été retenue, les couleurs choisies sont le bleu, le vert et le gris, référence est faite à l'île de Cythère, on parle de pont d'embarquement... La navigation hiérarchique (un chemin constamment affiché, indique la place de la page dans l'arborescence) a été retenue pour sa souplesse de mise à jour et d'utilisation.

L'accès aux thèses se fait de 3 manières possibles :

- par listes alphabétiques,
- par catalogue informatisé (Doris Web),
- par recherche en texte intégral.

Les documents PDF des thèses sont en effet indexés intégralement par Altavista Search Intranet 97, installé sur un serveur du CISM (Centre d'Informations Scientifiques et Médicales). Une page Web personnalisée, sur le site CITHER, permet alors de faire des recherches en texte intégral ou en utilisant des critères restrictifs.

Signalons pour finir que les thèses de CITHER sont diffusées dans le cadre du projet Webdoc [CARO-99] mis en place par l'ABES (Agence Bibliographique de l'Enseignement Supérieur) à la demande du MENRT (Ministère de l'Education Nationale, de la Recherche et de la Technologie).

Pour en savoir plus sur CITHER, n'hésitez pas à consulter la documentation en accès libre sur le site de CITHER, soit en HTML, soit en PDF à :

<URL : <http://csidoc.insa-lyon.fr/these/doc/index.html>>.

2.3 *Les besoins*

CITHER a été créé pour répondre à certaines attentes, certains besoins. Comme nous allons le voir, il permet déjà tel qu'il est maintenant de combler en partie les carences dont souffrait la diffusion du savoir et de la recherche française. Néanmoins, nous allons également constater qu'il a créé de nouveaux besoins auxquels il va maintenant falloir répondre.

2.3.1 Les missions remplies

La France souffre d'un manque de valorisation des résultats de sa recherche par rapport à ses concurrents que sont les autres grandes nations industrialisées [MERM-98]. De plus, les entreprises manquent de confiance dans les laboratoires publics et ne s'intéressent donc pas toujours à leurs travaux. Le résultat en est que de nombreuses découvertes et innovations technologiques « dorment » sans garantie d'évolution sur les rayons des bibliothèques. Une des missions premières de CITHER était donc d'accroître la visibilité de l'information scientifique produite par les organismes publics en effectuant une publication large et ouverte, et en espérant ainsi contribuer à ce que la recherche française soit mieux identifiée, plus souvent citée et plus souvent utilisée. Par la même occasion, les docteurs qui déposent leurs travaux sur le site de CITHER se font ainsi connaître de leurs collègues étrangers et participent à la promotion de leur laboratoire. Les statistiques de consultation des pages de CITHER sont très encourageantes et vont dans le sens que nous espérions. La mission est en partie remplie, mais il est évident qu'une politique de communication autour de ce projet est nécessaire auprès des doctorants, des directeurs de thèse pour obtenir la matière première et auprès du public pour obtenir une consultation encore meilleure, cet aspect est un des grands axes de mon stage.

Pour diffuser une thèse, il faut néanmoins faire attention à la propriété intellectuelle. D'un point de vue juridique, la législation ne reconnaît qu'un seul auteur au mémoire de thèse : le docteur. Il fallait donc prévoir un véritable contrat reliant l'auteur et le centre de documentation. Ici encore la mission a été remplie puisque Maître Alain Marter a réalisé pour

l'INSA, une autorisation de diffusion. Ce contrat de diffusion est en accès libre³ sur le site Web de Doc'INSA.

Le contrat de publication engage la responsabilité de l'INSA en cas de modifications ou altérations du document initial. Il était donc impératif de trouver des outils respectant le travail sur le fond et sur la forme et de procéder à un archivage fiable. Le choix du format PDF respectant tous les critères souhaités et le choix de l'archivage sur CD-ROM sont expliqués dans les rapports de M-E Huneau et J-M Mermet. Donc là encore la mission a été remplie.

Toutes ces précautions étant prises, on aurait pu penser que la montée en charge du serveur pouvait maintenant s'effectuer sans problème. Mais la réalité est tout autre car CITHER a créé de nouveaux besoins qu'il faut maintenant penser à satisfaire.

2.3.2 Les nouveaux besoins

L'utilisation du programme CEN s'est montré efficace à condition que le document initial soit correctement structuré. Le programme de conversion se servant des styles de Word pour repérer les titres, sous-titres ... il faut impérativement que la version électronique de la thèse soit correctement « stylée ». Or beaucoup d'étudiants ne maîtrisent pas l'emploi d'un traitement de texte car ils n'ont jamais suivi de cours de bureautique. En conséquence, de nombreuses thèses arrivent avec une construction manuelle fragile, sans utilisation de styles, de niveaux de titres ou de pagination automatique. Dans ces cas, l'opérateur du projet doit reprendre l'ensemble de la thèse et lui appliquer des feuilles de style avant d'entamer la conversion du document, ce qui entraîne une perte de temps considérable pour le centre de documentation. En effet, le temps passé à remettre en forme une thèse peut alors atteindre deux jours contre une demi-journée habituellement.

Cette nécessité d'obtenir des documents sources structurés démontre le besoin de former les étudiants à l'utilisation des feuilles de styles. Et la mise en place d'une formation complémentaire, fut-elle nécessaire, n'est pas facile à aménager. Il faut en effet trouver le matériel, les locaux, les formateurs et évidemment les fonds.

Paradoxalement, les étudiants utilisant LaTeX, langage très structuré, fournissent des fichiers qu'il est délicat de traiter de manière immédiate. Nous n'avions pas de programme nous permettant d'intégrer directement des fichiers issus de LaTeX dans le programme CEN.

³ <URL : <http://www.csidoc.insa-lyon.fr/these/doc/contrat.html>>

De plus, le projet devant s'étendre à Lyon1, le nombre de thèses en LaTeX devrait considérablement augmenter. Un élève ingénieur INSA en cinquième année d'informatique, Julien Tognazzi, également en stage à Doc'INSA a résolu le problème en créant une chaîne de conversion à partir de LaTeX. Le fichier de sortie de cette chaîne possède un format PostScript compatible avec les formats de fichiers sources acceptés par le programme CEN. Cette chaîne repose sur les 4 logiciels LaTeX, BibTeX, Dvips et Acrobat Distiller [TOGN-99].

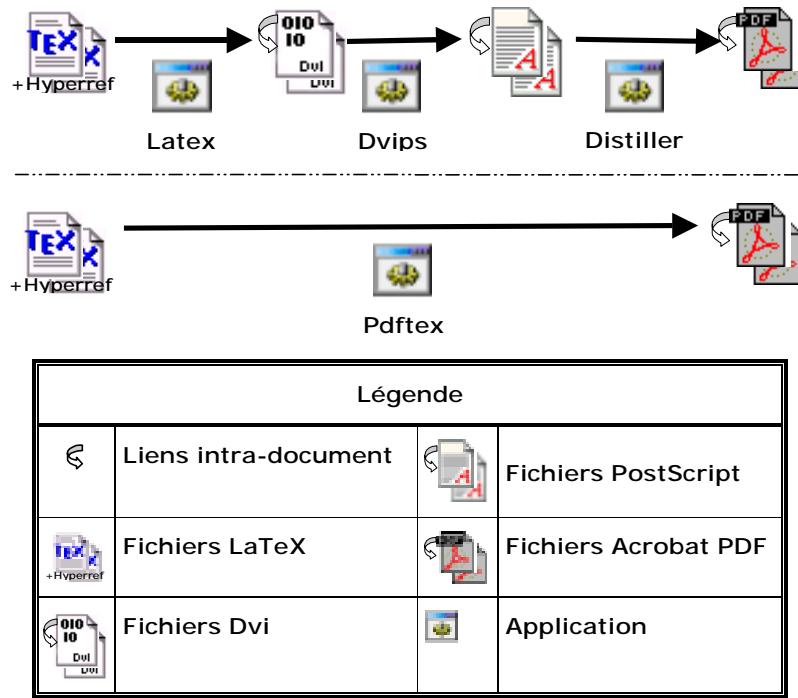


Fig. 2 : Les 2 chaînes de traitement possibles pour les documents LaTeX. Doc'INSA utilise la première chaîne.

Un besoin qui paraît annexe et qui est pourtant fondamental est celui de posséder des métadonnées prédéfinies par un système servant de référence. On peut en effet imaginer que, dans un futur plus ou moins proche, les différents établissements possédant des projets de thèses numériques regroupent leurs notices bibliographiques accompagnées des métadonnées permettant de les retrouver sur « la toile », sur un seul et même serveur. Ce serveur renverra ensuite vers chaque établissement pour la consultation de la thèse. Il apparaît clairement que pour aboutir à un tel serveur commun, il est obligatoire que tous les établissements utilisent le même système de métadonnées. Les recommandations du Dublin Core [DUBL-99] semblent

pour l'instant souvent suivies dans les projets de thèses électroniques. Pour plus d'information, il est possible de consulter la Request For Comment (RFC) à :

<URL : <ftp://ftp.isi.edu/in-notes/rfc2413.txt>>.

Un autre besoin apparaît comme latent : la promotion du site de diffusion. Créer un serveur de thèses ne suffit pas, il faut que ce dernier soit consulté. C'est une tâche de longue haleine et à laquelle il faut constamment penser car les étudiants qui fournissent la matière première changent chaque année et il faut donc les informer tous les ans. Quant au public, il faut constamment tenter de l'élargir en essayant en premier lieu de cibler les entreprises et les laboratoires de recherche.

Enfin, il faut penser à trouver un autre format d'archivage. Pour l'instant, les thèses sont en effet conservées sous forme de fichiers Word et PDF. Or, ces fichiers sont propriétaires et on ne connaît pas la politique que mèneront dans dix ans des sociétés comme Adobe ou Microsoft. Le format idéal envisagé est XML, mais il n'est malheureusement pas encore arrivé à maturité.

Voyons maintenant comment nous avons tenté de répondre à ces besoins.

3. Les développements en cours

Le plus important, pour permettre une bonne montée en charge du serveur et pouvoir espérer traiter un maximum des 130 thèses produites annuellement à l'INSA, était de s'occuper du problème de la structuration des documents et du besoin de formation des étudiants.

3.1 Les feuilles de styles

Les projets de thèses électroniques intéressent beaucoup d'institutions et celles qui ont commencé à mettre en place un prototype se heurtent au même problème commun : la nécessité d'obtenir des documents sources correctement structurés. C'est pourquoi nous avons créé un groupe de travail regroupant diverses institutions de divers horizons en prenant en compte les spécificités d'un maximum d'établissements.

3.1.1 Institutions impliquées dans le groupe de travail

D'une part, sont membres du groupe de travail les établissements possédant un projet de thèses électroniques, quel qu'en soit le degré d'avancement. Il s'agit des universités de Marne la Vallée, Lyon2, Grenoble, et l'INSA de Lyon.

D'autre part, sont présents tous les établissements intéressés par ce genre de projet. Il s'agit de toutes les universités de la Conférence Universitaire Rhône-Alpes, de l'ENSSIB, et de l'URFIST (Unité Régionale de Formation en Information Scientifique et Technique) de Nice [SANO-99].

Enfin, des membres du MENRT et plus particulièrement de la SDB (Sous Direction des Bibliothèques) et de l'ANRT (Atelier National de reproduction des thèses) de Lille nous honorent de leur présence.

3.1.2 Les objectifs à atteindre

Le groupe de travail s'est fixé deux objectifs principaux :

- Etablir des gabarits de documents :
 - en s'inspirant des travaux déjà publiés par d'autres bibliothèques universitaires,
 - attrayants grâce à l'expertise d'un typographe.
- Inciter les étudiants à les utiliser en leur montrant l'intérêt réciproque : rédiger à moindre effort des documents à la mise en page professionnelle grâce à l'aide d'un typographe.

Ces deux objectifs qui paraissent pourtant précis nous ont mené dans des réflexions très poussées et étonnantes, comme nous le verrons par la suite.

3.1.3 Les projets similaires

Il est bien évidemment impossible de recenser tous les projets de thèses électroniques qui existent de par le monde. Voici les quatre qu'il me semble le plus intéressant d'exposer car ils représentent quatre tendances différentes et sont tous les quatre déjà bien avancés. Il s'agit de :

- UMI (University Microfilms International),

- Virginia Tech et ses émules allemands et australiens,
- L'université de Laval,
- Les PUM (Presse Universitaire de Montréal) associées à l'université Lyon2.

3.1.3.1 *UMI*

La politique de UMI est avant tout commerciale. Les thèses ne sont pas en libre accès. Tous les services proposés sont payants que ce soit le dépôt ou la consultation d'une thèse. UMI diffuse les thèses principalement aux formats RTF de Microsoft et PDF d'Adobe. C'est un projet qui a débuté en 1997 et qui compte environ 1400000 références de thèses et plus de 100000 consultables en texte intégral [UMID-99]

3.1.3.2 *Virginia Tech (ETD-ML : Electronic Theses and Dissertations Markup Language)* :

Depuis janvier 1997, tous les travaux de maîtrise et de doctorat doivent être déposés à Virginia Tech sous forme électronique. Cette université a même créé son propre langage de balisage, le ETD-ML dérivé de SGML et possédant une DTD spécialement adaptée aux thèses. Virginia Tech possède pour le moment 1634 thèses en ligne en texte intégral dont 961 sont accessibles en PDF. Ce projet présente de nombreux avantages :

- Pratiquement tous les cas de figure de formats de fichiers sources sont traités : Word, LaTeX et ETD-ML.

