



HAL
open science

La Lettre de l'URFIST de Toulouse, n°29-30 - 2002

Jo Link-Pezet, Élisabeth Lacombe-Carraud, Carole Roger, Mounir Meddeb,
Colm Mannin, François Magnan

► **To cite this version:**

Jo Link-Pezet, Élisabeth Lacombe-Carraud, Carole Roger, Mounir Meddeb, Colm Mannin, et al.. La Lettre de l'URFIST de Toulouse, n°29-30 - 2002. 2002. sic_03250603

HAL Id: sic_03250603

https://archivesic.ccsd.cnrs.fr/sic_03250603v1

Submitted on 4 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La Lettre de l'UR *f*IST de Toulouse

N° 29-30

EDITORIAL

Et toujours un même souci ..., ce même défi....., ce même enjeu..... têtue : accéder à l'information (pertinente ...de préférence)? Comment surtout parvenir à maîtriser la surabondance de l'information? Comment accéder à une information de qualité et en faire bon usage dans des contextes différenciés?

Encore et toujours un seul objectif, essentiel. Il faut mettre l'accent sur la nécessité de ne pas fournir de simples données, une information vaguement structurée. Il s'agit de créer une véritable information permettant d'augmenter à la fois la connaissance humaine et surtout la compréhension que nous avons des choses et ce, dans différents environnements, qu'il s'agisse d'informations pour apprendre ou pour agir. Tous les moyens sont bons pour vous guider : les portails, les filtrages divers, les réseaux de préférence, les systèmes agents. Tous ces moyens participent d'une même envie : permettre à l'utilisateur, vous permettre d'atteindre une information pertinente (pour lui, pour vous). Cela nous ramène au caractère subjectif du sens et au problème majeur, mais toujours entier de la représentation du sens. Cela nous ramène à l'interaction nécessaire entre le travail d'analyse qui sert à décrire un contenu (et sa valeur). Cela reflète des connaissances mobilisées à travers leurs connexions à des concepts, dont la signification spécifique l'usage et le contexte d'exploitation des contenus. Cela revient à dire que je suis en permanence en train d'ajuster ce que je sais, la conscience que j'ai de ce que je peux (veux, sais) faire, ce que je comprends. C'est en principe un processus sans fin de transformation des connaissances : on vous en parle depuis un bon moment et en principe, cette transformation est irréversible. Il est difficile de revenir en arrière.

Eh bien! à cette heure-ci, je doute. Dans un état traumatique, je ne comprends plus grand chose. Pour tout vous dire, j'ai l'"honte-au-logis" en ville et ce simultanément dans les quatre dimensions et les cinq (voire seize) niveaux. Les fidèles lecteurs comprendront le jeu de mot... Sinon, voir la note de bas de page.

La modernisation des bibliothèques restera-t-elle un sujet d'actualité? quel espace et quels temps? Pour nous rassurer, dans cette lettre vous allez trouver nos rubriques habituelles autour de ces points essentiels qui gravitent autour de la formation et l'accès à l'usage de l'information documentaire et au sens.

A vous amis lecteurs, bonne lecture, bon courage, à bientôt.

On parle d'ontologie pour désigner une description formelle explicite (concepts et leurs relations) reflétant la

vision d'un groupe de pratique, un modèle abstrait exprimant la façon dont les gens considèrent les choses dans un champ particulier. Nous avons développé dans le [projet FoRSIC](#) une ontologie concernant le domaine de la recherche documentaire, qui comporte quatre dimensions et cinq niveaux.

Jo Link-Pezet / Elisabeth Lacombe-Carraud

Charte du formateur à la recherche documentaire

Les formateurs ayant suivi la formation de formateurs de ces trois dernières années à l'URFIST ont souhaité terminer le cycle en élaborant une charte du formateur à la recherche documentaire dont nous vous communiquons ici le texte. Cette charte est destinée à faciliter la mise en place des formations sur les campus en fixant une base de négociations entre les différents partenaires de la formation (formateurs et institutionnels).

Les personnes ayant participé à l'élaboration de la charte sont : Solange Cossou et Anne Delmas, (UTM), Yann Foucaud, Evelyne Gerbedoen et Catherine Lopez (UT1), Francis Pelata (UPS), Christine Vidal (Ecole Vétérinaire), Elisabeth Lacombe et Jo Link-Pezet (URFIST), Olivier Nunge (Croissance Formation). Ont participé à la critique et la relecture finale : Nicole Doan (CNRS), Christel Thébaud et Sophie Cancel (UT1), Claire Charles (IUT Blagnac).

Le rôle de la charte

Le but de cette charte est de définir la dimension éthique et stratégique de la formation à la recherche documentaire afin de clarifier les conditions de mise en œuvre et de réalisation des formations (organisation et contenus) et de faciliter les négociations avec les personnels chargés de formation des usagers des différents campus.

Ce document peut être considéré comme le résultat de différentes expériences analysées et comparées et doit permettre d'établir des règles afin de clarifier les différentes situations existantes (tant du point de vue de la demande que du point de vue de la réalisation). Il constitue une force de propositions à négocier avec les différents interlocuteurs de la formation documentaire en particulier avec la hiérarchie directe du formateur.

Le but corollaire poursuivi est la validation du travail réalisé par le collectif de formateurs au cours de leurs séances de travail et de réflexion, ainsi que l'identification d'un pôle d'enseignement à la recherche documentaire destiné aux usagers des universités.

Les points suivants vont être développés dans cette charte :

- identité du formateur à la recherche documentaire,
- mission du formateur,
- moyens, conditions matérielles et techniques d'enseignement,

- mise en œuvre et reconnaissance des enseignements,
- contrat administratif,
- contrat pédagogique,
- engagements du formateur.

Identité du formateur à la recherche documentaire

Il s'agit d'un membre du personnel d'un établissement d'enseignement supérieur et/ou de la recherche :

- ayant des connaissances théoriques et pratiques acquises dans le domaine de la documentation et ayant suivi ou s'engageant à suivre une formation de formateurs spécifique,
- dispensant des formations à des groupes d'étudiants, dans le cadre de cursus ou en réponse à des demandes ponctuelles d'enseignement.

Mission du formateur

Le formateur doit amener l'étudiant à maîtriser les outils et les méthodes de l'information et de la documentation pour l'aider à réussir dans sa vie universitaire. Il aura pour souci principal de rendre progressivement l'apprenant autonome.

Les conditions de mise en œuvre de cet enseignement doivent être déterminées dans un contrat avec les partenaires (institution, administration, enseignants des disciplines, autres formateurs...) pour définir :

- l'information dont l'apprenant a besoin,
- le niveau d'information à enseigner,
- les conditions optimales de la formation.

Moyens, conditions matérielles et techniques d'enseignement

La formation à la recherche documentaire repose sur une spécificité : l'étudiant doit acquérir la connaissance des lieux, des ressources et de leur organisation interne, mais aussi pouvoir manipuler les outils d'information pour effectuer les recherches nécessaires à son apprentissage.

En plus du temps indispensable à la préparation de toute formation, le formateur devra disposer de salles équipées et opérationnelles comportant :

- le matériel informatique en nombre suffisant,
- l'accès aux réseaux,

- le matériel audio-visuel,
- les mots de passe...

Cela nécessite une organisation préalable importante et des contacts avec de nombreux interlocuteurs pour une mise en place efficace de l'enseignement. Le nombre d'étudiants (minimal/maximal) devra être clairement défini pour chaque formation.

Mise en œuvre et reconnaissance des enseignements

Pour satisfaire et prendre en compte la demande des prescripteurs de formation, un certain nombre de paramètres et règles de fonctionnement devront être préalablement définis :

- intégration dans le cursus, nombre d'heures, type de cours (cm, tp, td),
- relation avec les autres enseignements, moment de l'année, durée de la formation,
- règle d'assiduité : inscription, feuille de présence,
- sanction : notation, contrôle,
- en ce qui concerne, plus directement le formateur, temps de préparation de la formation (concertation et réalisation de supports pédagogiques).

Ceci nécessite un double contrat : administratif et pédagogique.

Contrat administratif

Les modalités du contrat administratif de formation doivent être définies entre le formateur et l'institution dont il dépend. La mission du formateur doit être confirmée par sa hiérarchie (Service Commun de la Documentation ou direction de la formation des étudiants) afin de définir la place de cette mission par rapport au statut d'origine du formateur et à ses autres activités. Le SCD ou la direction de la formation des étudiants mandate le formateur pour mettre en œuvre ces formations.

La définition du contenu et du processus de ces formations se fait entre les différents partenaires du point de vue :

- de l'organisation et de la planification,
- de la logistique,
- de la communication (information des étudiants),
- des modalités de contrôle,
- de la définition des contenus d'enseignements.

Contrat pédagogique

L'enseignement de la recherche documentaire repose sur les aspects suivants :

- enseigner une méthodologie pour atteindre un résultat : il faut définir le but à atteindre, la tâche à réaliser,
- enseigner à l'apprenant à identifier les supports, les lieux d'information et lui apprendre à les utiliser,
- enseigner à :
 - rechercher l'information, l'analyser, l'exploiter,
 - utiliser efficacement les outils et l'information obtenue,
- enseigner par niveau d'apprentissage: objectifs de formation différenciés (3 niveaux : culture de la recherche documentaire, maîtrise d'un champ disciplinaire, utilisation experte),
- évaluer, pérenniser les enseignements.

Le contrat pédagogique est un contrat ouvert, qui nécessite cependant un partenariat avec les acteurs de la formation et un engagement personnel, en particulier de la part du formateur à la recherche documentaire.

→ Recherche

Journées du CELSA et de la SFIC

Jo Link-Pezet, Maître de conférences - URFIST de Toulouse

Les 18 et 19 Janvier se sont tenues au Celsa (Ecole des Hautes Etudes en sciences de l'information et de la communication-Paris-Sorbonne) les journées doctorales de la SFIC. La journée du 18 a été consacrée à une réflexion sur la recherche. G. Jacquimont a parlé de transdisciplinarité et de la nécessité de recontextualiser les concepts, Y. Reuter de l'écriture de la thèse en tant que convergence pratique et théorique pour la construction de connaissances. La matinée s'est achevée sur la présentation d'une étude sémiotique des rapports de thèse. L'après-midi a été consacré à des ateliers sur différents thèmes. Parallèlement à la mise en commun des résultats des ateliers, R. Boure et Y. Polity ont parlé de l'existence des laboratoires de recherche d'un point de vue historique et ont tenté de répondre à la question suivante: "qu'est-ce qu'un laboratoire de recherche?" et se sont interrogés sur la stratification sociale dans l'organisation scientifique.

Pour trouver un lien entre ces différentes interventions et les modèles théoriques sur lesquels reposent les sciences de l'information et de la communication, nous reprenons sous la forme d'un tableau récapitulatif (obligatoirement réducteur) les projets de recherche de trente doctorants présentés le 19. Les travaux présentés par ces jeunes chercheurs s'articulent autour d'un questionnement exprimé par Y. Jeannerey (organisateur des journées): les nouvelles formes de sociabilité sont-elles porteuses de sens? Définissent-elles des espaces de pouvoir symbolique, quelle représentation, quelle valeur?

