

Dossier de qualification
aux fonctions de
Maître de Conférences

Présenté par

Olivier Ertzscheid

Année universitaire 2003 - 2004

Table des matières.

REMARQUES INTRODUCTIVES.....	4
1. CURRICULUM VITAE.....	5
2. ACTIVITÉS DE RECHERCHE.....	6
2.1. VUE SYNOPTIQUE.....	6
2.2. THÈMES & MOTS-CLÉS DE RECHERCHE.....	7
2.3. RATTACHEMENT À DES ÉQUIPES DE RECHERCHE.....	8
2.3.1. <i>Equipe de rattachement</i>	8
2.3.2. <i>Chercheur associé</i>	8
2.3.3. <i>Chercheur invité</i>	8
2.4. ACTIONS ET PROJETS DE RECHERCHE.....	8
2.4.1. <i>Actions de recherche achevées ou en cours</i>	8
2.4.1.1. <i>Plateforme FoRSIC</i>	8
2.4.1.2. <i>SABRE (Système Auteur Basé sur la Réutilisation)</i>	10
2.4.1.3. <i>Usages experts & novices en recherche d'information</i>	11
2.4.1.4. <i>Groupe URFIST "Pratiques de Formation"</i>	12
2.4.2. <i>Projets de recherche</i>	12
2.4.2.1. <i>DYNAMO (Dynamic Ontology)</i>	12
2.4.2.2. <i>Documents électroniques pour les apprentissages humains</i>	12
2.4.2.3. <i>Scénarios de navigation et recherche d'information</i>	12
2.4.3. <i>Participation et affiliation à différents réseaux de recherche</i>	13
2.4.3.1. <i>RTP - DOC</i>	13
2.4.3.2. <i>ACM SIGWEB</i>	13
2.4.4. <i>Projets Industriels</i>	13
3. PUBLICATIONS.....	14
3.1. VUE SYNOPTIQUE.....	14
3.2. CHAPITRES D'OUVRAGE.....	15
3.3. JOURNAUX ET REVUES À COMITÉ DE LECTURE.....	15
3.4. COLLOQUES INTERNATIONAUX.....	15
3.5. COLLOQUES NATIONAUX (CN) & JOURNÉES D'ÉTUDE (JE).....	16
3.6. JOURNAUX ET REVUES SANS COMITÉ DE LECTURE.....	16
3.7. COLLABORATIONS.....	17
3.8. CONFÉRENCES ET PARTICIPATIONS INVITÉES.....	17
3.9. DIFFUSION DES CONNAISSANCES.....	17
3.10. SOUMISSIONS EN COURS.....	17
4. ACTIVITÉS D'ENSEIGNEMENT.....	18
4.1. VUES SYNOPTIQUES.....	18
4.2. D.E.S.S. IE & SIAD.....	19
4.3. LICENCE "INFORMATION & COMMUNICATION".....	20
4.4. URFIST DE TOULOUSE.....	21
4.5. ECOLE D'ORTHOPHONIE DE TOULOUSE (3 ^{ÈME} & 4 ^{ÈME} ANNÉES).....	21
4.6. CIES DE TOULOUSE.....	22
4.7. D.U.S.T.I.C.....	22
4.8. ECOLES DOCTORALES.....	22
4.9. QUELQUES POINTS À SOULIGNER.....	23
5. RESPONSABILITÉS COLLECTIVES ET ADMINISTRATIVES.....	24
5.1. FONCTIONS DE DIRECTION.....	24
5.2. ENCADREMENT DE TRAVAUX & MÉMOIRES DE RECHERCHE.....	24
5.3. RESPONSABILITÉS ADMINISTRATIVES.....	24
<i>Jurys d'examen et de recrutement</i>	24
<i>URFIST de Toulouse</i>	24
5.4. RESPONSABILITÉS PÉDAGOGIQUES.....	25
<i>Ecole d'orthophonie de Toulouse</i>	25
5.5. MANIFESTATIONS SCIENTIFIQUES.....	25

5.6. SITES WEB, FORUMS ET LISTES DE DIFFUSION	25
5.7. PARTICIPATION À DES PUBLICATIONS	25
5.8. ACTIVITÉS ASSOCIATIVES	25
6. VISIBILITÉ SCIENTIFIQUE.....	26
6.1. VUE SYNOPTIQUE	26
6.2. VALORISATION DE MES ACTIVITÉS DE RECHERCHE.....	27
6.3. PRIX & DISTINCTIONS.....	27
6.4. RELATIONS INTERNATIONALES.....	27
6.5. CONFÉRENCES & SESSIONS ORGANISÉES	28
7. ANNEXES.....	29
7.1. ARTICLES ET PUBLICATIONS	29
<i>Chercher faux et trouver juste, Serendipité et recherche d'information</i>	29
<i>Reflections on Information Literacy, Information Retrieval and Training</i>	41
<i>L'apport d'Internet à l'organisation des connaissances</i>	49
7.2. PIÈCES À JOINDRE AU DOSSIER.....	54
7.2.1. Carte d'identité	54
7.2.2. Attestation de diplôme	55
7.2.3. Contrat CDD avec le SICD de Toulouse 1	56
7.2.4. Direction des études et enseignements dans le DESS Intelligence économique.....	57
7.2.5. Enseignements CIES.....	58
7.2.6. Enseignements Licence InfoCom	59
7.2.7. Enseignements Ecole d'Orthophonie	60
7.3. RAPPORTS.....	61
7.3.1. Pré-rapport de JP Balpe	61
7.3.2. Pré-rapport de F Rastier.....	63
7.3.3. Rapport de soutenance	65

Remarques introductives.

Mes travaux de recherche peuvent être décrits par trois thématiques principales :

- premièrement, il s'agit de s'interroger sur les **nouveaux modes d'accès aux connaissances** et aux informations dans le cadre d'une société du savoir et/ou de la connaissance qui tente de fonctionner selon le **paradigme de l'hypertexte** ;
- deuxièmement, il s'agit de prendre en compte et d'étudier le **document numérique** comme la pierre angulaire de ces nouveaux modes d'accès, de ces nouvelles pratiques, de ces nouveaux usages de l'information et de cette organisation hypertextuelle et réticulée des **mémoires et des supports documentaires**
- troisièmement, l'étude de la **recherche d'information (Information Retrieval)**, c'est à dire l'ensemble des techniques, outils, et processus mis en œuvre lors de cette activité. Il s'agit plus précisément de déterminer et de mesurer les changements occasionnés par la **mise en réseau** de l'information et leur impact sur les **processus de recherche**, à différents **niveaux d'usage** (du plus novice au plus expert).

De manière plus générique et pour ancrer mes travaux dans le domaine de compétence de ma section, il s'agit d'étudier les **processus, les productions et les usages de l'information**, dans le cadre de **systèmes d'information non-exclusivement documentaires**. C'est pour ancrer mes travaux dans le champ et pour pouvoir échanger et construire autour d'eux des partenariats et des actions de recherche que je me suis tourné vers les laboratoires de recherche mentionnés ci-après.

S'agissant de ma seconde demande de qualification aux fonctions de Maître de Conférences, ces remarques introductives visent à présenter à mes rapporteurs les démarches entreprises pour tenir compte des remarques et critiques apportées à mon premier dossier. Ces remarques portaient sur trois points : la reprise d'éléments fournis dans le rapport de soutenance dans des publications, le rattachement à un laboratoire de recherche "identifié dans le champ", enfin la nécessité de me rapprocher davantage d'études et de méthodologies empiriques, "de terrain".

- *Concernant la prise en compte d'éléments mentionnés dans le rapport de soutenance :*
J'ai tenu compte de ces éléments dans l'ensemble de publications (dont certaines sont jointes à ce dossier) et des actions de recherche mises en œuvre depuis la soutenance de ma thèse.
- *Concernant l'intégration d'un laboratoire de recherche identifié dans le champ :*
 - J'ai contacté JM Salaün de l'**ENSSIB** pour intégrer le **DOCSI** (Laboratoire Document et Sciences de l'information.) Ce rattachement ne pourra être effectif qu'après la présentation de mes axes de recherche devant l'ensemble des chercheurs du laboratoire. Cette présentation se fera lors de la prochaine réunion plénière du laboratoire, au début du mois de février 2004.
 - J'avais également, dès avant de connaître le résultat de mon précédent dossier, contacté le **laboratoire Paragraphe du département Hypermédia de l'Université Paris 8**, et plus particulièrement Mr Jean Clément, qui dirige le **groupe "écritures hypertextuelles"**. Les préoccupations de ce groupe recourent effectivement la plupart de mes axes de recherche. J'ai ainsi pu intégrer comme chercheur associé ce groupe et ce laboratoire.
 - Je suis également entré en contact avec le **LERASS**, laboratoire toulousain en InfoCom, dont je connaissais déjà quelques-uns des membres. A la suite d'une entrevue avec sa directrice, j'acceptai sa proposition de ne pas entrer directement dans une équipe de recherche du LERASS mais de participer comme chercheur invité aux différentes journées thématiques organisées par ce laboratoire.
- *Concernant la nécessité de me rapprocher davantage d'études et de méthodologies empiriques, "de terrain" :*
 - L'une des parties de mon doctorat a servi d'amorce à ce travail. En dehors de développements propres au projet FoRSIC (cf point 2.4.1.), j'ai mis en place **une étude expérimentale portant sur l'analyse des usages en recherche d'information auprès de populations de novices et d'experts** (cf point 2.4.1.4.). Le premier volet de cette expérimentation n'a pu encore donner lieu à publication étant donné que le protocole qui la sous-tend doit être renouvelé cette année. Je ne peux donc pour l'instant produire de publication sur cette question. Néanmoins, ce point ayant été souligné par plusieurs des personnalités rencontrées suite à ma première demande de qualification, je m'autoriserai à entrer dans le détail du protocole expérimental et des premiers résultats produits pour attester de la manière dont je me suis efforcé de tenir compte de ces remarques et donner la trame d'une future publication.

1. Curriculum Vitae.

| Situation |

31 ans / Vie maritale / 2 enfants / Libéré des O.M.

| Adresse & Coordonnées |

10 impasse Jean Cheverry - 31 100 Toulouse
Tél : 05.61.44.04.24. - Courriel : o.ertzscheid@voila.fr

| Formation & Diplômes |

- 1997-02 **Doctorat** Sciences de l'information et de la communication / Sciences du langage. Sous la direction de J. Link-Pezet et de F.C. Gaudard. « **Les enjeux cognitifs et stylistiques de l'organisation hypertextuelle.** » Université des Sciences Sociales - Toulouse / Faculté des lettres du Mirail - Toulouse
- 1996-97 **D.E.S.S.** Bibliographie informatisée et informatique documentaire. « **De l'éveil des techniques à la veille technologique : description transversale d'un système d'information.** » Major de promotion | Mention Très Bien avec Félicitations du Jury / Faculté des lettres de Besançon
- 1993-94 **D.E.A.** « Marques et Masques de l'énonciation dans les ' Textes pour rien ' de S. Beckett » Mention Très Bien / Faculté des lettres du Mirail - Toulouse
- 1993-94 **Maîtrise** « L'infime, l'infini, l'infirme : trilogie théâtrale à l'oeuvre chez S.Beckett » Mention Très Bien / Faculté des lettres du Mirail - Toulouse
- 1992-93 **Licence** de Lettres Modernes / Faculté des lettres du Mirail - Toulouse
- 1990-92 **DEUG** de Lettres Modernes / Faculté des lettres du Mirail - Toulouse
- 1989-90 **Classe préparatoire** Hypokhâgne / Lycée St Sernin - Toulouse

| Expérience professionnelle |

[Activités professionnelles]

- **Octobre 2003 - Octobre 2004 :**
 - **Directeur des études** du DESS "Intelligence économique" de l'université des Sciences Sociales de Toulouse 1
 - CDD mi-temps. **Contrat** auprès du SICD de l'université de Toulouse 1. Poste basé à l'URFIST de Toulouse (Unité Régionale de Formation à l'Information Scientifique et Technique). Missions : Déploiement de la plateforme FoRSIC à l'échelon national.
- **Octobre 2001 - Septembre 2003 : ATER** (1/2 poste) en Sciences de l'information et de la communication
- **Avril 2000 – Septembre 2001 : Chargé d'étude** auprès de la société TRIVIUM (réalisation de logiciels de gestion des connaissances).

[Stages]

- **Avril - Décembre 1997 :** « Guide des compétences et des laboratoires de recherche en Midi Pyrénées. » Ouvrage imprimé répertoriant l'ensemble des laboratoires de recherche publics de la région, édité dans le cadre de l'ADEMAST (Association pour le Développement et la Maîtrise des Sciences et Techniques) sous le patronage du Conseil Régional. Missions : conception éditoriale du produit, travail rédactionnel, stratégie de recherche et de collecte d'information, relations avec les différents partenaires, vente et diffusion du produit, conception, réalisation et maintenance du site Internet correspondant.
- **Décembre 1997 - Juillet 1998 :** Contrat de type CDD dans le même cadre et avec les mêmes fonctions.

| Compétences informatiques |

Outils applicatifs et méthodologies dans les domaines suivants :

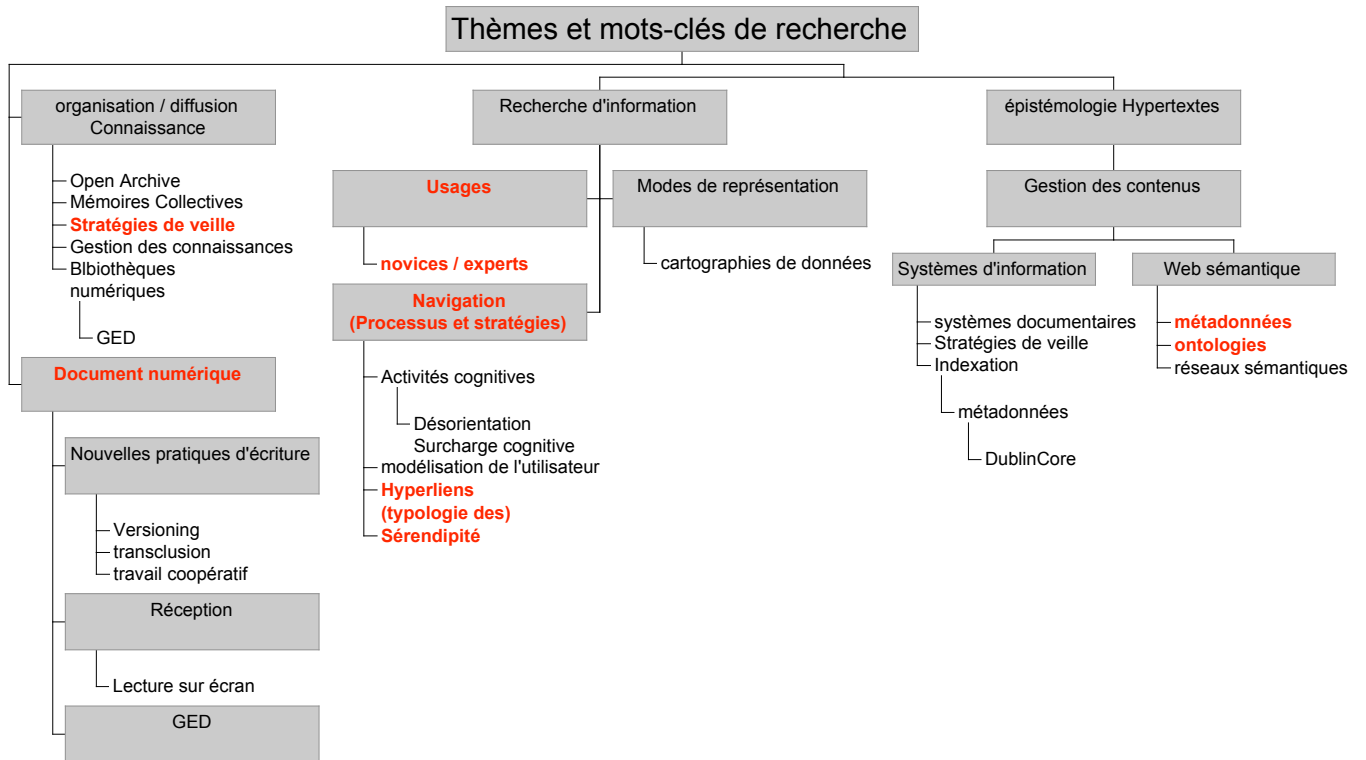
- | **Bureautique** | | Banques et **bases de données** (création & administration) | | **Veille** technologique & stratégique | | **Gestion des connaissances** | | Technologie **Internet** (conception, réalisation et maintenance) | | **Langages** (HTML, XML, Php) |

| Connaissances linguistiques |

- Anglais courant → Allemand (niveau licence)

2. Activités de recherche.

2.1. Vue synoptique.



_____ Thèmes correspondants à des activités de recherche en cours

2.2. Thèmes & mots-clés de recherche

| Actuellement |

Mes activités de recherche portent plus spécifiquement sur l'analyse des **usages en recherche d'information** au travers de scénarios de navigation ainsi que sur la facilitation **de l'accès aux documents numériques** par la construction d'**ontologies de domaine** dans une optique de **gestion et de diffusion de l'information scientifique et technique**.

| Mes axes de recherche |

Il est possible de décliner mes travaux, tels que décrits dans les remarques introductives du présent dossier, à partir de l'étude du **document numérique**. Celui-ci inaugure en effet :

- de nouveaux modes d'accès (recherche d'information) et de traitement (usages) de l'information et des connaissances ;
- de nouvelles pratiques d'écriture sensibles en réception (écrits d'écran, processus de navigation hypertexte) comme en production (niveau éditorial : balises Méta, Dublin Core);

| **Mon travail de doctorat**¹ | aborde ces questions (document numérique et organisation hypertextuelle) dans chacun de ses chapitres.

Le chapitre premier est une **approche épistémologique** qui s'efforce **d'analyser les mutations de l'écrit**, notamment au travers de l'analyse de pratiques et de théories littéraires (générateurs de textes, littérature combinatoire, théories du discours, théorie des genres, pratiques énonciatives).

Le second chapitre traite des questions de navigation et de **recherche d'information** sous l'angle d'une étude systématique des **liens hypertextes** qui propose 41 propriétés pouvant être mises en place dans une perspective de typage qui se veut exhaustive. La question de l'interfaçage de ces propriétés et de leur impact sur les **parcours de navigation** occupant une place centrale².

Le chapitre troisième rend compte de nouvelles **cartographies** qui se donnent à lire sur les **réseaux d'information** (cartographie des acteurs et des documents) et propose un cadre d'analyse (projet FoRSIC : cognition distribuée) pour la mise en place d'accès plus "pragmatiques" aux savoirs se déployant sur ces nouveaux **territoires de connaissance**.

| Mots-clés |

- **Décrivant mes recherches :**

Usages en recherche d'information | Document numérique | Typage des hyperliens pour l'aide à la navigation | Sérendipité | Syndrome d'Elpenor³ | Bibliothèque, édition numérique | Lecture sur écran | Stratégies de navigation | Surcharge cognitive | Cybertextes |

- **Décrivant mes enseignements :**

Veille (stratégies de) | Pratiques de formation en recherche d'information | Recherche documentaire | DataMining | Langages (SGML, HTML, XML)

- **Communs aux deux :**

Gestion des connaissances | Epistémologie des hypertextes / hypermédias | Cartographies d'information | Indexation > Métadonnées > Ontologies | Web sémantique | Recherche d'information | Gestion Electronique de documents |

¹ L'organisation hypertextuelle de la connaissance offre un angle d'approche inédit pour les notions de "document", "d'information" et de "système d'information" ainsi que pour la compréhension des pratiques et des usages qu'ils permettent d'appréhender. Mon travail de doctorat aborde l'hypertexte, non pas sous l'angle de la polysémie qui le fonde (il est en effet tour à tour texte, support, processus, réseau), mais sous celui des modes d'organisation que sa nature associative permet d'engendrer.

² Ces propriétés permettent en effet d'optimiser l'établissement de liens (en enrichissant leurs fonctionnalités), de mettre au point des interfaces adaptées aux usages réels ou attendus de l'information, de faciliter les procédures contextuelles de recherche d'information et d'écriture hypertextuelle, de limiter les problèmes de désorientation et de surcharge cognitive, de prendre en compte les différents contextes (cognitif, sémantique, symbolique) de chaque parcours, de chaque navigation, et ce dans chaque session, de déterminer (et de vérifier empiriquement) une adéquation entre des types de contenus et des systèmes de représentation les servant au mieux, et d'inférer des stratégies de navigation optimales prenant en compte la variabilité des contextes.

³ Ce syndrome est décrit dans la première publication jointe à ce dossier.

2.3. Rattachement à des équipes de recherche ⁴.

2.3.1. Equipe de rattachement

Le laboratoire DOCSI (Document et Sciences de l'Information), membre de l'équipe d'accueil URSIDOC EA 8074, deviendrait, à compter de Février 2004, mon laboratoire de rattachement. Les trois thématiques majeures de ce laboratoire sont les suivantes :

1. Modèles de production et de diffusion de document.
2. Echanges d'information et structuration des organisations.
3. Recherche d'information et appropriation des documents.

Mes travaux de recherche rencontrent principalement la première et la dernière de ces thématiques.

2.3.2. Chercheur associé

Laboratoire Paragraphe | Groupe "Ecritures hypertextuelles" | Dept Hypermédias | Université Paris 8.
Les quatre thématiques majeures de ce laboratoire sont les suivantes :

1. Méthodes et outils d'hypertextualisations et de recherche automatique
2. Méthodes et outils de génération automatique
3. Systèmes de documentation
4. Etude et analyse des impacts et des enjeux sociaux des technologies innovantes.

Mes travaux de recherche rencontrent principalement la première, la troisième et la quatrième de ces thématiques.

2.3.3. Chercheur invité

Laboratoire d'études et de recherches appliquées en sciences sociales (LERASS) | EA 827 | Université Paul Sabatier |

| Remarques |

Je tiens enfin à signaler que mon rattachement à l'URFIST⁵ (via le SICD de Toulouse 1 dans le cadre du contrat à mi-temps dont je bénéficie pour l'année en cours) comporte un volet recherche important autour du projet FoRSIC (cf point suivant) et a été l'occasion de nouer de nombreux partenariats universitaires détaillés ci-après (Laboratoire Travail & Cognition, Institut de Recherche en Informatique de Toulouse, IUFM ...).

2.4. Actions et projets de recherche.

2.4.1. Actions de recherche achevées ou en cours

2.4.1.1. Plateforme FoRSIC

| Descriptif et objectifs scientifiques. | FoRSIC (Formation et Recherche en Sciences de l'Information et de la Communication) est un projet de gestion collective des savoirs d'une communauté de formateurs et d'experts à la recherche documentaire⁶. Via une plate-forme Internet⁷ de mutualisation des connaissances (logiciel SEE-K), permettant la visualisation et la cartographie dynamique d'informations (Arbres de connaissance), nous observons la manière dont se créent des réseaux de coopération et d'échange de savoirs.

| 2003 - 2004 | Cette plateforme (voir le point suivant) a été validée auprès du Ministère par la SDBD (Sous-Direction des Bibliothèques et de la Documentation) pour être déployée à l'échelle nationale. Un plan d'action et de déploiement est actuellement à l'étude. Il s'agit de mettre en place des relais via les différentes URFIST (Rennes et Nice dans un premier temps, puis Bordeaux, Lyon, Paris et Strasbourg). Chaque URFIST utilisera cette plateforme auprès de son réseau local de formateurs et

⁴ Sur ce point, consulter également les "Remarques introductives" du présent dossier.

⁵ URFIST : Unité Régionale de Formation à l'Information Scientifique et Technique.