- Plusieurs procédés de conversion sont offerts :
 - Word \Rightarrow PDF,

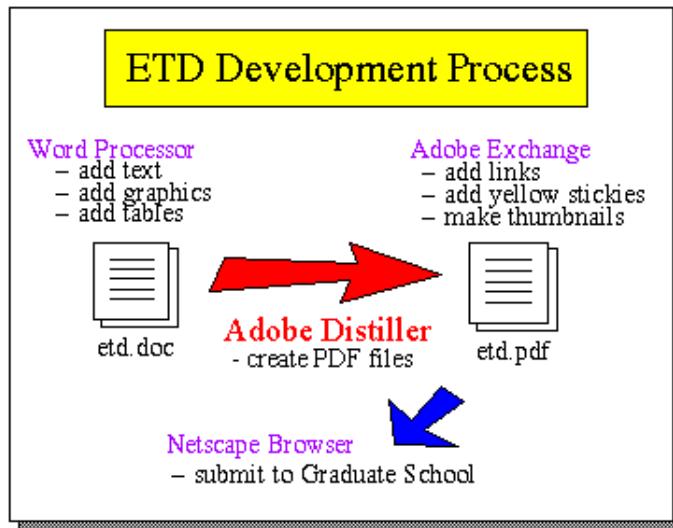


Fig.3 : Chaîne de traitement utilisé à Virginia Tech pour convertir les documents Word en PDF. Cette image est issue du serveur de Virginia Tech⁴

- TeX \Rightarrow DVI \Rightarrow PS \Rightarrow PDF (Un fichier TeX n'est pas exploitable immédiatement. Il faut d'abord le transformer en DVI, format indépendant du périphérique de sortie, (opération réalisée par le compilateur TeX) puis en PostScript, format d'impression et éventuellement en PDF (opérations réalisées par des logiciels extérieurs) pour le rendre accessible à un large public).

⁴ <URL : <http://etd.vt.edu/howto/slides/index.htm>>

- ETD-ML \Rightarrow HTML

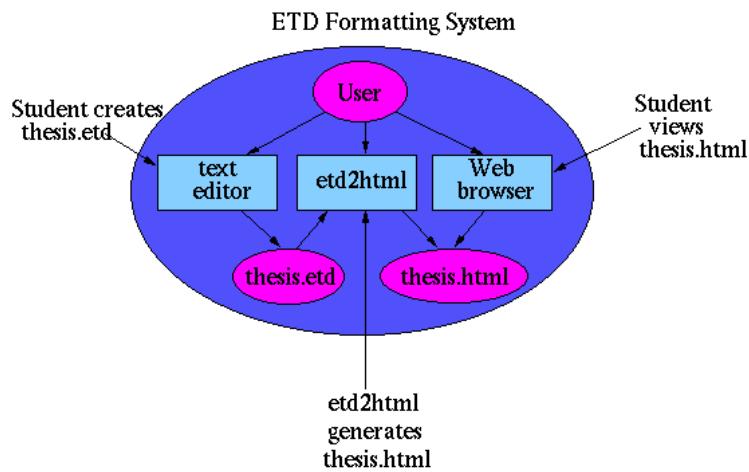


Fig. 4 : Chaîne de traitement utilisée à Virginia Tech pour la conversion des thèses rendues en ETD-ML. Cette figure est issue du serveur de Virginia Tech⁵.

- Pour la bibliothèque, l'économie de temps est conséquente, car c'est l'étudiant qui crée les fichiers publiables.

Mais les inconvénients sont également très importants :

- Il n'y a pas de cours ou de formation à destination des étudiants pour les aider à créer une version électronique de leurs travaux, mais juste quelques documents d'aide en ligne.
- Par conséquent, c'est un projet difficile à prendre en main quand on a des connaissances moyennes en informatique.
- Les présentations des documents d'aide et de la marche à suivre sont rébarbatives et peu conviviales.
- Le travail supplémentaire que doit fournir l'étudiant est conséquent.

Ce projet a été repris dans une version allemande à l'université de Berlin [PROJ-99]. Le Networked Digital Library of Theses and Dissertations (NDLTD) [NETW-99] est une extension du projet de Virginia Tech. Il a débuté en 1996 et a pour but d'établir un réseau de thèses électroniques entre plusieurs universités.

Pour plus d'informations vous pouvez consulter le site à <URL : <http://etd.vt.edu/>>

⁵ <URL : <http://etd.vt.edu/etd-ml/userguid.htm>>

3.1.3.3 L'université de Laval

Ce projet est plus récent (1998) et plus pédagogique. Voici ses avantages :

- C'est un projet avancé et pratique pour les étudiants.
- Il possède un guide clair et pédagogique.
- Les feuilles de style que peuvent utiliser les étudiants sont fournies pour Word et WordPerfect (prévues), et personnalisent l'environnement du traitement de texte pour plus de facilité d'emploi. Il suffit ainsi de cliquer sur le bouton « titre thèse », pour que le titre reçoive le style qui lui a été destiné, etc.
- Le processus de conversion proposé est Word (ou WordPerfect) \Rightarrow RTF \Rightarrow SGML

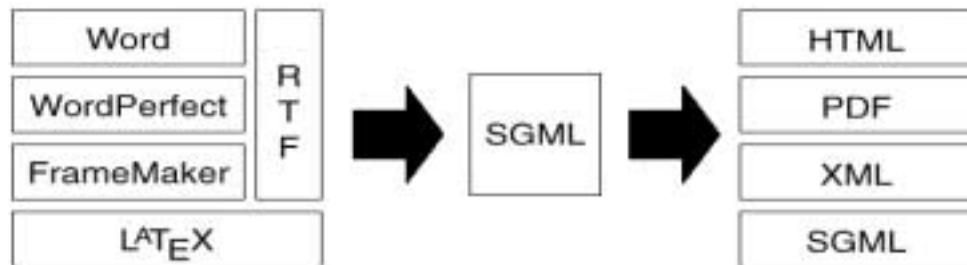


Fig.5 : Chaîne de traitement utilisée à l'université de Laval. Cette figure est issue du serveur de cette université⁶.

⁶ <URL : <http://www.bibl.ulaval.ca/doelec/theses/pr-pilote.pdf>>

- Enfin des cours sont prévus sur l'utilisation des feuilles de style

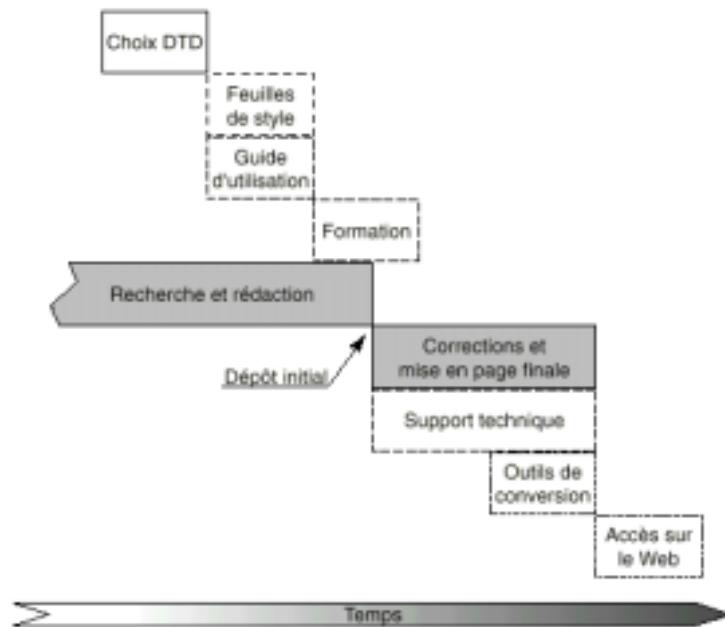


Fig.6 : Parcours de la thèse en fonction du temps. Ce dessin est issu du serveur de cette université⁷.

Néanmoins, ce projet a quand même quelques inconvénients :

- Les feuilles de style sont ternes.
- Il y a peu de diversité dans les feuilles de style proposées (un seul type) et le nombre de styles semble trop important (plus de 100).
- Le guide de l'étudiant est trop long pour espérer qu'il soit lu en entier.
- Les utilisateurs de LaTeX n'ont, pour l'instant, pas de solution adaptée à leur cas, mais un chercheur de l'université de Laval est en train de concevoir le convertisseur approprié.

Pour plus d'informations vous pouvez aller à :

<URL : <http://tracerbullet.bibl.ulaval.ca/doelec/theses/>>

⁷ <URL : <http://www.bibl.ulaval.ca/doelec/theses/pr-pilote.pdf>>

3.1.3.4 Les PUM et l'université Lyon2

Les PUM de Montréal et l'Université Lyon 2 travaillent ensemble sur un projet de thèses électroniques. Les feuilles de styles n'ont été disponibles que récemment mais ce projet est intéressant et nous semble prometteur.

En voici les avantages :

- Les feuilles de style sont fournies pour Word et WordPerfect (prévues), et personnalisent l'environnement.
- Un guide de présentation va prochainement être créé.
- Le catalogue présente déjà des thèses en SGML et en XML.
- Une formation est envisagée pour les étudiants de dernier cycle.

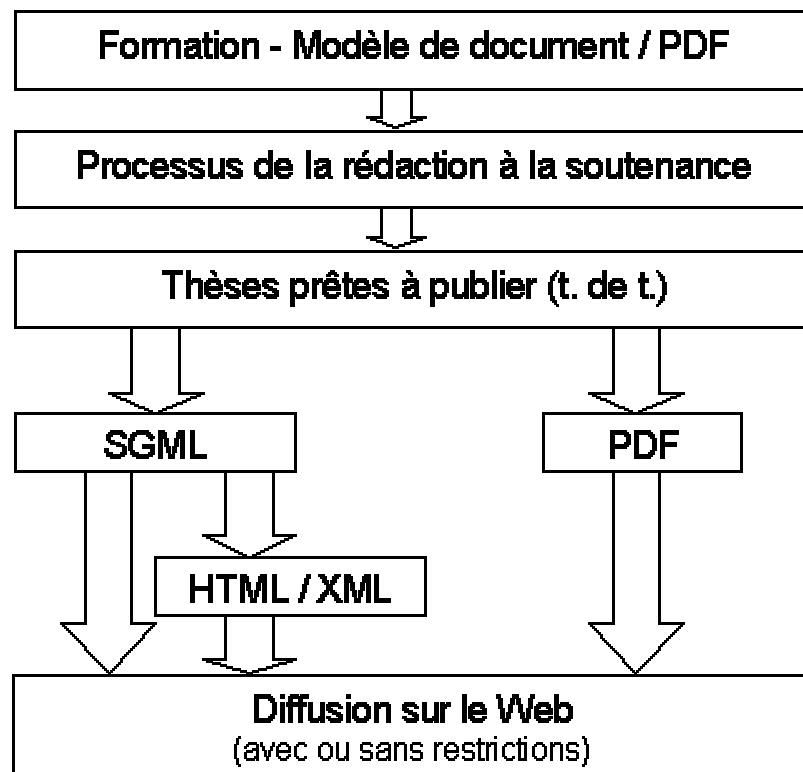


Fig. 7 : Chaîne de traitement utilisée pour la conversion des thèses aux PUM. Cette image est issue du serveur des PUM⁸.

⁸ <URL : http://www.pum.umontreal.ca/publ_electr/rapports/bea02/partie_5.html>

Et les inconvénients :

- Les feuilles de style sont ternes.
- Il y a peu de diversité dans les feuilles de style proposées (un seul type) et le nombre de styles semble là encore trop important (environ 80).
- Il n'existe pas encore de guide pour l'étudiant.
- Les utilisateurs de LaTeX n'ont, pour l'instant, pas de solution adaptée à leur cas. Mais le rapprochement des universités de Laval et Montréal pourrait fournir une solution.

Pour plus d'informations vous pouvez aller à :

<URL : <http://www.pum.umontreal.ca/theses/index.html>>

3.1.3.5 La France

On trouve en France beaucoup de manuels ou guides sur les règles d'écriture à respecter pour les thèses. Mais ces manuels sont trop longs, y compris celui du ministère (consultable à <URL : <http://www.abes.fr/teleactu.htm>>), et par trop de libertés laissées au rédacteur, manquent parfois de précision sur certains points.

Cette étude de l'existant a permis de prendre d'importantes décisions dès la première réunion du groupe de travail et d'y convier un typographe. Car il semblait d'ores et déjà établi que la réussite du projet reposait sur l'utilisation de feuilles de styles qu'il fallait attractives pour que les doctorants acceptent de les utiliser.

Mon rôle a été notamment d'animer ces réunions, de présenter à chaque séance l'avancée de mes recherches sur les feuilles de styles en fonction des remarques formulées à la réunion précédente et de fournir systématiquement un compte rendu détaillé.