Session A:	Investissements, projets, contextes, observer l'usage des TICs	Mots clés	Démarche	Théories
Y. Andonova Dir : N. d'Almeida	Les Tices au quotidien: atouts et limites d'une approche des usages en entreprise	Entreprise/usages Travail coopératif à distance	Questionnaires, enquête	Paradigme de la complexité Approche communicationnelle Coopération
H. Bourdeloie Dir : J. Jouët	Pratique des multimedia culturels dans la sphère domestique	Cadre de référence et pratiques culturelles Cédérom culturel Modes d'appropriation des medias interculturels et conditions sociales	Enquête, entretien qualitatif	Sociologie des usages et sociologie de la culture: usage, pratique, utilisation, consommation, appropriation P. Bourdieu, O. Donnat
A.H Le Cornec Ubertini Dir : R. Nadot	Internet et la démocratie. Les réseaux tels Internet sont-ils des outils susceptibles de renforcer la démocratie.	Lien science et société, sciences et sciences Rôle de l'état dans la fête internet, évolution des systèmes en entreprise	Expérimentations,	Approche plurielle Sources athéniennes des fondements de la démocratie Analyse du système politique Stratégie d'appropriation du pouvoir

C. Lukasiewicz Dir : Y.Jeanneret et H. Touzard.	L'identité, le territoire et Internet: approche interdisciplinaire de la relation entre territoire virtuel et identité réelle	Immaturité Auto-focus	Enquête, sociométrie, entretiens, : 50 questions (14 ouvertes)	Pluridisciplinaire
C. Martin Dir : J. Walter	Les représentations du téléphone portable dans les familles	Le portable accompagne l'évolution des familles vers l'autonomisation	Position cycle de vie et sexe Echantillon de 25 familles, entretiens	Représentations et signification sociale Sociologie de l'innovation (rejet du déterminisme classique) et imaginaire social, créativité. Sociologie de la famille : caractérisation des dynamiques familiales
D.Omrane Dir : R. Boure, MC. Monnoyer-Smith	Les enjeux relationnels d'une communication électronique entre les PME-PMI et leurs clients	Proximité relationnelle (communication)	Définition du paradigme relationnel de la communication	Anthropologie de la communication Appropriation des outils Pratiques communicationnelles Communication médiatisée par ordinateur
N. Pignard Dir : I. Pailliar	Les évolutions de la communication scientifique liée au développement d' Internet	Modes de publication de travaux de recherche		Brouillage de frontières Nouvelles formes de publication Rôle actuel de la revue (offre et pratiques) Stratégies des acteurs

Session B	Supports, langages et Interprétation : analyser les medias et la production de sens	Mots clés	Démarche	Théories
B. Amougou-Boone Dir : Y. Jeanneret	La (re)présentation des œuvres d'art sur Internet	Création artistique Internet Numérique Multimedia Nouvelles formes d'expression de publication	L'art et connaissance, espace de médiation; art et logique de communication	Représentations de l'art Analyse du discours (texte et métatexte)
M. Antunes Da Cunha Dir : Rémy Refel	Etude de la réception / Radio Télévision Portugaise Internationale auprès de la communauté portugaise	Ré-articulation de la réception et de la citoyenneté	Interviews Courier des spectateurs	Th. de la réception Théorie de l'immigration Ressources symboliques et cognitives
F. Bousquet Dir : G. Chapouillé	Cinéma Hollywoodien et sentiment national américain: choisir et analyser les films.	Industrie cinématographique	Sentiment national, systèmes de représentation	Analyse des images et contenus
C. Deprez Dir : Roger Odin et K. Feigelson	La télévision indienne : un modèle d'appropriation culturelle	Secteur audio-visuel indien : identité de l'Inde et mondialisation dans le domaine des mass media		Enquête sociologique fondée sur des questionnaires
F. Kristov Dir : Pascal Sanson	Archimed : Architecture et media	Monuments en tant que source d'information des us et coutumes des sociétés passées.	Relation à internet	Vocabulaire architectonique courant

I. Jura Dir : F. Jost	Représentations d'entreprise sur Internet : ces écrans à voir	Analyse modélisante des écrans observation	Grilles sémiotiques	Communication d'entreprise Ergonomie
C. Lenoir Dir:: R. Odin et C. Duchet	Convergence des medias, divergence des discours : Télévision et intégration des médias	Particularités communicationnelles de la télévision	Analyse idéologique Etude de la récurrence des discours	Perspective esthétique Emission télévisuelle en tant que point de départ d'une combinatoire
C. Recio Davila Dir : B. Lamizet	Le problème du temps dans la photographie	Temps du réel, de la réalisation, du regard	Esthétique, technique, culture	Barthes, Freund, Sontag
C. Roques Dir : P. Delcambre et Y. Jeanneret	La représentation d'Internet dans le discours de la télévision	Lien objet/media Relation à un objet technique nouveau	Exploration du terrain audio-visuel autour d'Internet	Dimension historique, diachronique, intermédiaire

Session C	Design informationnel et projets informatiques : comprendre et/ou modéliser les interactions autour de l'interactivité	Mots clés	Démarche	Théories
G.Bezon Dir : J. Walter	Représentations et processus de communication dans la conception de cédéroms éducatifs	Nouveaux dispositifs multimedia, auto-formation	Conception, élaboration, représentation d'un objet	Sociologie de l'innovation : négociations entre acteurs et action collective, différentes logiques en présence (éducative, technique, esthétique, communicationnelle)

N. Devos Dir : Y. Jeanneret	Evaluation des documents multimedia en ligne. Mise en concurrence des modèles : paradoxe et nécessité	Contexte éditorial	Evaluation	Ergonomie, évaluation
P. Jullia Dir : A. Mucchielli	Sites Web et processus de fidélisation des publics	Aide à la fidélisation des publics Typologie des sites : nature du document dispositif conversationnel et système d'édition	Enquête	Recherche fondamentale et appliquée Théorie de la complexité
A. Poulon Dir : A. Mucchielli	Processus d'émergence du sens dans les utilisations d'un système informatique de gestion documentaire	Idéologie des TIC Recherche et diffusion des informations scientifiques par les dispositifs informatiques	Consultation d'un service d'archives Méthodologie qualitative	Complexité: Morin, optique situationnelle Boîte à outils théoriques et méthodologique + analyse systémique et sémio-contextuelle
F. Renucci Dir : Philippe Dumas	Le spectateur et le film interactif face à la continuité de son histoire, au rythme de son énonciation et de l'harmonie de sa composition	Image numérique, notions de rupture et interactivité	La rupture provoquée par l'interactivité entre statut de spectateur et d'acteur	Rythme, dramaturgie, contexte global de la médiation Mabillot, Barboza, Weissberg, Balpe, Boissier, Genette, Benveniste
S. Vaillies Dir : A. Mucchielli	Apports de l'approche compréhensive aux méthodes de conduite des projets informatiques	Réorganisation du travail, des relations, du pouvoir	Enquête	Approche de la complexité: Morin Approche situationnelle et communicationnelle

L. Verclytte Dir : S. Leleu-Merviel	Vers des documents sur-mesure au moment opportun de l'approche compréhensive: application au document d'apprentissage	Développement des TIC à fin pédagogique Adéquation au besoin Critères de typage de documents : interactivité, interactionnalité, générativité	Documents pro-actifs Anticipation des attentes-besoins	Classification des états cognitifs et conatifs (besoins), styles d'apprentissage, profilage à but pédagogique
---	---	--	---	---

Session D	Médiations, transformations, incarnations: suivre les processus d'institution des ordres culturels	Mots clés	Démarche	Théories
J. Bonaccorsi Dir : D. Jacobi	Le devoir-lire: homogénéités, appropriations et métamorphoses du discours d'intentions dans les pratiques et dispositifs culturels de la diffusion de la lecture	Le livre et catégories de lecteurs et la ville Le réseau des bibliothèques Action culturelle	Textes de littérature grise, colloques professionnels, observations, entretiens	Politique de lecture Relation entre espace, territoire, place du livre, discours Analyse sémiotique
M.C. Bordeaux Dir : J. Davallon	L'art de la médiation : la médiation culturelle dans les arts de la scène	L'œuvre d'art : liens croisés entre production, mise en objet et réception	Récit des acteurs et observation de terrain,	Analyse des arts de la scène Description des genres
A. Dahlke Dir : N. D'Almeida	Identités et images d'entreprise Formes de médiation et impacts sur les représentations	Image et représentations : sociale, commerciale, institutionnelle	Analyse des pratiques de communication	Modèle d'analyse "accordéon relationnel-cristallin"

P. Fortin Dir : R. Reffel	Gouvernance de l'Internet: de quoi parle-t-on exactement?	La notion de mondialisation Gouvernance multi-niveaux (infrastructure, contenus, accès)	Analyse critique de la compréhension du développement de l'Internet	
S. Labelle Dir : Y. Jeanneret	Le rôle des technologies de l'information dans la redéfinition de la matérialité urbaine	Modalités du développement des médias informatisés dans la ville Démocratie locale Espace culturel multimedia	corpus de discours Pratiques des acteurs Dispositifs Analyse des condition de légitimation des discours	Analyse des processus Mise en relation de la sémiologie et de l'ethnologie
D. Miège Dir : D. Jacobi	Enjeux et modes de présence des artistes dans les médiations de l'art contemporain	Art contemporain, participation des artistes aux médiations de l'art contemporain (différentes formes)	Entretiens semi-directifs	Analyse des formes d'intervention des artistes dans les actes de médiation
J. Oheix Dir : P. Sanson	Modes de fonctionnement et paradoxes communicationnels des manifestations festives en marge de la légalité (free parties)	Etude des graffiti dans les structures urbaines, cultures populaires Les formes du détournement : les free parties	Analyse de discours croisés	Bricolages identitaires Notions de déviance et de marges
K. Tinat Dir : J.J. Boutaud	Vers une approche anthropologique de la communication d'un groupe juvénile urbain : les pijos de Madrid à la fin des années 1990	Identité	Anthropologie de la communication	Sociologie Interdisciplinarité : civilisation espagnole, Anthropologie, sociologie, Communication/sémiotique

Les sources en droit

Le 29 novembre 2001 a eu lieu à l'URFIST une réunion des juristes d'entreprises à l'initiative de Mme Peruzzetto, professeur à l'université des Sciences Sociales de Toulouse et de Mme Roger, déléguée régionale de l'Association des Juristes d'Entreprises. Il s'agissait de faire une

présentation des sites juridiques intéressants pour ce public. Les sites ont été présentés par M. Mounir Meddeb (doctorant à UT1), Mme Carole Roger et M. Mannin (Airbus).

Le 25 mars 2002, une nouvelle réunion, à l'initiative de l'Association Française des Juristes d'Entreprises et du Toulouse International Lawyers Group, a rassemblé les juristes spécialistes de droit international. Voici une liste des sites présentés au cours de ces deux réunions.