⁶ sur les campus de Toulouse et Montpellier

⁷ <http://forsic.univ-tlse1.fr> login : guest password : guest

d'experts. Dans la phase initiale de ce déploiement, l'URFIST de Toulouse assurera le pilotage et les actions de formation associées. A terme, il s'agit de mettre en place un portail national de la formation aux usages de l'information, complémentaire du dispositif existant FORMIST, permettant d'enrichir et de qualifier les ressources et supports de formation accessibles en ligne sur ce domaine, mais également de mutualiser les connaissances et les expertises disponibles dans le contexte d'une politique globale de formation des usagers à la recherche et aux usages de l'information.

| 2000 - 2003 |

| Dispositif opérationnel. |

- Constitution dynamique d'une base de connaissance : Les formateurs et les experts ont la possibilité de déposer des **ressources de formation** (tous types de documents, formalisés ou non), et de renseigner un certain nombre d'informations « administratives » les concernant, afin de mieux identifier leur parcours de formation et les réseaux de personnes avec lesquels ils travaillent habituellement. Il est également possible d'exprimer des "besoins" en terme de formation. En plus des différents formulaires disponibles, une grande place est faite aux *zones de saisie en texte libre* pour permettre à chacun de qualifier le plus finement possible les renseignements entrés.

- Cartographie des compétences : Chaque élément constituant la base de connaissance (ressources de formation, fiches individuelles, besoins exprimés), peut être associé par son déposant à une ou plusieurs compétences, à partir de la navigation dans un index facilitant l'accès au référentiel de 250 compétences. Comme pour l'entrée de données et de documents, il est possible pour chaque compétence associée d'ajouter un *commentaire en texte libre*.

Des cartographies d'informations sont ainsi automatiquement générées, en temps réel, sous forme d'arbres de connaissance qui permettent à l'ensemble des participants d'avoir une vue globale des connaissances du collectif. Différents scénarios d'usage⁸ peuvent alors voir le jour, suscitant l'évolution des informations existantes.

L'ensemble des informations entrées en texte libre dans le dispositif peut également être cartographié et visualisé. Il est enfin possible de "croiser" différentes cartographies pour des scénarios d'usage plus fins : ainsi, à partir de la carte des formateurs par compétence, et une fois "repérés" les formateurs pouvant répondre à un besoin déterminé, il est possible de "projeter" leur profil de compétence dans une autre carte associant cette fois les compétences aux ressources de formation déposées, pour distinguer les formateurs disposant "simplement" de telle ou telle compétence et ceux ayant également produit des documents de formation associés.

| **Volet Recherche** | Plusieurs aspects ont été mis en place dans le cadre de la partie recherche de cette plateforme. Ils ont fait l'objet de nombreuses publications que l'on trouvera mentionnées dans la partie idoïne de ce dossier. Nous reprenons ici les points principaux.

- *Référentiel de compétence en formation à la recherche d'information* : Ce référentiel a été élaboré en concertation avec les experts et les formateurs volontaires, enregistrés sur la plateforme. Il se compose de près de 250 compétences, accessibles par le biais d'un index générique (Actions de formation, Bénéficiaires de la formation, Environnement de la formation, Délivrables acquis) réparties en niveaux (culture de base, maîtrise disciplinaire, utilisation experte). Ce référentiel n'est pas "borné" et peut être enrichi sur simple proposition de tout utilisateur enregistré sur la plateforme. Le principal apport méthodologique de ce référentiel est qu'il est interrogeable par les experts ("Je sais" *faire telle chose*) aussi bien que par les novices ("Je veux apprendre à" *faire telle chose*). Il se démarque donc clairement des référentiels existants parce qu'il est **à la fois** un référentiel professionnel ET pédagogique.

- *Modélisation du champ de la formation à la recherche d'information sous forme de réseau sémantique*. Ce réseau comporte une liste de plus de 500 termes déployés sur une arborescence à 7 niveaux de profondeur. Deux types de relations permettent d'associer des termes : "*nécessite la connaissance de*" et "*est utile à la connaissance de*."

- *Sériation et validation d'invariants pour l'aide à la formation ("Information Retrieval Acts")* : Ces "fondamentaux" se déclinent différemment selon le niveau de formation auquel ils s'appliquent. La modélisation que nous proposons permet, pour chaque niveau de formation, de croiser ces fondamentaux avec des activités cognitives les caractérisant ainsi qu'avec des intentions pédagogiques permettant d'atteindre les objectifs du niveau susmentionné.

⁸ il est par exemple possible de voir l'arbre des formateurs par compétence, celui des ressources de formation par compétence, de visualiser son profil dans l'ensemble des formateurs, de repérer les compétences "fondamentales", les plus utilisées et/ou les plus expertes.

| Partenaires & financement | Projet national soutenu par le Ministère de l'éducation et de la recherche. N° de contrat 99-K-6403 | *Maître d'œuvre : URFIST de Toulouse.* | *Partenaires Universitaires :* IRIT (Institut de Recherche en Informatique de Toulouse, équipe SMAC - Systèmes Multi-Agents Coopératifs), LIHS (Laboratoire Interaction Homme-Systèmes) | *Partenaires industriels :* Société TRIVIUM |

2.4.1.2. SABRE (Système Auteur Basé sur la Réutilisation)

| Descriptif du projet |

SABRE est un **outil-auteur pour la production de ressources pédagogiques adaptatives** et modulables. Il a été développé dans le cadre du projet FoRSIC auquel il s'intègre comme application complémentaire. Il est actuellement disponible sous forme de prototype. Cet outil **fonctionne sur la base d'un réseau sémantique** permettant d'intégrer - notamment - les intentions pédagogiques de l'auteur de la ressource, ses objectifs pédagogiques, les pré-requis nécessaires, le type de cours souhaité. Ce **réseau sémantique** est composé d'une liste de plus de 500 termes (avec leurs définitions) représentatifs du vocabulaire du champ de la recherche documentaire.

| Dispositif opérationnel |

Il est possible d'entrer, au format texte, des ressources de formation complètes (ou des parties de ressources pédagogiques), de les enrichir (adaptation) en puisant dans la bibliothèque de ressources disponibles, et d'établir entre elles des liens sémantiques sur les intentions, les pré-requis, les objectifs, etc. Chaque ressource ou partie de ressource peut également être associée avec un ou plusieurs termes du réseau sémantique.

Au final, la ressource entrée est formatée en **XML** et affichable dans n'importe quel navigateur web (via une feuille de style). Il est alors possible de faire une recherche dans la bibliothèque des ressources, non plus seulement par mot-clé ou équation booléenne, mais également par « intention », par « objectif » ou par « pré-requis », tous ces critères apparaissant dans les métadonnées du document XML produit.

Développé comme une application indépendante à l'aide des résultats de recherche de la plateforme FoRSIC (réseau sémantique, "Information Retrieval Acts", etc.), l'outil-auteur SABRE est désormais réalisé et a été testé auprès d'une population de formateurs. Il sera intégré dans le projet FoRSIC comme module indépendant.

| Partenaires | URFIST, LIHS (Laboratoire d'Interaction Homme-Système) de Toulouse 1.

2.4.1.3. Usages experts & novices en recherche d'information

| **2002 - 2004** || **Partenaires** | SCD (Service commun de documentation) de Toulouse 1 ; André Tricot (psychologue-cogniticien), formateur IUFM, membre du LTC (Laboratoire Travail et Cognition) de Toulouse 2.

| **Descriptif** | Nous cherchons à disposer d'une observation fine des pratiques de recherche d'information de formateurs, de bibliothécaires et/ou d'experts ainsi que de novices pour valider plusieurs hypothèses concernant les compétences mises en œuvre dans ces pratiques. Il s'agit donc :

- d'analyser les pratiques d'utilisateurs non-experts (étudiants) (EXP 1) ;
- de faire « émerger » les meilleures pratiques dans un cadre qui est celui de l'analyse des activités expertes (EXP 2).

(EXP 1) Analyse des pratiques d'utilisateurs non-experts.

- *Population cible* : Etudiants en première année de DEUG de Droit.
- *Nombre d'individus entrant dans l'étude* : 250
- *Traits communs* : Novices⁹ dans leur domaine d'étude | Novices en recherche d'information.
- *Terrain* : Nous rencontrons ces étudiants lors du premier cours de "Méthodologie de la recherche documentaire"¹⁰ dispensé par les PRCE¹¹ de l'université de Toulouse 1.
- *Méthodologie* : Nous avons défini deux grandes catégories d'environnement pour des tâches de recherche d'information :
 - (EF) Environnements fermés / structurés : OPAC + Site web déterminé.
 - (EO) Environnements ouverts / non-structurés : Internet à partir d'un moteur de recherche.

Pour chaque type d'environnement, nous avons proposé aux étudiants, en binôme, de réaliser deux tâches prescrites en temps limité (15 minutes).

- (EF) :
 - Tâche 1 : "Constituez une liste des 5 ouvrages récents sur le droit de la propriété intellectuelle disponibles à la bibliothèque de l'Arsenal." | N réponses possibles.
 - Tâche 2 : " Sur le site web www.civilites.com, trouvez la page décrivant l'utilisation du code civil." | 1 réponse possible.
- (EO)
 - Tâche 3 : "Sur Internet, trouvez 5 sites vous permettant de vous perfectionner dans l'exercice de la dissertation juridique." | N réponses possibles
 - Tâche 4 : "Sur Internet, trouvez le texte de la circulaire Chevènement sur les sans-papiers." | 1 réponse possible.

L'ensemble de l'activité de recherche est enregistrée par le serveur de l'université. Nous récupérons ensuite les fichiers logs de l'ensemble des sessions de consultation. A l'issue de chaque session, une verbalisation de la tâche est demandée aux étudiants. L'ensemble du protocole est anonymé. Seules les adresses IP sont conservées.

- *Données d'analyse* : fichiers logs + verbalisation de la tâche (imprimés distribués)
- *Variables mesurées* : Performance¹² | Stratégies de recherche (utilisation des différentes fonctionnalités de l'interface) | Temps.

(EXP 2) Analyse des activités expertes.

- *Population cible* : Documentalistes et Bibliothécaires experts en recherche d'information.
- *Traits communs* : Experts en recherche d'information |
- *Composition du panel* : 3 experts dans trois domaines de spécialité différents (Droit, Biologie, Sociologie).
- *Tâche prescrite* : "Quel est l'impact de la crise de l'ESB dans votre domaine de spécialité".
- *Terrain* : L'activité des experts a été filmée sur leur lieu d'exercice. Elle a donné lieu à une verbalisation simultanée.

⁹ nous considérons comme novices les étudiants n'ayant jamais reçu de formation en droit et/ou en recherche d'information.

¹⁰ Ce cours est généralisé à tous les étudiants en première année de DEUG de l'université et comprend 10 heures de formation.

¹¹ Professeurs certifiés de Documentation.

¹² (Tâche 1) nb de références pertinentes notées sur papier | nb de références pertinentes sélectionnées | nb total de références sélectionnées | nb de références différentes sélectionnées.

- *Objectifs* : faire « émerger » les meilleures pratiques dans un cadre qui est celui de l'analyse des activités expertes, pour permettre de construire des scénarios de formation modulables en rapport avec les pratiques des usagers de l'information.
- *Résultats de l'étude* : "Etude des compétences mises en œuvre par les experts en recherche documentaire et application pour la formation." Mémoire de DESS Psychologie Cognitive, Juin 2003, Université de Toulouse le Mirail.

Cette étude a notamment permis :

- de valider un certain nombre de modèles mis en avant dans le projet FoRSIC (Actes documentaires, Référentiel de compétences professionnelles et pédagogiques);
- de caractériser très finement la notion d'expertise en recherche d'information au regard du cadre théorique de l'analyse des activités expertes en psychologie cognitive;
- de dégager des recommandations pour la construction de scénarios de formation et de ressources pédagogiques selon différents niveaux d'usage et de compétence.

2.4.1.4. Groupe URFIST "Pratiques de Formation".

| **2003 - 2004** | Un groupe de travail s'est mis en place cette année dans le cadre du réseau URFIST. Il est constitué d'enseignants-chercheurs et de conservateurs. Ce groupe réfléchit à la mise en place d'un **schéma d'enseignement commun** pour la formation des usagers à la recherche d'information dans le cadre du **LMD**. Voici quelques-unes des activités du groupe :

- Construction d'un référentiel de compétences commun.
- Analyse de ressources pédagogiques utilisées en formation.
- Schéma directeur d'un enseignement interdisciplinaire pour la "Formation à la maîtrise de l'information."

2.4.2. Projets de recherche

2.4.2.1. DYNAMO (Dynamic Ontology)

Le projet **DYNAMO** a comme objectif la "Conception et maintenance automatique d'une ontologie à partir de corpus de textes." Il s'agit d'automatiser le processus de conception d'ontologie à partir de corpus et le processus de maintenance de cette ontologie dans un contexte d'utilisation. Le domaine d'expérimentation choisi est celui des sciences de l'information et de la communication avec une application à la formation à la recherche d'informations.

Il a été présenté dans le cadre du programme interdisciplinaire **TCAN** (Traitement des Connaissances, Apprentissage et NTIC) du CNRS et retenu par le comité de pilotage du programme comme pré-projet¹³. L'IRIT est le porteur du projet. L'URFIST de Toulouse et de Nice sont partenaires.

2.4.2.2. Documents électroniques pour les apprentissages humains

J'ai été membre pendant deux ans de l'**atelier PRESCOT** "Documents électroniques et apprentissage". Cet atelier a débouché sur la proposition d'une équipe projet multilaboratoires (EPML) toulousaine, coordonnée par André Tricot, et centrée sur l'étude des "Documents électroniques pour les apprentissages humains", dans le cadre du **RTP 39** "Apprentissage, éducation, formation" du CNRS. Je suis impliqué dans l'un des axes de ce projet qui concerne l'étude de la "motivation, des buts d'apprentissage et de la recherche de feedback dans les documents électroniques".

2.4.2.3. Scénarios de navigation et recherche d'information.

| 2004 - ... | | **Descriptif du projet** |

Ce projet a comme base la typologie des liens proposée dans mon travail de thèse. A partir de celle-ci, des hypothèses ont été formulées concernant l'influence de certains types de liens sur certains comportements de navigation dans une tâche de recherche d'information. Ces hypothèses n'ont pour l'instant été validées que sur la base d'un corpus d'hyperfictions (hypertextes littéraires). Il s'agira donc cette fois de tester la robustesse de ces hypothèses sur un corpus plus générique.

¹³ <http://www.dr4.cnrs.fr/tcan/tcan/projets/2003/preprojetsacceptes.html>

Composition du corpus : il comprend trois catégories de sites webs permettant de tester trois archétypes de scénarios navigationnels :

1. Portails institutionnels : 15 sites.
 2. Environnements 3D ou flash : 10 sites.
 3. Bases de données « on-line » : 5 sites.
- *Opérationnalisation* : le corpus est pour l'essentiel constitué. Chacun des sites retenus sera analysé pour en extraire l'ossature des liens le composant. Ceci sera rendu possible par l'utilisation d'outils standard de création de pages web (éditeurs Wysiwyg), d'outils de création de pages flash et d'outils de création d'hypertextes (Storyspace). Sur la base de cette ossature, nous croiserons alors les différents types d'organisations et de hiérarchies de contenus avec notre typologie initiale. Il sera ainsi possible de valider le choix de tel ou tel type d'organisation des contenus en regard avec les objectifs navigationnels affichés par le site et à l'aide des modèles de comportements des usagers en situation de recherche d'information dont nous disposons.

2.4.3. Participation et affiliation à différents réseaux de recherche

2.4.3.1. RTP - DOC

Participation aux activités du RTP-DOC " Documents et contenu : création, indexation, navigation."
Contribution au texte collectif de réflexion sur le document numérique : "**Document : forme, signe et relation, les re-formulations du numérique.**"

2.4.3.2. ACM SIGWEB

"Special Interest Group on Hypertext, Hypermedia and the Web." Je suis depuis 4 ans membre de ce groupe.

2.4.4. Projets Industriels.

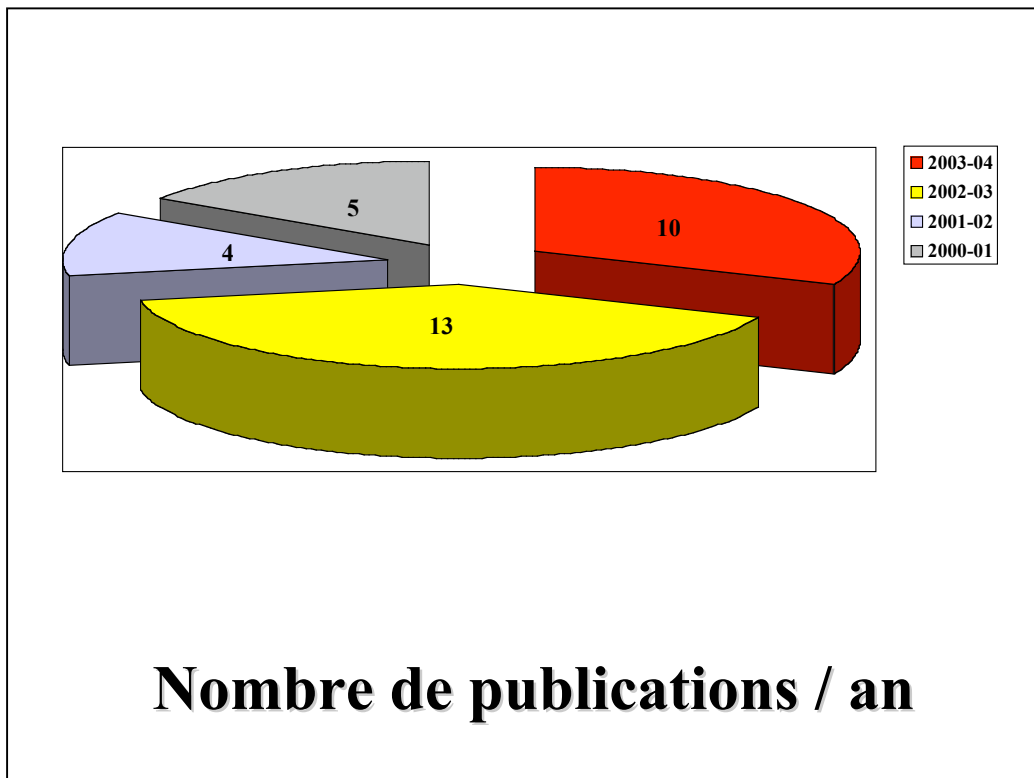
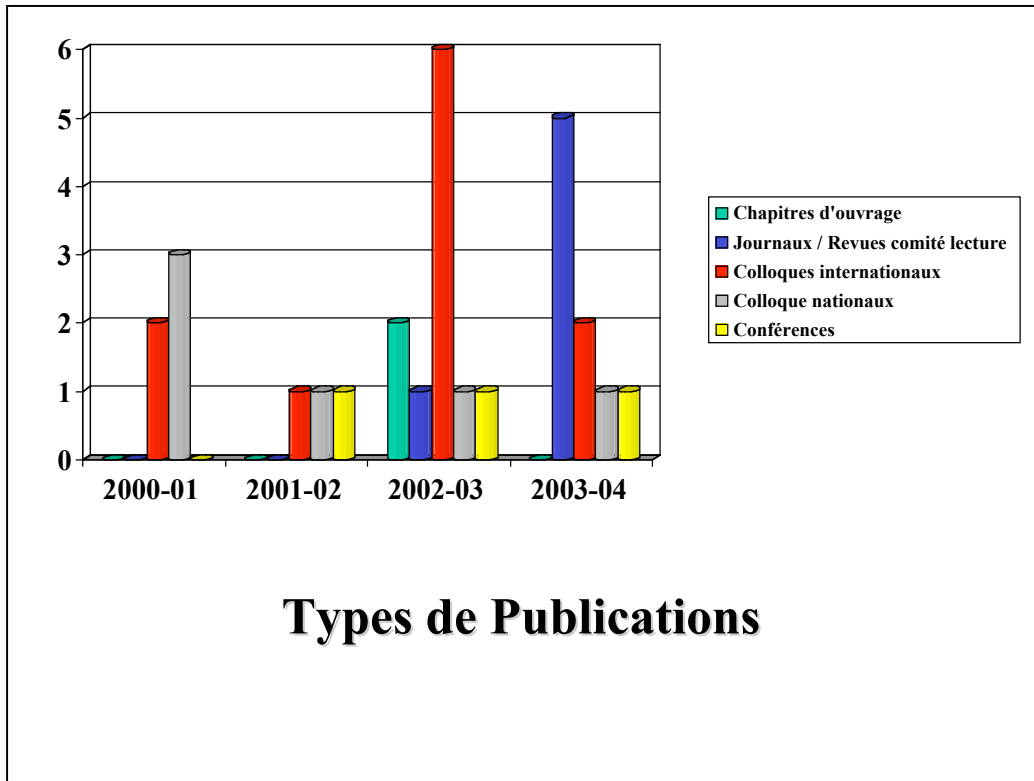
| 2000 - 2001 | **Projets C-2001 (Connaissance 2001)** : Ce projet porte sur l'amélioration du logiciel SEE-K dans un environnement X-Net. La mission confiée à l'URFIST concerne l'ergonomie de la plate-forme SEE-K, et la portabilité et la méthodologie générale de conception de la matrice d'indexation des compétences.

| **Partenaires** | URFIST de Toulouse, Société TRIVIUM, CEA.

| **Délivrables & Missions** | J'ai rédigé le rapport final de ce projet. La mission qui m'a été spécifiquement confiée en complément de celle confiée à l'URFIST reposait sur la mise en place d'une veille technologique sur plusieurs aspects du projet.

3. Publications.

3.1. Vue synoptique.



3.2. Chapitres d'ouvrage.

| 2002 - 2003 |

- Ertzscheid O. Link-Pezet J., "*Reflections on Information Literacy, Information Retrieval and Training.*", pp. 88-103, in *Information Literacy in Europe : A First Insight Into The State of the art of Information Literacy in the European Union*¹⁴, éd. par Basili Carla. Rome 2003. IT ISSN 0085-2309
- Ertzscheid O., Lacombe E., Link-Pezet J. , "*Chercher de l'information : outils, usages et pratiques*", in REPERE (Ressources Electroniques Pour les Etudiants, la Recherche et l'Enseignement.) | en ligne | <http://reperere.enssib.fr>

3.3. Journaux et revues à comité de lecture.

| 2003 - 2004 |

- Ertzscheid O., « *Stratégies de navigation multimedia : les parcours de l'errance.* » Proposition de communication acceptée pour les **Cahiers de Champs Visuels**. Numéro spécial "Le parcours multimédia."
- Ertzscheid O., "*Lisible, scriptible, visible : le rôle à jouer de l'image dans la communication de l'hypertexte.*" En cours de soumission. Pré-proposition acceptée avec modifications. Revue **COMPOSITE**.
- Ertzscheid O. « *Écran, image, interface : vers une sémiotique de la cognition.* », à paraître in **VISIO**, Revue Internationale de Sémiotique Visuelle, volume 8, 3-4 (vol. double), automne 2003-hiver 2004, "Le visuel à l'ère du post-visuel / The Visual in the Age of the Post-Visual", Marie Carani (sous la dir. de)
- Ertzscheid O., "*De la note de bas de page au lien hypertexte : philosophie de l'identique et stylistique de l'écart.*" à paraître in "*Espace de la note*", Revue **La Licorne**. ISSN : 0398-9992
- Ertzscheid O., "*Pratiques énonciatives hypertextuelles : vers de nouvelles organisations mémorielles.*" in **Archée**. | en ligne | <http://archee.qc.ca>

| 2002 - 2003 |

- Ertzscheid O., "*Du texte fragment à l'hypertexte fractal.*", pp. 125-137 in *Champs du Signe*, n°15. Presses Universitaires du Mirail.