3.1.4 La 1^{ère} réunion du groupe de travail

Lors de cette 1^{ère} réunion un bref exposé de l'état de l'art en ce qui concerne les feuilles de styles et les guides de rédaction a été réalisé. Puis le groupe de travail a adopté un cahier des charges préconisant la conception d'un guide de rédaction de thèses, prenant en compte non seulement des considérations techniques de présentation et d'utilisation des

feuilles de style, mais également les dispositions réglementaires en cours. Nous nous sommes ensuite attachés à décrire plus précisément la forme que doit prendre le document thèse (plan fourni en annexe). Enfin, le typographe a été chargé de proposer trois gabarits Word, de styles typographiques différents, de manière à satisfaire un maximum de personnes.

Pour plus de renseignements, vous pouvez consulter le compte rendu de la réunion fourni en annexe.

3.1.5 La 2^{ème} réunion du groupe de travail

Au cours de cette deuxième réunion du groupe de travail, le typographe Damien Gautier a présenté les 3 modèles de feuilles de styles de sa création. Certains de ces modèles étant anti-conformistes, mais fort intéressants, le typographe a motivé ses choix au cours d'un bref exposé (Le dossier et les feuilles de styles sont disponibles en annexe). La conclusion en était fort étonnante puisque qu'il nous a suggéré de nous faire créer une police pour le projet. En effet le principal défaut d'une thèse est que sa lisibilité n'est pas aisée. Pour les correcteurs, avoir à lire un document aéré, à la lisibilité et au rythme agréable est moins contrariant. L'idée d'avoir une police spécifique est donc bien pertinente. Néanmoins, il était impossible pour les membres du groupe de travail de se prononcer immédiatement sur la question. Pour juger de l'utilité et de l'apport bénéfique d'une nouvelle police, nous avons décidé de faire des tests. Une enquête auprès des directeurs de thèses va donc avoir lieu très prochainement dans le but d'avoir leurs appréciations sur ces feuilles de styles. Cette enquête est présentée dans la suite du document.

Le problème des métadonnées a ensuite été abordé. L'objectif final du projet de thèses électroniques est d'aboutir à une base de données commune. Il est donc impératif d'avoir un catalogue de métadonnées fédérateur. Lyon 2, les PUM, les universités de Laval et de Virginia Tech travaillent déjà en collaboration dans ce but. Pour l'instant les spécifications du Dublin-Core semblent faire l'unanimité. L'INSA a également utilisé les spécifications Dublin-Core pour toutes les thèses présentées en ligne.

Enfin, le plan détaillé du guide de rédaction proposé par l'INSA a été présenté (document fourni en annexe). Plusieurs remarques ont été formulées suite à la lecture de ce plan :

- La formation qui sera donnée aux étudiants sur la rédaction des documents structurés devrait commencer dès la première année de thèse avec un rappel en troisième année.
- Le multimédia et ses multiples formats (surtout en ce qui concerne les images et le son) pose un problème de relecture à long terme. Il y a beaucoup trop de "plugs ins" et de logiciels spécifiques pour assurer une certaine pérennité à l'information.
- Les mots clefs en anglais posent le problème de l'unification des métadonnées.

Le compte rendu complet de cette deuxième réunion est fourni en annexe.

3.1.6 L'enquête auprès des directeurs de thèses

Pour connaître l'accueil qui sera réservé à nos futures feuilles de styles, nous avons donc décidé de faire une enquête auprès des directeurs de thèses. Cette tâche m'est revenue. J'ai pris un texte scientifique et je lui ai appliqué les 3 modèles créés par le typographe. Ceci n'a pas été aisement en raison d'une incompatibilité de police entre Word pour PC et Word pour Macintosh. De la même manière j'ai pris un texte littéraire et lui ai appliqué les 3 modèles. Ensuite, j'ai repris les 6 documents ainsi créés et sans changer la mise en page, leur ai appliquée les polices systèmes libres de droit communes à PC et Mac : le Times et l'Arial. Ainsi, par un questionnaire adapté en cours de réalisation, nous pourrons tester à la fois l'impact de la police et des mises en page auprès des scientifiques et des littéraires. Ce test devrait être effectué auprès des enseignants de Lyon1, de l'ENSSIB et de l'INSA.

Cette enquête s'inscrit également dans le processus de communication autour du projet CITHER qui constitue le deuxième point très important de mon stage.

3.2 La promotion du projet, une communication tous azimuts

La promotion du projet peut-être envisagée dans plusieurs directions. De plus il ne faut pas oublier que si pour l'instant le projet ne concerne que l'INSA, il devrait très prochainement être étendu à Lyon1. L'information passera donc par :

- La formation, au cours de laquelle les intérêts communs de la bibliothèque et des étudiants seront présentés (avantages en terme de publicité et gain de temps qui peut être reporté sur

l'écriture de la thèse). Le doctorant devra sortir de la formation en étant convaincu que l'utilisation des feuilles de styles proposées ne lui apporte que des avantages d'une part immédiats en lui facilitant la conception de son mémoire de thèse, et d'autre part, à long terme en lui apportant les connaissances nécessaires à la rédaction de documents structurés. Ces connaissances lui permettront de mieux aborder la publication d'articles scientifiques.

- Les dossiers d'inscription où deux documents vont être ajoutés. Un présentant la bibliothèque et qui pourra être mis dans tous les dossiers et un présentant CITHER et qui ne sera mis que dans les dossiers d'inscription des doctorants (Ces deux pages sont fournies en annexe).
- Des pages Web destinées aux doctorants sur le site de CITHER. Ces pages contiendront notamment les feuilles de styles à télécharger, la liste des professeurs et de leurs grades, désormais tenue à jour par Doc'INSA, les dates des séances de formation ... Ces pages seront réalisées à la fin du stage si le temps le permet.
- La présentation du projet aux associations de doctorants (association Bernard Gregory, ADSL, GRADIENT...).
- La réussite d'une campagne d'information auprès des directeurs de thèses potentiels. Tout enseignant pouvant être directeur de thèse, c'est un public difficile à cerner et à atteindre.
- La présentation du projet auprès des groupes de discussion et journaux.

Ces trois derniers points ne seront pas traités pendant le stage par manque de temps, mais ils pourront faire l'objet d'un autre stage de DESSID. En effet un projet comme CITHER connaît des développements constants et même si au départ les objectifs sont clairement fixés, des besoins et de nouvelles urgences apparaissent au gré d'événements extérieurs. C'est en répondant à un besoin qu'on s'aperçoit en avoir créé de nouveaux et ainsi de suite... Il est difficile de maîtriser tous les tenants et les aboutissants d'un projet en cours de mise en place comme CITHER, mais c'est une expérience captivante.

3.3 L'extension du projet à d'autres établissements

Un projet de thèses électronique ne doit pas se limiter à un seul établissement, vu ses objectifs. La proximité physique de l'INSA et de l'UCBL a favorisé une collaboration entre les deux établissements et CITHER devrait bientôt être étendu à Lyon1. Quelques obstacles

restent encore à franchir et pour faciliter la communication, un annuaire du personnel de Lyon1 potentiellement impliqué dans le projet, fut réalisé.

Enfin, l'ensemble des membres du groupe de travail impliqué à divers degrés dans des projets similaires ou simples observateurs sont des partenaires potentiels.

Cet objectif de communication autour du projet CITHER nous a tout naturellement poussés à nous intéresser au futur standard du Web : Le XML. Celui-ci possède une grande interopérabilité et des avantages évidents pour un projet de thèses électroniques comme nous allons maintenant le constater.

4. XML (eXtended Markup Language)

4.1 Les avantages d'XML

XML est un langage informatique de description de structures de document recommandé par l'organisme de normalisation W3C (World Wide Web Consortium) [XML-98]. Il dérive de SGML (Structured Generalized Markup Language) auquel il apporte une extension avec l'intégration d'Hytime. Il conserve le vocabulaire standard de SGML et sa méthode de description de structure, tout en étant plus simple [MART-99]. XML est un langage puissant qui peut servir pour l'archivage sur ordinateur ou la recherche de n'importe quel type de document (technique, illustré, ...). Deux programmes qui veulent coopérer peuvent aujourd'hui le faire en échangeant des messages en XML.

Un document XML a deux parties : le document proprement dit, obligatoirement présent, et une description formelle optionnelle, appelée Document Type Definition (DTD) [MICH-99]. Cette description contient les règles syntaxiques que doit respecter le document. Tout se passe comme pour un programme que l'on écrit dans un langage de programmation : le langage a une grammaire, qui décrit les structures syntaxiques permises et les mots réservés [MART-99].

HTML ne se préoccupe pas de la hiérarchisation d'un document. Il permet d'afficher le contenu avec une mise en page sommaire. XML, lui, définit le fond indépendamment de la forme. Les balises d'HTML sont figées, alors qu'avec XML, au contraire, on peut définir n'importe quelle structure dans la mesure où elle est arborescente, avec notamment les balises que l'on veut. Ceci permet de définir des DTD pour des professions [MART2-99] ou des

domaines particuliers tels que l'automobile, la chimie, les banques, les mathématiques...et la documentation avec la Book DTD - ISO 12083 qui permet de baliser une monographie.

XML peut être utilisé pour créer des messages auto-descripteurs destinés aux échanges entre des applications différentes. Le codage de l'information est effectué en UNICODE-ISO 10646 (ASCII 16 bits) assurant ainsi une portabilité idéale et une relecture aisée. De tels messages sont indépendants des systèmes d'exploitation, des langages de programmation et des formats d'affichage. Des applications qui utilisent ces messages comprennent la structure et la sémantique des données qu'elles reçoivent.

- **XML est donc un standard d'archivage et d'échange de documents.**
- **XML permet des recherches de données puissantes et performantes.**
- **XML est une solution d'interopérabilité intéressante.**
- **XML est un format non propriétaire déjà utilisé par l'industrie.**

XSL (eXtended Style Language) [XSL-99] est le langage des feuilles de style. Une description XSL, écrite en XML, s'applique à un document comme le fait sa DTD, mais pour définir l'affichage des divers éléments. XSL est le complément « feuille de style » de XML comme CSS (Cascading Style Sheets) est le complément d'affichage de HTML.

Le standard DOM (Document Object Model) définit les méthodes d'accès à des documents XML par programme. Les opérations de conversion sur un document XML peuvent être effectuées sous le contrôle d'un programme Java ou Visual Basic, ce qui confère aux documents un caractère dynamique [MART-99].

En plus de XSL, un autre futur standard est en cours d'étude : XQL (eXtended Query Language). XQL permet la recherche dans une arborescence XML.

Autre standard en cours d'étude, XLL [XLL-99] définit des hyperliens puissants pour les documents XML. Les hyperliens définis par XLL sont :

- les liens bidirectionnels,
- les liens à destinations multiples,
- les liens typés (à sémantique précise),
- les liens d'inclusions ou de remplacement de contenu à partir d'un autre document,
- les bases de données de liens.

Par contre XML n'offre aucun verrouillage permettant d'assurer l'intégrité des données. Il n'est actuellement pas possible d'en interdire la modification ou le copié-collé. Mais malgré cet inconvénient, les avantages d'un tel langage et l'intérêt pour un projet comme CITHER sont évidents.

4.2 XML et CITHER

4.2.1 Analyse de l'existant des outils XML

Il n'existe pas encore à l'heure actuelle, de logiciels sachant utiliser correctement l'étendue des possibilités d'XML. La création d'une page XML avec sa DTD, mais sans style, se déroule sans problème, la norme étant fixée depuis plus d'un an (février 1998). Par contre, l'utilisation des feuilles de styles par XSL et des liens par XLL reste du domaine de l'expérimentation. Un autre problème est posé par les butineurs. Il n'existe pas pour l'instant de butineurs capables d'interpréter correctement XSL et XLL. Quant au seul "browser" capable de lire le XML, il s'agit d'Internet Explorer 5, mais il ne respecte pas toujours les recommandations du W3C. Cette carence de logiciels freine considérablement les développeurs de pages XML. A coté de grandes sociétés comme Microsoft, on trouve néanmoins de nombreux sharewares apportant chacun une pierre au logiciel futur qui intégrera toutes les facettes d'XML et de ses dérivés. En voici deux particulièrement intéressant :

4.2.1.1 Doc'Zilla

Doc'Zilla [DOCZ-99] s'installe pour l'instant sous les systèmes NT et Unix uniquement. C'est un logiciel intéressant dans le sens où les sources ne sont pas masquées complètement. Il est donc possible de voir en partie le déroulement du programme. Ce dernier fonctionne très bien avec les exemples fournis, mais "plante" quand on le confronte à d'autres exemples récupérés au travers du net. D'après la documentation fournie avec ce software, ce dernier est normalement capable de lire et comprendre XML et SGML. Doc'Zilla est produit par Citec, entreprise finlandaise.

4.2.1.2 *Hybrick*

Le grand intérêt d'Hybrick [FUJI-99] est que c'est un des rares logiciels à montrer les différents types de liens que permettent Xlink et Xpointer (de $1 \Rightarrow 1$ jusqu'à $1 \Rightarrow N$). Il inclut également DSSSL. Tous les exemples livrés avec ce programme fonctionnent très bien. Mais quand on fait à nouveau un test avec des exemples trouvés sur le net, le déroulement de la procédure se passe moins bien car, bien souvent, ces exemples ne possèdent pas de feuille de style DSSSL, ou cette dernière n'est pas valide. Hybrick s'installe sur Win95 ou sur NT. C'est un logiciel créé par la société Fujitsu.

Même si les logiciels ne sont pour l'instant pas encore au point, de nombreuses institutions s'essayent à la production de document XML pour être prêtes le moment venu.