SELECTION DE SITES GENERALISTES EN DROIT

Par Carole Roger

Le Village de la justice

Lieu de rencontre et d'échange entre les professionnels du Droit (avocats, juristes d'entreprises, notaires, experts, universitaires etc.).

Ce site propose des outils pratiques :

- Annonces, emplois
- Forum
- Actualités
- Annuaires
- Prestations de services à des tarifs préférentiels

Le journal du village propose l'Actualité du Droit, des dossiers juridiques avec le commentaire des lecteurs ainsi que des communiqués de presse.

Un espace de discussion est ouvert pour échanger des avis sur des sujets diversifiés, il s'agit des forums de discussion qui permettent de rechercher une information, donner un avis ou également trouver des précisions sur des traductions de termes juridiques.

Le journal du village est diffusé gratuitement par courrier électronique.

Site clair, convivial, facilité de navigation et riche en informations pratiques.

Juriguide

Annuaire des sites juridiques sur Internet.

C'est un service du village de la justice qui propose plus de 2000 sites classés et mis à jour quotidiennement. Catégories :

- Culture juridique
- Documentation juridique
- Edition juridique
- Formation
- Emploi

- Fournisseurs des juristes
- Organismes professionnels
- Juristes
- Gestion professionnelle
- Annuaire de sites juridiques
- Association et réseaux.

Un site incontournable, d'une grande qualité informative.

Juriforum

Site dédié aux professionnels du droit.

Juriforum propose :

- Une information juridique quotidienne :
- actualités juridiques (accessible gratuitement)
- actualités financières
- articles et chroniques
- la Tribune Libre (accessible gratuitement).
- Un annuaire de professionnels
- Une liste de fournisseurs
- Une bibliothèque : lien vers Lamyline
- Une liste des événements importants.

SELECTION DE SITES THEMATIQUES EN DROIT

Par Mounir Meddeb

I - DROIT DES AFFAIRES

1 - Site de l'association européenne de droit bancaire et boursier-France (AEDBF)

- Une rubrique actualités avec surtout une revue de presse de toute la bibliographie récente.
- Les publications de l'AEDBF Europe, France, Belgique et Luxembourg.
- La base de connaissances est très intéressante avec toutes les références en droit bancaire et financier depuis 1997. L'accès se fait par rubrique avec possibilité de recherche. Le problème est que le moteur de recherche est souvent indisponible.

2 - Site du Centre de recherche sur le droit des affaires (CREDA).

- Rubrique études : possible consultation des sommaires, index et bibliographies d'ouvrages parus dans le cadre de la CCIP. Possibilité de les commander.
- Consultation des colloques précédents et même possibilité de télécharger la plupart des

interventions.

- L'atout du site est la liste de diffusion de toute l'actualité soit en droit de la concurrence, soit en droit des sociétés soit en droit des procédures collectives.

3 - Site de droit fiscal.

- Site conçu par un fiscaliste toulousain.
- Une rubrique pratique avec différents domaines et possibilité de choix entre fiscalité interne, communautaire ou international (certains liens sont inopérants).
- Une rubrique théorie et une rubrique histoire assez complètes.
- Possible demande de consultation en ligne (service payant), de même le site propose un service d'études et de recherche documentaire payant.

4 - Site Juris International

- Offre l'accès à deux séries de documents : collections générales et collections spécialisées.
- Les collections générales donnent accès à des instruments juridiques en fonction de mots clés, à des modèles de contrats, des renseignements relatifs aux centres d'arbitrage, de médiation... ainsi qu'aux associations de juristes d'affaires de part de monde.
- Les collections spécialisées concernent trois thèmes : l'OMC, l'OHADA et le commerce électronique. Avec un accès aux instruments, documentation...

II - N.T.I.C.

1 - Site de l'autorité de régulation des télécommunications (ART)

- Site très complet qui offre accès à toutes les données en la matière.
- Rubrique textes de référence tant sur le plan national que communautaire.
- Rubrique Grands dossiers : boucle locale, UMTS... avec tous les textes concernant la question.
- La plupart des publications de l'ART à télécharger ou commander.

2 - Site de la direction générale société de l'information

- Site principalement en anglais avec accès aux thèmes de travail de la Direction.
- Possible accès aux archives de la direction.

3 - Site en matière de droit et technologies.

- Site très bien fait et très complet
- Les trois rubriques les plus intéressantes sont : Dossiers, législation et jurisprudence.
- Dossiers (59) avec accès par date ou mots clés.
- Législation avec un accès thématique qui couvre 5 zones : Belgique, France, Luxembourg, Europe et international.
- La jurisprudence couvre la Belgique, la France et les Etats-Unis.

4 - Site du programme d'action gouvernemental pour la société de l'information.

- Propose les textes de références les rapports ainsi que les discours officiels.

- Propose aussi un accès aux appels d'offres et à candidatures.
- Une rubrique spécifique est consacrée au commerce électronique et à la cryptologie.

5 - Site de la revue du droit des technologies de l'information

- Les lois et les arrêts les plus récents
- Les articles professionnels ou de doctrine à consulter en ligne
- Une rubrique actualité
- Par l'intermédiaire de la rubrique universitaire accès à une base d'articles, d'études et de rapport de recherche. De même pour la rubrique professionnels.

6 - Site sur le droit de l'informatique

- Site constitué d'une importante base de données très fournie en matière de NTIC
- Base plutôt doctrinale.
- Sont couverts : le droit de l'informatique, du multimédia, des réseaux...

III - Propriété intellectuelle et industrielle :

1 - Site de l'Office Européen des brevets ou <http://www.epo.co.at>

- Intéressant surtout parce qu'il aborde à la fois les aspects pratiques et les aspects recherches.
- Il propose dans la rubrique outils pour les déposants : les formulaires de dépôts, les logiciels de dépôts, les renseignements pour le paiement des taxes...
- Il propose aussi (surtout dans la rubrique Index et recherche) des textes de référence, la jurisprudence des chambres de recours, le droit national relatif à la Convention sur le brevet européen.

2 - Site legalis

- Site très intéressant en matière de droit d'auteur et internet.
- Rubrique jurisprudence très actuelle et très complète.
- Le plus intéressant c'est la rubrique contrats où on peut consulter des modèles de contrats (d'édition, licences d'exploitation...).
- A voir aussi le site apparenté : www.legalis.net/legalnet
- Possibilité d'abonnement (gratuit) à une lettre d'informations.

3 - Site de l'organisation mondiale de la propriété

Trois grandes rubriques :

- droits de propriété intellectuelle.
- rubrique informations avec l'accès au texte de lois, conférences, presses et divers documents.
- activités et services avec l'accès à un centre d'arbitrage, des répertoires de dessins et de modèles internationaux, de marques internationales...

IV - Jurisprudence, lois et codes

1. Les classiques :

- [Legifrance](#)
- [Jurisfrance](#), sur abonnement.
- [Jurisguide](#) : annuaire.
- [Curia](#)

2 - Site sur le droit aérien

- Présentation du guide du droit aérien, du guide pratique du pilote, du précis juridique du pilote de ligne...
- Une rubrique actualité
- Des dossiers spéciaux
- Pour la validité du contenu, avis aux spécialistes.

3 - Site des éditions Dalloz

- Une bonne partie du site est réservée aux abonnés.
- Le site reste intéressant dans la partie gratuite surtout pour les rubriques actualités et la rubrique emploi de juristes.
- L'abonnement est de 25 euros HT/mois.

4 - Site des éditions Jurisclasseurs.

- Rubrique dépêches (à recevoir par mail).
- La sélection du jour de la jurisprudence. Pour consulter les archives de la jurisprudence, un abonnement est nécessaire.
- Consultation de la une des revues "jurisclasseurs".

5 - Annuaire juridique francophone.

- Annuaire très complet et très intéressant
- Accès soit par rubrique, soit par thème, soit par mot-clés.
- Une rubrique emploi pour juristes.
- Inconvénient : généralement lent à se lancer et entre les pages.

6 - Portail de veille juridique

- Ce site se déclare comme le premier portail de veille juridique professionnel.
- Les rubriques les plus intéressantes (outre l'actualité) sont :
- Chronique dans différents secteurs (fiscal, Internet, social...)
- Des dossiers sur des thèmes particuliers (Loi NRE, euro...)
- Jurisprudence surtout revue de la jurisprudence.
- Des articles de doctrine : assez éclectique.

SITES POUR PRATICIENS INTERNATIONAUX

par Carole Roger

FREELANG :

Freelang est un site qui permet de télécharger gratuitement des dizaines de dictionnaires de langues.

De plus un service d'aide à la traduction, animé par des bénévoles est proposé pour répondre à la traduction d'un mot, d'une phrase ou d'un texte.

Vous trouverez également de nombreuses informations concernant la météo internationale, les ambassades de France à l'étranger... Un site vraiment intéressant.

LE DICTIONNAIRE DU DROIT PRIVE :

Ce site propose des définitions juridiques en Droit Civil, Droit Commercial, Droit Social, Droit International Privé. Vous trouverez également le vocabulaire de la doctrine juridique ainsi que le langage pratique juridique.

Ce dictionnaire est complet et contient des définitions qui sont complétées par des textes explicatifs. Des liens renvoient sur d'autres sites.

LE GRAND DICTIONNAIRE :

Ce dictionnaire est un ouvrage qui permet de trouver rapidement un terme spécialisé, il se compose de plus de 3 millions de termes français et anglais dans 200 domaines d'activité. Grande facilité de navigation et précision de la recherche.

ECLA :

L'ECLA représente la confédération européenne des associations des juristes d'entreprise soit plus de 31 000 juristes.

Ce site permet de recueillir de nombreuses informations sur le métier de juriste d'entreprise avec un accès aux sites juridiques des pays membres.

Des sondages, des études comparatives du statut et de la reconnaissance du Juriste d'entreprise dans les différents pays. L'ECLA publie également des propositions et des prises de positions sur des thèmes diversifiés

DICORAMA ONELOOK :

Dicorama est un annuaire des dictionnaires qui permet de rechercher un mot dans plus de 670 dictionnaires.

Ce site permet de trouver un dictionnaire dans l'annuaire mais surtout de rechercher un mot avec le metadictionnaire Onelook, la recherche du mot est effectuée dans plus de 600 dictionnaires, OneLook contient un index sur tous les mots définis dans ces dictionnaires.

Une liste de l'ensemble des dictionnaires qui proposant une définition du mot vous est proposée et les résultats sont classés par thèmes.

LAWYERS.COM :

Site qui référence des avocats au niveau international avec une recherche par pays.

Possibilité de poser des questions et de recueillir des informations sur des lois et des sujets juridiques diversifiés.

CORPORATE. MARTINDALE :

Site qui propose des services juridiques : listes d'avocats, des résumés de lois de différents pays, des informations juridiques ... Inscription nécessaire pour avoir accès à ce site.