3.4. Colloques internationaux¹⁵.

| 2003 - 2004 |

- Ertzscheid O., Gallezot G. "*Formalising the Concept of Serendipity in Web Searching.*" Conférence Internationale Search Engine Meeting 2004. La Hague. 19-20 Avril 2004.
- Ertzscheid O., Gallezot G., "*Des machines pour chercher au hasard (moteurs de recherche et recherche d'information).*" Proposition de communication retenue, 14^{ème} congrès de la SFSIC, Questionner l'internationalisation : Cultures, Acteurs, Organisations, Machines, Istanbul, Université Galatasaray, 3, 4 & 5 juin 2004.

| 2002 - 2003 |

- Ertzscheid O. "*Syndrome d'Elpenor et serendipité : deux nouveaux paramètres pour l'analyse de la navigation hypermedia.*" in *Créer du sens à l'ère du numérique*, Actes du Colloque Hypertextes et hypermédias, H2PTM'03, Editions Hermès, Paris, 2003.
- Ertzscheid O., Gallezot G., "*Chercher faux et trouver juste : sérendipité et recherche d'information.*" 1ère Conférence Internationale Francophone en Sciences de l'Information et de la Communication (CIFSIC). 10ème colloque bilatéral Franco-Roumain. Bucarest, 28 Juin - 3 Juillet 2003.
- Ertzscheid O., "*Génération de textes et hypertextes. De la logique de l'œuvre à celle de ses versions.*" Colloque international Textes en performance, Genève, 27-29 Novembre 2003.

¹⁴ L'objectif de l'ouvrage est de dresser un état des lieux des compétences et connaissances et recherche d'information pour chaque pays européen. Il est financé par le Conseil National de la Recherche italien dans le cadre de l'Institut d'études économiques sur l'innovation et la politique de recherche.

¹⁵ Sauf mention particulière, toutes les communications citées ont été ou seront publiées dans les actes de ces colloques.

- Ertzscheid O. "L'apport d'Internet à l'organisation des connaissances : vers un encyclopédisme d'usage." Colloque ISKO 2003, 4^{ème} Congrès du Chapitre Français de l'ISKO, Grenoble, 3-4 Juillet 2003.
- Ertzscheid O., Link-Pezet J., "De l'ingénierie des connaissances à celle des compétences : construction d'une ontologie pour la recherche d'information." Colloque ISKO 2003 4^{ème} Congrès du Chapitre Français de l'ISKO, Grenoble, 3-4 Juillet 2003.
- Ertzscheid O. "Comportements de navigation et documents électroniques : proposition d'invariants". pp. 137-148 in Actes du sixième Colloque International sur le Document Electronique : CIDE 6, 24 - 26 Novembre 2003, Université de Caen. Paris, Europa 2003.

| 2001 - 2002 |

- Ertzscheid O., Link-Pezet Jo "An ontology-based knowledge management system for a community of practice.", in **SCI / ISAS 2002**, 6th World Multi Conference on Systemics, Cybernetics and Informatics. Orlando (USA), 14-17 Juillet 2002.

| 2000 - 2001 |

- Ertzscheid O. « An attempt to identify and manage collective practices involved in information retrieval. », in **SCI / ISAS 2001**, 5th World Multi conference on Systemics, Cybernetics and Informatics. Orlando (USA), 22-25 Juillet 2001.
- Ertzscheid O., Lacombe E., Link-Pezet J. « FoRSIC : un dispositif innovant pour la formation de masse des usagers à la recherche documentaire. » in **Des ressources pour produire –** IIIèmes journées de l'innovation de Midi-Pyrénées – Colloque international francophone – 24 – 26 Janvier 2001, Foix.

3.5. Colloques nationaux (CN) & Journées d'étude (JE).

| 2003 - 2004 |

- (JE) Ertzscheid O., « A la recherche du document numérique : tentative de sériation des nouveaux modes d'accès aux documents. » **Journées d'étude Document Numérique.** Laboratoire Paragraphe, Université Paris 8. 15 Janvier 2004.

| 2002 - 2003 |

- (CN) Ertzscheid O., Link-Pezet Jo, . « La gestion des contenus pour l'aide à la formation. » Colloque National "Formation à la maîtrise de l'information et mutualisation de connaissances", dans le cadre des **Rencontres Nationales des URFIST**, Toulouse, 19 Novembre 2002.

| 2001 - 2002 |

- (CN) Ertzscheid O., Link-Pezet Jo, "Du savoir individuel à la connaissance collective le projet FoRSIC." Communication affichée, **IC 2002**, 13^{ème} journées francophones d'ingénierie des connaissances, Rouen, 28-30 Mai 2002.

| 2000 - 2001 |

- (CN) Ertzscheid O., Link-Pezet J. « FoRSIC : mise en œuvre d'un projet de gestion des connaissances. » pp. 307 à 320 in **Émergences et continuités dans les sciences de l'information et de la communication** – Actes du XIIème congrès national des sciences de l'information et de la communication – Paris, Unesco, 10-13 Janvier 2001.
- (CN) Ertzscheid O., Lacombe E., Link-Pezet J. « Les cycles de l'interaction pour l'accès aux connaissances en sciences de l'information et de la communication : l'exemple du dispositif FoRSIC. » in **Pratiques de situations de communication et N.T.I.C.** – IIIème colloque du Groupe de Recherche en Information et Communication, Montpellier, 30 Novembre, 1er Décembre 2001.
- (JE) Ertzscheid O., "Actes de lecture et cognition : nouvelles approches coopératives." In Les actes de lecture : nouveaux enjeux, nouvelles pratiques, XXXIème journées d'étude du **CEBRAL**, Cahors, 5-6 Mai 2001

3.6. Journaux et revues sans comité de lecture.

| 2002 - 2003 |

- Ertzscheid O., "L'hypertexte : haut-lieu de l'intertexte.", in **La revue des ressources**, [Création littéraire > Hyper-littérature], 31 Octobre 2002. [en ligne] <http://www.larevuedesressources.org/>

3.7. Collaborations.

| 1998 - 2004 |

- " La lettre de l'URFIST" ISSN 1257-2268. Lettre d'information trimestrielle. Collaboration régulière (Articles, Revue de presse, Compte-rendus de colloques ...)

| 1996 - 1997 |

- "Guide des compétences en recherche de Midi-Pyrénées." Guide pratique édité par L'ADEMAST (Association pour le Développement et la Maîtrise des Sciences et Techniques) Rédaction et mise en forme de textes transversaux.

| 1995 - 1996 |

- "Bibliographie de la philosophie.", revue trimestrielle sous le patronage du CNRS (VRIN éditeur). Critiques, Abstracts.

3.8. Conférences et participations invitées.

| 2003 - 2004 |

- « **L'hypertexte : domaine de recherche ou "simple" carrefour de problématiques ?** » Conférence invitée aux Rencontres Médias du CNAM. Paris, 17 Novembre 2003.

| 2002 - 2003 |

- « **Les défis de la publication sur le web : hyperlectures, cybertextes et méta-éditions.** » Colloque virtuel. Participation invitée.

| 2001 - 2002 |

- « **Internet sera-t-il l'abbaye de Thélématique du XXIème siècle ?** » Conférence invitée sur le thème : « L'enseignement à l'heure d'Internet. », Stage résidentiel du CIES de la région Centre. La Rochelle, 26 Mars 2001.

3.9. Diffusion des connaissances.

| 2002 - 2003 |

- Participation au SITEF (Salon International de l'Innovation et de la prospective) à Toulouse. Présentation du projet de recherche FoRSIC. Démonstration de la plateforme de gestion des connaissances associée au projet.

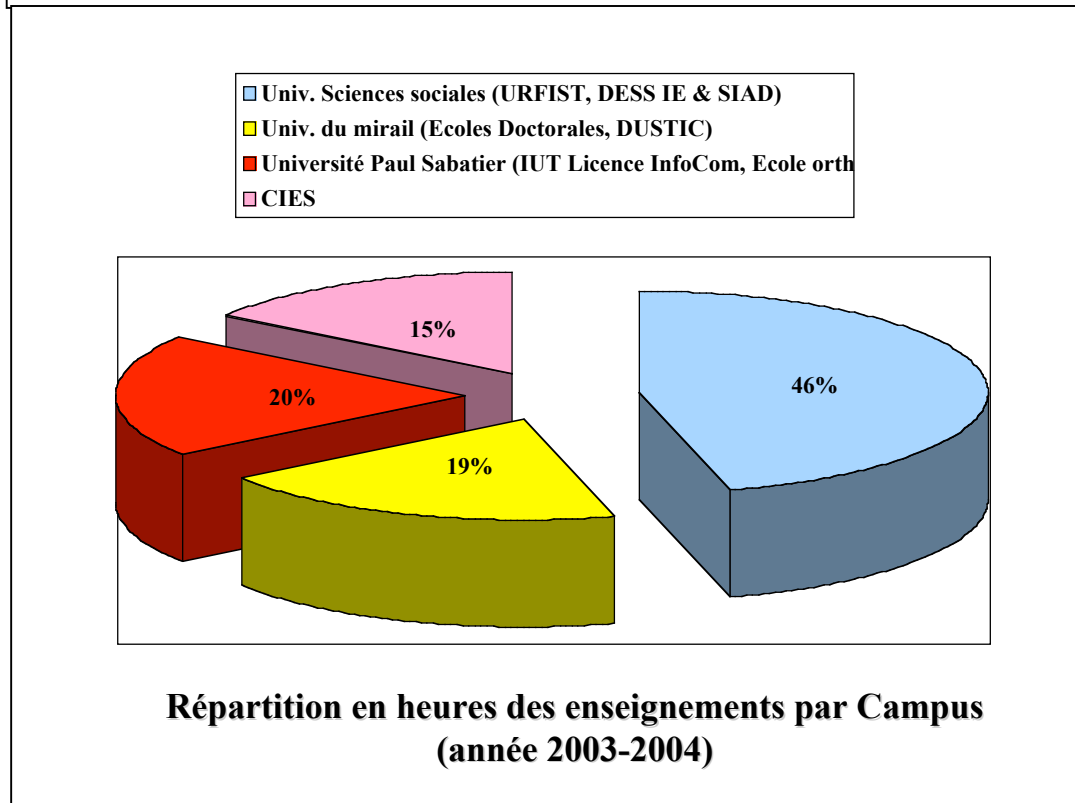
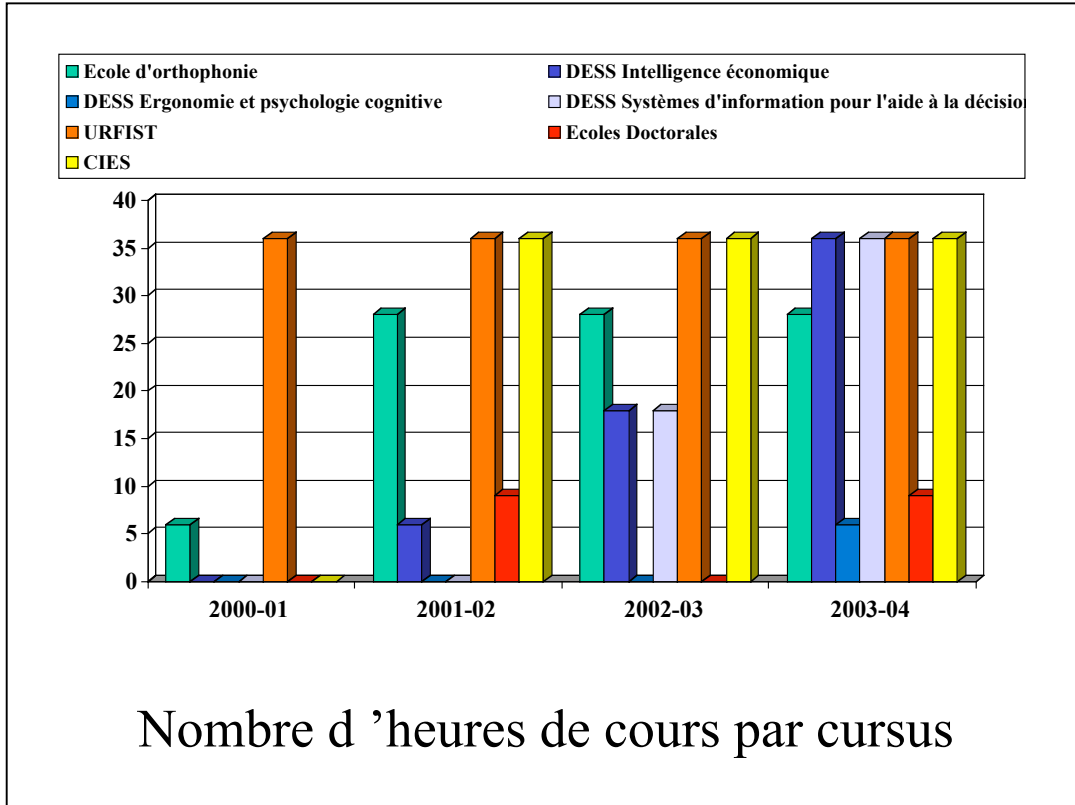
3.10. Soumissions en cours.

- | **Revue I3** | Revue Information-Interaction-Intelligence | Numéro sur la notion de document.
- | **CAIS / ACSI 2004** | Congrès annuel de l'Association Canadienne des Sciences de l'Information. | **Thème du congrès** | L'accès à l'information : technologies, compétences et contexte socio-politique.
- | **IC 2004** | 14^{ème} conférence francophone sur l'Ingénierie des Connaissances.
- | **Hypertext 2004** | 15^{ème} Conférence Internationale de l'ACM sur l'Hypertexte et les Hypermédias.

4. Activités d'enseignement

| DESS IE, SIAD | URFIST | Licence InfoCom | Orthophonie | CIES | DUSTIC | E. Doctorales | Points à souligner |

4.1. Vues synoptiques



4.2. D.E.S.S. IE & SIAD¹⁶

2001 - 2004	Volume Horaire annuel	36 heures IE, 36 heures SIAD.
	Nombre de formés / an	60 personnes (30 IE, 30 SIAD)
	Type d'intervention	2/3 Cours Magistraux, 1/3 TD

| **Objectifs Pédagogiques** | L'enseignement que je dispense dans ces deux formations varie sensiblement en fonction des profils d'étudiants qui les composent. Le cœur de l'enseignement a cependant été construit pour pouvoir être "porté" dans ces deux contextes (IE et SIAD). C'est ce cœur d'enseignement qui est ici détaillé, les différences principales se manifestant essentiellement dans le cadre des TP et des livrables demandés aux étudiants pour leur évaluation.

- Former (au travers de la création, de la montée en charge et de la maintenance de **sites web**) aux problématiques ayant trait aux **normes, aux langages** et aux architectures disponibles dans un cadre « réseau » (indexation, référencement, portails dynamiques, environnement X-Net, DSI, langages – html, xml).
- Présenter les **sources d'information** utilisables dans leur spécialité. Les former aux techniques et aux outils permettant d'initier et de conduire des **stratégies de veille** dans leur domaine.
- Former aux techniques avancées de représentation (**cartographie**) de l'**information** sur un plan pratique (outils et solutions logicielles) et technique (méthodologie de recueil, procédures d'indexation).
- Présenter quelques-unes des problématiques clés et des approches (méthodologies) utilisées en **gestion des connaissances** (approche compétence, vocabulaires métiers, conversion des connaissances ...).

| Contenus de formation |

- **"Du site web aux places d'affaires électroniques (portails)."**
 - *Objectifs* : Il s'agit de permettre aux étudiants de maîtriser les langages et processus de conception d'un site web institutionnel (de type Intranet), d'un point de vue aussi bien technique qu'ergonomique. Une grande place est faite aux différents moyens d'optimiser les aspects dynamiques des sites web (forums, référencement ...)
 - *Contenu* : Ce cours s'articule autour de trois axes principaux :
 1. Création de site et référencement. Conception, création et maintenance de sites portails (langages de développement – HTML, XML, PHP, My SQL), architecture de contenus dynamiques, ergonomie, référencement et positionnement, démarche marketing) ; *TP* : constitution d'un cahier des charges puis création du site web de la promotion.
 2. Mise en œuvre et déploiement d'un Intranet (feuille de route, montée en charge, pérennisation) - Gestion de l'information (sourcing, dsi ...). *TP* : Manipulation et reverse engineering d'une plateforme Intranet en ligne.
 3. Le rôle des places d'affaires électroniques (business-models, typologie des contenus, architecture logicielle et briques élémentaires ...)
- **"Exploitation des informations et stratégies de veille."**
 - *Objectifs* : Il s'agit de dresser un panorama complet des techniques et outils de recherche d'information et de veille en environnement X-net et dans les banques et bases de données.
 - *Contenu* : Ce cours s'articule autour de trois axes principaux :
 1. Historique et Méthodologie des différents types de veille (démarche, organismes, ressources, mise en place d'une cellule de veille en entreprise, cycle de l'IST, Cas d'étude sur des outils et méthodologies, Causes et dangers de la non-maîtrise de l'information dans l'entreprise.)
 2. Techniques de veille sur Internet (outils, sites ressource, moteurs de recherche, web invisible ...) Serveurs et Banques de données en ligne / Périodiques électroniques : de l'interrogation des sources à l'exploitation de l'information (présentation et fonctionnalités des outils, syntaxes avancées d'interrogation, équations de recherche, scénarios d'usage). Processus général de collecte d'information (identification des sources, collecte, évaluation des sources) ; Sources informelles (méthodes et outils d'identification, méthodes de recueil d'information, gestion et capitalisation) ;

¹⁶ DESS IE : "Intelligence économique" | DESS SIAD : "Systèmes d'Information pour l'Aide à la Décision."

3. Typologie des sources (Papier, électroniques – banques de données, cédéroms – Internet) ;
- TP (recherche d'information sur Internet, interrogation de serveurs [Questel & Dialog]). Fondements théoriques de la recherche d'information ; Tendances de l'industrie et demande du marché ; Informations économiques, politiques et administratives (sources) ; Prédire l'action d'une entreprise (indicateurs externes et internes) ; Présentation détaillée de l'organisation, des contenus et de l'interrogation de DialogWeb ; TP : interrogation de différentes sources.
 - **"De l'interprétation des informations à la connaissance de l'entreprise : Knowledge management et cartographies de connaissances dans les systèmes d'information distribués"**
 - *Objectifs* : l'objectif de ce cours est de présenter aux étudiants les derniers outils de cartographie et de gestion de l'information, et leurs spécificités au regard de leurs principaux usages en Knowledge Management.
 - *Contenu* :
 1. Historique et notions clés de la gestion des connaissances.
 2. Outils et méthodologies permettant d'obtenir des représentations structurées et dynamiques sur de vastes ensembles de données, concernant le capital humain et/ou documentaire de l'entreprise.
 3. Cas d'étude d'une solution logicielle et des ses applications (See-K, société Trivium). Indexation et systèmes de représentation : de l'indexation documentaire à la mise en œuvre d'une démarche de capitalisation des connaissances pour l'entreprise. Rapports existant entre différents niveaux d'indexation présents dans l'information structurée jusqu'à la construction d'ontologie pour exploiter de l'information non-structurée (capital humain). L'accent de cette partie du cours est mis sur l'exemplification d'une démarche de mutualisation et de capitalisation de connaissances au travers d'un cas concret (projet FoRSIC) et illustré par la manipulation des outils de la société TRIVIUM (SEE-K).
 4. Fondements théoriques (cognition distribuée, conversion de connaissance) ; Notions de bibliométrie ; Mapping ; DataMining ;
 5. L'information : processus stratégique pour l'action (outils dédiés – Henoach, Sampler, Umap, Arisem, Sinequa, Datops, Tétralogie ...)
 6. Méthodologie de la gestion des connaissances (approche compétence, cartographie de connaissance, objectifs, domaines d'application, technologies – indexation, ontologies, workflow, groupware ...) ; TP : élaboration et formalisation d'un réseau sémantique de l'aide à la décision ; Validation de ce réseau selon différentes méthodologies de capitalisation.

Remarques : Je prends en charge, depuis maintenant quatre ans, la semaine d'intégration des étudiants du DESS I.E., au cours de laquelle ils sont amenés à travailler en autonomie pour procéder à la validation des pré-requis demandés et amorcer un certain nombre de projets sur lesquels ils travailleront durant l'année.

4.3. Licence "Information & Communication".

2001 - 2004	Volume Horaire annuel	18 Heures.
	Nombre de formés / an	30 personnes.
	Type d'intervention	TD.

| **Contenus de formation** | « Réseaux et multimédia. »

Informatique en ligne (architecture client/serveur, réseau local, espace personnel, typologie des ressources en ligne) ; Architecture des services (banques de données, ftp ...) ; Ecriture multimédia (écriture hypertextuelle, notion d'interactivité, ergonomie, synchronisation de contenus dynamiques) ; Analyse des produits et services multimédia (constitution de cahier des charges) ;

4.4. **URFIST de Toulouse.**

1996 - 2004	Volume Horaire annuel	36 heures ¹⁷ .
	Nombre de formés personnes	(Professionnels, Personnels administratifs, Bibliothécaires, Enseignants)
	Type d'intervention	1/3 Cours Magistraux, 2/3 TD

| Contenus de formation |

« Langage HTML et création de sites web. »

Historique du langage (SGML). Contexte normatif (organismes de régulation). Norme HTML (balises, syntaxe, DTD). Langages associés (DHTML, JavaScript). Perspectives et évolutions (XML). Solutions logicielles dédiées (éditeurs de balises, éditeurs wysiwyg – FrontPage, Mozilla Composer – éditeurs professionnels – Dreamweaver, Golive). Démonstrations et réalisation d'une page web (TD).

« Publication sur Internet. »

Conception d'une architecture (arborescence). Règles d'ergonomie. Typologie des publications. Référencement (balises méta). Intégration de matériel multimédia (formats, plug-in, logiciels dédiés). Aspects juridiques (CNIL, gestion des liens externes, cookies, textes réglementaires, copyright).

« Powerpoint pour la création de ressources pédagogiques. »

Maîtrise des principales fonctionnalités du logiciel. Synchronisation multimédia. Réalisation d'un support de cours (TD).

| **Evolutions** | Le contenu du stage « Langage HTML et création de sites web » évolue dans le sens des dernières normes en vigueur sur le web. Ainsi, une partie plus importante est désormais consacrée à la norme XML. Je m'efforce également de répondre aux divers cahiers des charges exposés par les participants, en détaillant davantage les outils qu'il sont amenés à manipuler dans leur pratique quotidienne (le logiciel Dreamweaver est de plus en plus employé dans le monde de l'entreprise comme dans celui de l'université).

| **Remarques** | Ces formations m'ont permis de me confronter à des publics hétérogènes, tant par leur provenance (responsables administratifs, directeurs de laboratoires, écoles doctorales mais aussi personnels d'entreprises et industriels), que par leurs acquis. Ils m'ont, de ce point de vue, obligé à adapter constamment les contenus et l'approche pédagogique.

4.5. **Ecole d'orthophonie de Toulouse (3^{ème} & 4^{ème} années).**

1998 - 2004	Volume horaire annuel	2001 - 04 : 28 heures ¹⁸
	Nombre de formés / an	60 personnes (30 en 3 ^{ème} année et 30 en 4 ^{ème} année)
	Type d'intervention	2/3 Cours Magistraux, 1/3 TD

| Contenus de formation |

« Méthodologie de mémoire de fin d'études »

Construction d'une problématique de recherche ; règles et normes de présentation ; pratiques de la bibliographie (conception, recherche, normalisation) ; techniques orales de soutenance ;

« Méthodologie de la recherche d'information »

Principes méthodologiques généraux ; la recherche dans les bases de données ; veille scientifique dans le secteur bio-médical.

| **Evolutions** | Doté lors de sa création d'un volume horaire relativement confidentiel (4 heures en 4^{ème} année), cet enseignement est maintenant conforme au texte de loi régissant les études d'orthophonie (Arrêté du 16 Mai 1986).

¹⁷ Chacune des formations est d'une durée de 6 heures et est reproduite 2 à 3 fois dans l'année.

¹⁸ | 1998 - 99 : 4 heures | 1999 - 01 : 8 heures |

| Remarques | J'ai été recruté après entretien pour mettre en place cet enseignement. J'en assume désormais la responsabilité pédagogique et administrative. Initialement centré sur la « méthodologie de mémoire », j'ai été amené à développer un volet « méthodologie de la recherche d'information ». Les étudiants étant encadrés par un ou plusieurs maîtres de mémoire, j'ai également mis en place un cycle annuel de réunions pour les 4^{ème} années, en présence de l'ensemble de la promotion et des maîtres et tuteurs de stage pour leur permettre d'exposer leur problématique de recherche et de bénéficier du retour d'expérience et des conseils méthodologiques des participants.