4.3 *Produire du XML à partir du RTF, l'expérience de Lyon2 – les PUM*

Le projet de Lyon2 – les PUM étant plus avancé que le nôtre sur l'aspect XML, et une politique de rassemblement des acquis vis à vis de ce genre de projet étant en cours, nous avons cordialement été invités par Lyon2 pour un exposé sur leur chaîne de conversion [SERV-99]. En voici le déroulement :

- Le document Word ou WordPerfect est épuré de ses images, remplacées par des liens. La feuille de style opérateur est substituée à la feuille de style appliquée par l'étudiant.
- Le document ainsi préparé est converti en RTF.
- Un script du langage Omnimark (équivalent de Perl) surnommé "Omlette" [GEIM-99] transforme le fichier RTF en fichier XML à plat (Fichier avec balisage XML où on retrouve chaque chapitre RTF sous la forme d'une balise de chapitre, accompagnée des descriptions de police, style, etc.).
- Le fichier XML est ensuite converti en fichier SGML qui est le format d'archivage et le format pivot de ce projet.
- A partir du fichier SGML, on peut obtenir les formats MARC, SGML stylé, HTML, XML par l'application de divers scripts Omnimark.

Si le document source est un fichier LaTeX, il sera également converti en SGML par un programme en cours de création à l'université de Laval. Ce script s'appuie sur un ensem-

ble de macros LaTeX et MathML. La sortie de ce programme sera un fichier DVI-SGML auquel un script Omnimark sera appliqué ainsi que la DTD Teilite [TEI-99] pour obtenir du SGML correct.

Il nous semblait, dans un premier temps, intéressant d'étudier le résultat du programme Omlette donnant du XML à plat, pour évaluer si le format de sortie pouvait être intéressant à garder tel quel. Après avoir vu la conversion effectuée à Lyon 2, on peut conclure qu'une telle chaîne de traitement est inadaptée aux thèses scientifiques en raison d'un lourd traitement manuel à réaliser sur les formules et les figures. Néanmoins, le fichier XML obtenu à l'issu du premier script Omnimark [OMNI-99] est valide et peut donc avoir son intérêt. Le seul complément qui lui est apporté par la suite est l'application de la DTD TeiLite (DTD permettant le codage de document structuré comme une monographie). La conception d'un programme permettant d'automatiser le traitement des images et des figures rendrait alors la conversion jusqu'au XML à plat fort intéressante. Voici le schéma complet de la chaîne d'édition numérique de Lyon2 :

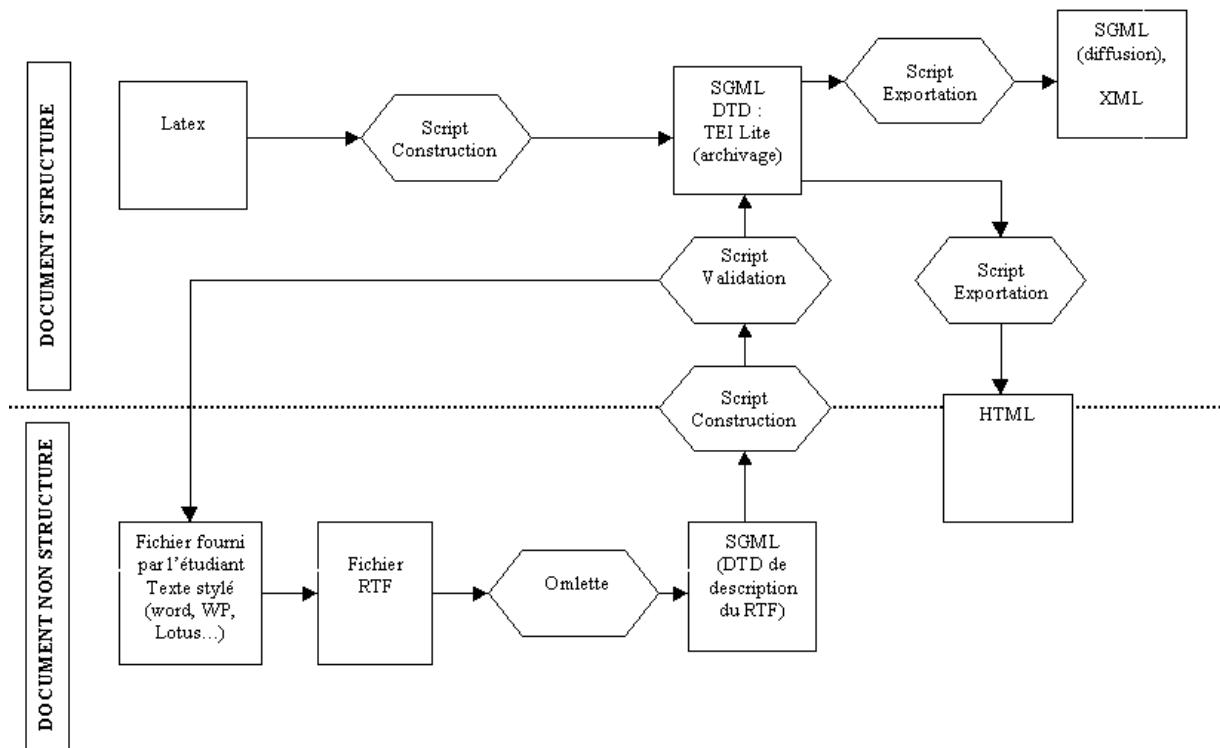


Fig.8 : Chaîne de traitement détaillée utilisée aux PUM et à Lyon2 pour la conversion des thèses. Cette images est issue du serveur de Lyon2⁹.

⁹ <URL : <http://phebus.univ-lyon2.fr/sentiers/editon/theses/ressources/SchemaChaine.gif>>

La mise en place de tels projets et la construction de bibliothèque numérique est une aventure passionnante. Un des facteurs qui nous rappelle à une réalité bien matérielle est le coût de mise en place de telles infrastructures.

5. Bilan financier

Il est évident que monter un serveur de thèses électroniques et tout ce qu'il implique a un coût. Il faut évidemment acheter du matériel (ordinateurs, programmes), mais également trouver les fonds pour la réalisation des feuilles de styles (rétribution du typographe), la formation du personnel (coût des conférences), la formation des étudiants (indemnité du formateur, achat du matériel d'enseignement), la conversion des thèses (paye de l'opérateur), l'entretien du serveur de thèses (paye de l'ingénieur). Et il ne faut surtout pas oublier le coût du développement d'un tel projet. Les stagiaires y travaillent à temps plein, une partie du personnel à temps partiel, une autre partie est mobilisée pour les réunions. Et le développement d'un tel projet est continu car les avancées technologiques permettant de l'améliorer sont permanentes.

6. Bilan et futur des projets de thèses électroniques

6.1 Aspect ressources humaines

Sur le plan des ressources humaines, travailler avec autant de personnes d'horizons différents est très enrichissant. Il n'est pas toujours aisé de mettre tout le monde d'accord, mais ce sont ces échanges d'idées, ces confrontations d'opinions, la réunion des besoins de chacun qui permettent d'obtenir des documents les plus complets et les plus pertinents possibles. Et il est réconfortant de constater qu'un travail mené à plusieurs dans la concertation peut aboutir à une belle réalisation.

6.2 Aspects technologiques

6.2.1 Les métadonnées

Comme nous l'avons vu précédemment, une unification des métadonnées est nécessaire. Il n'est pas envisageable de créer un serveur fédérateur à plusieurs si chacun adopte un système de métadonnées différent. Les métadonnées sont très importantes. C'est elles qui servent à l'indexation des documents dans les systèmes documentaires. Pour l'instant les recommandations du Dublin Core [DUBL-99]semblent faire l'unanimité. Lyon2, les PUM et Virginia Tech travaillent ensemble sur ce problème et voici les schémas qu'ils proposent actuellement sur leur site.

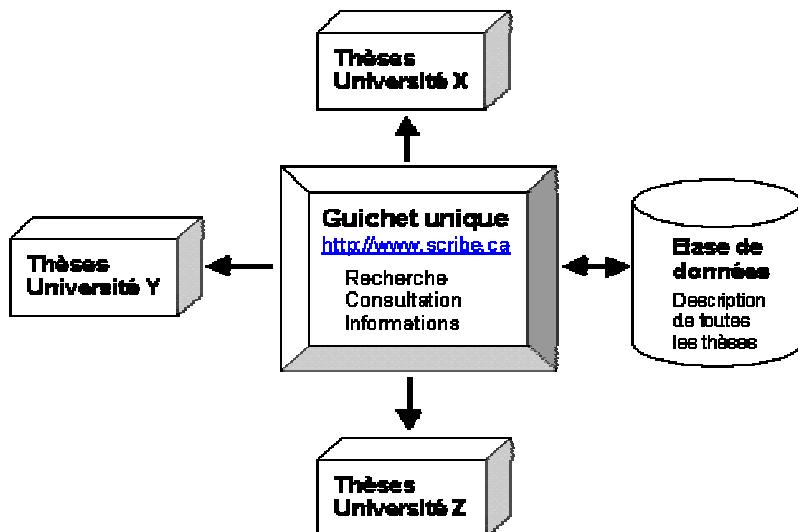


Fig. 9 : Serveur de thèses unique : 1^{ère} idée. Cette image est issue du serveur des PUM¹⁰.

¹⁰ <URL : http://www.pum.umontreal.ca/publ_electr/rapports/bea02/partie_5.html>

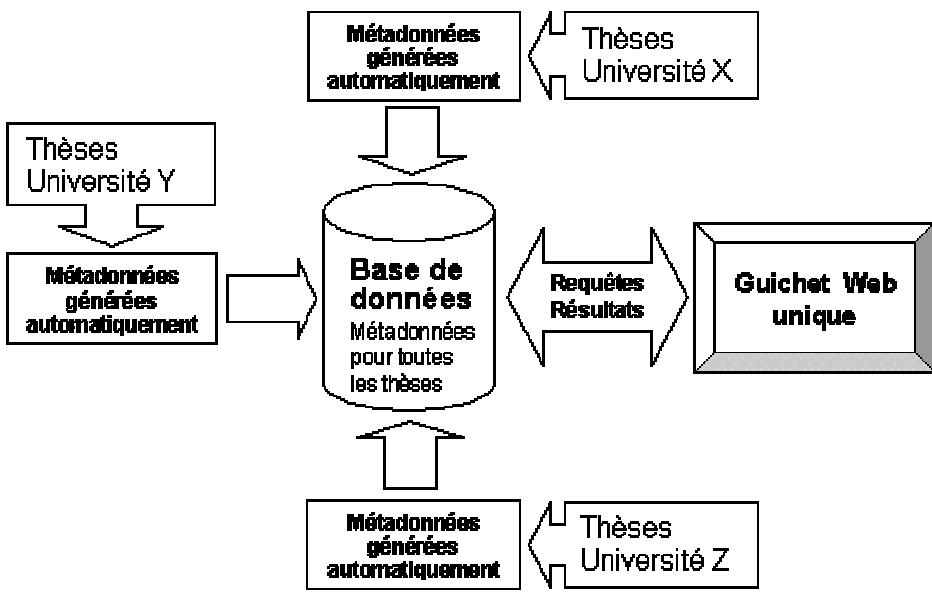


Fig. 10 : Serveur de thèses unique : 2^{ème} idée utilisant les métadonnées. Cette image est issue du serveur des PUM¹¹.

6.2.2 Le multimédia

L'avantage de posséder des documents électroniques est que l'on peut y inclure du multimédia. Son, photographie, dessin, vidéo peuvent maintenant apporter un complément d'information et de compréhension à une thèse. Mais les formats multimédias trouvés sur le net sont innombrables. Des dizaines de formats différents existent pour les images et le son. En ce qui concerne les animations ou les vidéos, les formats avi et mpeg se taillent la part du lion. Toutes ces données au formatage multiple posent le problème de la pérennité de l'information à long terme. Rien ne nous garantit que dans dix ans, les programmes permettant de lire ces fichiers existeront encore. L'évolution de ces formats est donc à surveiller en espérant que de grands standards émergeront et feront l'unanimité.

6.2.3 XML

Enfin, il faudra bien évidemment guetter les avancées [COVE-99] du langage XML et de ses dérivés. L'avenir de ce standard paraît assurément bon, mais en attendant d'en avoir la confirmation, l'INSA a choisi de continuer l'exploitation et le développement de sa chaîne de traitement actuelle, parallèlement au développement d'une solution d'archivage en XML.

¹¹ <URL : http://www.pum.umontreal.ca/publ_electr/rapports/bea02/partie_5.html>

Ainsi, quand l'emploi d'XSL et XLL sera banalisé et que ces standards seront intégrés, nous serons prêts à proposer nos thèses au format du nouveau standard non propriétaire du Web.

6.2.4 Les prochains développements de CITHER

L'aspect communication autour du projet est encore à développer. Une bonne information résulte d'un effort constant vers un public d'étudiants qui se renouvelle chaque année et un public d'entreprises et d'institutions scientifiques qu'il faut tenter d'accroître. Il faudra engager des actions de communication auprès des éventuels directeurs de thèse, faire connaître le serveur sur les listes de discussion et les groupes de news, établir des relations avec les associations de doctorants.