SITES EN ANGLAIS POUR PRATICIENS INTERNATIONAUX

par Colm Mannin

1 - The Legal Information Institute

Créé en 1992, ce premier site juridique sur Internet est le fruit de l'activité bénévole de l'Ecole de droit de Cornell et reçoit neuf millions d'appels par semaine (en provenance de plus de 70 pays). D'abord source d'informations sur l'Amérique, la section internationale donne de plus en plus de liens matériels vers des systèmes juridiques locaux de 40 nations européennes. Les responsables du Web du Legal Information Institute ont mis l'accent sur le caractère convivial de l'interface qui se reflète dans le découpage de sujets juridiques. Pour chaque sujet, qu'il s'agisse d'un contrat, d'une loi commerciale ou antitrust, ou de n'importe lequel des nombreux sujets couverts, un court résumé apparaît sur un côté de la page et de l'autre côté on trouve différentes rubriques ainsi que les liens vers les documents pertinents.

2 - World Legal Information institute

L'accès à une législation sélectionnée (cliquer sur View All Databases) ainsi que sur le cas juridique dans plusieurs juridictions du Commonwealth (cliquer sur Case Law). Les autres sections comme les traités et les publications sont d'un intérêt limité.

3 - Hieros Gamos

Contient des sections sur la loi américaine ainsi que les lois de 220 pays dans des domaines particuliers. Certains ne sont pas très intéressants comme les traités sur l'aviation.

4 - International Centre For Commercial Law

Réalisé par les éditeurs anglais Legalese. Ce site est caractérisé par différentes rubriques surtout intéressantes pour des praticiens britanniques. Il y a deux sections internationales, l'une donne la liste des cabinets juridiques dans 70 pays, et l'autre contient des articles et des caractéristiques des pratiques sélectionnées, ainsi que la liste des pays dans lesquels cette information est disponible. Ce site peut être utile pour trouver l'information de base des systèmes juridiques nationaux sur différents sujets comme les types d'entreprises et de régulations nationales, les agents commerciaux etc...

5 - International Bar Association

Le site Web de l'association internationale des avocats donne un accès en ligne à des journaux spécialisés qu'on peut télécharger. Site parfois payant.

6 - Findlaw

Un des sites les plus complets sur Internet. Spécialisé dans la loi américaine mais avec une section internationale importante. Il est relié à 50 000 sites (juridiques, thématiques, cours, universités, consultants, annuaires, publications, fournisseurs de logiciels). On y trouve des

exemplaires de contrat. Ce site présente deux défauts : il est trop riche et génère trop de liens (perte de temps possible) ; le contenu est parfois trop mince.

7 - [New Law Journal](#)

Site d'une publication américaine prestigieuse. Contient des informations très utiles (litiges, conseils domestiques, courrier électronique associé).

8 - [Lawlinks](#)

Décrite comme l'annuaire expert (formulaire de loi, articles par sujet, groupes de discussion, recherche d'avocats).

9 - [Infolaw](#)

Portail de ressources légales au Royaume Uni.

10 - [Institute of Advanced Legal Studies](#)

Site anglo-saxon donnant accès à une information juridique globale par pays et par type de loi. Ressources officielles et non officielles, organisation internationale et inter gouvernementale (Conseil de l'Europe OMC). Il comprend des informations à propos des institutions académiques, des annuaires et dictionnaires juridiques, des journaux en ligne.

11 - [Law & Legal Research](#)

Existe depuis 1998. -Annuaire indépendant fait par une personne. Liste sur les sources d'information en ligne ou commentaires ; moteurs de recherche juridiques.

12 - [International Law Office](#)

Annuaire de cabinets réputés pour sa lettre d'information (distribuée par e-mail avec souscription gratuite). Les abonnés sélectionnent les demandes qui les intéressent et ne reçoivent que l'information pertinente. Les contributions sont celles des cabinets juridiques les plus importants dans le monde.

Colloque de Compiègne

Jo Link-Pezet, maître de conférences

Nous avons assisté les 21 et 22 janvier 2002 au séminaire interdisciplinaire sur la cognition organisé par l'Université de Technologie de Compiègne qui se tenait du 21 au 25 Janvier sur le thème suivant : Technologies cognitives : technologies numériques, systèmes d'aide et normes. Le thème de ces journées fut présenté par C. Lenay qui a rappelé que la connaissance est devenue source principale de valeur de la société de l'information. Il a posé un certain nombre de questions fondamentales : "Comprend-on vraiment ce qu'on fait et les technologies de l'information et de la communication (dispositifs matériels, logiciels ou organisationnels participant à la production du sens par leur utilisateur) sont-elles cognitives ?" Comment analyser le rôle social et cognitif de ces mêmes techniques et répondre parallèlement à cette question : en quoi ces technologies réalisent-elles une industrialisation ou un marché des connaissances ? Ce séminaire au cours de cinq journées a tenté de répondre à ces interrogations en accueillant des

chercheurs de différents champs disciplinaires.

Nous allons rendre compte d'une communication qui nous a intéressés plus particulièrement au regard du travail que nous accomplissons dans le cadre de Forsic à savoir la mise en œuvre d'ontologies et d'un modèle de domaine (de la recherche documentaire) en XML, ainsi que la production semi-automatique de ressources pédagogiques par ré-ingénierie de ressources hétérogènes existantes.

"Le numérique et les conditions a priori de l'interprétation"

Bruno Bachimont (UTC)

Introduction

Le numérique impose de représenter formellement le sens pour paramétrer l'exploitation des contenus (interchangeabilité des contenus et intercompréhension) entre concrétisation interprétative et abstraction formelle et de : · penser l'usage de l'outil a priori; · fixer a priori l'intelligibilité des contenus (pas de manipulation directe; · permettre la pluralité interprétative de l'usage à partir d'une même forme syntaxique (ensemble d'usages). Autant d'étapes qui permettent la mise en place d'une industrie des contenus entraînant de fait une standardisation et une banalisation qui s'opposeraient à la création et à l'originalité.

I- Le paramétrage du contenu

1. 1- Le cadre traditionnel

Le traitement sur les contenus est indépendant de la signification des contenus. Le transport des informations (si on prend l'exemple de la voix) ne dépend pas de la signification des informations. La transformation des contenus qu'il s'agisse de compression ou de restauration, s'effectue indépendamment de la signification. On peut dire sans se tromper qu'il y a clivage entre le traitement du contenu, sa transmission et l'étude de son sens.

1.2- Un changement

Actuellement, des réseaux universels autorisent la présence de toutes sortes de données qui jusqu'alors étaient disséminées sur des réseaux distincts: voix, sons, images, textes, ce qui se traduit par la complexification des données sur un même réseau. Pour conserver une certaine qualité, le réseau doit savoir ce qu'il transmet afin d'optimiser sa transmission (routeurs intelligents) et adapter le dispositif aux conditions d'une restitution maximale (adaptation). Les réseaux se vantent de plus en plus fréquemment de pouvoir offrir des services fondés sur les contenus et leur signification pour l'utilisateur. Dans ces conditions, la valeur ajoutée se traduit souvent par un verrouillage informationnel. Le traitement des contenus est basé sur leur signification pour optimiser la transmission et la mise en place de services.

1.3- Un nouveau cadre

Toutefois, si le traitement des contenus repose sur la prise en compte de la signification : comment caractériser la signification d'un contenu et la prendre en compte dans un traitement opérationnel ? Si les documents correspondent à une pratique qui leur confère sens et pertinence, ils sont formatés par des codes sémiotiques, linguistiques, sociaux-culturels. En effet, ils reflètent une pratique qui leur donne sens et pertinence. La caractérisation de la signification relève de la description de ces différents codes qui permettent d'exploiter formellement le sens : la formalisation, la déconstruction peuvent être exploitées informatiquement; la reconstruction est faite par les utilisateurs (fonctionnalités des outils).

La représentation formelle du sens doit permettre de reconstruire les documents . On peut citer 3 types de systèmes

a- technique : ensemble d'outils et de matériaux organisés pour transformer un objet (énergie, mouvement, information);

b- sémiotique : ensemble de signes organisés fonctionnellement pour structurer la manière de "faire sens" ;

c- d'usage : ensemble des comportements organisés culturellement et socialement pour constituer et réguler les relations interindividuelles.

Les technologies cognitives sont conçues pour favoriser l'interaction entre ces 3 types de système: elles sont fondées sur des procédés de transformation pour exploiter la valeur sémiotique des objets manipulés ancrés sur des systèmes d'usage constitués ou co-constitués. Elles exploitent une représentation explicite et formalisée du sens. Elles naturalisent l'esprit et la cognition au moyen d'une théorie relevant des sciences de la nature et outillent l'esprit pour faciliter la cognition et des systèmes d'usage particuliers.

Les technologies de la convergence sont des technologies cognitives (sémiotiques) fondées sur les télécommunications, l'informatique, qui doivent faire sens dans un système d'usage pour que des documents soient interprétables par des utilisateurs en fonction de leurs usages à partir d'une représentation formelle du sens des contenus.

Sont décrits 3 niveaux d'exploitation des contenus :

- Interchangeabilité des données et des formats d'échange : XML, SGML, MPEG-7,
- Interopérabilité des outils d'exploitation pour unifier la chaîne de production, transmission et consultation,
- Intercompréhension des données : sémantique explicite des données pour comprendre les traitements effectués.

II- L'intelligibilité a priori ?

Les dimensions des inscriptions reposent sur un enregistrement (l'information est consignée via un codage). Le numérique fournit un accès conditionnel au contenu en tant que forme matérielle interprétable, reconstruite à partir d'un instrument de lecture. Les fonctionnalités de l'outil conditionne l'intelligibilité du contenu à partir des :

- unités de manipulation sur lesquelles sont effectués les calculs définis à priori,
- unités de sémiotisation concernant le signifiant du contenu (moins les parcours interprétatifs), qui se définissent en général à posteriori, même si elles se construisent à partir des unités de manipulation (Rastier).

Ce que je manipule et qui fait sens, est géré par l'usage. La manipulation pour faire sens repose sur des actions qui sont effectuées sur le document comme lire, annoter, gloser, sélectionner. Les actions sont effectuées sur le document du point de vue des contenus (spatial et temporel) et déterminent ce que fait le document la sémiotisation et la manipulation. Si on prend l'exemple de la lecture, quelles actions sont en jeu : sélection (choix dans le temps et l'espace du contenu, accès au contenu sans modification, construction d'une vue qui n'est pas dans le contenu en sélectionnant, composant, comparant, communicant.

L'acte de lecture donne deux types d'accès au contenu :

- Accès synthétique : appréhension globale du document dans sa linéarité et son rythme.
- Accès analytique : appréhension par segmentation, comparaison, élaboration.

L'audiovisuel numérique, par exemple, permet un travail analytique interactif sur le contenu audiovisuel pour véhiculer, élaborer et transmettre des connaissances et un savoir, d'où l'implication de communautés scientifiques en marge de l'audiovisuel comme le monde éducatif, l'univers professionnel pour interagir au niveau des contenus. La représentation du sens doit permettre l'interaction nécessaire au travail analytique pour décrire la valeur du contenu du point de vue de la connaissance mobilisée, associer les contenus aux concepts, ajuster ce que je peux faire et ce que je comprends. La signification des concepts spécifie l'usage et l'exploitation des contenus. Le numérique modifie profondément les conditions d'intelligibilité du contenu.

L'instrumentation technique est la condition à priori du faire sens.