4.6. CIES de Toulouse.

 1999 - 2004 	 Volume Horaire annuel 	36 Heures.
	 Nombre de formés / an 	90 personnes (Moniteurs)
	 Type d'intervention 	1/2 Cours Magistraux, 1/2 TD

| Contenus de formation | "Langage HTML et création de sites web" (cf plus haut [Urfist])

| Remarques | Le CIES organise une évaluation systématique de l'ensemble des stages proposés aux moniteurs et en tire un indice de satisfaction qu'il remet à chacun des intervenants. Le stage que j'anime a toujours bénéficié d'un très bon retour, les moniteurs le plaçant en tête sur plusieurs points d'évaluation (réponse aux attentes, répartition entre cours et TP, pertinence par rapport à leur formation, etc.).

| Evolutions | J'ai proposé une nouvelle intervention pour cette année autour des pratiques, outils et techniques de veille scientifique dans le cadre d'une thèse." La directrice du CIES m'a donné son accord oral pour la conduite de cette formation.

4.7. D.U.S.T.I.C.¹⁹

 2002 - 2004 	 Volume Horaire annuel 	12 Heures.
	 Nombre de formés / an 	12 personnes ²⁰
	 Type d'intervention 	Cours Magistraux

| Contenus de formation |

« *Gestion électronique des documents* »

Fonctions et usages possibles d'un GED ; Documents issus de ressources externes et internes ; Problématique d'indexation ; Mise en œuvre d'un mode de diffusion.

« *Management de l'information* »

Cinématique des contenus ; Outils de partage et de capitalisation des connaissances.

| Remarques | Ces enseignements prennent place dans le cadre du module « information numérique » de l'option « e-documentaliste ». Chacun d'entre eux comporte une partie réalisée en distanciel et une autre en présentiel. J'ai participé à l'élaboration de l'ensemble des parties du cours en distanciel, et j'ai effectué, en présentiel, la partie consacrée au « Management de l'information ».

4.8. Ecoles Doctorales.

 2001 - 2004 	 Volume Horaire annuel 	33 Heures.
	 Nombre de formés / an 	60 personnes
	 Type d'intervention 	1/2 Cours Magistraux, 1/2 TP

| Contenus de formation |²¹

¹⁹ DUSTIC : Diplôme Universitaire des Sciences et Techniques de l'Information et de la Communication. Le DUSTIC fait partie du D.A.M. (Département Archives et Médiathèque).

²⁰ (Professionnels, Personnels administratifs, Personnels des bibliothèques. Formation continue)

²¹ Ces cours n'ont pas été reconduits pour les années 2001-02 et 2002-03 du fait des limitations horaires de mon demi-poste d'ATER. Ils reprennent tous cette année.

[DEA de philosophie]. « *Recherche d'information en philosophie.* »

Présentation d'Internet : historique, architecture client-serveur ; Présentation des outils de navigation et des principaux modes de communication (navigateurs, ftp, courrier électronique ...) ; Méthodologie de la recherche d'information sur Internet (typologie des outils, moteurs, annuaires, technologies push ...) ; La recherche d'information en philosophie (panorama des ressources existantes, moteurs et annuaires disciplinaires, bases de données dédiées ...) ; TP : interrogation d'outils en ligne sur les sujets de DEA des étudiants.

[DESS Ergonomie & Psychologie Cognitive] *"Recherche d'information dans un domaine de spécialité."*

La recherche d'information en psychologie cognitive (panorama des ressources existantes, moteurs et annuaires disciplinaires, bases de données dédiées ...) ; TP : interrogation d'outils en ligne sur les sujets de DESS des étudiants.

[Ecole Doctorale Lettres, Langages et Arts] *"Langage HTML et création de sites web"* (cf supra [Urfist])

| Remarques | Le principal enjeu – et la principale difficulté – de ces cours, fut de prendre en compte l'hétérogénéité du public concernant la connaissance d'Internet et la maîtrise de l'outil informatique. Pour y remédier, des sous-groupes ont été constitués lors des TP, sur la base de questionnaires préalablement distribués.

4.9. Quelques points à souligner.

| *Evaluation* | *Implication* | *Retour exprimé* | *Enseignement & recherche* | *Réseau personnel* |

| Evaluation | Tous les enseignements dont j'ai la charge et qui sont délivrés dans le cadre de cursus universitaires font l'objet d'une évaluation notée qui leur est propre.

| Implication | Pour chacune de mes interventions, je me suis efforcé de m'impliquer et de me rendre disponible autant que possible. Cela m'a amené à encadrer des stages d'étudiants, à participer aux jurys d'examen et de recrutement, à collaborer aux réunions pédagogiques.

| Retour exprimé | Cet engagement s'est vu « récompensé » par le retour globalement positif dont bénéficient mes enseignements et par la confiance renouvelée de leurs responsables, m'accordant la responsabilité pédagogique et administrative de certains modules.

| Enseignement et recherche | Je me suis toujours efforcé de maintenir un lien fort et visible entre mes enseignements et mes thématiques de recherche. Je transmets les principes et contextes théoriques sur lesquels se fonde mon activité de recherche (gestion des connaissances, environnements distribués, organisation hypertextuelle, outils et méthodes de recherche d'information) et fais bénéficier les étudiants de mon acquis professionnel (chargé d'étude auprès de la société TRIVIUM).

| Réseau personnel | Enfin, mes interventions sur l'ensemble des trois campus universitaires toulousains m'ont permis de créer des liens forts et transversaux avec de nombreux enseignants chercheurs et de renforcer la pertinence et la reconnaissance de mes enseignements.

5. Responsabilités collectives et administratives.

5.1. Fonctions de direction.

| 2003-2004 | **Direction des études**²² du DESS "Intelligence économique". Université des Sciences Sociales de Toulouse 1.

J'ai depuis cette année, suite à la proposition d'Alain Alcouffe, directeur du DESS, pris en charge la direction des études du DESS "Intelligence économique". J'assume donc un certain nombre de tâches administratives et pédagogiques afférentes à cette fonction : organisation de conférences (intervenants extérieurs du DESS), mise en place, suivi et coordination des stages et projets tutorés avec les entreprises partenaires, mise en place pédagogique des enseignements dans le cadre des quatre modules composant le DESS, gestion du budget de la formation (suivi des dépenses, taxe d'apprentissage, etc.)

5.2. Encadrement de travaux & mémoires de recherche

| 2003-2004 |

- Thèse professionnelle²³ Darell-Allen SKINNER (auditeur Master MiSTE / ESIEE)
- Mémoire professionnel Amandine Agrech²⁴. DESS Sciences Cognitives et Ergonomie. Université du Mirail. Toulouse.

| 2001-2004 |

- Mémoires professionnels dans les DESS IE et SIAD²⁵. Université des Sciences sociales de Toulouse.
- Mémoires de fin d'étude. Ecole d'orthophonie de Toulouse.

5.3. Responsabilités administratives

Jurys d'examen et de recrutement

| 1998 – 2004 | Je participe aux **jurys d'examen** (Licence InfoCom, DUSTIC) **et de recrutement** (DESS IE, DESS SIAD, Licence InfoCom, DUSTIC) de la plupart des cursus dans lesquels j'interviens.

URFIST de Toulouse

| 1998 – 2004 | **Budget** (affectation et suivi des dépenses, engagements, DBM, etc ...). | **Fonctionnement administratif** (activités de gestion, inscriptions, suivi des stages, convention avec le ministère, etc.) | **Programme des stages**. | **Organisation** des journées d'étude, colloques nationaux. | **Encadrement et formation** des trois stagiaires intégrés sur le projet FoRSIC (DESS psychologie cognitive, IUT documentation, IUP informatique).

²² Je dispose du titre et de la fonction mais ne perçois aucune rémunération (prime) n'étant pas agent titulaire de l'éducation nationale.

²³ Titre de la thèse professionnelle : "Développement de l'atelier d'analyse SPOTTER pour l'analyse du discours des médias internationaux."

²⁴ Titre du mémoire : "Etude des compétences mises en œuvre par les experts en recherche documentaire et application pour la formation." Ce mémoire a pris place dans le cadre de l'action de recherche sur les usages en recherche d'information (cf point 2.4.1.4.). Il a été réalisé auprès des experts qui sont membres du projet FoRSIC (cf point 2.4.1.1.)

²⁵ Ces mémoires portaient principalement sur des thématiques de veille dans différents secteurs économiques et sur la mise en place de portails de type Intranet dans des composantes de plusieurs grandes entreprises.

5.4. Responsabilités pédagogiques.

Ecole d'orthophonie de Toulouse

| 1996 – 2004 | Mise en œuvre et **supervision pédagogique et administrative** d'un cycle d'enseignement à la méthodologie de mémoire de fin d'étude et à la méthodologie de la recherche d'information pour professionnels de santé. Cette mise en œuvre s'est effectuée dans le cadre du texte de loi (arrêté) régissant l'organisation des enseignements pour l'obtention du certificat de capacité en orthophonie. Sur ce point voir la rubrique « Activités d'enseignement » du présent dossier.

5.5. Manifestations scientifiques

| 2003 - 2004 | Colloque International "Littératures et Linguistiques", Université Toulouse le Mirail | **Valorisation** : diffusion et promotion du colloque auprès de listes de diffusions et d'associations.

5.6. Sites web, forums et listes de diffusion.

| 1998 – 2004 | Au titre de mes compétences en réalisation et conception de sites web, j'ai été amené, depuis plusieurs années, à mener une **activité de conseil et d'audit** (ergonomie, architecture, contenus dynamiques, référencement) auprès de diverses personnes, **entreprises et responsables universitaires**.

| 1998 – 2004 | Création et administration de nombreux forums et listes de discussion, dont :

- "IRSW2002" (Information Retrieval and Semantic Web) | **Modérateur** | Groupe de discussion créé en mars 2001 | *43 Membres* | Ce groupe de discussion est rattaché à la communauté du web sémantique et débat des futurs outils et méthodologies applicables au web sémantique. Adresse : <http://groups.yahoo.com/group/irsw2002/>
- "FoRSIC'NeWs" Lettre d'information autour du projet FoRSIC | **Modérateur** | Créée en juin 2000 | *215 abonnés* | "Cette lettre est destinée à tous les acteurs impliqués dans le projet FoRSIC mais également à tous ceux travaillant autour des thématiques suivantes : gestion collective des savoirs, développement des compétences documentaires, formation des usagers, processus d'apprentissage, travail coopératif, e-learning / campus virtuels" | Adresse : <http://www.urfist.cict.fr/forsic> | **Abonnement** : urf11@cict.fr
- LaLiF (Langue et Littérature Française) | **Administration temporaire** | Liste de diffusion créée en février 2000 | *600 abonnés* | "Prenant pour objet privilégié les textes littéraires, cette liste poursuit l'objectif de resserrer les liens entre les arts et les sciences du langage et œuvre pour la modernisation et l'élargissement des études littéraires et des sciences du langage. Elle entend favoriser le débat entre courants : rhétorique, stylistique, sémiotique, analyse du discours, pragmatique, sémantique textuelle, herméneutique, sémantique textuelle" | **Adresse abonnement et contact** : Calvet@univ-tlse2.fr ; o.Ertzscheid@voila.fr

5.7. Participation à des publications.

| 1998 – 2004 | Collaborations régulières à la « **Lettre de l'URFIST** ».

| 2002 – 2004 | Collaboration ponctuelle à la revue « **Champs du signe** », Presses Universitaires du Mirail (relecture, ligne éditoriale).

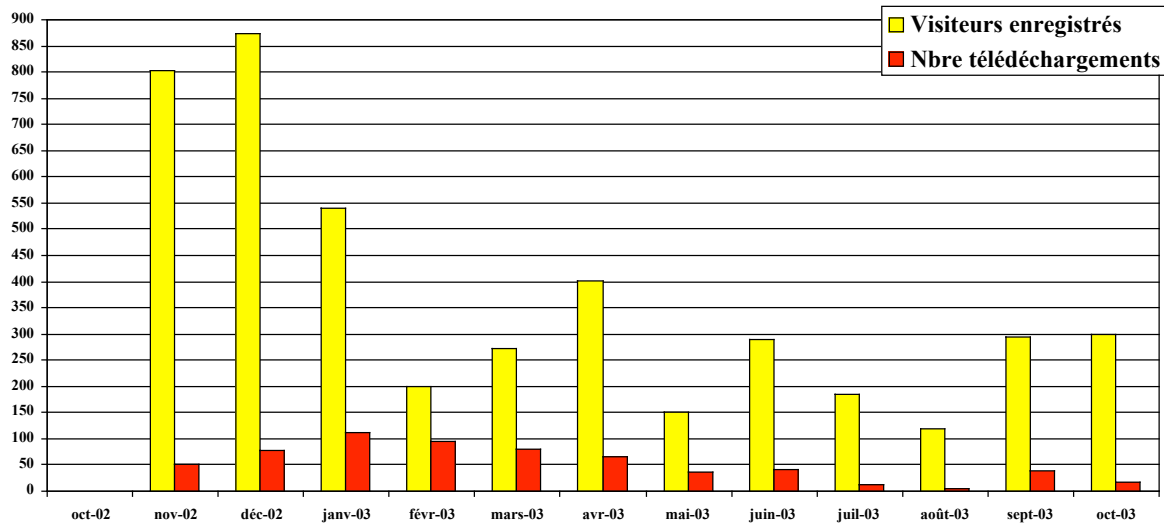
5.8. Activités associatives.

| 1996 – 1998 | Membre de l'ADEMAST (Association pour le développement et la maîtrise des sciences et techniques). C'est dans le cadre de cette association que j'ai effectué mon stage de DESS. Durant ces deux années, j'ai également été impliqué dans divers projets de **valorisation et de promotion de l'information scientifique et technique**.

6. Visibilité scientifique.

| Valorisation de mes recherches | Prix et distinctions | Relations internationales |

6.1. Vue synoptique.



Visiteurs enregistrés et téléchargements de ma thèse
(source : baromètre  [www.xiti.com])
Site : www.Ertzscheid.net

6.2. Valorisation de mes activités de recherche.

Dès sa soutenance, j'ai souhaité que mon travail de thèse soit disponible en téléchargement depuis mon site web. Si des extraits peuvent être récupérés directement, j'ai en revanche choisi de mettre en place un protocole de demande par mail pour ceux souhaitant récupérer l'intégralité de ce travail. Cela m'a permis de mieux connaître un certain nombre de personnes et d'organismes s'intéressant à ces questions et d'entamer avec certains d'entre eux un dialogue fructueux sur des thématiques de recherche communes. Un an après la mise en ligne et en accès libre de ce travail voici un rapide bilan de cette action.

- Ce travail a été référencé par de nombreux sites (Formatic2000 – <http://www.formatic2000.com> , Thot – <http://thot.cursus.edu> - , Zazie web <http://www.zazieweb.fr> - Fabula <http://www.fabula.org>, ...).
- Il a également été référencé sur de nombreuses listes de diffusion et groupes de discussion dédiés aux NTIC (NTIC.org), à la recherche d'information (Motrech), au monde des bibliothèques et de la documentation (ADBS-info), à la littérature électronique (e-critures).
- Il a fait ce jour l'objet de près de 1000 demandes de télédownload.
- En sus des demandes individuelles de chercheurs, j'ai également reçu de nombreuses demandes pour qu'il soit intégré dans le cadre d'Intranets publics et privés au titre de ressource pédagogique. Ces demandes émanent principalement d'organismes français, suisses, belges et canadiens s'intéressant aux questions d'écriture hypertextuelle, de formation, de partage des connaissances, de document numérique, de navigation et de représentation (cartographie) dans de vastes ensembles de données.

Voici une liste indicative des contacts que j'ai pu nouer à l'occasion de ces demandes de télédownload. Suite à ces demandes, je continue d'échanger régulièrement autour de thématiques et/ou de projets de recherche avec un dizaine d'interlocuteurs²⁶ issus de ces entreprises et/ou organismes.

- ➔ **En France** : CEA (Cellules de documentation) | *Centre Paris-lecture* | Ecole des Mines d'Alès (DESS « Auteur rédacteur Multimédia ») | *ENSSIB* (centre de documentation et BBF – Bulletin des Bibliothèques de France) | France-Telecom R&D | Groupe Desybel (consulting) | INRA (Cellules de documentation) | *INTD* | IUFM (Cellules de documentation) | Laboratoire communication et politique (LCP) . CNRS | Laboratoire d'intelligence des organisations. (Jeune équipe). Université de Haute-Alsace. | Missions départementales TICE (Rectorats)
- ➔ **À l'étranger** : *Agence universitaire de la francophonie (AUF), Montréal* | Bibliothèque nationale du Burkina Fasso | Département de l'instruction publique de Genève (Suisse). | *École supérieure neuchâteloise d'informatique et de gestion (ESNIG)* | Université Aberta. Lisbonne (Portugal) | Université catholique de Louvain (Belgique) | Université de Dubrovnik (Croatie). | Université de Moncton (Nouveau Brunswick, Canada) | Université de Ratisbonne (Allemagne) | Université de Sherbrooke (Canada) | *Université du Québec à Montréal (UqAM)*, Laboratoire d'analyse cognitive de l'information (LACI) | Université fédérale Sainte Catherine (Brésil), *Équipe de recherche sur les hypertextes* | Université Macerata (Italie). Faculté des sciences politiques. | Université pontificale catholique de Sao Paulo (Brésil). | Université Tel-Aviv (Israël) |

6.3. Prix & Distinctions.

- ➔ 1^{er} prix des doctoriales²⁷ (édition 1998-1999)

6.4. Relations internationales.

| Autour de l'hypertexte |

Je suis en contact régulier avec **Ted Nelson**. J'ai été amené à échanger avec lui durant la conduite de mon doctorat, alors même qu'il était en train de terminer le sien à l'université de Keio (Japon). Il a accepté de me faire bénéficier de son expertise sur quelques points clés de mon travail (transclusion, versioning).

²⁶ En italique dans la liste qui suit.

²⁷ Créées en France par l'ancienne Direction de la Recherche et de la Technologie du Ministère de la Défense (DRET) et l'Association Bernard Gregory, les Doctoriales sont des stages de préparation des doctorants à l'après-thèse. La formule a été étendue au plan national en 1997 par le Ministère chargé de la Recherche.

Je suis en contact avec **Robert Kendall**, responsable du site « **Electronic Literature Directory** » (<http://directory.wordcircuits.com>), sous l'égide de l'Electronic Literature Organization » (<http://www.eliterature.org>), site américain qui constitue à ce jour la plus grosse banque de donnée consacrée à la littérature électronique hypertextuelle. Une collaboration pour développer le volet francophone de cette base est à l'étude.

J'ai récemment pris contact avec Mark Bernstein de la société Eastgate, éditrice du logiciel Storyspace pour étudier l'implémentation de fonctionnalités de liaison présentées dans ma thèse.

| Autour de la recherche d'information |

Correspondant pour la France du réseau européen ENIL, « European Thematic Network on Information Literacy. » Le travail de ce réseau, qui compte autant de correspondants que de pays membres de l'union est publié dans une monographie dont j'ai co-rédigé l'un des chapitres avec Jo Link-Pezet (voir la rubrique « Publications » du présent dossier).

6.5. Conférences & Sessions organisées

"L'avenir des usages et des modèles de la recherche d'information sur le web." J'organise cette Session invitée à la 8ème conférence mondiale Systémique, Cybernétique et Informatique (SCI 2004) | <http://www.iiisci.org/sci2004/website/> |

7. Annexes

7.1. Articles et publications

Chercher faux et trouver juste, Serendipité et recherche d'information

CIFSIC – Bucarest 2003 – Atelier D2 - «Communication et complexité »

Animation : C. LE BŒUF / N. Pelissier

Chercher faux et trouver juste, Serendipité et recherche d'information

Olivier Ertzscheid

Laboratoire Paragraphe - Université Paris 8 -

Gabriel Gallezot

Urfist - Université de Nice-sophia Antipolis

Résumé

Merton définit la serendipité ainsi : "*découverte par chance ou sagacité de résultats que l'on ne cherchait pas*". L'idée mise en avant dans ce texte est de montrer comment dans un effort de maîtrise de l'entropie informationnelle, (2) l'essor des technologies intellectuelles de représentation et d'accès aux informations (3) fait chaque jour une place plus grande au phénomène de sérendipité comme adjuvant précieux de la recherche (4).

Abstract

Merton defines the serendipity as follows : "*the faculty or phenomenon of finding valuable or agreeable things not sought for*". We propose to show how, in an effort to control informational entropy,(2) the rise of intellectual technologies for representation and access to information (3) makes each day a larger place to phenomenon of serendipity as a precious additive of research (4).

Mots clés : sérendipité, recherche d'information, technologies intellectuelles, entropie informationnelle

Keywords : serendipity, information retrieval, informational entropy, intellectual technologies

1. Introduction

1.1 - Définition et origine de la sérendipité

Le terme de serendipity apparaît avec Walpole dans un conte oriental "Voyages et aventures des trois princes de Serendip" (Ceylan), où ceux-ci, "*ayant d'abord été formés avec soins, dans toutes les sciences, se tiraient toujours d'affaire grâce à leur talent exceptionnel pour remarquer, observer, déduire, à toute occasion.*" (source : www.granddictionnaire.com). Ce terme apparaît en sciences et se conceptualise avec Merton qui le définit ainsi : "*la découverte par chance ou sagacité de résultats que l'on ne cherchait pas*". La sérendipité ("fortuité" pour nos amis québécois) est une problématique qui n'a fait que récemment son entrée dans le champ des sciences de l'information – francophones – sous la plume de Perriault : "*L'effet "serendip" (...) consiste à trouver par hasard et avec agilité une chose que l'on ne cherche pas. On est alors conduit à pratiquer l'inférence abductive, à construire un cadre théorique qui englobe grâce à un "bricolage" approprié des informations jusqu'alors disparates.*"[Perriault, 00].

Pour cerner ce concept et appréhender le phénomène nous indiquons ci-dessous quelques exemples célèbres. Tout le monde a appris comment Christophe Colomb, cherchant la route occidentale des Indes découvrit en fait l'Amérique. Nombre d'autres découvertes tout aussi essentielles pour l'humanité ont partie liées avec la sérendipité. En voici une liste non-exhaustive : le principe de champagnisation (Dom Pérignon), la pasteurisation (L. Pasteur), la pénicilline (A. Fleming), les rayons X (W. Röntgen), la vulcanisation du caoutchouc (Ch. Goodyear). Citons également d'autres découvertes moins "essentielles" comme le "post-it" (où comment répondre à la question : "que faire d'une colle qui ne colle pas ?"), le Caprice des dieux, le Coca-Cola, le Zyban et le Viagra...

1.2 – Rechercher ou Rechercher

Compte tenu du contexte francophone dans lequel nous présentons ce texte, nous voulons d'abord introduire la distinction et le parallèle entre le " rechercher " de la recherche d'information (*Information Retrieval* en anglais, IR) et le " rechercher " de l'épistémé, la recherche (*Research* en anglais). Ce signifiant unique en français introduit un signifié commun " trouver l'information " mais aussi des signifiés distincts qui se trouvent quelque part dans les moyens et l'objectif. Pour l'IR, les moyens ce sont des outils du traitement de l'information sur un corpus documentaire. Pour l'épistémé ce peut être l'empirisme ou des outils de traitement de l'information dont l'objet change et devient " naturel ". L'objectif, pour l'IR, c'est de repérer et de ramener des infos pertinentes. Pour l'épistémé c'est de découvrir, de produire de nouvelles connaissances. Enfin, pour être complet signalons aussi dans notre domaine disciplinaire la recherche en IR qui désigne notamment la modélisation et les études d'usages informatiques dont l'objet est constitué par des corpus de texte, de connaissance.