Le circuit de la thèse mériterait d'être repensé. Les exemplaires à destination des bibliothèques et centres de documentation sont fournis par le doctorant avant la soutenance, et il n'existe actuellement aucun moyen d'obliger le docteur à fournir un exemplaire mis à jour de sa thèse à la bibliothèque si le jury à demandé des corrections mineures. En cas de corrections importantes, le docteur ne peut de toute façon pas prétendre à une mention et à une publication.

La création d'un catalogue commun de thèses électroniques reste encore à l'étude et sa mise en place entraînera sûrement de nombreux développements et de nombreux projets annexes, notamment autour des métadonnées.

L'accueil qui sera réservé aux feuilles de styles par les étudiants et leur directeur de thèses imposera peut-être quelques bouleversements sur la manière de présenter une thèse et l'aboutissement à un nouveau "Guide de rédaction des thèses".

Enfin l'adaptation des chaînes d'édition numérique à des programmes et langages gratuits tel que StarOffice et XML dépendra de l'engouement du public pour ces derniers.

7. Et en dehors de CITHER ?

En plus de la possibilité qui m'a été offerte de participer au projet CITHER à l'INSA, j'ai également eu l'occasion d'intervenir dans chaque réunion inter-établissements de ce projet et l'opportunité d'exposer les premiers résultats de mon stage lors d'un club de la Doua, à l'ENSSIB, portant sur les projets de thèses électroniques.

L'INSA a également contribué à ma formation dans le domaine de la documentation en m'envoyant aux salons IDT au palais des congrès de Paris, au salon Rank Xerox à la cité

internationale de Lyon ainsi qu'aux journées Gutenberg [ASSO-99] à l'Institut de Physique Nucléaire de Lyon. Ces salons ont, pour moi, été l'occasion de prendre contact avec les grands éditeurs de logiciels documentaires et de mieux connaître les produits et les noms qui sont respectés dans le monde de la documentation.

Enfin, j'ai eu l'occasion d'expérimenter la vie d'une bibliothèque en participant à l'inventaire de Doc'INSA. Mon stage, qui possède mille et une facettes, aura donc été un enrichissement fabuleux pour moi sur toutes les activités que peut mener un centre de documentation.

8. Glossaire

ABES : Agence Bibliographique de l'Enseignement Supérieur

ADBS : Association des professionnels de l'information et de la documentation

ADSL : Association du Doctorat Scientifique Lyonnais

ANRT : Atelier National de Reproduction des Thèses

ASCII : American Standard Code for Information Interchange

BackOffice : « Zone » du projet CITHER hors de la vue du lecteur

CEN : Chaîne d'Edition Numérique

CISM : Centre d'Informations Scientifiques et Médicales de l'UCBL

CITHER : Consultation en texte Intégral des THèses En Réseau

CNRS : Centre National de Recherche Scientifique

GRADIENT : Groupe Rhône-Alpes des Docteurs en sIENces

CRIG : Centre de Ressources Informatiques de Gestion pour l'UCBL

CURA : Conférence Universitaire Rhône-Alpes

DED : Département des études doctorales de l'INSA de Lyon

DAED : Direction Administrative des Etudes Doctorales de l'UCBL

ENSSIB : Ecole Nationale Supérieure des Sciences de l'Information et des bibliothèques

FrontOffice : « Zone » du projet CITHER visible par le lecteur

GED : Gestion Electronique de Documents

HTML : HyperText Markup Language

IN2P3 : Institut National de Physique Nucléaire et de Physique des Particules

INSA : Institut National des Sciences Appliquées

MENRT : Ministère de l'Education Nationale, de la Recherche et de la Technologie

OPAC : Online Public Access Catalogue

PDF : Portable Document Format (format propriétaire de la société Adobe)

PFE : Projet de Fin d'Etude

Pont d'embarquement : Fichier HTML. Point d'entrée normalisé vers la thèse

Postscript : Langage d'impression développé par la société Adobe

PUM : Presses Universitaires de Montréal

RTF : Rich Text Format : format développé par Microsoft pour faciliter l'échange de documents entre traitements de texte.

SDB : Sous Direction des Bibliothèques

SGML : Standard Generalized Markup Language. Norme ISO 8879

UCBL : Université Claude Bernard Lyon

UMI : University Microfilms International

URL : Uniform Resource Location

XML : eXtended Markup Language

9. Références bibliographiques

Les références bibliographiques sont rédigées selon le guide proposé par Doc'INSA.

Ce guide repose sur les normes :

- Z44-005 « Documentation, références bibliographiques : contenu, forme et structure » de décembre 1987 qui elle-même reproduit intégralement la norme internationale ISO 690-1987.
- ISO/DIS 690-2-1995 « Information et documentation – Références bibliographiques – Documents électroniques ou parties de ceux-ci ».

[ASSO-99] *L'association Gutenberg*. [On-Line]. Juin 1999. [Visité le 15 septembre 1999] Available from Internet : <URL : <http://www.gutenberg.eu.org/>>

[BIEN-99] *Bienvenue sur le site du prototype de diffusion électronique des thèses de l'Université de Montréal*. [On-Line]. 1999. [Visité le 15 septembre 1999] Available from Internet : <URL : <http://www.pum.umontreal.ca/theses/>>

[CAFE-99] *Cafe con Leche XML News and Resources*. [On-Line]. Sept.1999. [Visité le 15 septembre 1999] Available from Internet : <URL : <http://metalab.unc.edu/xml/>>

[CARO-99] **CARON, Y.** *SU : système universitaire de documentation présentation synthétique*. [On-Line]. 1999. [Visité le 15 septembre 1999] Available from Internet : <URL : <http://13av01.univ-lille3.fr/www/idist/dfmld/SUWEB2.htm#z 39>>

[COMI-99] *Comité de travail sur les thèses électroniques - page d'accueil* . [On-Line]. Jan. 1999. [Visité le 15 septembre 1999] Available from Internet : <URL : <http://www.bibl.ulaval.ca/doelec/theses/>>

[CONS-98] *Conseils techniques pour le dépôt de thèses électroniques*. [On-Line]. Juil. 1998. [Visité le 13 septembre 1999] Available from Internet : <URL : http://csidoc.insa-lyon.fr/these/doc/tdepot_conseils.html>

[CONT-98] *Contrat de diffusion d'un travail universitaire*. [On-Line]. Mars 1998. [Visité le 13 septembre 1999] Available from Internet : <URL : <http://csidoc.insa-lyon.fr/these/doc/contrat.html>>

[COVE-99] **COVER, R.** *The SGML/XML Web Page*. [On-Line] Sept. 1999. [Visité le 10 septembre 1999] Available from Internet : <URL : <http://www.oasis-open.org/cover/sgmlnew.html>>

[DOCZ-99] *DocZilla XML/SGML Module Alpha 3*. [On-Line]. 1999. [Visité le 15 septembre 1999] Available from Internet : <URL : <http://www.doczilla.com/download/index.html>>

[DUBL-99] *Dublin Core Metadata Initiative*. [On-Line]. Sept. 1999. [Visité le 15 septembre 1999] Available from Internet : <URL : <http://purl.oclc.org/dc/>>

[FLYN-99] **FLYNN, P.** *XML FAQ*. [On-Line]. Juin 1999. [Visité le 14 septembre 1999] Available from Internet : <URL : <http://www.ucc.ie/xml/>>

[FUJI-99] *SGML Tool*. [On-Line]. Juin.1999. [Visité le 14 juillet 1999] Available from Internet : <URL : <http://www.infotek.no/dgmltool/vendors.htm#Fujitsu>>

[GEIM-99] **GEIMER, R.** *The OmniMark LE Tips and Trading Exchange, RTF2XML*. [On-Line]. 1999. [Visité le 15 septembre 1999] Available from Internet : <URL : <http://www.xmeta.com/omlette/>>

[GUIL-99] *La guilde des doctorants*. [On-Line].[Visité le 15 juillet 1999] Available from Internet : <URL : <http://garp.univ-bpclermont.fr/gilde/Alire>>

[HUNE1-98] **HUNEAU, M-E.** *Rapport de Projet de Fin d'Etude - Serveur de thèses en texte intégral*. [On-Line]. Juin 1998. [Visité le 13 septembre 1999] Available from Internet : URL:http://csidoc.insa-lyon.fr/these/doc/rapport_synthese.pdf

[HUNE2-98] **HUNEAU, M-E.** *Rapport de Synthèse - Serveur de thèses en texte intégral*. [On-Line]. Juin 1998. [Visité le 13 septembre 1999] Available from Internet : <URL : http://csidoc.insa-lyon.fr/these/doc/rapport_synthese.pdf>

[LUDW-99] **LUDWIG, J.** *An Investigation of XML with Emphasis on Extensible Linking Language (XLL)*. [On-Line]. Mars 1999. [Visité le 15 septembre 1999] Available from Internet : <URL : <http://pages.wooster.edu/ludwigj/xml/thesis.html>>

[MART-99] **MARTIN, D.** *Transmission des connaissances: une solution prometteuse en XML*. [On-Line]. 1999. [Visité le 15 septembre 1999] Available from Internet : <URL : http://worldserver2.oleane.com/dmartin/Transmission_des_connaissances.htm>

[MART2-99] **MARTIN, D.** *Stockage et interopérabilité en XML*. [On-Line]. 1999. [Visité le 12 septembre 1999] Available from Internet : <URL : http://worldserver2.oleane.com/dmartin/xml_livre_banc.pdf>

[MERM-98] **MERMET, J-M.** *Coordination et mise en place d'un serveur de thèses en texte intégral à l'INSA de Lyon. Conception du FrontOffice*. [On-Line]. Sept. 1998. [visité le 13 septembre 1999] Available from Internet : <URL : <http://csidoc.insa-lyon.fr/these/doc/rapportprojet.html>>

[MICH-99] **MICHARD A..** *XML. Langage et applications*. 2^{ème} tirage. Paris : Eyrolles, 1999. 361 pp.

[NETW-99] *Networked Digital Library of Theses and Dissertations*. [On-line]. Juil. 1999 [Visité le 20 septembre 1999] Available from Internet : <URL : <http://www.ndltd.org>>

[OMNI-99] *Omnimark.com*. [On-Line]. 1999. [Visité le 15 septembre 1999] Available from Internet : <URL : <http://www.omnimark.com>>

[PETI-99] **PETITJEAN, A.** *Le langage XML*. [On-Line]. Août 1999. [Visité le 10 septembre 1999] Available from Internet : <URL : <http://www.citeweb.net/apetitje/xml/index.html>>

[PROJ-98] *Project "Digitale Dissertationen" at the Humboldt University of Berlin.* [On-Line]. 1998. [Visité le 11 septembre 1999] Available from Internet :
<URL : http://dochost.rz.hu-berlin.de/epdiss/index_en.html>

[SANO-99] **SANOUILLET, A.** *Thèses et styles de citations.* [On-Line]. Fevr. 1999. [Visité le 8 septembre 1999]. Available from Internet :
<URL : <http://www.unice.fr/UrfistDEH/Dissertations.html>>

[SANO2-99] **SANOUILLET, A.** *Guide de présentation des thèses et mémoires.* [On-Line]. 1999. [Visité le 9 septembre 1999]. Available from Internet :<URL :
<http://www.unice.fr/UrfistDEH/Theses1.html>>

[SERV-99] *Service des Nouvelles Technologies pour l'Information Et la Réalisation de Serveurs (SENTIERS).* [On-Line]. Septembre 1999. [Visité le 13 septembre 1999]. Available from Internet :<URL : <http://phebus.univ-lyon2.fr/sentiers/edition/theses/projettheses.html>>

[TEI-99] *Text Encoding Initiative.* [On-Line]. Sept. 1999. [Visité le 13 septembre 1999] Available from Internet :
<URL : <http://www-tei.uic.edu/orgs/tei/>>

[TOGN-99] **TOGNAZZI, J.** *CITHER. Etude XML, intégration de LaTeX.* Lyon : INSA, 1999. 6pp.

[UMID-99] *UMI Dissertation Service.* [On-Line]. 1999. [visité le 15 septembre 1999] Available from Internet :<URL : <http://www.umi.com/hp/Support/DExplorer/>>

[XML-98] *EXtensible Markup Language (XML) 1.0.* [On-Line]. sept.1999 [Visité le 14 septembre] Available from Internet :
<URL : <http://www.w3.org/TR/1998/REC-xml-19980210>>

[XML2-99] *XML francophone.* [On-Line]. 1999. [Visité le 10 septembre 1999] Available from Internet :<URL : <http://www.chez.com/xml/index.htm>>

[XML3-99] *XML Page at TECFA.* [On-Line]. Aout 1999. [Visité le 15 septembre 1999] Available from Internet :<URL : <http://tecfa.unige.ch/guides/xml/pointers.html>>

[XLL-99] XML Linking Language (Xlink) working draft. [On-Line]. Sept. 1999 [Visité le 13 septembre 1999] Available from Internet :
<URL : <http://www.w3.org/TR/1998/WD-xlink-19980303>>

[XSL-99] eXtensible Stylesheet Language (XSL) working draft [On-Line]. Sept. 1999 [Visité le 13 septembre 1999] Available from Internet :
<URL : <http://www.w3.org/TR/WD-sxl/>>

10. Bibliographie

Documents Papiers:

AHONEN H., HEIKKINEN B., HEINONEN O. Analysis of document structures for element type classification. In *PODDP' 98 : principles of digital document processing : Saint Malo, 29-30 March 1998. Lecture notes in computer science*, 1998, vol 1481, p 24-42.