III- La concrétisation interprétative

L'instrumentation numérique doit penser a priori de faire du sens pour le rendre possible à travers les actions de lecture, sachant que le sens d'un contenu n'est pas déterminé, mais seulement conditionné par son instrumentation. On articule le sens d'un contenu via l'instrumentation. On détermine l'existence de 3 couches :

forme (matériel/physique), technique (computationnel/syntaxique), usage (intentionnel/interprétatif).

En fonction du contexte, l'instrumentation prend tel ou tel sens. L'instrumentation doit contraindre l'interprétation par son organisation et sa structure. Le processus de concrétisation interprétative de Simondon s'étend à l'information et se fait en fonction des outils intégrant l'ensemble des fonctions. La forme syntaxique, sur laquelle repose la manipulation, doit permettre les différentes interprétations. Au lieu d'une forme pour une interprétation, on a une forme pour toutes les interprétations. Les différentes valeurs interprétatives ont une même forme syntaxique, matérielle et varient en fonction des utilisateurs.

IV- L'abstraction formelle ou devenir ingénieur de contenus.

La concrétisation interprétative est rendue possible par une formalisation et abstraction accrues de la forme syntaxique au niveau formel (méta) et interprétatif (les différents langages de l'usage). Citons l'exemple de XML qui permettra selon les situations d'engendrer différents

usages, DTD ou schémas.

Traditionnellement, le numérique renvoie à l'analytique et au formel :

- Analytique : qui répond au principe de non contradiction.
- Formel : qui est indépendant du sens (contenu).

En instrumentant le contenu, le numérique prescrit le sens, le contenu a priori car prescriptif: s'agit-il d'une synthèse intuitive (Kant) ? conceptuelle (Bolzano) ? Le numérique prescrirait-il un formatage de l'intuition et de l'action ? S'agit-il de concevoir des abstractions permettant des interprétations plurielles (renégociables dans des séquences d'usage).

Penser l'usage a priori, représenter formellement la sémantique pour paramétrer l'exploitation, concevoir l'abstraction pour permettre une pluralité interprétative pose problème. Le risque est de standardiser les contenus par des chaînes opératoires identiques, en passant de l'artisanat à un système technique industriel, car le numérique est en passe de constituer un système technique global. Toutes les étapes du cycle de vie des contenus sont pris en charge par les outils numériques, inter-opérables, avec des données inter-échangeables.

Le cycle de vie se formalise en un processus systématique et industriel, qui lui-même se banalise, outils accessibles et intégrés.

V- De l'artisanat à l'industrie des contenus

Le passage de l'artisanat (interprétation des outils par un schéma d'usage où chaque production est une création) à l'industrie (système technique où l'interprétation des outils est externalisée dans les méthodes et processus de production) pose la question de la standardisation industrielle. L'externalisation de l'interprétation des outils implique la répétitivité de la production qui permet de produire de l'identique, ainsi que la banalisation du travail de l'opérateur, qui devient un simple exécutant. Appliquée aux contenus la standardisation industrielle, elle permet de produire des contenus homogènes (possédant une valeur culturelle analogue) sur de grandes séries (industrie documentaire étendue aux loisirs et contenus culturels) et d'accéder aux contenus et à leur production. Le coût en est la réduction de la diversité et de la création (banalisation des valeurs culturelles). La notion d'art numérique introduit une logique de déviance dans les outils: faire sens autrement, faire usage autrement en sortant de la répétition pour instaurer l'objet ou le contenu dans une autre logique du sens non conventionnelle : C'est le contre-poison !

Selon la tradition du séminaire UTC, l'auteur de la communication est sorti pour permettre aux auditeurs de discuter des questions qu'ils souhaitaient lui poser. Ces questions ont été synthétisées avant d'être posées : elles ont surtout porté sur l'accès aux contenus, sur les routeurs intelligents, sur la représentation du sens (représentations formelles et contraintes opératoires), sur la forme de l'information et le processus éditorial : structuration, contrôle des mots-clés, l'art en tant que déviance.

Ce texte a été rédigé à partir des notes prises lors du séminaire et des transparents que B. Bachimont a bien voulu nous transmettre, qu'il en soit remercié.

Dialog

Elisabeth Lacombe, Conservateur en Chef

Xavier Jaunet, de la société Dialog Corporation, a animé une formation à l'Urfist le 13 décembre 2001. Nous avons emprunté son support de cours pour faire le compte rendu de ce stage.

paris_customer_service@dialog.com

téléphone 01 55 23 52 52

Introduction

Le groupe Thomson, auquel appartient Dialog depuis un an, regroupe plusieurs entités de plus de 35 000 personnes sur les 5 continents :

- Thomson Scientific, Reference & Healthcare
- Thomson Financial
- Thomson Learning (USA)
- Thomson Legal & Regulatory

Dialog Corporation gère 200 000 sources d'informations multidisciplinaires (plus de 6 milliards de pages texte et 3 millions d'images, ce qui représente 6 fois l'information gratuite du web):

- Economiques
- Techniques
- Scientifiques
- Brevets et Marques

Dialog Corporation regroupe trois serveurs : DIALOG (Palo Alto) - DATASTAR (Berne)- PROFOUND (Londres). Elle dispose de 60 bureaux répartis dans 40 pays et possède 20 000 sociétés clientes réparties dans 120 pays.

Les utilisateurs de Dialog Corporation sont des professionnels de l'information et des utilisateurs finaux, dans le cadre de la recherche et développement, de la veille et de la stratégie, du marketing, du service financier.

Adresses des serveurs :

- <http://www.dialog.com>
- <http://www.dialogweb.com>
- <http://www.datastarweb.com>
- <http://www.profound.com>
- <http://library.dialog.com> accès aux "blue sheets" (fiches techniques des banques de données)

Les services de Dialog Corporation

Les Producteurs de l'Information appartiennent aux domaines suivants :

- Presse : les banques de données sont actualisées deux à trois fois par jour
- Cabinets d'études, (comme Frost et Sullivan) : ils sont à 60% américains
- Organismes Institutionnels / Gouvernementaux
- Organismes Professionnels
- Sociétés privées
- Instituts de Recherche
- Bibliothèques

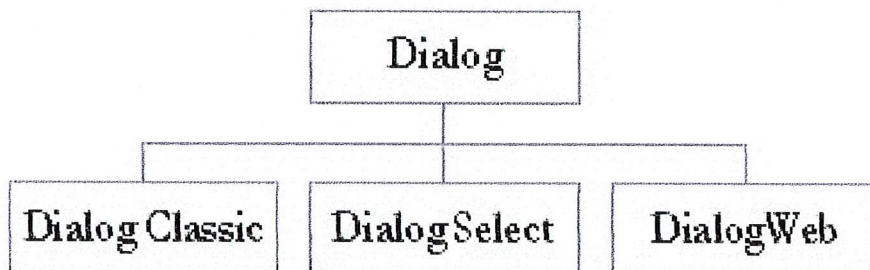
Les banques de données de Dialog correspondent à la typologie classique:

- Texte Intégral
- Références Bibliographiques
- Répertoires
- Images
- Tableaux Statistiques
- Dialog

Dialog met 470 banques de données à disposition de ses utilisateurs dans les domaines suivants :

- Brevets - Marques : 80% de l'information brevets n'est disponible nulle part ailleurs
- Sciences & Techniques
- Chimie
- Energie
- Agro-alimentaire
- Biomédicale - Pharmaceutique
- Economie - Actualité internationale
- Législation

Dialog offre plusieurs interfaces d'interrogation :



Dialog Select est guidée et destinée à l'utilisateur final.

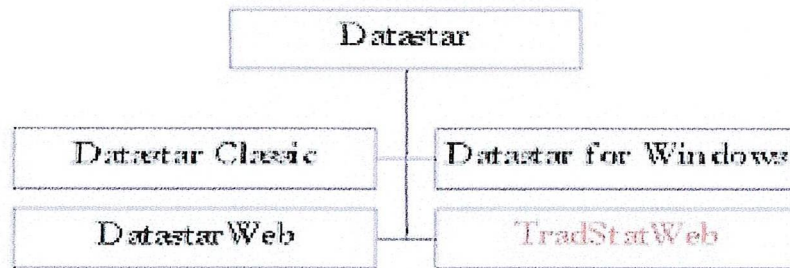
DialogWeb offre une interface conviviale pour tout type d'utilisateur.

DataStar

DataStar gère 350 banques de données

- Biomedical - Pharmaceutique
- Sciences & Techniques
- Energie & Environnement
- Agro-alimentaire
- Législation Européenne
- Economie - Finances
- Actualité européenne et mondiale

Les interfaces de DataStar sont :



Tradstatweb, interface web conviviale et rapide, est une compilation mensuelle des données statistiques douanières et représente 90% du marché mondial des échanges : import/export de 27 pays, plus de 600 000 produits répertoriés.

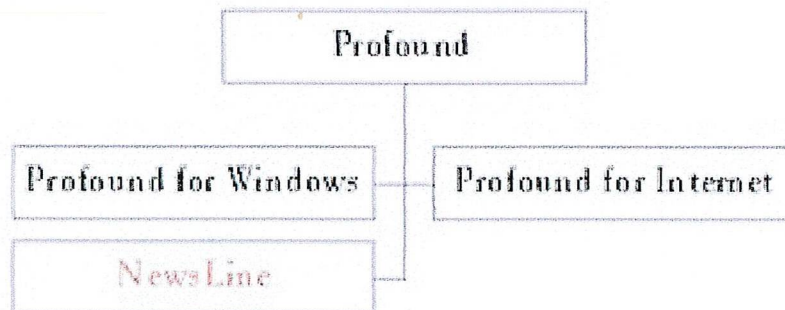
Les résultats sont présentés sous forme de tableaux sectoriels et de tableaux sur des produits spécifiques.

Profound est spécialisé dans :

- l'actualité internationale
- les études de marchés
- les études par pays
- les rapports de brokers
- l'information sur les sociétés

Il est relié aux données de Dialog par des liens (300 Banques de Dialog, notamment sur la propriété intellectuelle et sur les répertoires de sociétés)

Les interfaces de Profound sont :



Newsline, interface web conviviale de Profound offre un accès illimité à des articles de la presse économique et professionnelle, une recherche simultanée sur 5500 sources issues de 190

pays et des systèmes d'alertes.
L'indexation est unique pour toute l'information.

Les nouveautés 2001

Dialog IQ (interface unique pour les trois serveurs)

Selon une récente étude du cabinet 'Outsell', plus de deux tiers des utilisateurs finaux préfèrent obtenir de l'information en la recherchant eux mêmes.

Selon des estimations, plus de 90% des recherches sont faites par mots clés. Mais 57% des utilisateurs ne trouvent pas ce qu'ils cherchent.