1.3 - Complexité

Pour appréhender le champ de l'IR et ses objectifs nous proposons de fouiller la métaphore de « l'aiguille dans la botte de foin » qui, de façon triviale, signifie la difficulté (l'impossibilité) de trouver quelque chose. Chercher une aiguille dans une botte de foin peut s'appréhender de différentes manières et ainsi correspondre à plusieurs scénarios de recherche :

- *trouver une aiguille connue dans une botte de foin connue*
- *trouver une aiguille connue dans une botte de foin inconnue*
- *trouver une aiguille inconnue dans une botte de foin inconnue*
- *trouver n'importe quelle aiguille dans une botte de foin*
- *trouver l'aiguille la plus pointue dans une botte de foin*
- *trouver la plupart des aiguilles contenues dans une botte de foin*
- *trouver toutes les aiguilles contenues dans une botte de foin*
- *pouvoir affirmer qu'il n'y a pas d'aiguille dans une botte de foin*
- *trouver des choses qui ressemblent à des aiguilles dans une botte de foin*
- *connaître chaque nouvelle aiguille qui apparaît dans la botte de foin*
- *trouver où sont les bottes de foin*
- *trouver des aiguilles et des bottes de foin, quelles qu'elles soient." [Koll, 00]*

Une vision plus globale est proposée par [Toms 00], pour qui il existe trois grandes manières de chercher de l'information :

- *"chercher de l'information sur un objet bien défini ;*
- *chercher de l'information sur un objet incomplètement décrit mais qui sera reconnaissable dès qu'un le rencontrera ;*
- *trouver de l'information de manière fortuite."*

Ce troisième et dernier cas, fait écho à l'une des déclinaisons non citées de l'aiguille et de la botte de foin : " Mal chercher l'aiguille dans la botte de foin et la trouver quand même ". L'idée mise en avant dans notre article est de montrer comment dans un effort de maîtrise de l'entropie informationnelle (2) l'essor des technologies intellectuelles de représentation et d'accès aux informations (3) fait chaque jour une place plus grande au phénomène de sérendipité comme adjuvant précieux de la recherche (4).

2. Maîtriser l'entropie informationnelle ?

Appréhender la complexité d'un phénomène nécessite le repérage et la gestion d'un corpus de documents pertinent et volumineux pour extraire des informations et ensuite les transformer en nouvelles connaissances. Ce cycle de la production scientifique ce juxtapose avec le cycle du document. Ainsi, le " bouclage " d'un cycle produit de nouvelles connaissances et conséquemment de nouveaux documents. Les idées ne peuvent se former que sur des constructions cognitives antérieures présentes sous forme d'information dans les documents. Nous inscrivons donc la construction de connaissance dans un processus de transformation de l'information où :

- la connaissance est la formation des idées,
- l'information est la mise en forme des connaissances (in-formation) et
- l'information inscrite sur un support constitue un document.

L'itération de ce processus conduit à une somme d'informations toujours plus importante que les chercheurs de toutes disciplines cherchent à maîtriser. Internet accroît la rapidité de ce cycle et contribue plus encore à l'entropie informationnelle. [Gallezot, 02a]

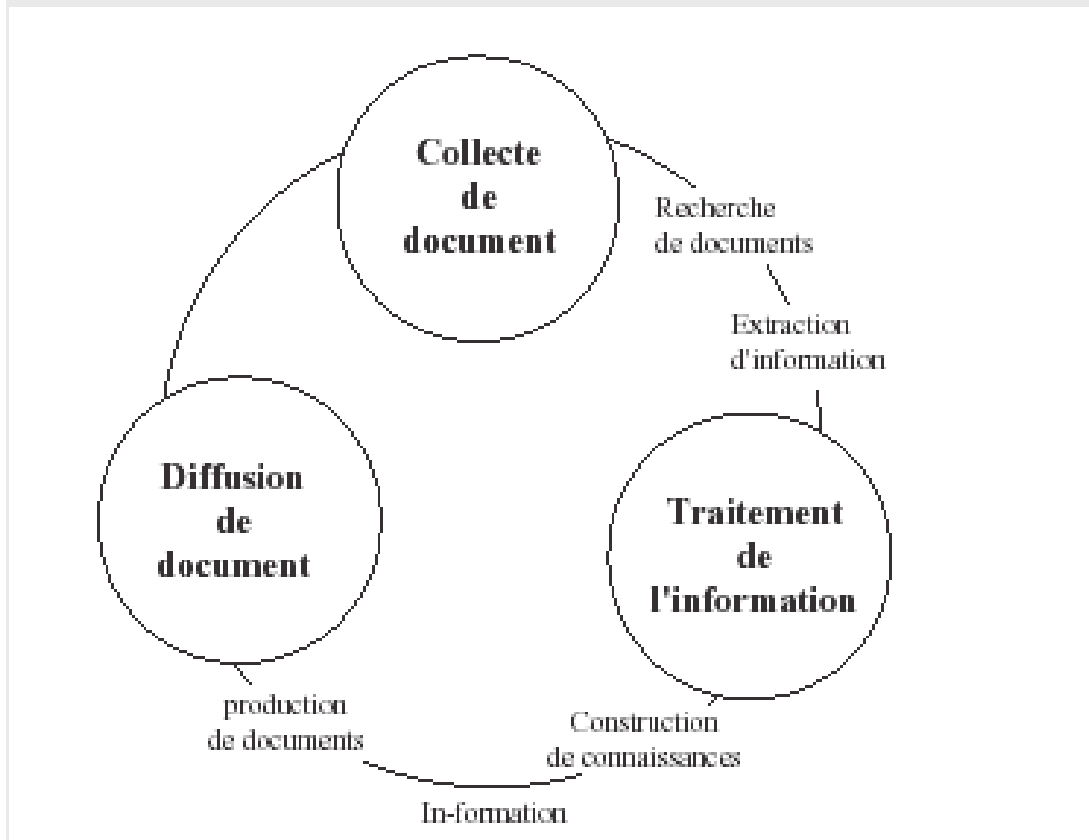


Fig. 1 : le cycle de l'IST

Entropie et homogénéité.

Cette course à la maîtrise de l'information continue à alimenter la somme potentiellement disponible de documents, Internet contribuant dans une large mesure à ce phénomène. Comme "support" il est celui qui assume et assure la plus grande partie de cette croissance exponentielle de l'information disponible sur les réseaux. Comme mode de diffusion et d'accès, Internet garantit à cette masse d'information une "homogénéité" jusque-là jamais atteinte.

Nombre de facteurs objectifs tendent à étayer cette thèse *a priori* surprenante de l'homogénéité de l'information sur les réseaux quand l'habitude veut que l'on considère plutôt comme essentiel le caractère hétérogène de cette information. Reprécisons donc ce que nous entendons ici : la nature de l'information disponible sur les réseaux est effectivement profondément hétérogène (validité scientifique, "fraîcheur" éditoriale, qualité graphique, etc.) Pourtant, cette hétérogénéité s'efface complètement du fait du niveau de relation entre unités d'information, qui, selon le niveau d'échelle et de granularité auquel on se place (site web, système d'information en ligne, web dans son ensemble), permet d'affirmer que tout est lié à tout²⁸.

Nombre de facteurs objectifs viennent appuyer cette thèse parmi lesquels les études qui ont tenté de mesurer le "diamètre" du web : la dernière en date fait état d'un diamètre de dix-neuf liens [Barabasi, 99]. Cela signifie, que quelles que soient les unités d'information choisies (en l'occurrence des pages web), elles se trouvent connectées par une chaîne d'au plus dix-neuf liens.

Au delà de chiffres qui, du fait de la nature même du web ne sauraient être stabilisés, ces études ont surtout permis de construire une topologie de l'espace informationnel tel qu'il se déploie sur les réseaux, en faisant émerger certaines zones "obscurées" (web invisible), déconnectées d'autres zones mais tout aussi connectées entre elles, et en ce sens homogènes.

²⁸ <http://www.almaden.ibm.com/cs/k53/www9.final>, Cette étude conjointe d'Altavista, Compaq et IBM fait état d'une topologie du web en forme de nœud papillon : le nœud est constitué de pages hyperconnectées, la partie gauche comprend les pages qui permettent d'y accéder et la partie droite celles vers lesquelles pointe ce nœud. Même s'il demeure, au vu de cette étude un certain nombre de pages déconnectées, cela ne fait que renforcer l'hypothèse d'une connection optimale pour la partie sinon la plus dense, du moins la plus visible du web.

La recherche d'outils capables de maîtriser l'information dans cet espace réticulaire constitue pour beaucoup un enjeu majeur. S'il est évident qu'il faille tendre vers une appropriation informationnelle exhaustive pour édifier l'épistémé, la tâche est incommensurable. Que reste-t-il au chercheur devant cette entropie informationnelle ? Se servir des outils *ad hoc* pour repérer au mieux l'information pertinente, borner son référentiel documentaire, expérimenter, observer, évaluer et produire ses résultats à l'aide de méthodologies éprouvées, de protocoles heuristiques... passer des achoppements aux paradigmes scientifiques. Il existe un raccourci : la serendipité. Elle s'offre et se révèle lors de découvertes informationnelles.

3. Découvertes informationnelles et technologies intellectuelles.

Les pratiques informatives sont multiples et constituées des usages différents de l'ensemble : "repérage/collecte/traitement/diffusion" de l'information. Les actions de cet ensemble sont réalisées par un binôme indissociable d'outils et de méthodes : les technologies intellectuelles [Fayet-Scribe, 00]. Celles-ci permettent des découvertes informationnelles qui se transformeront, après un processus créatif, en connaissance. Ainsi nous distinguons la recherche d'information de l'épistémé mais soulignons leur appartenance au même processus.

Le processus de création conventionnel fonctionne sur le principe de divergence/convergence où la reconnaissance d'un problème est introduite par une divergence pour converger vers une nouvelle solution. Le processus de création par serendipité est le contraire : bien que la solution à un problème soit attendue il y a divergence de parcours, lesquels conduisent à un problème différent ou plus fréquemment, à la solution d'un problème dont nous n'avons aucune connaissance [Figueirado, 01] (cf. infra, point 4.2.).

Comment dès lors garder une emprise sur ce phénomène ? Si l'on considère que les découvertes n'arrivent jamais par chance, il faut donc insister sur le rôle de la préparation intellectuelle et/ou l'intensité de l'observation et de la recherche [Van Andel, 92]. On peut aussi penser que l'IR prenne en compte les phénomènes de serendipité en complément des requêtes (*querying*) et de la navigation (*browsing*) pour stimuler la curiosité et encourager l'exploration [Toms, 01]. Les outils qui permettent, le *cross-matching*, la *clusterisation*, la *percolation* sont des facilitateurs de serendipité. Ils proposent souvent des visuels qui « donnent à voir » aux chercheurs. Nous retenons ainsi le vocable de technologies *procognitives* employé par Licklider pour signifier l'importance des outils de traitement de l'information qui servent la connaissance et passons ci-dessous en revue les techniques intellectuelles qui nous semblent favoriser la création d'éléments stochastiques.

3.1 Naissance de la bibliologie.

Paul Otlet (1868-1944) peut être considéré l'un des pères de la documentation et de la science de l'information. Il est le co-fondateur avec Henri La Fontaine, en 1895, de " l'office international de bibliographie ". Son Traité de documentation (1934) est la première approche systématique de ce que nous appelons aujourd'hui la (les) science(s) de l'information. C'est le premier à comprendre le problème que posera à terme l'augmentation considérable du nombre de livres et de documents, et à proposer, pour y répondre, la création de la bibliologie, à la fois comme science et comme technique générale pour la documentation. La plupart des idées qui fondent ce que l'on appellera l'hypertexte/hypermédia sont déjà explicitement présentes chez Otlet, qu'il s'agisse d'offrir un accès automatisé aux documents ou de relier chacun d'eux avec d'autres, tout en conservant leur individualité, dans le cadre d'un " cerveau collectif ".

3.2 Indexation associative : le MeMex.

[Bush 45] est unanimement reconnu comme le pionnier de l'hypertexte/hypermédia sous sa forme actuelle. Il est – tout comme Otlet – confronté à l'explosion de la masse documentaire. Il imagine alors un système automatisé de microfiches, baptisé MeMex (Memory Extender) lequel ne sera jamais effectivement réalisé, mais contient déjà la plupart des idées de l'hypertexte. Celle-ci sont exposées dans l'article " **As we may think** " qui commence par ces mots : " *Consider a future device for individual use which is a sort of mechanized private file and library.* ".

L'idée de base est de reproduire le fonctionnement caractéristique de l'esprit humain en imaginant des machines capables de fonctionner par association et non plus selon le modèle classique de l'indexation. " *Human mind (...) operates by association. (...) Selection by association, rather than by indexing, may yet be mechanized.* "

Son système est défini comme suit : “ *A memex is a device in which an individual stores all his books, records and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. an enlarged intimate supplement to his memory* ”. L'évolution de l'approche de Bush peut se résumer à deux idées principales :

- il est possible de "mécaniser" le fonctionnement associatif de l'esprit humain
- les parcours de navigation (“ *trails* ”) et d'accès dans un tel environnement associatif sont des éléments de construction du sens.

3.3 De Nelson à Engelbart : l'avènement de l'hypertexte comme système informatique et comme mode d'accès et d'organisation des connaissances.

C'est Ted Nelson, philosophe de formation, qui le premier forge le terme “ hypertexte ” dans un article éponyme donné lors de la conférence de la Fédération Mondiale de Documentation. Personnage contesté, il n'en demeure pas moins l'un des visionnaires les plus actifs et il est à l'origine de nombre de concepts aujourd'hui au cœur de problématiques importantes (“transpublishing” pour les questions de droit d'auteur, “versioning” pour celles des archives ouvertes et des nouveaux modes de publication, etc.). Tous ces concepts prennent place dans le cadre de son projet XANADU (<http://www.xanadu.net>).

Douglas Engelbart, chercheur au mythique SRI (Stanford Research Institute), est non seulement l'inventeur du système actuel de fenêtrage et de la souris, mais également le concepteur d'un système baptisé “ **Augment** ” destiné à faciliter l'augmentation des capacités de l'intelligence humaine. “ *By “ augmenting human intellect ” we mean increasing the capability of a man to approach a complex problem situation, to gain comprehension to suit his particular needs, and to derive solutions to problems.* ” [Engelbart 62 p.1].

Augment peut être considéré comme le premier système hypertextuel effectivement réalisé :

“ *As part of the Augment Project, primarily designed for office automation, Engelbart (...) developed a system called NLS which had hypertext-like features. This system was used to store all research papers, memos and reports in a shared workspace that could be cross-referenced with each other. In 1968, he demonstrated NLS as a collaborative system among people spread geographically.* ” [Balasubramanian 94]

3.4 Hypermédias générés

L'abondance de l'information nécessite un repérage accru et efficient des connaissances, qui explique l'intérêt pour les techniques de visualisation [Shneiderman, 97]. L'exploration informationnelle s'inscrit dans cette démarche. Plusieurs techniques peuvent bénéficier de cette approche, de l'analyse statistique sur du texte à la structuration et l'organisation de données dans des bases de connaissances (*knowledge bases*). Ce qu'il faut noter, c'est la généralisation de ce processus. L'extraction d'information sur un seul type de données en vue d'obtenir un résultat précis, ne suffit pas à rendre compte de situations complexes. La globalisation d'informations sur un sujet et la visualisation sous forme graphique des résultats d'un traitement réalisé par des techniques d'information offrent des *machines de vision* [Virilio, 88] capables de générer de nouvelles connaissances, de nouveaux projets de recherche ou d'autres éléments de réflexion... de créer. Ce sont des artefacts informationnels qui offrent une vision heuristique des résultats de la recherche. Les formes nouvelles de documents, que l'on peut qualifier de tertiaires, deviennent des adjuvants prépondérants et essentiels pour lire l'ensemble des documents disponibles. Ces construits sont la synthèse de résultats expérimentaux (un ensemble de données factuelles) et de conceptions théoriques (à travers la modélisation de la base de données et la génération de liens). En revanche, les interprétations, les créations sont liées à des perceptions, des appropriations personnelles des représentations. Cette appréhension repose sur la culture technique et informationnelle de chaque individu et sa capacité à l'intégrer dans son activité quotidienne, pour produire du sens sur les objets qu'il manipule [Gallezot, 02b].

Il ne s'agit pas spécifiquement de naviguer ou déambuler, mais de fouiller les sédiments cognitifs accumulés depuis quelques années à la recherche d'information. Les visualisations proposées mettent en lumière des liens qui n'auraient pas pu être perçus autrement et peuvent faire sens auprès d'un expert. Les hypermédias générés qui aident la lecture d'information, ne relèvent pas exactement d'un choix, d'une sélection d'information, mais d'une composition aléatoire dirigée par des solidarités annotationnelles [Bachimont, 99]. Le renouvellement de ces hypermédias est lié à l'ajout de documents dans les entrepôts d'informations.

Un agrégat d'informations intégré dans un artefact informationnel, n'étant pas connu *a priori*, la recherche d'information, *a posteriori*, peut retourner des documents auxquels le chercheur ne pensait pas. Plus encore, la mise en relation des unités informationnelles peut permettre de découvrir des

corrélations insoupçonnées soit par lecture directe d'un résultat, soit par inspection d'un visuel. Le chercheur détecte des faits de façon quasi fortuite. Cette réécriture aléatoire et cette relecture fortuite relève de la serendipité.

3.5 Vers de nouvelles logiques implicites d'accès et de représentation des connaissances : le cas des moteurs de recherche.

Les moteurs de recherche, dans l'utilisation qu'ils font des liens comme principes de classification, ne sont pas de simples interfaces de recherche, au même titre que celles que l'on trouve sur des cédéroms : ces dernières ne prennent exclusivement en compte que les mots (clés ou non) et les occurrences de ces mots. A l'inverse, faire le choix des liens comme principe de classement, de tri et d'organisation de l'information²⁹, c'est revendiquer clairement le choix de l'immatériel ou à tout le moins le choix de l'information comme mesure " *d'une différence qui produit une autre différence* " [Bateson 77 p.231].

Quand nous consultons une page de résultat de Google ou de tout autre moteur utilisant un algorithme semblable, nous ne disposons pas simplement du résultat d'un croisement combinatoire binaire entre des pages répondant à la requête et d'autres n'y répondant pas ou moins (matching). Nous disposons d'une vue sur le monde (watching) dont la neutralité est clairement absente. Derrière la liste de ces résultats se donnent à lire des principes de classification du savoir et d'autres encore plus implicites d'organisation des connaissances. C'est ce rapport particulier entre la (re-)quête d'un individu et la (re-)présentation d'une connaissance qui était présente dans les bibliothèques de la Haute-Egypte, pour en être évacuée avec l'arrivée des principes de classement alphabétiques.

Une nouvelle logique se donne à lire. Moins " subjective " que les principes classificatoires retenus par une élite minoritaire (clergé, etc.) elle n'en est pas moins sujette à caution. Les premières étaient douteuses mais lisibles, celles-ci le sont tout autant parce qu'illISIBLES³⁰, c'est-à-dire invisibles : l'affichage lisible d'une liste de résultats, est le résultat de l'itération de principes non plus seulement implicites (comme les plans de classement ou les langages documentaires utilisés dans les bibliothèques) mais invisibles et surtout dynamiques, le classement de la liste répondant à la requête étant susceptible d'évoluer en interaction avec le nombre et le type de requêtes ainsi qu'en interaction avec le renforcement (ou l'effacement) des liens pointant vers les pages présentées dans la page de résultat.

Ainsi, à mesure que se tissent, à chaque instant de nouveaux liens entre les nouvelles entités (documentaires ou non) composant le réseau, à mesure que ceux-ci n'ajoutent pas simplement à une complexité existante mais la reconfigurent à chaque instant, et à mesure que s'affirment comme les plus efficaces des algorithmes de recherche, ceux systématisant la part faite à l'objectivation de phénomènes subjectifs (" best practices ", pages pivots et d'autorité ...) l'horizon qui se dessine pour la contribution des sciences de l'information à l'organisation de la connaissance est désormais celui pointé par [Carr et al. 99], qui indiquent, en conclusion de leur article :

" Le challenge est désormais de construire des systèmes capables d'extraire ou d'apprendre la sémantique des connaissances implicites dans le média et de construire des associations entre ces représentations liées au média et la sémantique, sans qu'il y ait pour cela besoin de lourdes entrées manuelles de données. Rechercher et naviguer plus directement à partir des concepts, plutôt qu'à partir de leurs représentations variées, sera alors une réalité. "

Dans cette perspective là le rôle fondamental des ontologies s'affirme chaque jour davantage (www.semantic-web.org).

4 La serendipité

4.1 Définition dans le contexte de l'IR

Nous plaçant du point de vue de l'IR dans des environnements distribués (Internet) nous définissons la serendipité comme la propagation d'un **style cognitif stable** (mis en place au début de la session de navigation) dans un **environnement différent** mais contenant de **l'information pertinente** pour l'utilisateur dans **le contexte initial de sa navigation** et au vu de la tâche qu'il s'était assigné. Constatant alors que cette procédure donne des résultats permettant de satisfaire ses besoins de manière non prévue il va mettre en place de nouvelles stratégies de navigation lui

²⁹ comme ce fut le cas pour la révolution entraînée par l'algorithme PageRank du moteur Google (www.google.com) qui considéra que la pertinence d'une page était liée en priorité au nombre de pages la référencant (liens entrants) et non plus exclusivement à des mesures d'occurrence linguistique. Ce critère (inspiré de Garfield et de la bibliométrie) est actuellement pris en compte par la plupart des outils de recherche.

³⁰ pour les utilisateurs non spécialistes.

permettant d'amorcer un nouveau cycle, soit en assignant un nouvel objet-cible à sa recherche, soit en initiant un nouveau parcours permettant d'atteindre l'objet-cible initial

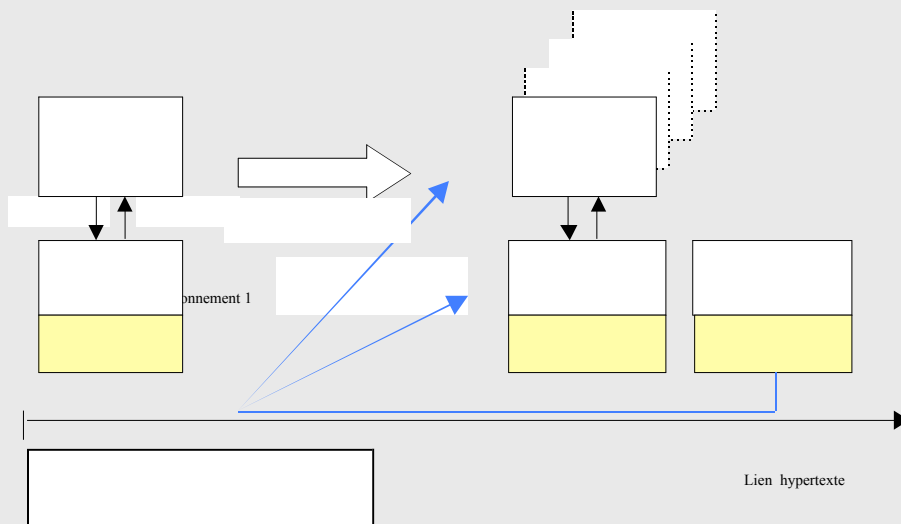


Fig. 2 : Sérendipité et cycle de l'IR

4.2 Tentatives de sériation de la sérendipité.

Une première sériation de la serendipité peut être observée à partir des quatre équations de Figueiredo et Campos [Figueiredo, 01] (Cf. Fig 3). Les auteurs formalisent sous forme d'équations (simplification) la serendipité en distinguant un problème (P), le contexte du problème (KP), la métaphore (M), le contexte de la métaphore (KM), la solution (S), le contexte de la solution (KS) et le gain de connaissance dans le processus de formulation du problème (KN). Ces quatre équations reposent en fait sur un « déclencheur », la métaphore comme moyen de provoquer la perspicacité :

- métaphore,
- métaphore inattendue
- absence de métaphore
- métaphore de l'ignorance

Dans la première équation la métaphore inattendue inspire la solution. Dans la seconde équation la métaphore inattendue conduit à un nouveau problème puis à une nouvelle solution. Dans la troisième équation, l'absence de métaphore impose un pragmatisme, un problème trouve écho à un autre problème et propose ainsi une nouvelle solution. Dans la quatrième équation, la métaphore de l'ignorance introduit l'erreur dans le contexte de la description du problème, elle implique un nouveau problème, puis une nouvelle solution.