ANDRES F., BUFORD J., ONO K. An application-oriented approach for HyTime structured document management. In *DEXA '98 : database and expert systems applications, Vienna, 24-28 August 1998. Lecture notes in computer science*, 1998, vol 1460, p 260-269.

BAUWENS B., EVENEPOEL F., ENGELEN J. SGML as an enabling technology for access to digital information by print disabled readers. SGML into the nineties. *Computer standards & interfaces*, 1996, vol 18, n°1, p 55-69.

BAYRAKTAR M., ZHANG C., VADAPALLI B. A Web art gallery. In *Digital 98 Libraries. Third ACM Conference on Digital Libraries, Pittsburgh PA, 23-26 juin 1998*. New York : ACM, 1998, p 277-278.

BURLAT J.M., PRUDHOMME B. *REFERENCES BIBLIOGRAPHIQUES – Rédaction et lecture.* 5^{ème} éd. [On-line]. Villeurbanne (Fr) : Inst. Nat. Sci. Appl., Doc'INSA, Sep. 1997 [visité le 15/09/1998] Available from internet :
<URL : <http://csidoc.insa-lyon.fr/docs/refbibli.html>>

CHESNUTT D.R. SGML and the digital libraries of tomorrow. *The Journal of academic librarianship*, 1998, vol 24, n°3, p 232-236.

CLEMSON P.A. An inside approach to networked document cataloging. *Journal-of-Internet-Cataloging*, 1997, vol 1, n°2, p 57-64

CORTHOUTS J., PHILIPS R. SGML : a librarian's perception. *Electronic-Library*, 1996, vol 14, n°2, p 101-110.

ERICKSON J., STOEFFLER M. An SGML/HTML electronic thesis and dissertation library. *Computers-and-the-Humanities*, 1999, vol 33, n°1-2, p 175-184.

FOX E.A., McMILLAN G., EATON J.L. The evolving genre of electronic theses and dissertations. In *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, 1999.* [CD-ROM], Los Alamitos : IEEE Comput. Soc., 1999, liii+341 pp. 20 pp.

FREW J., FREESTON M., FREITAS N. The Alexandria Digital Library Architecture. In *Research and advanced technology for digital libraries, Heraklion, 21-23 September 1998. Lecture notes in computer science*, 1998, vol 1513, p 61-73.

GOOSSENS M., LE MEUR J.Y., Afficher les documents scientifiques sur le Web. In *EuroTEX'98, Saint Malo, 29 mars-1 avril 1998. Cahiers GUTenberg*, 1998, n°28-29, p 181-196.

GORDON, T.F. qwertz synthesis of SGML and LaT//EX. *Computer Standards & Interfaces*, 1995, vol.17, n°1, p 25-33.

KLUGKIST A., LUYT-PRINSEN Jola GB van. Case study :The Eldoradoc project (supply by authors). In *International summer school on the digital library, Tilburg, 4-16 August 1996*. Tilburg : Ticer, 1996, p 269-277.

McMILLAN G. Electronic theses and dissertations : merging perspectives. *Cataloging-&-Classification-Quarterly*, 1996, vol 22, n°3-4, p 105-125.

NIJSSEN T ,LUYT-PRINSEN Jola GB van. Libraries and SGML. In *International summer school on the digital library, Tilburg, 4-16 August 1996*. Tilburg : Ticer, 1996, p 363-375.

NIJSSEN T; LUYT-PRINSEN Jola GB van. Encoding formats for electronic documents. In *International summer school on the digital library, Tilburg, 4-16 August 1996*. Tilburg : Ticer, 1996, p 349-359.

PRESCOD, P. Multiple media publishing in SGML. In *Proceedings of the 1996 ACM SIGDOC 14th Annual International Conference on Computer Documentation, Triangle Park NC, 20-23 October 1996*. New York : ACM, 1996, p 3-9.

ROISIN C., ROVAN B. Authoring structured multimedia documents. In *SOFSEM '98 : theory and practice of informatics : Jasna, 21-27 November 1998. Lecture notes in computer science*, 1998, vol 1521, p 222-239.

SORGAARD P., SANDAHL T.I. Problems with styles in word processing: a weak foundation for electronic publishing with SGML. In *Proceedings of the Thirtieth Hawaii International Conference on System Sciences*, vol 6. Los Alamitos : IEEE Comput. Soc. Press, 1997, p 137-146.

SPERBERG-MCQUEEN C.M. XML and the future of digital libraries. *The Journal of academic librarianship*, 1998, vol 24, n°4, p 314-317.

WEDGEWORTH R., MISCHO W.H., SHAW D., ed. Issues affecting the development of digital libraries in science and technology. In *Joint ICSU Press / UNESCO Expert conference, Electronic publishing in science, Paris, February, 1996*. Paris : ICSU, 1996, p 174-181.

WEIBEL S.L. Metadata the foundation of resource description. *Annual review of OCLC research*, 1995, p 52-56.

Documents électroniques :

eXtended Markup Language. [On-Line]. Sept. 1998.[Visité le 2septembre 1999] Available from Internet : <URL : <http://www.cern.ch/WebOffice/Doc/Courses/XML/>>

BRYAN, M. *OII Guide to XML and related standards.* [On-Line]. Juin 1999. [Visité le 15 septembre 1999] Available from Internet :
<URL : <http://www.echo.lu/oiic/en/xmlguide.html>>

DeveloperWorks : XML. [On-Line]. 1999. [Visité le 14 septembre 1999] Available from Internet : <URL : <http://www.ibm.com/developer/xml/>>

GARSHOL, L.M. *Free XML Software*. [On-Line]. Sept. 1999. [Visité le 15 septembre 1999] Available from Internet :
<URL : <http://www.stud.ifi.uio.no/~larsga/linker/XMLtools.html>>

Gottinger Digitalisierungs Zentrum. [On-Line]. Juin 1999. [Visité le 13 septembre 1999] Available from Internet :
<URL : http://www.sub.uni-goettingen.de/gdz/en/gdz_main_en.html>

PEPPER, S. *SGML & XML TOOLS - By Tool Category*. [On-Line]. Juin 1999. [Visité le 15 septembre 1999] Available from Internet :
<URL : <http://www.infotek.no/sgmltool/products.htm>>

XML Industry Portal. [On-Line]. 1999. [Visité le 15 septembre 1999] Available from Internet
<URL : <http://www.xml.org>>

XML : Langage et Applications. [On-Line]. 1999. [Visité le 15 septembre 1999] Available from Internet :
<URL : <http://www.eyrolles.com/livres/michard/logiciels.asp?ID=121480>>

XML Software. [On-Line]. Sept. 1999. [Visité le 15 septembre 1999] Available from Internet:
<URL : <http://www.xmlsoftware.com/>>

Annexes

Table des matières des annexes :

1. Compte rendu de la première réunion du groupe de travail
2. Compte rendu de la deuxième réunion du groupe de travail
3. Guide de rédaction des thèses proposé par l'INSA de Lyon à la date du 13/09/99
4. Mise en page que prendra le guide de rédaction des thèses
5. Modèle de feuilles de styles "classique"
6. Modèle de feuilles de styles "basique"
7. Modèle de feuilles de styles "moderne"
8. Documents ajoutés aux dossiers d'inscription

1

Compte rendu de la première réunion du groupe de travail

Structuration électronique des thèses
Réunion du 18/06/99 à Doc'INSA

Compte-rendu rédigé par Carole Clerc (clerc@cismsun.univ-lyon1.fr), Monique Joly (Monique.Joly@insa-lyon.fr), Jean-Michel Mermet (Jean-Michel.Mermet@insa-lyon.fr) et Brigitte Prudhomme (Brigitte Prudhomme@insa-lyon.fr)

Étaient présents à la réunion :

Universités et Pôle européen de Grenoble : E. Cherhal, N. Serrano
Université Lyon 1 : E. Iwema
ENSSIB : L. Miremont
INSA-Lyon : D. Boudia, C. Clerc, M. Joly, J.M. Mermet, B. Prudhomme, L. Quinson, J. Tognazzi

Société Trafik : Damien Gautier

Étaient excusés ou représentés :

Université Lyon 1 : A.M. Denis, M. Jouineau, A. Nivet
Université Lyon 2 : J.P. Ducasse, C. Micol, M.J. Tinland, T. Sabin
Université Lyon 3 : F. Molline
ENS : A. Mauries
ENSSIB : D. Roger, J.M. Salaun
Universités et Pôle européen de Grenoble : C. Guitart, M.N. Icardo, M.F. Rochard
MENRT-SDB : C. Okret
Université de Marne la Vallée : D. Schibler
Université de Savoie : C. Commanay
Université de St-Etienne : M.C. Achard
INSA-Lyon : S. Calabreto, J.M. Pinon

A / Résumé des objectifs du groupe de travail :

Faciliter la montée en charge d'un serveur de thèses en texte intégral par l'obtention, en amont, de mémoires électroniques correctement structurés.

Pour cela, il est impératif :

- d'établir des " gabarits " de documents :
 - en s'inspirant des travaux déjà publiés par d'autres bibliothèques universitaires
 - attrayants grâce à l'expertise d'un typographe
- d'inciter les étudiants à les utiliser en leur montrant l'intérêt réciproque : rédiger, à moindre effort, des documents à la mise en page professionnelle grâce à des modèles informatiques, un guide de rédaction et une formation adaptée.

B / Thèmes abordés :

1. Gabarits proposés par d'autres projets

Nous avons recherché, de la manière la plus exhaustive possible, quels sont les différents projets en cours, en France et à l'étranger qui proposent des aides à la rédaction structurée. Nous les avons évalués. En voici les principaux :

11 / Le projet de l'université Laval :

Avantages :

- Projet avancé et pratique pour les étudiants.
- Guide clair et pédagogique.
- Feuilles de style fournies pour Word et WordPerfect (prévues), personnalisant l'environnement.
- Processus de conversion Word (ou WordPerfect) ==> RTF ==> SGML
- Cours sur l'utilisation des feuilles de style

Inconvénients :

- Les feuilles de style sont ternes.
- Peu de diversité dans les feuilles de style proposées (un seul type).
- Le guide de l'étudiant est trop long pour espérer qu'il soit lu en entier.
- Les utilisateurs de LaTeX n'ont, pour l'instant, pas de solution adaptée à leur cas.

Pour plus d'informations :

<http://tracerbullet.bibl.ulaval.ca/doelec/theses/aide.html>

12 / Le Projet de Virginia Tech (ETD-ML : Electronic Theses and Dissertations Markup Language) :

Avantages :

- Pratiquement tous les cas de figure sont traités, Word, LaTeX et ETD-ML.
- Format de conversion Word ==> PDF ou TeX ==> DVI ==> PS ==> PDF ou ETD-ML ==> HTML
- Pour la bibliothèque, économie de temps, c'est l'étudiant qui crée les fichiers publiables.

Inconvénients :

- Pas de cours ou de formation pour les étudiants mais juste des documents d'aide en ligne.
- Projet difficile à prendre en main quand on a des connaissances moyennes en informatique.
- Présentation rébarbative et peu conviviale.
- Demande du travail supplémentaire à l'étudiant.

Pour plus d'informations : <http://etd.vt.edu/>

Projet similaire : http://dochost.rz.hu-berlin.de/epdiss/index_en.html

13 / Projet des PUM de Montréal :

les PUM de Montréal et l'Université Lyon 2 travaillent ensemble sur un projet de thèses électroniques. Les feuilles de styles n'ont été rendues disponibles que récemment mais ce projet est intéressant et nous semble prometteur.

Avantages :

- Feuilles de style fournies pour Word et WordPerfect (prévues), personnalisant l'environnement.
- Un guide de présentation va prochainement être créé.
- Une formation est envisagée pour les étudiants de dernier cycle.

Inconvénients :

- Les feuilles de style sont ternes.
- Peu de diversité dans les feuilles de style proposées (un seul type).
- Pas encore de guide pour l'étudiant.
- Les utilisateurs de LaTeX n'ont, pour l'instant, pas de solution adaptée à leur cas.

Pour plus d'informations : <http://www.pum.umontreal.ca/theses/index.html>

14 / En France :

- Beaucoup de manuels ou guides sur les règles d'écriture à respecter pour les thèses
- Manuels trop longs et trop flous, y compris celui du ministère (<http://www.abes.fr/teleactu.htm>)

2. Cahier des Charges adopté par le groupe de travail à la date du 18 juin 1999 :

Après examen de l'existant, il nous apparaît impératif de concevoir un guide de rédaction de thèses, prenant en compte non seulement des considérations techniques de présentation et d'utilisation des feuilles de style, mais également les dispositions réglementaires en cours en proposant des amendements au besoin.