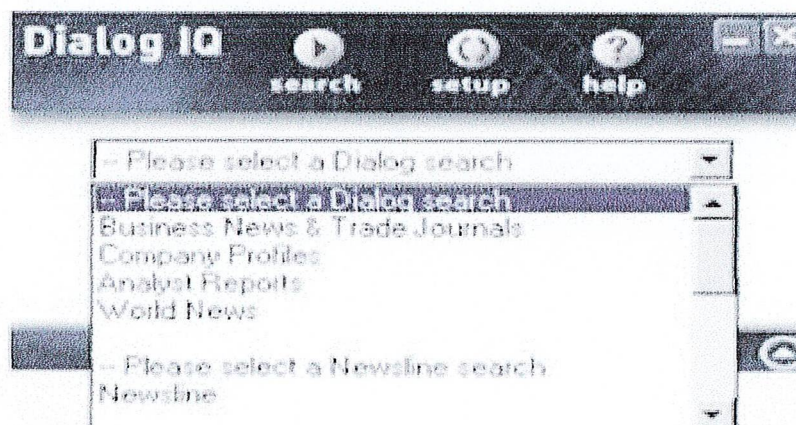
Dialog IQ résout ces problèmes en présentant aux utilisateurs novices un contenu sélectionné, un accès facile à l'information par une recherche guidée et un seul champ de recherche, enfin une visualisation claire.

Dialog IQ leur fournit un format de document et une indexation unique.

L'interface d'interrogation est commune aux quatre plate formes (une à la fois avec un affichage séparé sur plusieurs écrans). Le contenu est le suivant :

- DataStar: profils de sociétés avec D&B, World Reporter
- Dialog: presse économique, Répertoire des sociétés américaines avec D&B, World Reporter, Investext
- NewsLine: actualité avec plus de 5500 sources internationales
- Profound: rapports, news, statistiques pays, investissement et rapports de brokers.

Voici comment se présente un écran de recherche sur Dialog IQ :



Profound offre de nouvelles interfaces et de nouvelles fonctionnalités.

La nouvelle interface présente les liens suivants : Worldsearch, Quicksearch, Market briefings, Company briefings (environ 12000 entreprises), Investment briefings et Country briefings. Les briefings sont des documents donnant une vue générale du marché (parts de marché, acteurs-clés du marché, entreprises...) Ils sont visualisables en format PDF.

1. Worldsearch : permet de choisir les banques de données ResearchLine (études de marchés), NewsLine (presse), BrokerLine (études de brokers), CountryLine (informations macro-économiques).

Le navigateur InfoSort offre une sélection de termes à l'intérieur du domaine choisi. On peut élargir ou préciser la recherche à l'aide des boutons "broader", "narrower", "related terms".

2. Quicksearch

Quicksearch permet de faire la recherche dans les 300 banques de données les plus utilisées sur Dialog : entreprises, propriété industrielle, recherche et développement, nouveautés.... La "TableBase" (base n°93 dans Dialog) donne des tableaux chiffrés (classement des meilleures entreprises, entreprises par chiffres d'affaires...)

Dans cette base, on peut faire une sélection par masques à choix multiples pour se focaliser sur les sociétés, les concepts, les types de documents.

Quicksearch établit un lien direct vers Investext pour obtenir les documents PDF complets. Pour faire une recherche sur les marchés, Profound dispose des banques de données suivantes : Datamonitor, Reuters Business Insight, Global Industry Analysts, Informa Global

Pharmaceuticals, IBIS World Market Research, Asian Information Resources, IDC, Gobi International.

Nouveautés Dialog

Les nouvelles banques de données commercialisées par Dialog sont : Reuter Business Insight report (759), Investext en PDF (745), IBISWorld Market Research (753), EIU magazines (622), Daily/Sunday Telegraph (756), Mirror Group Publications (757), Financial Time Fulltext (476), Mergent Company (555, 556, 557), Trademarkscan Japan (669), Harris Business Profiler (537).

Nouveautés DataStar

Les nouvelles banques de données commercialisées par DataStar sont : Reuter Business Insight report (RBIR), Kompass: Europe, M.East, Africa, Asia Pacific, Americas, BBC International Reports (BBCM), MEDLINE in process (MEIP), MEDITEC: Medical Engineering (BMED), Worldbox Swiss Company Financial Records (SWFR), NewsRx Healthcare Newsletters (RXNS), Cochrane Database of Systematic Reviews (CDSR), Pharmaprojects (PHAR), Adis Clinical Trials Insight (ACTI).

Documents primaires

Dialog Co. a établi des partenariats avec 12 fournisseurs de documents primaires qui permettent à l'utilisateur d'obtenir le texte intégral en format PDF . Ces partenaires sont : Academic press, American Institute of Physics, CatchWord, High wire press, Ideal, Information Quest, Institute of Physics, Karger, Siam, Springer, Swets, Synergy.

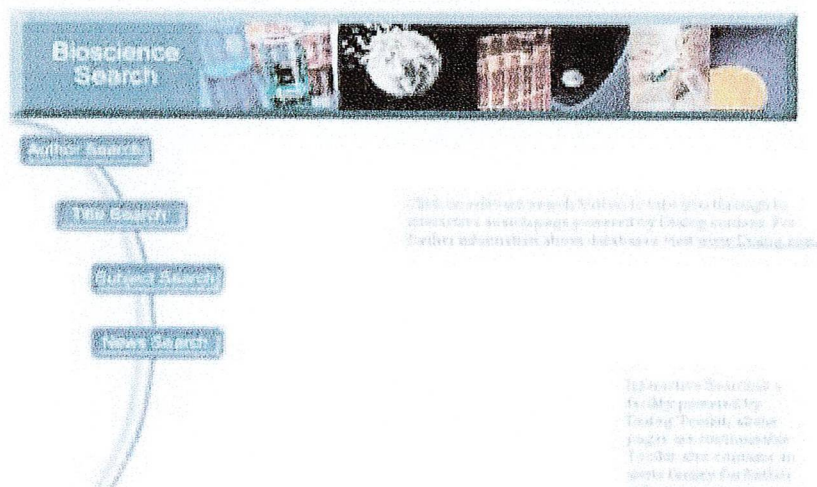
Solutions Intranet

<http://www.intraintelligence.com>

Il s'agit de proposer une solution personnalisée conforme aux besoins du client, c'est-à-dire :

- des accès aux serveurs personnalisables: la création d'interfaces peut être prédéfinie ou personnalisée (Toolkit de Dialog ou Datastar personnalisé).
- l'intégration et la diffusion de profils de surveillance au sein de l'Intranet :
- système d'alerte (email, courrier, Lotus), Monitor et Intranews (dans l'Intranet), toolkit (hot topics = profils automatiques pour les sujets sensibles, ou alerte manuelle)
- l'hébergement et le partage des données internes.

Dialog Toolkit (exemple d'interface pré-définie):



Monitor est une application propriétaire qui diffuse au sein de l'Intranet l'information spécifique venant d'alertes automatiques et de recherches sur Dialog. Les informations sont envoyées par e-mail vers l'Intranet : liste de titres, article en texte complet, gestion du copyright. Il couvre tous les fichiers Dialog.

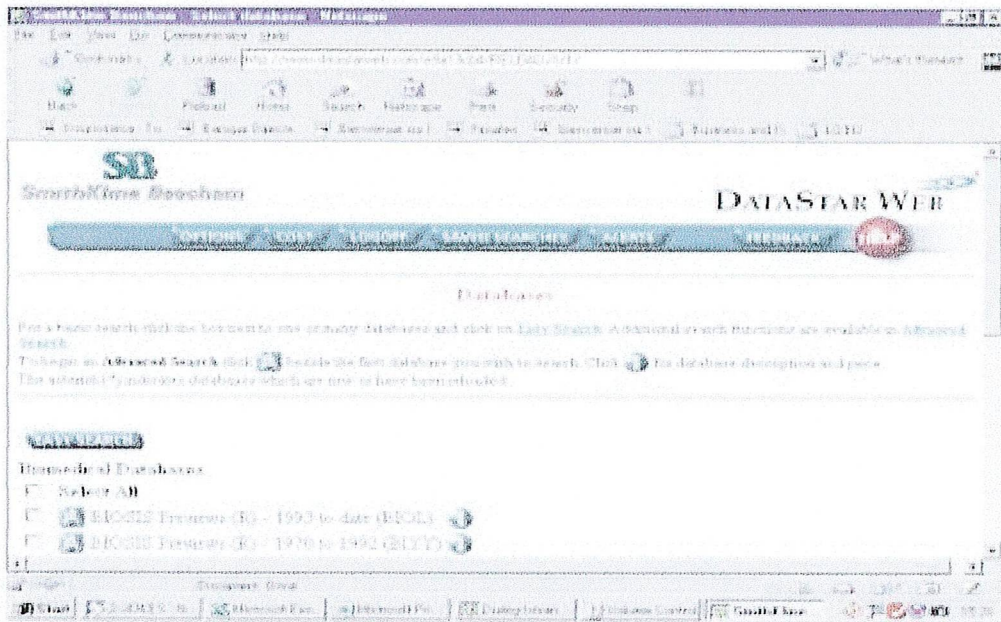
Le client se connecte sur Dialog pour télécharger les documents obtenus à partir de recherches simples ou d'alertes. Monitor effectue alors automatiquement le routage vers les utilisateurs de l'Intranet. Des documents externes à Dialog peuvent être rajoutés, scannés par Monitor et distribués.

Intranews, de même, délivre en continu l'information presse au sein de l'Intranet, avec des profils personnalisés, en XML ou HTML, en utilisant la catégorisation INFOSORT. Cette fonction s'intègre avec les systèmes existants.

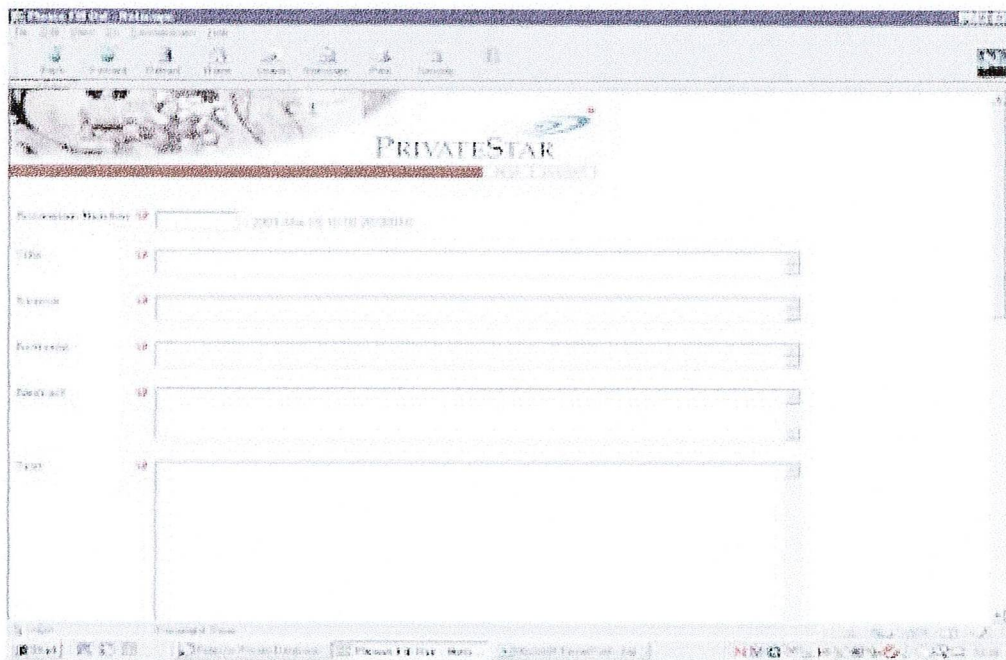
DataStar personnalisé

L'accès personnalisé à DataStar (customise) présente un choix des sources d'informations ainsi que des critères de recherches. Il permet l'intégration de profils et établit des liens avec e-journal pour le texte intégral en format PDF. Il autorise aussi l'intégration de bases de données privées. Grâce à DataStar personnalisé, il est possible de se constituer une banque de données privée (PrivateStar) avec les contenus issus de Datastar et des ressources Clients. Les documents privés sont chargés par le Graphic User Interface. L'interrogation est simultanée avec les autres sources présentes sur Datastar et une indexation supplémentaire est possible (constituée des termes des clients).