<p>1. Pseudosérendipité, exemple d'Archimede</p> $P1 \subset (KP1) \Rightarrow S1 \subset (KP1, KM, KN)$ $M \subset (KM)$	<p>2. sérendipité avec Méthaphore, exemple de Rontgens (Xray)</p> $P1 \subset (KP1) \Rightarrow P2 \subset (KP2)$ $M \subset (KM) \Rightarrow S2 \subset (KP2, KM, KN)$
<p>3. Sérendipité sans Méthaphore , exemple de la 2cv Citroën</p> $P1 \subset (KP1) \Rightarrow P2 \subset (KP2)$ $S2 \subset (KP2, KN)$	<p>4. sérendipité avec métaphore de l'ignorance, exemple de Christophe Colomb</p> $P1 \subset (KP1, EP1) \Rightarrow P2 \subset (KP2)$ $S2 \subset (KP2, KN)$

Fig 3 : les équations de la serendipity

De son côté, [Toms 00] propose de distinguer entre le raisonnement par analogie (favorisant la sérendipité) et ce qu'elle nomme "blind luck" où seul le hasard est à l'origine d'une découverte informationnelle. Ce type de sérendipité peut se rencontrer dans le cas de générateurs aléatoires de nœuds (graphes) hypertextuels. Elle rappelle également l'importance du "principe de Pasteur" selon

lequel "*le hasard ne favorise que les esprits préparés*". Apparaît ainsi l'idée qu'il est nécessaire d'amorcer le processus de sérendipité, c'est à dire de l'inclure dans un cycle initial de recherche.

[Boursier & Van Andel, 92] proposent de "qualifier" ce que [Figueirado 01] ont mis en équation et parlent de "*sérendipité positive*" pour désigner l'"*observation d'un fait non anticipé suivi d'une abduction correcte*", de "*sérendipité négative*" pour désigner l'observation d'un fait ou accomplissement d'une tâche sans interprétation juste (équation n°4, C. Colomb) et enfin de "*pseudo-sérendipité*" lorsque l'on cherche quelque chose qui avait déjà été conceptualisé mais qu'on le trouve par un autre chemin que celui initialement prévu (équation n°1)

La dernière tentative de sériation que nous avons repéré vient du monde de l'intelligence économique dont les problématiques de "veille" garantissent le lien avec les sciences de l'information et de la communication. [Marti, 02] se focalise sur les aspects volontaristes ou proactifs de la sérendipité et cite trois contextes possibles :

- celui du groupe Bourbaki qui avait pris comme habitude d'inviter à ses conférences de jeunes confrères en leur demandant d'intervenir sur des domaines où ils n'avaient aucune expérience, pariant ainsi sur leur fraîcheur d'esprit pour apporter idées neuves ;
- celui de la technique du "Pot de Miel" quand par exemple IBM offre un accès gratuit à des brevets (www.delphion.com) mais qu'il s'en sert pour repérer des tendances technologiques et observer le comportement et les requêtes de ses concurrents (traçage adresses IP) ;
- celui enfin du groupe 3M (leader de l'innovation), le plus proactif de tous, qui "oblige" ses scientifiques à consacrer 15% de leur temps à des axes de recherche en dehors de ceux définis par la R&D.

4.3. L'apport des modèles de l'IR : vers une sériation plus globale.

Ce paragraphe tente de mettre en avant une vue globale du processus de recherche d'information, du point de vue des différents types de sérendipité qu'il autorise (ou interdit). Il existe 3 "états initiaux" de l'IR, auxquels sont associés 3 processus, trois types de tâches, qui font eux-mêmes référence à 3 grands types de modèles.

Etat initial	Processus	Modèles
[Je sais] [ce que je cherche]	Querying / Browsing Sérendipité nulle	Computationnel
[Je ne sais pas] [ce que je cherche]	Searching Sérendipité structurelle	Utilisateur
[Je sais] [que je ne sais pas ce que je cherche]	Learning Sérendipité associative	Environnementaliste

Le premier cas représenté dans ce tableau (Je sais ce que je cherche) repose sur l'idée que dans la majorité des démarches de recherche d'information, l'utilisateur sait déjà (partiellement) ce qu'il cherche. Il lui reste alors à mettre en place une série de requêtes (querying) correspondant au modèle computationnel classique autorisé par les systèmes documentaires (booléens, langages documentaires, etc.). L'utilisateur est dans une logique de consultation et cherche à savoir ce que peut lui apporter comme résultats (matching) le système d'information qu'il est en train d'utiliser (browsing). Cet utilisateur met en place un raisonnement de type hypothético-déductif. La sérendipité est alors quasi-nulle ou ne relève en tout cas d'aucune démarche volontariste ou consciente.

Le second cas (Je ne sais pas ce que je cherche) correspond à l'objectif de l'IR selon [Belkin, 00], à savoir : "*Helping people find what they don't know.*" Le processus alors appelé est de type exploratoire (searching). L'utilisateur va, à partir de ce qu'il sait, raisonner par inférence et abduction en fonction de son but ou de son "profil". La sérendipité qui se met ici en place est de type structurelle (cf. infra)

Le dernier cas (Je sais que je ne sais pas ce que je cherche) est celui qui peut le plus bénéficier du phénomène de sérendipité. L'utilisateur ayant formalisé et explicité qu'il "ne sait pas ce qu'il cherche" se met alors consciemment en situation d'adopter le comportement le plus simple, le plus intuitif et associatif possible, et ce quel que soit la complexité des systèmes qu'il consultera. Nous sommes alors dans le cadre d'un authentique processus "d'apprentissage périphérique" tel que défini par [Lave & Wenger, 91]. Dans ce processus, l'information qui sera prioritairement "captée" par l'utilisateur et servira de base aux associations qu'il échaffaudera pour aller au bout de sa quête, cette information donc, relève en premier lieu des propriétés invariantes de l'environnement : de la même manière que je peux utiliser un stylo comme un stylo si je veux écrire, ou comme un marteau si je veux planter un clou, je peux utiliser la liste des 10 premiers résultats d'un moteur de recherche de manière systématique (et aller voir chacune des pages vers lesquelles ils pointent) ou de manière

associative pour repérer aléatoirement (dans le texte de description fourni pour chaque page par exemple) de nouveaux mots-clés, de nouveaux noms de personnes qui vont m'engager sur une autre piste de recherche ou vont en l'état constituer une réponse/solution à ma question/problème³¹. La sérendipité est ici de type associative (cf. infra).

4.1.1 Sérendipité structurelle.

Admettons pour l'exemple, que nous nous trouvions dans une bibliothèque, à la recherche d'une thèse déjà repérée, pour construire un état de l'art sur une question donnée. Non loin de la thèse recherchée, sur le même rayonnement, figure une autre thèse dont le titre est évocateur et dans laquelle, après lecture rapide, nous trouvons effectivement des informations intéressantes. La sérendipité ici à l'œuvre est de type structurelle : elle est liée à une identification, à un parallélisme formel, structurel (de fait on est dans le rayonnement des thèses et non dans celui des journaux qui eût été moins approprié pour l'objectif de notre recherche initiale).

Admettons maintenant que nous effectuions la même recherche, dans la même bibliothèque, mais cette fois en consultant l'une des bases de données dont elle dispose : on utilisera alors les champs structurés de la base de donnée pour exprimer notre requête (mots du titre, nom de l'auteur, etc.). Selon la règle de " matching " applicable à tout type d'information structurée, l'échelle du phénomène de sérendipité se réduit considérablement, même si elle reste possible (un même auteur ayant pu rédiger deux thèses différentes par exemple) et demeure de nature structurelle.

4.1.2 Sérendipité associative.

Sur Internet, et plus généralement dans tout système distribué d'information non-structurée, ce phénomène change de nature et se donne à lire avec une acuité déterminante dans les stratégies de navigation choisies par les utilisateurs. Si l'on interroge un moteur de recherche en entrant une série de mots-clés (qui peuvent être les mêmes que ceux utilisés pour l'interrogation de la base de donnée), deux cas se présentent :

- le moteur de recherche dispose, dans sa base de donnée ou dans sa base d'index, d'informations présentant un relatif niveau de structuration (c'est par exemple le cas des annuaires de recherche si on les interroge en utilisant les catégories qu'ils proposent) : le phénomène de sérendipité structurelle reste opérant. Au vu du nombre de résultats possibles, dans ce cas comme dans les deux premiers évoqués (interrogation du rayonnement des thèses ou d'une base de donnée), le facteur déterminant consiste à limiter le silence (absence de résultats) ;
- le moteur de recherche ne dispose pas d'information structurée – ce qui demeure le cas le plus fréquent – et les listes de résultats qu'il présente à la requête de l'utilisateur sont alors considérables. La sérendipité se manifestant cette fois dans l'affichage possible d'un résultat pertinent bien que ne correspondant pas aux termes exacts de la requête est alors de nature associative. Le facteur déterminant dans les stratégies de navigation qui seront alors mises en place par l'internaute est celui qui lui permettra d'éviter le bruit et non plus le silence.

Notons ici que la sérendipité associative résulte de la conjugaison de phénomènes sémantiques, algorithmiques, individuels (usages) et techniques (référencement, balises méta³², spam³³ ...). On peut par ailleurs constater, avec la dernière génération de moteurs de recherche que le facteur déterminant redevient celui du modèle classique, c'est-à-dire éviter le silence³⁴.

5. Conclusion

5.1. Sérendipité et recherche d'information.

La sérendipité dans le cadre d'un processus de recherche d'information peut-être passagère (le temps que les modèles mentaux adéquats soient appelés) ou devenir un mode privilégié d'accès à l'information dans le cadre d'un processus de recherche ou de l'une de ses itérations. Elle se décline sous deux formes exclusives (structurelle et associative) qui dépendent principalement de variables d'environnement (structuré ou non).

³¹ c'est ce type de processus qui est systématisé par la plupart des outils de recherche ayant fait le choix de représentations graphiques (Kartoo, Mapstan, etc ...) pour optimiser l'instrumentalisation de ce type de sérendipité.

³² en HTML, ces balises permettent aux auteurs de contrôler l'indexation de leurs documents.

³³ le " spam " désigne les pratiques frauduleuses qui permettent de " fausser " l'indexation d'un document (faux mots-clés ...)

³⁴ certaines pratiques sont à ce titre tout à fait éclairantes du point de vue d'une " sociologie " de la recherche d'information, comme celle du " GoogleWhacking " (<http://www.googlewhack.com>) : Google étant le moteur de recherche le plus en vogue et celui disposant de la plus grande base d'index, cette pratique consiste à formuler des requêtes ne ramenant qu'une seule réponse.

Cette sérendipité a comme mérite méthodologique d'attester - s'il en était encore besoin - qu'il n'est pas nécessairement plus facile de trouver de l'information dans un système ordonné, structuré et formaté que, comme cela semble être le cas pour le web, dans un système d'information caractérisé par une forte entropie et ne disposant en tout cas d'aucun niveau de contrôle unique³⁵. Pour autant, il nous semble essentiel de se donner les moyens de penser la diffusion d'information et la structuration de contenus numériques en des termes qui prendront en compte, à la source, les sauts conceptuels et autres ruptures d'arborescence dont se nourrit la sérendipité. Les principales voies de recherche œuvrant actuellement dans ce domaine sont celles du web sémantique, des hypermédias d'apprentissage et bien entendu des approches théoriques de la recherche d'information (IR).

Parallèlement, une étude globale des scénarios de navigation disponibles sur le web [Ertzscheid 02] doit permettre d'identifier quelques invariants qui, à leur tour, constitueront un recours précieux permettant d'aller dans le sens d'une plus grande adéquation entre les objectifs visés par l'hypertexte, les habitus techniques sollicités et les styles cognitifs à l'œuvre chez l'utilisateur.

5.2. Recherche d'information et complexité

Au vu des quelques éléments décrits dans cet article (contexte réseau, diachronie des techniques de gestion de l'information, nouvelles approches de l'indexation et nouveaux biais, nouveaux comportements et nouvelles pratiques), il nous semble que le champ d'étude que constitue la recherche d'information doit être défini comme un **processus d'apprentissage dynamique**. C'est à la lumière de ce processus que devient chaque jour plus perceptible le **renouvellement des technologies intellectuelles** de classement, de représentation et d'accès aux connaissances. Dans l'utilisation qui est faite des ontologies, dans les thématiques qui émergent de domaines connexes à l'InfoCom (ingénierie documentaire, ingénierie des connaissances, hypermédias pédagogiques, etc ...), dans les perspectives ouvertes par le web sémantique et la main mise sur le web de technologies agents de plus en plus sophistiquées et transparentes, il semble clair que ce renouvellement des technologies intellectuelles passe par la combinaison - au sein de systèmes d'informations (eux-mêmes de plus en plus complexes et distribués) - de modèles formels hérités et de modèles plus ouverts, c'est à dire intégrant l'entropie comme partie intégrante du processus. Alors, au prix de la mise en œuvre d'un cycle cohérent et repérable de gestion des connaissances, l'émergence peut être prise en compte, de nouveaux éléments d'information voir le jour et le processus de recherche bénéficier de cet enrichissement constant en le déclinant sur plusieurs niveaux dépendant de l'acculturation de l'utilisateur à ces phénomènes. C'est, nous semble-t-il, à ce prix que l'adéquation nécessaire entre les méthodologies de recherche et le monde et/ou les objets qu'elles veulent cerner et décrire demeurera pérenne.

5.3. Sérendipité et hypertexte

Si l'hypertexte et plus globalement les hypermédias générés constituent un terrain d'observation et d'expérimentation privilégié pour l'étude de la sérendipité c'est parce qu'ils sont l'unique moyen d'organisation et de classification des connaissances qui *"offre comme capacité inhérente la création de classifications latérales."* [Balasubramanian 94] C'est cette dimension de "latéralité"³⁶ que tentent en permanence d'implémenter différents outils de recherche pour offrir des pistes d'accès à des mondes des plus en plus complexes³⁷.

5.4. Sérendipité et créativité

Appréhender dans leur complexité les phénomènes, les objets de recherches semble une tâche impossible. Pour sortir de la "boucle récursive"³⁸ ou graphe complexe le chercheur doit entreprendre différentes stratégies et opter pour les choix qui s'offrent à lui. La troisième voie est introduite par la sérendipité. Elle est une approche socio-cognitive de la recherche d'information et impose l'abduction comme heuristique.

³⁵ Un protocole expérimental est actuellement en cours à l'Université de Toulouse 1, auprès d'étudiants en première année de DEUG de droit pour évaluer les usages novices en recherche d'information. Le phénomène de sérendipité peut ainsi être "expérimentalement" observé.

³⁶ On parle de "latéralité" en recherche documentaire à propos de la reformulation de requêtes. De Bono, "Lateral Thinking", Penguin Books, 1990

³⁷ Voir notamment le mouvement initié par les "Folders" de Northernlight, repris actuellement par des outils comme Vivissimo ou constituant le cœur de technologie de sociétés (Exalead).

³⁸ Morin E, « La méthode »

Pour tenter d'expliquer l'influence de la serendipité en matière de construction de connaissances, il semble que deux dimensions soient à retenir : l'importance du contexte et le transfert de compétences dans une situation nouvelle (métaphore). Le contexte est notamment composé de la connaissance et des technologies intellectuelles qui la manipulent. Le transfert de compétences dans une situation nouvelle est lié à l'appropriation d'une culture technique et informationnelle, de savoirs par les chercheurs et leur capacité à transposer, transfigurer des phénomènes, des problèmes.

La serendipité se réalise alors par l'appropriation individuelle du contexte socio-technique, une *lecture* spécifique, créative du réservoir cognitif et instrumental. Les chercheurs les plus en phase avec le contexte socio-technique favorisent ainsi leur perspicacité et la mise en œuvre d'artefacts informationnels qui permettent de faire apparaître des éléments stochastiques.

Lors d'achoppements du processus de production scientifique ou de surcharge cognitive [Ertzscheid 03], quand il est impossible de rendre compte de phénomènes, un saut qualitatif doit être réalisé... la serendipité guette.

6. Bibliographie

Bachimont B., « Du texte à l'hypotexte les parcours de la mémoire documentaire », Technologie, Idéologies, Pratiques, n° spécial « Mémoires collectives », 1999.

Balasubramanian V., « **State of the Art on Hypermedia Issues And Applications.** » [en ligne] http://www.isg.sfu.ca/~duchier/misc/hypertext_review/, consulté le 26/10/2001.

Barabasi, A.-L., Jeong H., Albert R., " The Diameter of the World Wide Web ", pp.130-131 in Nature, 401, 1999. [en ligne] http://xxx.lanl.gov/PS_Cache/cond-mat/pdf/9907/9907038.pdf, consulté le 05/07/2002.

Bateson G., **Vers une écologie de l'esprit, T. 1.** Paris, Seuil, 1977.

Belkin N., *Helping People Find What They Don't Know*, in Communications of the ACM, August 2000, Vol. 43, No. 8.

Boursier & Van Anandel, « Serendipity : expect also the unexpected », creativity and innovation management, vol 3, p.20-32, 1992.

Bush V., « *As We May Think.* », pp. 101-108, in **The Atlantic Monthly**, vol.1, n°176, Juillet 1945. [en ligne] <http://www.isg.sfu.ca/~duchier/misc/vbush>, consulté le 07/02/1998.

Carr L., Hall W., Lewis P.H., De Roure D., « *The significance of Linking.* », in **ACM Computing Surveys**, vol. 31, n°4, Décembre 1999. [en ligne] http://www.cs.brown.edu/memex/ACM_HypertextTestbed/papers/20.html, consulté le 22/03/2002.

Engelbart D.C., « **Augmenting Human Intellect : a Conceptual Framework** », Summary Report, AFOSR-3233, Stanford Research Institute (SRI), Contract AF49(638)-1024, SRI Project N° 3578, Octobre 1962. [en ligne] <http://www.histech.rwth-aachen.de/www/quellen/engelbart/ahi62index.html>, consulté le 03/03/2002.

Ertzscheid O., Les enjeux cognitifs et stylistiques de l'organisation hypertextuelle, Thèse en Sciences de l'information et de la communication, Université de Toulouse 2, sous la dir. de FC Gaudard & J. Link-Pezet, 450 pages. [en ligne] <http://www.ertzscheid.net>, consulté le 10/06/03.

Ertzscheid O., "Syndrome d'Elpenor et sérendipité : deux nouveaux paramètres pour l'analyse de la navigation hypermédia." in Actes du colloque H2PTM'03. Editions Hermès, septembre 2003

Fayet-Scribe Sylvie "histoire de la documentation en France : Culture science et technologie de l'information", CNRS éditions, 2000.

Figueirado A. Dias de, Campos J., "The Serendipity Equations". *Proceedings of the Workshop Program at the Fourth International Conference on Case-Based Reasoning*, ICCBR 2001, Technical Note AIC-01-003. Washington, DC: Naval Research Laboratory, Navy Center for Applied Research in Artificial Intelligence [en ligne] max.ipv.pt/pub/AdeFigueiredo01.pdf

Gallezot G., « La recherche in silico » In : Chartron G. (dir.) *Les chercheurs et la documentation électronique : nouveaux services, nouveaux usages*, Edition du cercle de la Librairie, Coll. Bibliothèque, juillet 2002.

Gallezot G., "Exploration informationnelle et construction des connaissances en génomique", *Les Cahiers du numérique*, Hermès, vol.3, n°3, novembre 2002.

Kleinberg Jon Lawrence Steve, « The structure of the web », *Science*, vol 294, 30 nov. 2001, p 1849-1850.

Koll Matthew, "Information Retrieval", bulletin de Jasis vol. 26, N°2 Dec/jan 2000.
<http://www.asis.org/Bulletin/Jan-00/track_3.html>

Kolmayer Elisabeth, Peyrelong Marie-France, « Partage de connaissances ou partage de documents », *Document numérique* . vol 3(3/4):283-299. 01 décembre 1999. et
http://archivesic.ccsd.cnrs.fr/documents/archives0/00/00/01/00/index_fr.html

Lave G., Wenger E., *Situated Learning : Legitimate Peripheral Participation*. New-York, Cambridge University Press, 1991.

Marti Y.-M., "Dirigeants : quelle posture de combat ?" <<http://www.Egideria.fr/posturecombat.html>>

Perriault J., " Effet diligence, effet serendip et autres défis pour les sciences de l'information. " [en ligne] <http://www.limsi.fr/WkG/PCD2000/textes/perriault.html>, consulté le 15/02/01.

Shneiderman, B. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley Publishing Company, Reading, MA, 1997.

Toms Elaine G. « Serendipitous Information Retrieval » < http://www.ercim.org/publication/ws-proceedings/DelNoe01/3_Toms.pdf >

Virilio, P., *La machine de vision*, Ed. Galilée, Paris, 1988.

Reflections on Information Literacy, Information Retrieval and Training

Keywords: Information literacy - Information seeking - Information retrieval - Systemic information processing - Cognitive information processing - Information seeking as a learning process -

Introduction

The field of information retrieval (I.R.) has been since the second world war the catalyst of many problems caused by the exponential growth of information, which has reached a critical threshold with the arrival of networks. These ones make obvious the fundamental transversalities with other scientific fields, such as cognition, man-machine interfaces and knowledge engineering, because of – primarily but not exclusively – a complexity due to the information hypertextual organization.

Since 1997³⁹, we (as Urfist and training organism) were confronted (with others) to the need to find solutions to train the university students to bibliographic research. We discovered the complex nature of the difficulty to train a large number of students to information literacy. We also found out how difficult it is to find the relevant teachers (trained in information retrieval and information use). It became rapidly obvious that the solution to train the students of the various campuses was either to train teachers from a discipline to the techniques and methods of information retrieval or information specialists (librarians, documentalists, research engineers) to pedagogical methods. That's why we started organizing trainings of trainers in IR. Among the difficulties encountered, the most important one was the absence of a commonly shared representation of what IR is (which is extremely amazing as this "science" is supposed to produce classification schemes for various disciplines, but seemed unable to produce its own representation, despite a few attempts⁴⁰).

After training various groups of trainers, from what had been said during the workshops and using Takeuchi and Nonaka's [1] models of knowledge conversion, we reached a few results that we are presenting in this contribution. In this paper, we present four main aspects:

Information retrieval characteristics :

- 1.1. Information processing in the information system
- 1.2. Information seeking as a learning process
- 1.3. A description of information seeking acts

Information literacy, as a cooperative model domain (existence of various tools at various levels of use by various persons with various outcomes)

A knowledge management platform as a collaborative attempt of knowledge and pedagogical resources mutualization : FoRSIC⁴¹

The major role of ontologies to represent knowledge within those knowledge management systems to encourage and sustain the evolution of the domain model.

1. Information retrieval characteristics

This question has long been of interest to us and we participated IMIS [2],[3],[4],[5],[6], a European project aiming at the creation of an expert system to retrieve information from commercial databases for non-expert users. This study was based on the observation of information retrieval experts while searching information on a very specific question in a given field. Instead of coming out with such an expert system, we could only reach a few recommendations (as a course of actions) to help non-experts to use these databases. We finally came to consider the existence of two different models at work simultaneously in the information retrieval process. One of them concerns the information and its transformations (organization) in the information system, the second concerns the cognitive action and solution oriented activity. Both models can be considered as a process of transformations that have to be learnt, the second one evolving according to the expression of the user's need and his/her searching abilities (i.e. his/her information culture). We present in this section a brief state of the art on both of these aspects of the information retrieval process and a point of view considering both aspects present in information retrieval as the constitutive parts of a learning process.

1.1. Information processing in the information system

The information memorization in a system corresponds to various transformations : computational, semantic and technological. This transformational process is based on four main activities

³⁹ arrêté ministériel du 9 Avril 1997.

⁴⁰ see point 2. « Information literacy as a cooperative domain model. »

⁴¹ FoRSIC : Formation et Recherche en Science de l'information et de la communication

(a) collecting, (b) organizing, (c) analyzing, (d) communicating. (use of procedures (system), and interaction between the system and the user).