Pour établir ce guide, on s'inspirera :

- du guide du ministère proposé sur le site de l'ABES
- du guide de l'étudiant de l'Urfist de Nice s'appuyant sur les normes ISO 7144 et AFNOR Z44-005 (<http://www.unice.fr/UrfistDEH/Dissertations.html>)
- des guides canadiens et américains

Nous nous sommes ensuite attachés à décrire plus précisément la forme que doit prendre le document thèse :

21 / Le plan d'une thèse, avec ses différentes composantes, d'après les préconisations ministérielles et les pratiques en cours :

- 1 Couverture
- 2 Page de titre, copie exacte de la couverture
- 3* Liste des professeurs de l'établissement, des formations doctorales, des écoles doctorales
- 4 Remerciements, dédicaces, avant-propos
- 5 Résumés longs
- 6 Table des matières ou sommaire
- 7 Liste des abréviations, figures, tableaux
- 8 Introduction
- 9 Chapitres
- 10 Conclusions
- 11 Références Bibliographiques
- 12 Annexes
- 13* Folio administratif
- 14 Quatrième couverture comportant résumé court et mots-clefs, en anglais et en Français.

* Les points 3 et 13 sont propres à l'INSA.

22 / En-têtes :

Rappel du numéro de chapitre et du titre de chapitre

23 / Pieds de page :

Nom auteur / type de document / année / établissement de soutenance

24 / Nouveau chapitre :

- Page de présentation à part (le rendre optionnel)
- Rappel, sur la page suivante, de la table des matières du chapitre (intéressant : voir s'il faut le rendre optionnel)

25 / Format du document :

Les feuilles de style seront conçues pour le format A4, même si le groupe de travail est bien conscient du fait que d'autres formats sont parfois utilisés.

26 / Mode de reliure préconisé:

Dos carré collé

27 / Numérotation :

- Pagination continue en bas de page centré.
- De cinq à sept niveaux de titre.

3. Mission attribuée au graphiste :

Damien Gautier va proposer trois gabarits Word, de styles typographiques différents, de manière à satisfaire un maximum de personnes. L'INSA se chargera par la suite de convertir ces gabarits en LaTeX.

Tous les guides de rédaction de la thèse ont été fournis à D. Gautier. Ont

également été données à D. Gautier, sur disquette, les différentes feuilles de style commentées ici, ainsi que des exemplaires de thèses de factures diverses (très textuelles, très graphiques, en version papier et version électronique).

C / Prochaine réunion :

Objectifs :

Remise des propositions de D. Gautier : modèles et guide de rédaction

Validations par le groupe de travail.

Date : 12 juillet à 14 heures

Lieu : Doc'INSA, Bât. 220, 20 avenue Albert Einstein, 69100

Villeurbanne

Pour en savoir plus, vous pouvez contacter :

Jean-Michel Mermet - Tél : 04-72-43-85-64

Jean-Michel.Mermet@insa-lyon.fr

Brigitte Prudhomme - Tél : 04-72-43-83-59

Brigitte.Prudhomme@insa-lyon.fr

2

Compte rendu de la deuxième réunion du groupe de travail

Structuration électronique des thèses

Réunion du 12/07/99

Rédigé par Carole Clerc (clerc@cisimsun.univ-lyon1.fr)

Etaient présents à la réunion :

MENRT-SDB : C. Okret

LISI : S. Calabretto

Urfist de Nice : A. Sanouillet

Université Lyon 1 : M. Jouineau

Université Lyon 2 : V. Bouletreau

ENSSIB : L. Miremont

Société Trafik : D. Gautier + 1 stagiaire

INSA de Lyon : D. Boudia, M. Joly, J-M Mermet, B. Prudhomme, L. Quinson, J.

Tognazzi, F. Charbonnier

Etaient excusés : E. Cherhal, N. Serrano, A. Nivet, A-M. Denis, J-P Ducasse, J-M Pinon

Résumé des objectifs du groupe de travail :

- Présentation des modèles de thèses créés par D. Gautier
- Plan détaillé du guide de rédaction des thèses

Thèmes abordés :

Présentation des modèles de thèses créés par D. Gautier

Introduction

Les modèles vont définir la structure de la thèse (enchainement et ordre des parties, niveaux de titre, marges, ...). La police utilisée va conditionner les modèles réalisés. C'est pourquoi le choix d'une police de qualité est fondamental.

1.1 Etude approfondie des polices de caractères appropriées pour la thèse.

D. Gautier a présenté les 5 polices systèmes disponibles à la fois dans les environnements PC et Mac. Il s'agit de trois polices de texte (Times, Georgia, Palatino (ou Book Antiqua)) et deux de titrages (Arial, verdana). Ces polices ont été jugées en fonction de leur lisibilité et de leur résistance à des tests de distorsion simulant les diverses reproductions que subira une thèse. Ces polices systèmes résistent mal en général aux tests de distorsion. Trois polices propriétaires, correspondant mieux à nos attentes, ont ensuite été présentées : Swift, Oranda, The Sans

Constatant que les polices qui offrent la meilleure lisibilité et la meilleure résistance aux tests de distorsion sont des polices propriétaires, D. Gautier nous a proposé deux solutions :

- La création d'une police avec cession des droits pour le projet
Le budget serait ici de 125 000 à 150 000 francs pour six polices (caractères de texte et de tirage avec pour chacun d'eux le roman, l'italique et gras).

Franck Jalleau de l'Imprimerie Nationale pourrait réaliser un caractère ayant les caractéristiques souhaitées de lisibilité (écran et papier) et de résistance à la reproduction, sachant tout de même qu'il n'existe pas de caractère idéal à l'écran et à l'impression.

- L'achat d'une police de caractère

Le coût de cet achat revient à 25,00F. par thèse et par an.

Comparaison des deux solutions pour l'année 1997 :

L'achat d'une police revient pour l'INSA à un coût de 3500F. par an pour les 140 thèses soutenues, et pour Lyon1 à un coût de 10000F. par an pour les 400 thèses soutenues en sciences. Soit un coût total annuel de 13500F pour les deux établissements considérés. On peut alors constater qu'en 10 ans le coût de création des polices de caractère serait amorti.

Si l'on considère l'ensemble des thèses francaises (sciences, lettres, sciences humaines) pour lesquelles ces polices pourraient potentiellement s'appliquer, soit 8953 thèses par an, le coût d'achat s'élève à 223825F. On s'aperçoit alors que dans ce cas, la création des police est bien plus avantageuse financierement que l'achat. En effet, la somme à investir est alors divisé par 2.

Les enjeux financiers étant importants, aucune décision n'a pour l'instant été prise en ce qui concerne le mode d'obtention des polices. Il a été toutefois bien compris que les polices constituent un enjeu important du projet.

1.2 Les feuilles de styles

D. Gautier a présenté, selon ce qui lui avait été demandé lors de la dernière réunion, trois feuilles de styles de trois genres différents :

Classique :

Réalisée à l'aide de la police Oranda

Proche de la présentation habituelle des thèses et des préconisations du ministère en ce qui concerne la page de couverture.

Basique :

Réalisée en The Sans avec une page de titre alignée à gauche, et un rappel de la table des matières en début de parties (cf. les feuilles de styles distribuées)

Moderne :

Réalisée en The Swith, au design moderne, avec une page de titre pour moitié alignée à gauche et pour l'autre moitié centrée, et un rappel de la table des matières en début de parties (cf. les feuilles de styles distribuées)

Remarques formulées sur les feuilles de styles :

A. Sanouillet : Ces feuilles de styles ne respectent pas forcément la norme, mais cette dernière a certainement besoin d'être rajeunie.

V. Bouletreau : Le texte paraît compact. L'espace laissé entre les lignes ne permet pas d'annotations or la thèse est un document de travail. Il faudrait faire des tests auprès des directeurs de thèses pour savoir si une telle mise en page leur conviendrait.

M. Joly : La numérotation n'est pas centrée.

Suite à ces remarques, la décision de faire une enquête auprès des directeurs de thèse a été prise. L'ENSSIB, Lyon2, et L'INSA de Lyon vont chacun fournir un chapitre de thèse de leur domaine à D. Gautier qui leur appliquera ses feuilles de styles. Ces chapitres stylés devront être jugés par les directeurs de thèses. Compte tenu des vacances, ces tests ne pourront se dérouler qu'en septembre selon le planning suivant :

- fin de la 1^{ère} quinzaine de septembre, D. Gautier fournit les chapitres stylés qui serviront pour les tests.
- pendant 3 semaines, les tests s'effectuent auprès des directeurs de thèses.
- le 7 octobre, mise en commun des dépouillements de l'enquête.

Les Métadonnées

L'objectif final du projet de thèses électroniques est d'aboutir à une base de données commune. Il est donc impératif d'avoir un catalogue de métadonnées fédérateur. Lyon 2, les PUM, les universités de Laval et de Virginia Tech travaillent déjà en collaboration dans ce but. Pour l'instant les spécifications du Dublin-Core semblent faire l'unanimité. L'INSA a également utilisé les spécifications Dublin-Core pour toutes les thèses présentées en ligne.

Le plan détaillé du guide de rédaction proposé par l'INSA

(fichier GuideDeRedaction.doc fourni en document attaché)

Plusieurs remarques ont été formulées suite à la lecture de ce plan :

- La formation qui sera donnée aux étudiants sur la rédaction des documents structurés devrait commencer dès la première année de thèse avec un rappel en troisième année.
- Le multimédia et ses multiples formats (surtout en ce qui concerne les images et le son) pose un problème de relecture à long terme. Il y a beaucoup trop de "plugs ins" et de logiciels spécifiques pour assurer une certaine pérennité à l'information.
- Les mots-clefs en anglais posent le problème de l'unification des métadonnées.

Prochaine réunion :

Objectifs :

- Dépouiller les résultats de l'enquête réalisée auprès des directeurs de thèse et concernant l'appréciation des feuilles de styles.
- Etudier le guide de rédaction des thèses.

Date : 7 octobre à 14 heures

Lieu : Doc'INSA, Bât. 220, 20 avenue Albert Einstein, 69100 Villeurbanne

Pour en savoir plus, vous pouvez contacter :

Jean-Michel Mermet - Tél. : 04-72-43-85-64

Jean-Michel.Mermet@insa-lyon.fr

Brigitte Prudhomme - Tél. : 04-72-43-83-59

Brigitte.Prudhomme@insa-lyon.fr

3

Guide de rédaction des thèses proposé par l'INSA de Lyon à la
date du 13/09/99

Guide de rédaction des thèses proposé par l'INSA

Introduction

Vous êtes sur le point de rédiger votre mémoire de thèse. Il sera pour vous l'aboutissement de trois années de labeur, le témoignage de votre passage au sein d'une équipe et d'un laboratoire de recherche. Ce mémoire est votre première publication, une sorte d'intronisation dans la communauté scientifique.

La mise en valeur de ce travail est donc fondamentale.

Cette valorisation commence avec une bonne lisibilité et une bonne structuration. Une thèse est un document long, qui est appelé à être lu, parcouru et reproduit. Il est donc impératif qu'il soit construit de manière correcte, et que sa lecture soit facilitée par une mise en page étudiée et une typographie agréable et permettant la reproduction. Une bonne valorisation passe également par une bonne diffusion auprès d'un public le plus large possible. Nous nous sommes engagés à vous aider sur ces points.

La lisibilité

Pour valoriser votre travail et en faciliter la consultation en le rendant plus attractif, nous avons fait appel à un typographe. De par son expertise, il est le plus à même de savoir comment améliorer la lisibilité de vos travaux. Nous lui avons demandé d'établir trois modèles de trois styles différents de manière à satisfaire le plus grand nombre d'entre-vous. Voici le plan qu'il développera dans une future version de ce document :

- 1.1 Présentation du travail typographique réalisé
- 1.2 Présentation des choix typographiques
- 1.3 Présentation des trois modèles et de leurs caractéristiques principales

- 1.3.1 Gabarits
- 1.3.2 Niveaux de lecture et feuilles de style

- 1.4 Sensibilisation à la typographie et à la mise en pages. Présentation des principes typographiques et de mise en pages fondamentaux

- 1.4.1 Choisir un caractère
 - 1.4.1.1 Proportions, axe, rythme
 - 1.4.1.2 Caractère de titrage/caractère de lecture
 - 1.4.1.3 Contraintes propres au projet "CITHER"
- 1.4.2 Variantes typographiques (caractères romain, italique et gras)
- 1.4.3 Alignements (lisibilité respective des différents type d'alignement)
- 1.4.4 Rapport Caractère/Corps/Justification
- 1.4.5 Interlignage (parcours de l'œil, règles fondamentales)
- 1.4.6 Déformation des caractères (incidence sur la lisibilité)
- 1.4.7 Césures (lisibilité des titres et textes courts)
- 1.4.8 Capitales (lisibilité, lettres accentuées)
- 1.4.9 Ponctuation (espaces à respecter, apostrophe et guillemets typographiques)
- 1.4.10 Abréviations et symboles internationaux
- 1.4.11 Marges
 - 1.4.11.1 Rapport entre les marges
 - 1.4.11.2 Reliure
- 1.4.12 Hiérarchie des informations (principe de base, possibilités typographiques)
- 1.4.13 Cohérence des documents (de la page de titre à la dernière page)
 - 1.4.13.1 Alignements
 - 1.4.13.2 Choix typographiques
 - 1.4.13.3 Choix graphiques
- 1.4.14 Incorporation des visuels
 - 1.4.14.1 Placement
 - 1.4.14.2 Légendes
- 1.4.15 Lexique

La performance

En vous apportant par une formation courte tout ce que vous devez savoir sur la rédaction de documents scientifiques structurés, l'équipe de CITHER espère vous faire gagner du temps sur votre planning de rédaction. Cette formation vous sera utile pour votre thèse mais également pour votre future carrière de diffuseur de l'information scientifique lorsque vous aurez des articles à rédiger.