Exemple d'interface DataStarWeb :



Les accès sont sécurisés : le système reconnaît les numéros IP des machines. La banque de données Privatestar est hébergée sur DataStar et utilise le moteur de recherche du serveur.



Conclusion :

Les solutions proposées par Dialog qui allient contenus et technologie, sont des solutions personnalisées de recherche, de diffusion et de partage, d'hébergement, de développement et d'enrichissements des Intranets.

XML et les Bibliothèques électroniques

Elisabeth Lacombe, Conservateur en Chef

Le 12 décembre 2001 a eu lieu à l'URFIST un stage sur l'utilisation de XML, notamment dans le domaine de l'information et de la documentation. Cette formation était animée par M. Jacques Ducloy du CNRS. Le compte rendu de ce stage est écrit à partir du [support de cours](#) de M. Ducloy.

On peut aussi lire une présentation de XML dans la Lettre de [l'URFIST n° 26](#).

I-Introduction

XML représente vraisemblablement la future norme, aussi bien dans le domaine de l'information scientifique et technique que dans celui du commerce électronique, car elle apporte beaucoup par rapport à HTML :

- elle peut représenter n'importe quel type d'information
- de nombreux outils sont développés autour de XML
- elle s'applique aux métadonnées.

Internet et la révolution technologique ont bouleversé le monde de l'information qui est maintenant à la portée de tout le monde. L'évolution de l'informatique en 20 ans a été fulgurante dans la rapidité des traitements, du temps de programmation et de la baisse des coûts. Les conséquences sont importantes pour les banques de données : leur mise en ligne sur un serveur qui coûtait 200 MF en 1976 ne revient plus qu'à 50000 F ! Chaque entreprise a maintenant les moyens autrefois dévolus à l'état et aux grandes centrales de recherche.

Les relations entre l'informaticien et le professionnel de l'information se sont resserrées dans un dialogue qui évite l'élaboration de cahier des charges. XML joue un rôle fondamental, c'est un moyen d'utiliser autrement l'information. Certaines recherches d'information sur Internet sont faciles, par exemple "trouver les travaux de chercheurs sur l'utilisation de composés pour la thérapie du SIDA". Par contre une question comme "qu'y a-t-il de nouveau sur le SIDA depuis un an ?" est très complexe : se pose le problème de comment découvrir ce que nous ne connaissons pas encore et aussi comment produire une information synthétique. Aucun système intégré n'existe, par contre un ensemble d'outils permettent des réponses à ces questions qui sont hors du champ de la recherche documentaire basique : data-mining, linguistique, statistique... Afin de manipuler ces logiciels différents ainsi que des données hétérogènes (en provenance de banques de données, de bases de connaissance, de documents multimedia...), un langage existe : XML.

II-Langages de balisage

II.1 Principe du balisage

Il s'agit d'un mécanisme pour repérer des éléments dans un texte, le structurer et rappeler le type

des éléments. Tout élément est précédé d'une balise de début : ex : `<titre>` et suivi d'une balise de fin : `</titre>`

Entre ces deux balises se trouvent le texte avec éventuellement d'autres balises imbriquées. XML permet d'exprimer la logique booléenne :

```
<and> <texte> <texte> </and>
```

II.2 DTD et Langage de balisage

La structure logique d'un document est une structure arborescente : un livre comporte un titre et des auteurs, il est divisé en chapitres qui ont eux-mêmes des titres, chaque chapitre est divisé en paragraphes ...

La description logique du document s'appelle une DTD : document type definition., sorte de règle de grammaire

```
<!DOCTYPE book [
<!ELEMENT book -- (title, author*, chap+) >
<!ELEMENT title -- (#PCDATA)>
<!ELEMENT author -- (#PCDATA)>
<!ELEMENT chap -- (title ?, p+)>
<!ELEMENT p -- (#PCDATA)>
]>
```

Remarques :

* désigne un élément optionnel mais répétable

+ désigne un élément obligatoire et répétable

? désigne un élément optionnel

#PCDATA désigne une donnée textuelle

Le balisage est un balisage descriptif symbolique qui reprend les identifiants de la DTD et représente une garantie de lisibilité à long terme.

II.3 Origines des normes SGML, HTML, XML.

- SGML : Standard Generalized Markup Language (langage de balisage généralisé) Son origine prend racine chez IBM, dans la documentation militaire, l'avionique. Le besoin était de pouvoir utiliser une séparation selon une structure logique (chapitre, paragraphe...) et une présentation (gras, italique, rouge, clignotant...). SGML est devenu le système normatif " générique " pour les documents techniques, il fait l'objet de la norme ISO 8879.
- HTML : HyperText Markup Language (langage de balisage pour documents hypertexte) correspond à un besoin de langage pour des documents simples à échanger sur Internet. C'est une DTD particulière de SGML . Son origine est le CERN pour l'échange de documents de physique nucléaire. Il est actuellement maintenu par W3C (INRIA, MIT, KEIO University). Cependant HTML est un langage trop limité pour couvrir toutes les applications du web .
- XML : eXtensible Markup Language correspond aux besoins d'élargir la nature des documents du web, de faciliter l'interopérabilité entre applications, de faciliter la coopération entre communautés, de permettre des descriptions plus précises. XML a gardé l'essentiel de SGML, avec des extensions pour prendre en compte la modularité,

l'interopérabilité, le multilinguisme.

II.4 Le langage XML : comparaisons avec SGML et HTML

- La minimisation est une disposition qui permet un balisage réduit (ex : dans les balises imbriquées, on omet d'indiquer la balise de fin). La machine peut passer de la forme réduite à la forme complète. Les règles de minimisation sont décrites dans la DTD pour SGML.

Il n'existe pas de minimisation pour XML : tout ce qui était particulier au monde de l'édition avec SGML a été " élagué ".

- La codification dans SGML et HTML : les caractères spéciaux (& < >), les caractères accentués, les lettres grecques etc... sont repérés par des entités : ex & ' etc ... Par contre, XML adopte une norme qui affecte un code numérique à chaque caractère spécial.
- XML permet une plus grande rigueur syntaxique que SGML. De légères modifications simplifient la programmation et la portabilité des applications. Majuscules et minuscules sont significatives dans les balises.
- Il y a possibilité dans XML de mélanger des éléments déclarés dans plusieurs DTD à l'intérieur d'un même document, sans conflit de nommage.

La déclaration du domaine nominal se fait par un attribut commençant par `xmlns` ou `xmlns` :

Exemple :

```
<html xmlns :html='http://... '>
```

Ceci favorise le travail normatif en sous-groupe. L'intérêt est grand pour les organismes universitaires qui doivent respecter des standards de sources venant de communautés multiples pour la production de document.

Pour le codage, on peut :

- utiliser une DTD pré-existante : HTML, ISO 12083 -utilisée par les éditeurs- ou TEI, text encoding initiative
- adapter une structure existante : notices bibliographiques, formats MARC
- créer une nouvelle DTD

HTML permet d'échanger des documents sur le web. Les balises principales sont les suivantes :

- Pour les titres ou changements de fonte : balises `h1`, `h2`, `h3`, `h4`

- Pour les listes : balises `` ``

- Les ancres et les liens : balise `a`

`<a` repère la partie active ; la notion d'attribut qui complète l'information associée à la balise est exprimée par `href=`

ex : Olivier travaille à l'URFIST

Pour créer un lien avec le site de l'URFIST, le codage est le suivant :

```
<A href=http://www.urfist.cict.fr>URFIST</A>
```

III-Le langage XML

III.1 Codage

XML utilise un langage de codification universel : TEI (text encoding initiative), dont l'objectif est de décrire tous les textes existants, tel que l'avait imaginé l'auteur.

Exemples de codage d'une adresse :

```
<address>
<addrLine> texte
<addrLine> texte
<addrLine> texte
</address>

<address>
<street>texte</street>
<name type=city n=OX>Oxford</name>
<postCode>texte</postCode>
<name type=country n=UK>United Kingdom</name>y </address>
```

Le codage XML peut aussi adapter des notices bibliographiques, par exemple Medline :

```
<medline>;
<AN>numéro</AN>
<TI>texte</TI> <AU><e> nom d'auteur</e><e> nom</e></AU>
<CS> texte</CS>
<JN> nom du périodique</JN>
<PY> date</PY>
</medline>
```

III.2 XML et les différentes normes

XML peut adapter des structures plus complexes comme les formats MARC (machine readable cataloguing). Il s'agit de formats d'échange entre des opérateurs et des structures de documents, correspondant à la norme de base ISO 2709. Différents formats existent :

USMARC et les variantes LC MARC et MARC de l'OCLC (qui s'unifient)

UKMARC

UNIMARC utilisé par de nombreuses bibliothèques, par BN-Opale, l'Inist

PASCAL (Inist) , norme particulière

CCF (Unesco)

Il y a des différences énormes dans les formats, mais un accord sur le mode de structuration.

La norme ISO 2709

Elle comporte deux niveaux de description :

Une zone avec un élément bibliographique, une étiquette de longueur fixe pour un format, une sémantique très différente suivant les formats.

Une sous-zone, comportant un élément de données (contenu, langue) et une étiquette de longueur fixe.

En général, les logiciels de bibliothèques respectent assez bien la norme. Il faut développer un programme (cela prend d'une semaine à un mois) pour convertir le format MARC du catalogue en XML.

XML peut aussi être utilisé :

- sur des données internes à un système informatique, par exemple le projet DILIB (Nancy) ou des spécifications des bibliothèques de composants,

- sur des langages de programmation (opérateurs booléens et champs ; feuilles de style XSL),
- sur des données issues de SGBD, par exemple l'application UMLS, méta-thésaurus du MESH
- dans des domaines comme la chimie (Chemical Markup Language), les mathématiques (MathML, W3C), les graphiques (Scalable Vector Graphics, W3C, WebCGM), le développement international (International Development Markup Language), les documents multimedia (Synchronized Multimedia Integration Language, W3C), l'édition (ISO 12083)

Comment transformer des documents XML ?

Des normes sont liées aux opérations sur les documents XML :

CCS (Cascading Style Sheet), XSL (Extensible Stylesheet Language)

XSLT (Extensible Stylesheet Language Transformation)

DOM. (Document Object Model)

Pour transformer les documents XML , on peut utiliser le langage de programmation XSLT (sous-ensemble de XSL).