- **From a computational point of view:** the information results from the processing of data into stored, organized information, structured into searchable indexes, according to a specific architecture and to a specific program.
- **From a linguistic point of view:** there are two different transformations : one is descriptive, the other is analytical. The first representation deals with the identity of the document. It is made by using the paratextual elements of the document : authors, title, edition, publication, date, source, author's affiliation. The second representation concerns the semantics of the text, its construction and meaning. The transformation deals mainly with the information contents and its representation in a semantic perspective through formal languages, classifications and indexed searchable fields. The key-words act as structuring indicators for the scientific field they belong to. They correspond to the historic evolution of knowledge in a specialized domain or discipline. The formal languages used reflect the historical evolution of a given field: controlled vocabularies (index, thesauri, syndetic structures). They represent at the same time keys to understand the type of processing the memorized information has received in the information system and to move into the represented data for selecting (identifying), locating and fetching the relevant information [7].
- **From a communicational point of view (man-machine interface):** Buckland [8] calls the first two operations: representation-making and partitioning (selecting, sorting) and the last operation matching, which takes place through the choice of words (matching the representation of information in the system with the user's internal knowledge). This moment concerns the interaction between the system, its architecture, its structure and the user or group of users. This happens according to an exchange pattern when the system and the user are alternatively emitters and receivers. Belkin insists on the user's necessary familiarity with the underlying operations (architecture, programming, languages, commands) and also qualifies them as matching operations. He speaks for example of "exact match" and "partial match" to express the user's satisfaction. The familiarity with the functions of the system, as well as with the vocabularies describing the information objects, is necessary. Belkin shows how difficult language is to use in this type of communication situation with synonyms and/or homonyms adding to the natural ambiguity of language. As a result, databases are quite uneasy to search. For a long time, information seeking has been performed by experts like reference librarians or documentalists: the mediator's presence being an additional difficulty especially on the question of relevance.

The cognitive processing of information

This aspect deals with the interaction between the information system, the user and the emergence of meaning. It concerns the psycho-cognitive activity implied in the information-seeking process and in the conversion of information into knowledge. This occurs through various mental operations involved in information seeking itself and in the use of the retrieved information for intellectual activities like reading, writing, producing new pieces of knowledge. The theoretical perspectives have tremendously changed these last few years. This important evolution leads to reconsider the way we think, the way we remember, the way we reason. What do people want when they initiate an information-seeking process? A possible answer to that question could be: they hope to address some kind of problem or achieve some goal, for which their knowledge is inadequate.

Generally speaking, concerning students, it is amazing to find out that they have at first very little information need. This need has to be roused (curiosity or motivation) and this is part of the teaching challenge. It is all the more challenging as most of the teachers tend to start by teaching the importance of a clear demand expression, whereas most of the time, the students have no demand at all. This does not indicate a lack of motivation, it essentially indicates that the process has to start somewhere, has to be bootstrapped and then and only then, learning can start.

Coding/decoding, information interpretation theories

From the beginning, information seeking has been considered in a problem-solving perspective [9]: the problem to solve is divided into sub-problems, which must be solved the one after the other. Courbon, for example [10], sees information as constituted by the sign and the code. This interpretative model mainly concerns the procedures and processing necessary to judge the value of things and transform them into knowledge after they have been mentally recognized. The world described here is predefined, does not evolve and the problem-solving activity requires a long process, numerous steps, a division of the problems into sequential operations (sub-problems).

To this point of view, Simon [11] prefers the notion of limited rationality, which admits the existence of psychological preferences (more or less self-organizational approach introducing motivation, affect) to guide the problem-solving process. Though he considers that man reasons in a sequential way, he also thinks that he sorts the problems to be able to deal with the quantity of information. He considers that the mind works according to four different, iterative phases: **Intelligence** (understanding the problem), **Design** (organizing and modeling the various possible solutions), **Choice** (choosing a solution), **Control** (checking on the user's satisfaction with the results). Here, rationality includes the process, but it does not only correspond to a formal (analytical and deductive) reasoning process.

Both Courbon and Simon agree on the fact that in the information seeking process and interpretation process, a representation of the problem to solve is built, from which depend both the realization and the success of the information-seeking process (to the user's satisfaction).

In an equivalent perspective, Cole [12] uses K. Popper's scheme of conjectures and refutations to represent the users' various cognitive states leading from an initial problem state (unorganized, but containing some ideas about concepts) to progress through interaction with the contents of the information system towards a satisfying solution to the problem. According to Cole, the interaction results in the user's reconceptualization of the problem. Cole sees information-seeking as a three-level model resting on perception (identify), comprehension (design and choose), application (produce). His cognitive definition of information-seeking is that the "user grasps bits of understanding from the information system, he creates a new message through feedback and interaction" [12].

We consider the information-seeking process as a learning process taking place through the evolution of the mental representation one has of the problem in more or less familiar situations.

1.2. Information-seeking as a learning process

From a cognitive point of view, J. Piaget [13] considers the outcome of information-seeking as a "disequilibrium", as the acquisition of new information unbalancing a person's former knowledge. Learning, a highly dynamic process, occurs when information is transformed into knowledge. The INRC model (Identification, Negation, Reciprocal, Correlative) shows the dynamic steps of transformations conducted through mental operations like assimilation, accommodation and transduction. This happens with each individual when he/she receives a new piece of information, before it becomes part of his/her knowledge (differentiation, integration, application).

The user, through the information he receives after the query is made, learns about the query itself, about the (constantly evolving) perception of his/her information need, in a succession of transformations, to which he adds strategic perspectives to embody this knowledge into an action to perform. In this situation, information is meaningful, goal and action oriented. One of the remaining difficulties concerns the high dynamics and permanent evolution both of the environment and of the user's need (time-oriented environment, actions).

Information-seeking as a goal-oriented, situated, learning process

The cognitive environment we just described is an open, unstable, complex construct of events, interaction, retroaction, action. The mental organization necessary to perform an action is constantly modified by new pieces of information which creates a new mental organization. Knowledge is created through the flow of information, anchored on the beliefs and commitments of its holder (subjective nature of knowledge in social context).

Once again, we wish to underline the importance of language in those matching activities:

- Formulation of the question in proper concepts (explicit versus implicit knowledge (subjective aspect))
- Know-how to use the information system to reach goals
- Knowledge creation process from the obtained information
- Teleonomic (intentional and strategic) aspects of the action to perform, from information to knowledge in regard to action and social construction

The question of culture, experience, mind/body relationship, memory [14] is very important, as the experiential and cultural aspects are part of the understanding of the world and of the social context [15]. The more complex the situation is, the more precise the goals must be to make sense, to understand a given situation the meaning of which is linked to the context of use. Culture and experience provide use scenarios that allow to anticipate the various solutions from the semantics of the situations. Here we try to show how to search for information in classical information systems the user must use at the same time the knowledge he has of the available tools, and a cognitive activity, because of the relationship between information organization, cognitive strategy, action context and

action outcome. Therefore, we try to express the dynamic process of information-seeking behavior in a given intellectual environment. We will study these various levels of use.

Levels of use, levels of learning.

We take into consideration the information systems, the user's mental activity, the environment and the action to take. A. David describes four various uses of information seeking based on evocative cognitive theory [16]. We consider information seeking as the organization of ideas and concepts according to an action performed in a given context or environment. We distinguish three levels of information-seeking literacy, each step of which corresponds to a transformation taking place during the information-seeking process.

According to Hutchins' [17] theory of propagation of representational states, we consider that in information seeking the user activates a general information-seeking model (based on hierarchical, sequential, explicit models), then propagates it into different contexts to perform more and more specific (tacit) types of searches in various environments more or less distributed, more or less cooperative. We consider information-seeking as a process of usable knowledge acquisition, allowing access to more and more complex situations (time, space, uncertainty) based on the goals (work with the end in mind), the relevance of the retrieved information depending on the user's knowledge of the context, the relevance of the interpretation he makes. This interpretation is based on his former knowledge (culture and experience) and strategy of action (which in turn will create more knowledge for the user, more information to the environment).

We consider information-seeking as a three-stage learning process:

1 - The level of culture and socialization: The information-seeking concerns essentially the description of an event, of a domain, of a subject. The question to answer is "What it is about?", the intention and goal are to understand, to know from what exists. It can concern the information environment itself: the available tools, uses and procedures to use them, the techniques to analyze the results. It is a social construct and an understanding of the laws of the systems, the ability to solve well-defined problems, to understand the principal concepts.

2 - The experience of the context: Collecting information is a mean to observe and understand the situation and the variety of the elements involved there. From the observation stage, this level takes the user to a more involving, embodied use of information based on the cultural level of information previously acquired. The user understands the various technologies implied by the situations and the level of description of the situation, the categories. He is able to analyze them as separate components, specific to the field they belong to. Relevance depends on the level of knowledge the user has of the environment and the perception he has of his own needs based on a sense-making approach of the situations.

3 - The strategy level: The question is no longer what it is, but what to do, what to choose, what to decide ? The information-seeking activity does not concern the objects, but the events, the situations, the occurrences, the multidimensional aspects of a situation. The final goal is action, whether it is to write a doctoral dissertation or to take decision in a strategic field. The challenge is to understand the types of relationship between elements, people, situations and to understand the interactions between those various elements constitutive of a given context to detect and find answers to uncertainty, to be able to build scenarios to understand the future.

Information seeking depends on the mental representation of the thinking and reasoning strategies necessary to retrieve and process information. At an even more generic level, we can say that each level of use corresponds (from explicit to tacit) to four mental activities that also correspond to the knowledge conversion steps: reasoning with (socialize and conceptualize), acting on (match words to tools, analyze and combine the results), producing a result (lectures notes, reports, strategic decision).

	Culture	Experience	Strategy
Reasoning with	laws	casual stories	casual models
Acting on	symbols	situations	conceptual situations
Resolving	well-defined problems	emergent problems	ill-defined problems, dilemmas
Producing	fixed meaning immutable concepts	social construct negotiable meaning	network understanding, future oriented multidimensional

Three levels of learning

1.3. Information seeking acts.

To summarize what we mean by information literacy training, we have defined information seeking as a social, cultural, contextual learning process and have identified what we call Information Seeking Acts:

- Identify the information tools and systems (i.e. the transformations that data received in the information system).
- Express a clear query.
- Choose information systems: select and categorize them according to the expressed need.
- Match the query, the tools and the access procedures.
- Analyze the retrieved information (according to the action to perform) which defines the relevance of the retrieved information.
- Embody the information into a new product(ion) or action or new knowledge.

Training people to train students to information searching is not easy, especially when they are practitioners, since they have a different view of the activity (based on tacit know-how). That's why we tried to produce an explicit representation (common language to share) of information literacy as a model domain, hoping it would help them to create relevant didactic resources.

Information literacy as a cooperative domain model

As we have said in the introduction, the IR field does not possess a language representing it as a whole, taking into account all its formalisms and specificities. After browsing the field, we tried to isolate a list of relevant terms that could be used as a basic classification. We consulted various linguistic tools : online dictionaries, online thesaurus like WordNet or the ASIS thesaurus of information science⁴² . We also used the expertise of the various members of the teaching team. We have gathered a list of 500 terms organized around three main entrances : « types of tools », « searching procedures » and « use scenarios ». Each of them can be extended to a seven level depth. Each term is illustrated by a generic definition, a dedicated pedagogical resource, an exercise or an example, plus a more generic pedagogical resource to recontextualize the concerned term.

As an example, here is the first level of this model :

- types of tools
 - supports and places of information
 - organization and internal structuration of information
 - nature of the stored information
- interrogation or searching procedures
 - operators
 - languages and classifications
 - information system architecture
- use scenarios
 - analysis technics for information evaluation
 - scientific knowledge production
 - production scenarios

This domain model has been implemented in XML to facilitate the navigation into the tree and also to make it portable onto the web. It can be consulted at the following adress : <http://www.urfist.cict.fr/forsic/domaine.html>

It is not a thesaurus properly speaking, it can be considered as a semantic network. Each term is related to the others by different types of relations like : « requires the knowledge of » (corresponds to a prerequisite notion) or « is useful to the knowledge of ». We tried to make this model as complete as possible. We tried also to limitate the vocabulary to the training practices (to information retrieval). We did not integrate any of the socio-organizational aspects specific to the field of documentation or libraries as it is the case with ASIS thesaurus, which is really very complete but too library-oriented.

Our wish has been to match this model to the notion of the trainers' competences, as each term deals with specific training competence such as « know how to use Boolean operators » in relation

⁴² WordNet : <http://www.cogsci.princeton.edu/~wn>, ASIS Thesaurus : <http://www.asis.org/Publications/Thesaurus/isframe.htm>.

with the various levels of knowledge described beforehand (culture, experience, strategy). It is possible to read this domain model in two different ways :

- used horizontally : the further, one goes into the arborescence, the more sophisticated levels we reach (the concepts concern a highly specific practice).
- used in a vertical way, the first item found is « types of tools » , then « searching techniques» and finally the « use scenarios » which is also a progression towards a more expert knowledge and more expert levels of use⁴³ and intellectual practice.

A COLLABORATIVE ATTEMPT OF KNOWLEDGE MUTUALIZATION : FORSIC

We apply those principles in FORSIC [18], a cooperative platform concerned with the mutualization of the knowledge of a community of trainers in information literacy. It is a virtual hypertextual system, in which each person's knowledge (knowledge, know-how, professional preferences and expertise, partners) is described and where every single item of information is both linked to other items (context creation) and to other persons (networking). Our goal is to build a collaborative situation and to describe the abilities of an intellectual workteam able to perform a collective task like training the students of campuses to information literacy and participate to the realization of this common goal. We use several technologies in a cooperative way (a knowledge base, a knowledge representation software, an authoring-tool and multi-agent software). This aim is to create and animate an intelligent community of practices [19], [20] through the cooperative interweaving of links allowing :

- to identify the trainers and their capabilities,
- to find the pedagogical resources they have created in order to add value to their knowledge,
- to extend their possibilities and create new resources by learning and adaptation,
- to identify the needs in correlation to the social context.

Basically, trainers know-how expression, their needs and the pedagogical resources they use or produce are indexed with the same competence index. This index is multidimensional (four dimensions: action, goal, environnement, result and five steps – one term for each – that express the transformational process of information retrieval). Though the index may seem a constraint, it is very useful and helpful to give a common representation and expression to each trainer who can identify himself/herself in clear terms of personal practice (as a teacher in relation with a professional practice).

FORSIC 's theoretical bases.

Though we work on the basis of knowledge management and knowledge conversion, distributed cognition and situated learning, here the central topic is training and defining information literacy in various contexts. We want to be able to represent the competences of the trainers (i.e. of the persons who have accepted to mutualize, index and comment them as well as their pedagogical resources and needs).

In this platform we also use an author-tool that allows people to create personalized pedagogical resources adapted to the various levels of training, and using the model of domain as a structuring basis. Using the trainer's personal competence, the level of competence the student is supposed to reach, the level of teaching, the discipline or domain, the pedagogical intentions, we build a guideline for training to answer specific needs. We hope to be able to open this service to the students soon (self-training). FoRSIC allows the optimization of cooperative situations by putting people in contact the one with the others without any centralized control and hierarchy offering them the possibility to work together and to exchange by knowledge sharing and reciprocal teaching. It allows situations of knowledge creation and co-construction by the validation (through use) of the resources. This network of interactions that we tried to build allows us to take into account the most adaptative and emerging aspects of information searching and information techniques by the re-creation of the context of use, of a teaching environment through the dynamics of situations.

Conclusions

Are ontologies important for information literacy ? Indexation, as well as the notion of ontologies, is very important for information management. We use ontologies as defined by Gruber [21] "*An ontology is a formal, explicit specification of a shared conceptualization*", where "formal" means "machine-readable", "explicit specification" means that concepts, properties, relations, functions,

⁴³ This is a general consideration. The good reading path should neither be horizontal or vertical, it should be transversal, as this transversality translates the very nature of information retrieval as a learning process different from the former sequential approaches.

constraints and axioms are explicitly defined, “shared” evokes consensual knowledge and “conceptualization” is an abstract model of some phenomenon in the world.

Ontologies are not dealing with the reality of the world but are built upon social conventions of representation : concepts, what they mean, the way they can be interpreted are drawing their meaning from the community where they are used and produced. They are the best way to enlarge a field of knowledge in a generic manner, to offer a common understanding of this field and to share or retrieve information (which is often the same), essentially because they allow the users to share a vocabulary but not necessarily a global vision of knowledge involved. Everybody knows things that others don't know : using ontologies in a multidimensional way (action, goal, environnement, result) can « force » retrieval and engages the sharing processes, by the explicit expression of what is commonly tacit.

Ontologies and collaborative I.R. in FoRSIC.

We have created a task-oriented ontology of competences in IR which allows learners to express their needs (“I want to ...”) and at the same time allows trainers to express their own competences (“I know how ...”). This particular ontology allows us to establish and implement an explicit link between :

- the domain model and the trainers competences (the terms composing the domain model can also be found in the competence label)
- the competences and the levels of use (each competence corresponds, can be indexed by a dedicated level of use)
- the pedagogical resources and the competences : each resource can be linked to one or many competences.

The urge for a dedicated teaching

FoRSIC is a device intending to collectively manage trainers' skills in the field of I.R. The task-ontology described is only the emerging part of a more global process of apprenticeship and knowledge sharing.

Most of the problems met when teaching information literacy are linked to the absence of a common representation of this field (tension between Library Science, Archives, Information and Communication Science, Technologies of Information and Communication). It is urgent to organize a course-schedule dealing with the information organization and retrieval, training a critical mass of people skilled in accessing and locating, organizing, filtering information, taking into account the impact of information technologies on uses and practices (cf SIMS, School of Information Management of Berkeley).

It is also important to organize a research field on those topics to give this emerging field the possibility to develop a real cognitive ecology.

References

- [1] Takeuchi, H., Nonaka, I. *The Knowledge-Creating Company*. New York : Oxford University Press, 1995.
- [2] IMIS Rapport d'activités pour 1990. “*Intellectual Multilingual Interface to European Databases*”, EEC Impact 1024. MITI, 1990.
- [3] Barthes-Régis, C. *Spécification du contrôle d'un raisonnement en monde évolutif : application à la recherche documentaire en ligne*. Thèse de l'Université Toulouse III, 1989.
- [4] Barthelme, S., Rolland-Godfriz, A. *Expertise et la familiarité de la situation dans les modèles de mise en œuvre des connaissances : le cas de l'interrogation des banques de données en ligne*. Maîtrise de l'Université de Toulouse III, 1992.
- [5] Duphil, B. *Une Interface intelligente pour la recherche documentaire*. Mémoire de DEA. CNAM, 1992.
- [6] Link-Pezet, J., Glize, P., Régis, C. “*A Cognitive approach to intelligent databases*”. On line, London : Learned Information, 1992.
- [7] Belkin, N. J., Croft, W. B. “Information Filtering and Information Retrieval: Two Sides of the Same Coin?”, *Communications of the ACM* Vol. 35 n°12, 1992
- [8] Buckland, M. *Information and Information Systems*. New York : Praeger, 1991.
- [9] Davis, R., Buchanann, B. G. “*Meta-Level knowledge : Overview and applications*”, IJCAI, ACM SIGIR, n° 5, Cambridge : 1984.
- [10] Courbon, J.C. *Système d'Information : modélisation, structuration et communication*. Paris : Interéditions, 1993.
- [11] Simon, H.A. “Applying Information Technology to Organization Design”. *Public Administration Review*, 32, pp. 268-278, 1964.

-
- [12] Cole, C. Interaction with an Enabling Information Retrieval System: Modeling the User's Decoding and Encoding Operations, *Journal of American Society for Information Science*, vol 51, number5, 2000.
- [13] Piaget, J. *L'Épistémologie cognitive*. Paris : PUF, 1978.
- [14] Varela, F. *Connaître : les sciences cognitives tendances et perspectives*. Paris : Seuil. Coll. La couleur des Idées, 1989.
- [15] Lave, J., Wenger, E. *Situated Learning: Legitimate peripheral participation*. ,Cambridge: Cambridge Press,1991.
- [16] David, A. Bueno, D. " Experimenting user model in cooperative IRS". Colloque SAS/SIC, Orlando, 2000.
- [17] Hutchins, E. *Cognition in the Wild*. Cambridge : MIT Press, 1995.
- [18] Link-Pezet, J, Gleizes, M.P., Glize,P. "Forsic: a self-organizing collective management system". Colloque SAS/SIC, Orlando, 2000
- [19] Levy, P. *L'Intelligence collective : Pour une anthropologie du cyberspace*. Paris : La Découverte, 1994.
- [20] Link-Pezet,J. Mémoires et intelligence collectives. Habilitation à Diriger des travaux de recherche, Toulouse 1, Juillet 2000.
- [21] Gruber T., "What is an ontology ?" <http://ontoserver.aifb.uni-karlsruhe.de/objectives.htm>

L'apport d'Internet à l'organisation des connaissances.

Vers un encyclopédisme d'usage.

Olivier Ertzscheid

Résumé :

Autour de la figure centrale que représente la bibliothèque (et ses derniers avatars numériques) par rapport à l'organisation du savoir et des connaissances, nous montrons en quoi Internet conditionne la mise en place et en œuvre d'un encyclopédisme d'usage, déjà largement perceptible dans les usages et les comportements standards des usagers des réseaux d'information.

Mots-clés : Bibliothèque ; Bibliothèque numérique ; Bibliothèque virtuelle ; organisation hypertextuelle ; Encyclopédisme ; Archive ; Archivistique topologique ;

Abstract :

Considering the central figure of the library (and its last numerical changes) from a knowledge organization point of view, we discuss the way Internet initiates the logic of an encyclopedism of use, already largely perceptible in the standard behaviors of the users upon information networks.

Keywords : Library ; Digital library ; Virtual Library ; hypertextual organization ; Encyclopedism ; Archive ; Topological Archive Science ;

1. Introduction

Au-delà de l'inventaire des modalités d'existence des bibliothèques (liées à l'archivage, à la conservation et à l'accès qu'elles permettent, aux supports dont elles ont la charge), la bibliothèque pose de manière obsédante la question du lieu : lieu de consécration de l'objet-livre, lieu de consultation et d'accès au(x) savoir(s), lieu qui organise la connaissance pour la distribuer selon les infrastructures technologiques disponibles à une époque donnée. Classique, électronique puis virtuelle, la bibliothèque est le premier de ces lieux collectifs dans lesquels l'organisation hypertextuelle des connaissances se donne à voir.

2. Bibliothèques : classique, électronique, virtuelle.

Les mutations des bibliothèques sont de deux ordres : celles liées à l'organisation des savoirs qu'elles permettent de conserver, et celles des traces, des supports sur lesquels sont inscrits ces savoirs. C'est cette double articulation qui permet de distinguer la bibliothèque classique, de la bibliothèque électronique (« *l'informatisation de la bibliothèque classique et la numérisation des textes* ») et de la bibliothèque virtuelle (« *la bibliothèque électronique plus le réseau, plus l'appropriation individuelle.* » [Giffard, 1996]).

Le passage à la bibliothèque électronique relève d'un changement de support quand la bibliothèque virtuelle procède d'une dynamique plus générale liée aux nouvelles formes d'organisation du savoir sur les réseaux. Ambitions intellectuelles et contraintes technologiques sont par ailleurs, et depuis toujours, les deux faces d'une même réalité bibliothéconomique : dès 1576, « *Alde Manuce réussit à fusionner le programme intellectuel de l'édition humaniste et le programme technologique de la typographie.* » [Giffard, 1996]. L'avènement du numérique et de l'organisation hypertextuelle dont il procède, pose cependant de nouveaux problèmes.

Premièrement, les bibliothèques virtuelles, tout en maintenant leur vocation de diffusion/conservation, « *reçoivent de l'information et non seulement des œuvres* » [Bougnoux, 1993 p.658], ce qui, deuxièmement, « (...) conduit à poser en d'autres termes la dialectique du « *périssable* » et du « *conservable* ». Il nous permettra peut-être de constituer et de transmettre un « *patrimoine du temporaire* ». » [BPI, 2001]. Sur ces aspects, le problème du versioning⁴⁴ constitue le principal champ d'investigation.