Cette formation comprendra deux aspects :

- l'utilisation efficace des feuilles de styles
- la rédaction correcte des références bibliographiques.

Les feuilles de styles

Pour vous permettre de mieux structurer votre mémoire de thèse, tout en y incluant certains critères informatiques dont nous avons besoin, pour traiter votre document et le rendre publiable sur le Net, nous allons vous fournir des feuilles de styles. Ces dernières sont prêtes à l'emploi et vous permettent par simple clic de souris de définir un style pour chaque partie de votre thèse. De plus elles vous apportent une mise en page professionnelle car leur typographie a été étudiée par un professionnel.

Les références bibliographiques

Dans votre mémoire de thèse ainsi que dans vos futures publications, vous aurez à rédiger une bibliographie. L'objectif majeur d'une bibliographie est de permettre au lecteur de replacer votre travail au cœur du domaine de recherche considéré. Par la consultation de vos références bibliographiques, n'importe qui dans le monde doit pouvoir retrouver les documents dont vous vous êtes servis.

Constituer une bibliographie relève donc en premier lieu du domaine de la communication. C'est pourquoi l'écriture des références est normalisée. **La rédaction correcte de ces références est donc fondamentale.** La formation a également pour but de vous expliquer comment citer un article de périodique, une monographie, un congrès... en respectant les normes. Vous pourrez, pour compléter cette formation, vous aider du guide à votre disposition à :

<http://csidoc.insa-lyon.fr/docs/refbibli.html>

La structure

La présentation de votre mémoire de thèse doit suivre certaines règles proposées par le Ministère de l'Education Nationale et de la Recherche. Les **recommandations** pour la présentation des thèses en vue de faciliter leur lisibilité, leur identification et leur diffusion sont disponibles à :

<http://csidoc.insa-lyon.fr/these/pagedocteur.htm>

Toutes les thèses, sauf les thèses "d'exercice" dans les sciences de santé, sont reproduites sur microfiches.

Elles sont reproduites telles qu'elles sont reçues : il n'y a ni mise en page ni nouvelle frappe. C'est pourquoi il est nécessaire de veiller à la bonne présentation comme à l'absence de fautes.

Pour vous éviter une lecture de ces normes, nous avons inclu dans les feuilles de styles que nous vous proposons les critères indispensables de lisibilité et de mise en page que nécessitent une bonne reproduction et une bonne reliure.

Pour vous aider à mieux organiser l'information à rédiger, nous avons essayé de constituer le plan de thèse idéal. N'importe quel lecteur doit comprendre votre démarche scientifique d'un simple coup d'œil sur le plan. La structuration de votre travail reflète la structuration de votre pensée, il est donc très important qu'elle soit cohérente. Nous vous exposons en annexe, un modèle d'agencement des différentes parties d'une thèse.

La diffusion

Une bonne valorisation des travaux de recherche passe évidemment par une bonne diffusion. Le projet CITHER (Consultation en texte Intégral des THèses En Réseau) vous offre la possibilité d'une diffusion optimale, puisque planétaire, par le biais d'Internet. En plaçant votre thèse sur le Net, vous lui offrez le public le plus large possible.

De même, en exposant votre travail aux yeux de la communauté scientifique internationale, vous assurez votre promotion et celle de votre laboratoire.

Diffuser votre travail sous forme électronique vous permet de plus l'utilisation du multimédia et l'inclusion de séquences vidéo, audio, ou de photographies hautes définitions, apportant un complément d'information non négligeable.

Enfin, en conservant une version électronique de votre thèse, nous assurons une plus grande pérennité à l'information scientifique que vous nous confiez.

Annexe : Le plan type d'une thèse¹²

A. Préliminaires

1. Couverture de couleur claire
2. Page de titre, identique à la couverture
 - 2.1 Numéro d'ordre
 - 2.2 Année de soutenance
 - 2.3 Etablissement de soutenance
 - 2.4 Grade postulé
 - 2.5 Nom de l'arrêté (Lyon1), nom des formations et écoles doctorales, ainsi que la spécialité (INSA Lyon)
 - 2.6 Nom de l'auteur
 - 2.7 Grade précédent la soutenance de la thèse (INSA Lyon)
 - 2.8 Titre de la thèse
 - 2.9 Date de soutenance
 - 2.10 Composition du jury
3. Liste des professeurs et formations doctorales pour l'INSA, présidence et directeurs d'UFR pour l'UCBL (facultatif).
4. Remerciements, dédicace, avant propos
5. Résumés longs (français, anglais, toute autre langue)
6. Table des matières ou sommaire
7. Listes des illustrations et des tableaux
8. Listes des abréviations et des symboles
9. Glossaire (facultatif)

B. Texte principal

1. Texte principal structuré en parties, chapitres, sections, ...
2. Références bibliographiques

C. Annexes

Les annexes sont placées après le texte principal de la thèse.

D. Parties finales

1. Index
2. Pages 3 de couverture vierge
3. Quatrième de couverture
 - 4.1 Titre de la thèse en français
 - 4.2 Résumé en français
 - 4.3 Mots-clés en français
 - 4.4 Titre de la thèse en anglais
 - 4.5 Résumé en anglais
 - 4.6 Mots-clés en anglais
 - 4.7 Discipline
 - 4.8 Intitulé et adresse du laboratoire

Rappelons que la pagination (même sur plusieurs volumes) est continue et commence à la page de titre (= p.1).

¹² Ce plan s'inspire des préconisations du Ministère de l'Education Nationale, de la Recherche et de la Technologie, reprises par l'Agence Bibliographique de l'Enseignement Supérieur, et des pratiques de l'Université Claude Bernard Lyon1 et de l'INSA de Lyon.

4

Mise en page que prendra le guide de rédaction des thèses :

Page de gauche : exemple

Page de droite : explication typographique

5

Modèle de feuilles de styles "classique"

6

Modèle de feuilles de styles "basique"

7

Modèle de feuilles de styles "moderne"

8

Documents ajoutés aux dossiers d'inscription



Doc'INSA Centre de Documentation Scientifique et Technique

INFORMATIONS PRATIQUES :

Coordonnées :

Doc'INSA
Bâtiment 220
20 avenue Albert Einstein
69621 Villeurbanne Cedex (France)

Tél. (+33) 04 72 43 81 40

Fax (+33) 04 72 43 85 02

e-mail : doc@insa-lyon.fr

Web : <http://csidoc.insa-lyon.fr/>

Horaires d'ouverture : De 8h30 à 19h00 du lundi au vendredi et de 8h30 à 12h00 le samedi en période scolaire. Se renseigner à l'accueil pour les périodes de vacances.

LES UTILISATEURS de DOC'INSA :

Doc'INSA est accessible :

- aux élèves ingénieurs de l'INSA et aux étudiants de l'UCBL.
- aux enseignants et chercheurs de l'INSA et de tous les établissements universitaires de la région.
- aux ingénieurs des entreprises

LES SERVICES DE Doc'INSA :

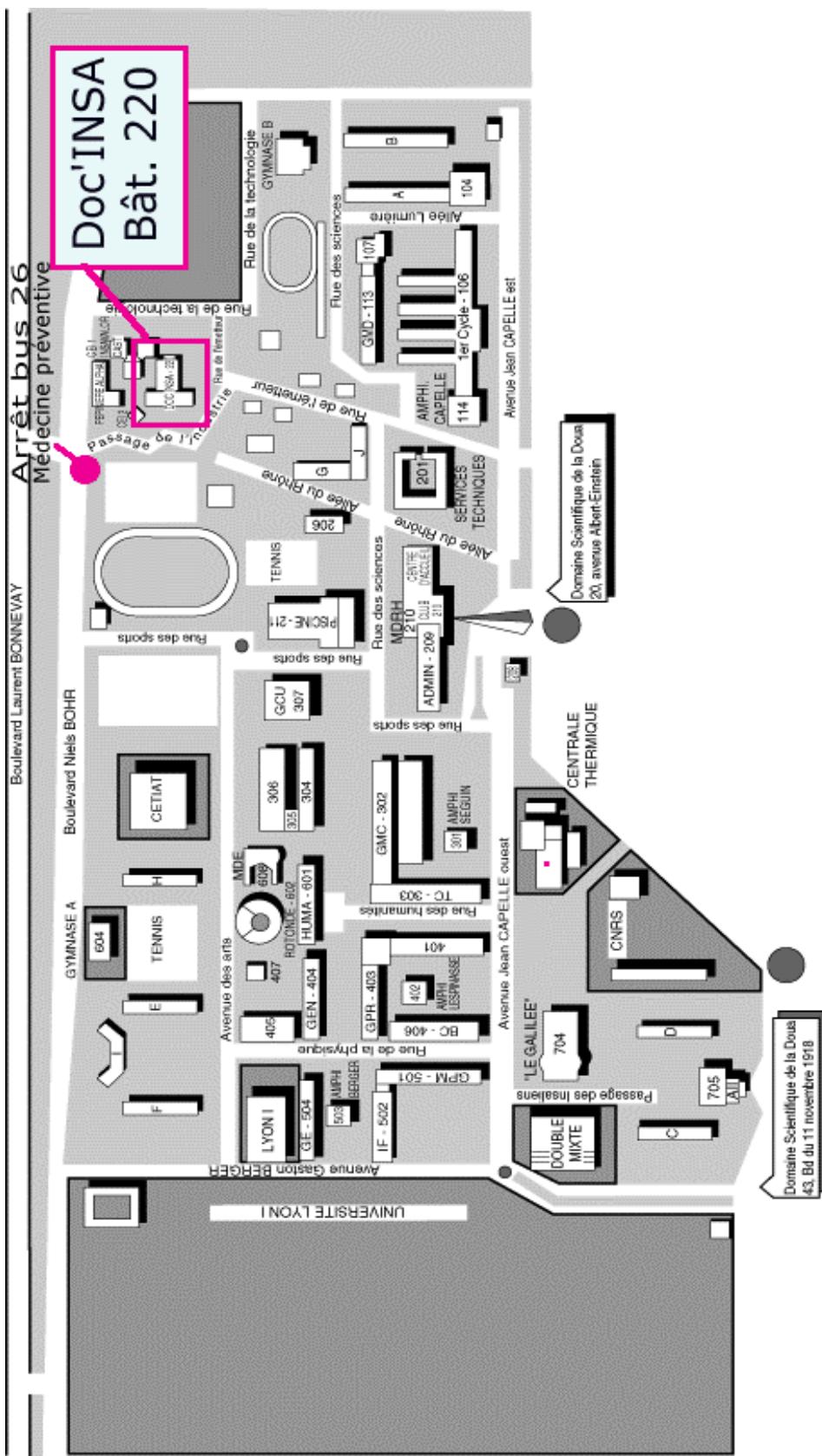
Doc'INSA vous propose :

- la lecture sur place
- le prêt à domicile (pour 15 jours renouvelables, jusqu'à 5 ouvrages simultanément par personne pour le personnel et les étudiants INSA ; jusqu'à 3 ouvrages simultanément pour les étudiants extérieurs)
- le prêt interbibliothèques
- l'accès en intranet à une trentaine de banques de données scientifiques, techniques et économiques.
- un service de veille technologique pour les entreprises
- la Consultation en texte Intégral des Thèses en Réseau (CITHER), assurant ainsi par une mise en valeur des thèses de l'INSA sur Internet, la promotion du jeune chercheur et de son laboratoire.

Un peu d'histoire :

En 1914, le principal site de radio français est celui de la Tour Eiffel à Paris. Devant l'imminence d'un conflit avec l'Allemagne, le Gouvernement français décide la mise en place d'un deuxième grand site émetteur afin d'assurer les communications radio françaises, notamment avec les alliés russes. La ville de Lyon, éloignée du front, est retenue pour l'installation de ce site radio. Et c'est ainsi qu'en 1916, les bâtiments actuels de Doc'INSA sont construits et abriteront des émetteurs de haute technologie.

Localisation de Doc'INSA





CITHER

Consultation en texte Intégral des Thèses en Réseau

CITHER est un nouveau service proposé par Doc'INSA. Il assure, par une **mise en valeur des thèses** de l'INSA sur **Internet**, la **promotion du jeune chercheur** et de son laboratoire. La diffusion sous forme électronique permet de plus, l'utilisation du **multimédia** et l'inclusion de séquences vidéo, audio, ou de photographies haute définition, apportant un complément d'information non négligeable.

Pour cela, avec l'accord du docteur, une version électronique de la thèse est déposée à Doc'INSA après la soutenance. Pour que nous puissions traiter sans problème votre travail de recherche, il est indispensable que la version remise possède certains critères informatiques.

Pour vous aider à appliquer ces particularités informatiques sans difficulté et vous préparer à votre future carrière de diffuseur de l'information scientifique, une **formation** vous sera proposée. Son but est de vous former à la rédaction de documents scientifiques structurés, notamment par l'apprentissage des feuilles de styles. Ces dernières vous seront remises sous forme électronique accompagnées d'un guide de rédaction pour les thèses.

Le projet CITHER est avant tout, une formidable opportunité de mieux faire connaître vos travaux de recherche.

Pour plus de renseignements sur CITHER et pour voir les thèses déjà diffusées, rendez-vous à :

<http://csidoc.insa-lyon.fr/these/index.html> ou écrivez-nous à : cither@insa-lyon.fr