Le DOM permet de spécifier l'interface pour organiser une bibliothèque de base XML, grâce à des opérations élémentaires.

Il existe des mécanismes de désignation d'un élément dans un document : Xpath et Xpointer, ainsi que des mécanismes de gestion des liens sur le web : Xlink.

III.3 XML et les métadonnées

Les métadonnées constituent le catalogage moderne. Le cadre générique en est le Resource Description Framework (RDF). Différentes applications existent : le catalogage de ressources multimedia, le Dublin Core (OCLC), la sélection de contenus, PICS (Platform for Internet Content Selection)

Le Dublin Core

La direction du Dublin Core est l'OCLC. Le comité éditorial (Dublin Core Policy Advisory Committee) comprend des représentants des USA, de la CE, de l'Allemagne, de la Grande Bretagne, de la Suède, du Canada, de l'Australie, du Japon.

Le groupe de travail du [Dublin Core](#) a pour objectif de faciliter la recherche de documents électroniques. Au départ, il s'agissait d'utiliser un catalogage simplifié dans un contexte HTML. Puis s'est faite une ouverture vers XML, ainsi qu'un élargissement de la communauté, avec deux tendances contradictoires : l'une maximaliste (plus compliqué que l'USMARC), l'autre minimaliste.

Les éléments META ont pour attributs : name, content, scheme , dans HTML.

Les 15 éléments du Dublin Core sont :

Title, author, subject, description, publisher, contributors, date, type, format, identifier, source, language, relation, coverage, rights.

Caractéristiques :

- Les mêmes informations figurent dans différents systèmes de métadonnées

Exemple :

```
<meta NAME= "keywords"  
CONTENT= "oclc,online, computer, library, ...">  
<meta NAME= "DC.subject.keyword"  
CONTENT="oclc,online, computer, library, ...">
```

- On peut utiliser ou non les qualifieurs

Encoded Archival Description : EAD

Cet autre type de métadonnées est destiné aux descriptions d'archives, aux collections et aux catalogues collectifs. Les sponsors sont la Library of Congress et la Society of American Archivists.

Les projets sont les suivants :

- [Library of Congress Finding Aids Project](#)
- [American Heritage Virtual Archive Project](#)
- [Online Archive of California](#)

Resource Description Framework : RDF

Les objectifs sont de décrire de manière plus précise les ressources du web, de s'adapter au contexte XML, de constituer un support pour les ontologies, les thésaurus.

Un document est décrit à l'aide d'élément : c'est la description RDF. Les éléments sont contenus dans un ou plusieurs schémas RDF. RDF permet de prendre en compte des métadonnées d'origine différente, en utilisant plusieurs domaines nominaux. Mais il y a un manque de stabilité par rapport aux normes, aux outils, aux usages.

IV-Ingénierie XML

Il existe trois catégories d'outils XML :

- l'outil SGML-XML indépendant des applications
- l'application d'XML au web
- les adaptations d'applications existantes au web

Les outils du marché

- Il existe plusieurs catégories de parsers qui permettent de vérifier la conformité d'un document avec la DTD, ou d'associer une action à un élément SGML, des parsers DOM (Document Object Model), des éditeurs structurés, des systèmes de recherche d'informations SGML.

- Xpath, Xpointer, Xslt, Xlink

Ils permettent de repérer et manipuler des portions de documents XML : reconnaissance et sélection. Xpath permet de faire une sélection par nom de balise, par position, par attribut. Xpointer est une généralisation de Xpath.

Xlink généralise la notion de liens HTML.

- Les feuilles de style : CSS et XSL
- La transformation de documents : XSLT

V-Conclusion : les enjeux d'XML

- Future norme du web

Les navigateurs intègrent XML et les feuilles de style : XML et CSS est en cours de distribution, bientôt suivi par XML et XSL. Les moteurs de recherche utilisent la description RDF (de même que le commerce électronique) et le Dublin Core pour les métadonnées.

- Nouvelle base pour le document électronique

XML permet une description fine du document multimedia. Le balisage sémantique, indépendant d'un langage de programmation particulier, est une garantie de pérennité.

XML mélange les éléments de plusieurs DTD.

- Cœur de l'ingénierie Intranet et Internet

XML apporte des éléments de réponse à la modularité et l'interopérabilité, permettant ainsi la conception d'applications complexes. Exemples : la plate-forme DILIB

Ce prototype, réalisé par le LORIA et l'INIST, est une plate-forme pour l'ingénierie du document et de l'information scientifique et technique. Elle fonctionne sur des données XML et des données hétérogènes. Elle développe des outils pour l'analyse de l'information. Une des applications est MedExplore, exploration de bases biomédicales (Medline, Pascal, Scisearch, Internet...)

- Ciment des applications collectives

XML offre un support pour la normalisation propre à un réseau et permet aux acteurs d'appartenir à plusieurs réseaux. L'ingénierie est adaptée aux différents éléments informatiques d'un réseau.

Le web invisible

François Magnan, formateur indépendant

François Magnan a assuré à l'URFIST une formation sur la recherche sur Internet pour des utilisateurs avancés, dans laquelle il a montré comment trouver les informations et les documents du " Web invisible ". Il nous donne ici une définition de cette notion et cite les outils d'accès aux informations que ne savent pas récupérer les moteurs de recherche.

Lorsque nous recherchons de l'information sur Internet, trois types d'outils nous servent pour la plupart de nos recherches :

les annuaires de sites comme Yahoo, pour lesquels travaillent des équipes de netsurfeurs qui sélectionnent et parfois commentent des sites sur lesquels il sera possible de faire une recherche par catégorie ou par mots-clefs,

les moteurs de recherche, outils purement automatiques, dont les robots explorent le Web, soit en passant d'un site à l'autre, soit à partir des adresses déclarées par les Webmasters;

quant aux *métamoteurs de recherche*, ce sont en quelque sorte des outils parasites qui se contentent de chercher à l'intérieur des bases constituées par les autres outils.

Ces outils de recherche étant très nombreux, on pourrait penser qu'ils recensent tout le Web. Hélas, il n'en est rien ! Selon une étude parue dans la revue " Nature " en juillet 1999, les plus grands moteurs de recherche réunis n'indexaient alors que la moitié du Web. Pour quelle raison ? Si certains webmasters négligent le référencement de leurs sites, d'autres causes tiennent à la nature des documents mis en ligne. L'ensemble des documents qui échappent aux moteurs de recherche s'appelle le "Web invisible".

De façon schématique, on peut dire que les robots des moteurs de recherches dressent des cartes du Web : ils circulent sur le réseau, ramassent les adresses des pages Web ainsi que les informations qu'elles contiennent afin de répondre à vos requêtes. Lorsqu'ils arrivent jusqu'à un document au format bureautique ou une base de donnée, c'est un peu comme s'ils arrivaient à la porte d'une bibliothèque dont la porte serait cadenassée : les robots peuvent noter l'adresse, mais ils ignorent les documents qui sont à l'intérieur ! (1)

Pour des raisons pratiques, certains auteurs d'articles scientifique joignent à leurs sites leurs travaux sous un format bureautique, inaccessible à la plupart des moteurs de recherches. Puisqu'un site sur la toxicomanie est plus pertinent avec des informations sur chaque drogue, de nombreuses bases de données sont rajoutées aux sites Internet. Or, les moteurs de recherches ont quelques problèmes à repérer les pages qui sont " générées dynamiquement " comme celles qui apparaissent après avoir interrogé une base de données.

Pourtant, ces bases de données présentes sur le Web sont de précieuses ressources et constituent le cœur de ce qu'on appelle le "Web invisible" : de plus en plus d'Universités, de Laboratoires de recherche les proposent en accès gratuit sur Internet. Elles peuvent aussi contenir de l'information bibliographique, comme les catalogues de bibliothèques en lignes. Leur contenu est structuré : il est plus facile de rechercher de l'information sur un médicament dans une base de donnée spécialisée qu'en faisant appel à un moteur de recherche comme Altavista. Avec le Web invisible, les moteurs de recherche montrent leurs limites : comment faire alors pour trouver ces bases de données ? Il existe heureusement des répertoires spécialisés, des outils qui vous permettront de repérer quelques ressources essentielles :

- [Le CISMEF](#) ou Catalogue et Index des Sites Médicaux Francophones propose un répertoire de bases de données médicales.
- Une société Américaine, [Brightplanet](#), propose de partir à la découverte du Web Invisible : [Completeplanet](#) est un outil de recherche en ligne, qui permet de trouver des bases de données spécialisées. Il est possible d'y effectuer des requêtes assez complexes. [Lexibot](#) est un métamoteur de recherche à télécharger (la version d'essai, gratuite, durera 30 jours) qui permet de faire des recherches dans le Web invisible et dans le Web traditionnel.
- Ce répertoire des Bases de données gratuites sur Internet, tenu par Jean-Pierre Lardy, de [l'URFIST de Lyon](#), est une source à retenir pour partir à la découverte du Web invisible. Ce site signale près de 600 bases de données.
- [Scirus](#) est un outil de recherche spécialisé dans l'information scientifique. Le propriétaire de cet outil n'est autre qu' Elsevier Science, un important éditeur de revues scientifiques, et c'est Fast, la société qui produit le moteur de recherche alltheweb qui a conçu sa technologie. Le formulaire de recherche avancée permet d'affiner les requêtes, et de ne rechercher que des articles.
- Comme Scirus, [Search4science](#) est un nouvel outil de recherche conçu pour les

scientifiques. Son formulaire de recherche avancé - ici appelé "dynamic search" - permet aussi d'affiner les requêtes.

- Certains agents intelligents permettent de rechercher l'information directement dans les bases de données. Ainsi, [Strategic Finder](#) permet à la manière d'un métamoteur d'interroger aussi bien le Web visible (via les moteurs de recherche les plus connus) que le Web "invisible". L'un de ses "plugins-in" ou module additionnel, Biognome, permet d'interroger près de 500 sources dans le domaine des biotechnologies, un autre, Transport, permet d'interroger près de 180 sources dans le domaine de l'économie du Transport.
- Kevin Elliott du site [About.com](#) maintient à cette adresse un métarépertoire de moteurs ou répertoires consacrés aux Sciences : Sciences de l'Ingénieur, Chimie, Mathématiques, etc.
- Le métamoteur de recherche [Profusion](#), les répertoires [Webdata](#) et [Infomine](#) répertorient aussi de nombreuses bases de données interrogeables
- [Invisible-web.net](#) est un site américain sur le Web invisible.
- Sur le site de [l'URFIST de Toulouse](#), la revue de presse trimestrielle présente une sélection d'informations et de sites-ressources, élaborée à partir d'une vingtaine de revues spécialisées.

Si ces adresses permettent une première approche du Web invisible, vous trouverez certainement d'autres liens intéressants !

(1) Notons toutefois les récents progrès du moteur de recherche [Google](#) qui indexe désormais des documents aux formats bureautiques, comme les fichiers .xls, .doc, .ppt, etc.