⁴⁴ Le versioning désigne l'ensemble des manières de gérer, à l'échelle de l'hypertexte planétaire, les procédures permettant de rattacher un texte à un auteur, tout en permettant à chacun de s'appropriier tout ou partie de documents produits par d'autres ou par eux-mêmes afin, premièrement, de limiter la prolifération « bruyante » des versions différentes d'une même information sur le réseau et deuxièmement, d'identifier la nature et les origines de ces modifications dans l'optique d'une gestion cohérente de l'ensemble des documents électroniques actuellement disponibles, indépendamment de leur format, de leur statut et en dehors de tout institution centralisée. Voir la thèse de [Whitehead, 2000].

Enfin, troisièmement, « *La technologie permettant de délivrer d'immenses quantités de données ne délivre pas simultanément une raison légitime d'accumuler ces informations.* » [Friedlander, 1995 p.75] Et avant même de songer à trouver une justification à la conservation de ce « *patrimoine du temporaire* », il faut organiser et homogénéiser les systèmes de classification permettant d'y accéder (Dublin Core, Métadonnées, TEI⁴⁵ ...).

2. Pour un nouvel encyclopédisme.

2.1 Encyclopédisme savant.

Confrontée à la masse d'information potentiellement recevable par les bibliothèques, aux difficultés à organiser cette masse de données et aux choix d'archivage qui sont en permanence opérés et remaniés, c'est la perspective encyclopédiste, demeurée intacte depuis la renaissance, qui se trouve remise en question.

« *Au catalogue actuel, ne faut-il pas substituer ou au moins adjoindre un système de cartographie intelligente qui permette de restituer chaque donnée dans son « arbre de savoir » ? (...) On observera que c'est aussi le chemin obligé d'une redéfinition de l'encyclopédisme, question qui est aujourd'hui massivement posée.* » [BPI 01]

L'encyclopédisme n'est en effet valide que tant qu'il peut prétendre à une certaine stabilité, au maintien de certaines linéarités :

- stabilité temporelle qui permet de maintenir un décalage minimal entre les avancées du savoir et les encyclopédies en rendant compte, dans une perspective cumulative ;
- linéarités croisées, transversalités permettant d'offrir des points d'entrée communs à un ensemble de domaines, de techniques.

Or, comme le note [Varet, 1956, p.69] et ce avant même les premières entreprises encyclopédistes : « *Le savoir affecte forcément une forme circulaire : c'est en effet la seule manière de se représenter un ensemble de données diverses tel que chacune renvoie à toutes les autres et ait perspective sur toutes les autres. (...) Ce savoir n'est pas simplement cumulatif (...) mais circulaire parce qu'il y a une circulation du savoir d'un point quelconque à tout autre point possible. Sans doute cette circulation se fait elle le long de certains axes perspectifs qui seront par habitude plus fréquentés que d'autres à l'intérieur du tout, mais dont la commodité ne tient jamais finalement qu'à un état momentané du savoir, à un équilibre météorologique métastable.* »

S'il est effectivement question de linéarités comme autant « *d'axes perceptifs momentanés* », elles ne sont que des commodités de l'analyse et ne témoignent en aucun cas de la nature des objets dont elles peuvent momentanément permettre de rendre compte.

Notre travail de doctorat [Ertzscheid, 2002] montre qu'à l'inverse de ce qu'écrit [Zeldin, 2001] pour qui « *L'Internet en est encore à l'âge de Bouvard et Pécuchet dont l'ambition était de copier toutes les connaissances.* », nombre de manières d'organiser ces connaissances sont déjà opératoires⁴⁶ sur l'ensemble des réseaux de communication et d'échange de savoirs, et que de nombreux et significatifs efforts⁴⁷ sont mis en place pour harmoniser et relier ces différentes approches, tendant à l'instauration d'un encyclopédisme d'usage(s).

2.2 Encyclopédisme d'usage.

La perspective historique dressée par Goody et reprise par Parrochia nous permettra de comprendre la nature du changement qui affecte l'idée encyclopédiste, l'encyclopédisme savant cédant la place à un encyclopédisme d'usage qui s'avère seul pertinent pour rendre compte des modes de transmission et d'organisation de la connaissance sur les réseaux.

Le premier type de classification fut celui introduit par l'usage des listes. « [les listes] *permirent non seulement de fixer en l'état un savoir mais aussi de poser des problèmes de classification au point d'atteindre les limites extrêmes d'un certain type de compréhension du monde.* » [Goody, 1979 p.169]. Une fois cette limite atteinte, d'autres types de classification furent proposés et prirent place en complémentarité des premières : « *Historiquement, d'abord on énumère (...), ensuite on rassemble (...), enfin on classe et l'on ordonne (...).* Dès

⁴⁵ Le Dublin Core, les métadonnées, et la Text Encoding Initiative (TEI) sont des projets permettant d'harmoniser et de structurer l'indexation des données sur le réseau en balisant les documents à l'aide de titres, de mots-clés, de descriptions bibliographiques, etc. La plupart de ces spécifications sont accessible depuis le site du consortium W3 (<http://www.w3c.org>)

⁴⁶ liens hypertextes dynamiques, adaptatifs, typage des liens, hypermédia éducatifs, etc.

⁴⁷ web sémantique, ingénierie des connaissances, ré-ingénierie documentaire, etc.

lors, trois grands types de structuration du savoir se font jour : le séquentiel linéaire (la liste), le central-circulaire (l'organisation systématique – encyclopédique), enfin le relationnel-réticulé (la base ou banque de données, le néo-fichier). » [Parrochia, 1993, p.181]

Ainsi la bibliothèque, quelle que soit sa dénomination (numérique, électronique ...) ne change pas : elle propose et dispose de ces trois invariants, s'inscrit pleinement dans leur cycle. Ce qui change c'est « l'accès » compris comme synthèse du niveau de connaissance et de structuration mentale (styles cognitifs) de l'utilisateur. Quand les formes traditionnelles de la bibliothèque supposaient un accès par l'un des trois modes (procéduraux) présentés ci-dessus pour pouvoir accéder à la connaissance sous forme déclarative, la dynamique s'inverse doublement avec la bibliothèque virtuelle.

Un nouveau scénario voit le jour : l'utilisateur accède (essentiellement grâce à du « *relationnel réticulé* » – voie privilégiée mais non-exclusive) à des connaissances déclaratives (données) à partir desquelles il va tenter de comprendre, d'inférer comment il peut avoir accès à d'autres données liées et considérées comme les plus pertinentes de son point de vue (et non plus de celui des classifications établies *a priori*). L'accès aux documents n'est plus subordonné aux modes de classement et d'organisation. Ce sont les modes de classement et d'organisation qui sont inférés de l'accès au document et de l'analyse de son contenu.

On parlera d'un encyclopédisme d'usage à partir du moment où ces parcours de recherche, d'accès et de consultation sont récupérés et réinjectés dans l'organisation de la bibliothèque virtuelle (Internet en est évidemment le premier exemple) pour organiser, à l'aune de ces parcours, les nouvelles données devant être classifiées⁴⁸.

3. De la bibliothèque-mémoire à la bibliothèque neuronale.

L'organisation, les modes de classification qui émergent ainsi de la prise en compte simultanée des parcours associatifs individualisés d'accès à l'information, augure d'une nouvelle étape dans les processus complexes de mutation qui travaillent les bibliothèques. Les conditions semblent réunies pour que s'opère un nouveau transfert de compétences depuis les capacités mnésiques individuelles vers celles, collectives, sur lesquelles s'érigent les lieux de conservation et de diffusion des savoirs.

Comme le note [Atlan, 1995] à propos des premières : « *On a affaire à une mémoire qui est, comme on dit, « a content address » : l'adresse est déterminée par le contenu, et non plus, comme précédemment par un lieu. (...) Ce qu'on va chercher n'est pas défini par le lieu où on l'a mis (...), mais par le contenu même de ce qu'on va chercher. On pourrait objecter : « Cela veut dire qu'il faut savoir à l'avance ce que je veux me rappeler. » Non ! Il suffit que je sache un tout petit morceau, plus ou moins grand, de ce que je veux me rappeler pour déclencher le rappel de la totalité. Cela ressemble plus à notre façon de nous souvenir ...* »

On retrouve ici le précepte fondateur de [Bush, 1945] : « *l'esprit humain fonctionne par associations* ». La bibliothèque du « *Neuroworld* » dont parle Leary⁴⁹ est désormais à portée de connaissance : ses modes de classement et d'organisation sont ceux qui, indépendamment du support permettent de « *déclencher le rappel de la totalité* ». Avant d'en arriver au stockage systématique et universel d'unités mémoires toutes interconnectées, les réalités informatiques – ou en passe de le devenir – que constituent déjà la transclusion⁵⁰, le versioning ou certains types de navigation (DataScapes - paysages de données), les recherches menées sur les réseaux neuronaux et synaptiques et leurs applications technologiques, le formidable déploiement de langages universels de balisage (XML), l'essor de l'ingénierie des connaissances et ses implications dans le domaine de la ré-ingénierie documentaire, le web sémantique et la place qu'il accorde à la construction d'ontologies⁵¹ en sont autant d'indices prometteurs.

⁴⁸ certains algorithmes de moteurs de recherche fonctionnent sur ce principe (google, clever) ; nombre de systèmes établissant des métadonnées se fondent également sur des études de profiling et plus généralement sur des études comportementales des usagers en recherche d'information.

⁴⁹ « Timothy Leary parle de « *Psybernétique* ». Selon lui « *la prochaine frontière est notre propre cerveau, le 'neuroworld'* ». » [Donguy, 1999]

⁵⁰ « *Mécanisme qui permet à un document d'être à plusieurs endroits simultanément. Un même paragraphe (...) pourra se trouver dans plusieurs contextes. Il ne s'agira pas de copies mais d'un original unique : lorsque l'on créera un lien avec ce document dans un environnement donné, on retrouvera ce lien dans tous les autres contextes. Le document ne sera donc pas dupliqué mais transclus, c'est à dire inclus simultanément dans divers environnements.* » Ted Nelson dans une entrevue avec [Eudes, 1996]

⁵¹ voir <http://www.semantic-web.org>

4. La bibliothèque sans livres : pour une archivistique topologique.

Le principal argument, tant philosophique que technique qui est développé comme horizon limitatif des mutations possibles de la bibliothèque actuelle, est celui dont fait état [Lepers, 1997 p.315] « *L'idée structuraliste d'une universalité des moules de la pensée humaine serait évidemment merveilleuse pour la constitution des hypertextes planétaires ; en réalité, cela relève plutôt de la fable, ou d'un mythe local, reflétant le fantasme universaliste de la rationalité occidentale. Il n'est pas question, bien entendu, de contester les bénéfices évidents de ce fantasme, mais de trouver les moyens d'une autre négociation entre les rationalités locales et la raison universelle.* »

Nous voulons ici montrer en quoi l'organisation hypertextuelle est une manière originale et pérenne de questionner les bénéfices possibles de ce fantasme, par la mise en place d'une archivistique « topologique » du savoir, à la confluence du rêve borgésien de la bibliothèque universelle et de la notion foucauldienne d'archive.

A cette fin, il nous faut d'abord poser la dimension oxymorique mais non contradictoire de ces deux approches : la bibliothèque borgésienne contient tous les livres quand l'archive foucauldienne est une bibliothèque sans livres.

« *Quand on proclama que la Bibliothèque comprenait tous les livres, la première réaction fut un bonheur extravagant. Tous les hommes se sentirent maîtres d'un trésor intact et secret. Il n'y avait pas de problème personnel ou mondial dont l'éloquente solution n'existât quelque part, dans quelque hexagone. L'univers se trouvait justifié (...) mais les chercheurs ne s'avaient pas que la probabilité pour un homme de trouver sa justification, ou même quelque perfide variante de la sienne, approche de zéro.* » [Borges, 1965 p.76]

Pour [Foucault, 1994 p.499] la perspective est inversée et c'est parce qu' « *On devrait tout lire, tout étudier (...) [qu'il faut] avoir à sa disposition l'archive générale d'une époque à un moment donné. Et l'archéologie est, au sens strict, la science de cette archive.* »

La finalité utopiste de ces deux approches est aujourd'hui au cœur de problématiques bien réelles.

Concernant le projet d'une bibliothèque universelle de Borges, on sait que depuis Paul Otlet et son "Mundaneum"⁵², les projets d'élaboration d'une bibliographie universelle⁵³ furent nombreux et envisagés tout à fait sérieusement par des organismes comme l'Unesco. S'ils furent pour la plupart progressivement abandonnés, ce n'est pas tant à cause de l'ampleur de la tâche et des moyens à mettre en œuvre – problèmes auxquels des solutions matérielles auraient pu être apportées – qu'à cause d'un singulier changement de perspective dans les mentalités. En effet, ces projets ne pouvaient jusqu'alors que difficilement se départir de la consonance utopiste qu'ils véhiculaient. Mais avec l'arrivée de l'hypertexte, l'utopie devient – sous l'effet d'une alchimie mal maîtrisée – l'a-topie. L'omniprésence du livre, sa désormais évidente ubiquité est réalisée hors de la matérialité des murs de la bibliothèque, en un lieu qui n'en est pas un, un lieu qui est l'absence même de lieu puisqu'il est constitué simultanément et instantanément de tous les autres lieux possibles et existants⁵⁴. Même si l'esprit demeure inapte à concevoir la notion d'infini, il ne peut qu'assister à la mise en œuvre désormais effective des réseaux infinis du savoir et de la connaissance.

Si la bibliothèque borgésienne demeure une vue fantasmée de la bibliothèque universelle, ce n'est pas tant par sa finalité que par son caractère a-topique. Depuis sa création la bibliothèque est en effet, par essence, un « *dispositif topographique* » :

« *Grâce à ce dispositif topographique de la bibliothèque d'où personne n'est exclu et où rien n'est interdit au regard analytique, se peuvent produire ces « expériences mentales » qui seront l'axe majeur de la révolution scientifique. La bibliothèque par son rassemblement ordonné multipliera la possibilité de synthèses partielles, en diffusera les compte-rendus pour élaborer une adhésion toujours rémissible et rehaussée de sa propre critique.* » [Damien, 1995 p.111]

Et, une fois admise la disparition du Livre, la bibliothèque s'efface devant la figure de l'archive, qui, parce qu'elle se développe conjointement et en adéquation parfaite avec l'ensemble des connaissances présentes à un moment donné (et non plus à un endroit donné comme dans les rayonnages de la bibliothèque), offre ainsi à cet ensemble de connaissance un espace topologique nécessairement adapté : la carte que dessine l'archive est à l'échelle du territoire de connaissance dont elle permet de rendre compte.

« *Par archive, j'entends d'abord la masse des choses dites dans une culture, conservées, valorisées, réutilisées, répétées et transformées. Bref toute cette masse verbale qui a été fabriquée par les hommes, investie dans leurs techniques et leurs institutions, et qui est tissée avec leur existence et leur histoire. Cette masse de choses dites, je l'envisage non pas du côté de la langue, du système linguistique qu'elles mettent en œuvre, mais du côté des*

⁵² <http://www.mundaneum.be/content/mundaneum/qqsots.html>

⁵³ préalable indispensable à une bibliothèque universelle.

⁵⁴ semblable donc à l'aleph borgésien.

opérations qui lui donnent naissance. Mon problème pourrait s'énoncer ainsi : comment se fait-il qu'à une époque donnée on puisse dire ceci et que jamais cela n'ait été dit ? C'est, en un mot, (...) l'analyse des conditions historiques qui rendent compte de ce qu'on dit ou de ce qu'on rejette, ou de ce qu'on transforme dans la masse des choses dites. » [Foucault, 1994 p.786]

5. Conclusion.

L'hypertexte, l'étude des modèles sur lesquels il s'érige, le mode de distribution des agencements collectifs d'énonciation qui le parcourent, est un moyen d'analyser (en fait de spatialiser, et d'analyser les répartitions qui se donnent alors à lire) cet ensemble d'opérations donnant naissance à l'information et à son ingénierie (« conservées, valorisées, réutilisées, répétées et transformées »). Voilà pourquoi la notion foucauldienne d'archive – à la condition que la concordance qu'elle établit entre la carte qu'elle trace et le territoire dont elle rend compte soit vérifiée – nous semble être la clef de voûte de la mise en œuvre d'une pragmatique de la connaissance, c'est à dire d'un mode d'accès aux connaissances qui prenne en compte non seulement la disponibilité de ces connaissances mais également la volonté et les stratégies individuelles et collectives d'accès, de diffusion et d'appropriation de ces connaissances.

6. Références.

- Atlan H., « Mémoire : entretien avec Roger-Pol Droit. », reprise d'une émission radiophonique donnée sur France Culture. [en ligne] <http://www.trivium.fr>, consulté le 15/03/1997.
- Bibliothèque publique d'information (texte collectif), « Babel ou le choix du caviste : la bibliothèque à l'heure du numérique. », in Colloque virtuel « Text-e » de la BPI Georges Pompidou. Année 2001. [en ligne] <http://www.text-e.org/>, consulté le 14/01/2002.
- Borges J-L., *Fictions*. Paris, Gallimard, 1965.
- Bougnoux D., *Sciences de l'information et de la communication*. Paris, Larousse, 1993.
- Bush V., « As We May Think. », pp. 101-108, in *The Atlantic Monthly*, vol.1, n°176, Juillet 1945. [en ligne] <http://www.isg.sfu.ca/~duchier/misc/vbush>, consulté le 07/02/1998.
- Damien R., *Bibliothèque et état*. Paris P.U.F, coll. « Questions », 1995.
- Donguy J., « Vers une esthétique du cyberspace. Ou une poésie du quantum linguistique. », in *Web-Doc(k)s* série 3, n°21-24, avec CD-ROM, 1999. [en ligne] <http://www.costis.org/x/donguy/cyberspace.htm>, consulté le 08/04/2002.
- Ertzscheid O., "Les enjeux cognitifs et stylistiques de l'organisation hypertextuelle." Thèse de doctorat. Université de Toulouse 2. Octobre 2002.
- Eudes Y., « Nelson, amiral de l'hypertexte. » Entretien avec Théodore Nelson, in *Le Monde Multimédia*, semaine du 1^{er} Avril 1996. [en ligne] <http://www.lemonde.fr/multimedia/sem1496/textes/act14961.html>, consulté le 28/11/1997.
- Foucault M., *Dits et écrits - 1954-88*. Tome I (1954-69). Paris, Gallimard, 1994.
- Friedlander L. « Du savoir à l'information : concevoir en pensant à l'utilisateur. », pp.75-84, in *Hypertextes et hypermédias : réalisations, outils, méthodes*. Balpe J.-P., Lelu A., Saleh I. (coordinateurs), Paris, Hermès, coll. « Techniques de l'information », 1995. Actes de la conférence « Hypertextes et hypermédias », 11-12 Mai 1995, Université Paris VIII.
- Giffard A., « Mnémoniques : la bibliothèque virtuelle. », in *Résonance*, n°10, Mai 1996. [en ligne] <http://varese.ircam.fr/articles/textes-publics/Giffard96a/>, consulté le 12/06/1997.
- Goody J., *La raison graphique*. Paris, Minuit, 1979.
- Lepers J.M., « Des simulacres rationnels locaux. », pp.311-317, in *Hypertextes et hypermédias*, Balpe J.-P., Lelu A., Nanard M., Saleh I. (sous la dir.de), H_PTM'97, vol.1, n°2-3-4 /1997, Paris, Hermès, 1997.
- Parrochia D., *Philosophie des réseaux*. Paris, P.U.F, 1993.
- Varet G., *Histoire et savoir - Introduction théorique à la bibliographie : les champs articulés de la bibliographie philosophique*. Paris, Les Belles Lettres, 1956.
- Whitehead E.J., An analysis of the Hypertext Versioning Domain, Doctor of Philosophy in Information and Computer Science, Université de Californie, Irvine, 2000. En ligne : http://www.cs.ucsc.edu/~ejw/papers/whitehead_diss.pdf, consulté le 04/06/2001.
- Zeldin T., « Le futur de l'Internet : une conversation avec Théodore Zeldin. » Entrevue menée par G. Orrigi. Colloque virtuel « Text-e » de la BPI Georges Pompidou. Année 2001. [en ligne] <http://www.text-e.org>, consulté le 14/12/01.

7.2. Pièces à joindre au dossier

7.2.1. Carte d'identité

7.2.2. Attestation de diplôme

7.2.3. Contrat CDD avec le SICD de Toulouse 1

7.2.4. Direction des études et enseignements dans le DESS Intelligence économique

Toulouse, le 13 septembre 2006

Attestation

Réf. :

Alain Alcouffe

Tél. : 05 61 63 38 73

Fax : 05 61 63 38 60

Mail : alain.alcouffe@univ-tlse1.fr

Objet : M. Olivier Ertzscheid

Je sous-signé Alain Alcouffe, professeur des universités, directeur du DESS intelligence économique (en cours de transformation en Master) certifie que M. Olivier Ertzscheid enseigne depuis trois ans dans le DESS. Ces enseignements y ont connu un fort succès et compte tenu de sa forte implication, il assure, à ma demande, depuis la rentrée 2003 (et dans la maquette du Master Intelligence économique) la direction des études.

Fait à Toulouse, pour valoir ce que de droit,



INTELLIGENCE ECONOMIQUE

Alain Alcouffe
Directeur du DESS

7.2.5. Enseignements CIES

Centre d'Initiation à l'Enseignement Supérieur

39, allées Jules Guesde

31000 Toulouse

Tél. 05.61.55.80.95

Fax. 05.61.55.80.96

E-mail : ciesmon@cict.fr

<http://www.univ-tlse1.fr/CIES>

Je soussignée, Christiane Alcouffe, directrice du CIES de Toulouse, déclare connaître Olivier Ertzscheid depuis trois ans, années au cours desquelles il a réalisé la formation des moniteurs du CIES à différentes techniques de l'Information et de la Communication, et ce à deux titres : en tant qu'intervenant dans de nombreux stages de formation de l'URFIST ouverts, par convention, aux Moniteurs, et également en tant qu'intervenant direct pour le compte propre du CIES, cadre dans lequel, O. Ertzscheid a élaboré et dispensé une formation spécifique à l'HTML et à la réalisation de pages Web à l'intention des futurs enseignants chercheurs que sont les moniteurs.

O. Ertzscheid nous a proposé à cette rentrée, en outre, une nouvelle formation aux principaux outils de la recherche scientifique.

Dans ces activités, O. Ertzscheid est extrêmement apprécié des moniteurs qui lui ont attribué l'une des meilleures évaluations jamais donnée au CIES (voir, dans l'extrait joint de notre « Bilan » annuel des stages de formation, d'une part le descriptif des stages proposés par O. Ertzscheid et les évaluations que les moniteurs en ont faites, et d'autre part, à titre de comparaison, dans le tableau récapitulatif, les évaluations obtenues par les autres formations).

O. Ertzscheid témoigne à la fois d'un grand souci de rigueur (apprécié de ce public exigeant constitué de doctorants particulièrement avare de leur temps) et de qualités pédagogiques certaines.

Pour ces différentes raisons, le CIES est très attaché aux prestations d'O. Ertzscheid et lui est reconnaissant d'accepter d'allouer une partie importante de son temps à la formation de nos futurs collègues.

Fait à Toulouse, ce 12 décembre 2003, pour valoir ce que de droit.

Christiane Alcouffe

Professeur des Universités,

Université des Sciences Sociales, UT 1,

Place Anatole France

31042 Toulouse cedex

e-mail: christiane.alcouffe@univ-tlse1.fr

Tel 05 61 12 87 58

7.2.6. Enseignements Licence InfoCom

7.2.7. Enseignements Ecole d'Orthophonie

7.3. *Rapports*

7.3.1. Pré-rapport de JP Balpe

7.3.2. Pré-rapport de F Rastier

7.3.3. Rapport de soutenance

