

Modélisation des méta-données bibliographiques : une contribution instrumentée destinée à améliorer la circulation des usagers dans les connaissances de la bibliothèque

Fabrice Papy

► To cite this version:

Fabrice Papy. Modélisation des méta-données bibliographiques : une contribution instrumentée destinée à améliorer la circulation des usagers dans les connaissances de la bibliothèque. Mounira Harzallah, Jean Charlet, Nathalie Aussenac-Gilles. IC - 17èmes Journées francophones d'Ingénierie des Connaissances, Jun 2006, Nantes, France. pp.171-180, 2006. <hal-01026328>

HAL Id: hal-01026328

<https://hal.inria.fr/hal-01026328>

Submitted on 21 Jul 2014

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

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

Modélisation des meta-données bibliographiques : une contribution instrumentée destinée à améliorer la circulation des usagers dans les connaissances de la bibliothèque

Fabrice Papy

Document numérique & Usages (DnU), Université Paris 8, 2 rue de la Liberté, 93526 Saint-Denis
<http://doc.univ-paris8.fr>
fabrice.papy@univ-paris8.fr

Résumé

La gestion des connaissances au sein des bibliothèques universitaires et de recherche se manifeste pragmatiquement par des opérations métacognitives de la part des professionnels de la bibliothèques qui ont pour but d'organiser un flux continu d'œuvres intellectuelles hétérogènes en correspondance avec les champs de recherches et les disciplines scientifiques enseignés dans les universités.

Les modèles conceptuels sur lesquels se construisent les processus d'acquisition et de mise à disposition des collections sont orientés système et sont dissimulés aux yeux des usagers. Ainsi, les catalogues en ligne (OPAC), ne permettent pas aux utilisateurs d'exploiter les collections aussi efficacement qu'ils le pourraient dans leurs tâches de recherche d'information.

Les modèles conceptuels des OPAC exploitent mal la richesse des meta-données résultant du travail taxinomique des professionnels et qui révèlent un modèle conceptuel d'organisation des connaissances puissant destiné d'une part à structurer des collections couvrant les domaines de savoirs et de connaissances scientifiques et d'autre part de favoriser la recherche d'informations au sens large (interrogation, exploration). Rendre ce modèle visible par l'utilisateur, favoriserait le repérage dans l'univers complexe de la bibliothèque et l'exploitation des richesses documentaires mises à sa disposition. L'interface de visualisation des données du catalogue que nous avons développée vise à restituer aux utilisateurs le modèle conceptuel orchestrant l'organisation intellectuelle sous-jacente. Le modèle que nous proposons pour ce dispositif tend à améliorer la connaissance de l'organisation intellectuelle de la bibliothèque et donc son appropriation par les usagers. Cette amélioration passe notamment par la mise en correspondance immédiate des classifications, de la distribution pluridisciplinaire des ouvrages et des éléments conceptuels RAMEAU.

Mots-clés : interface, OPAC, classification, concept RAMEAU, représentation, visualisation, navigation, modèle conceptuel de référence, modèle de données, paradigme système, paradigme utilisateur.

1 Modélisation orientée système

La modélisation joue un rôle important dans l'analyse et la conception des systèmes d'information [Axiome 90] [Meinardier98]. Elle rend possible le déploiement de solutions mieux adaptées aux contraintes des processus d'organisation. Mais ces modèles qu'ils soient issus de la recherche ou des organisations, associées ou non à des méthodes, sont délibérément tournées vers une approche système plutôt qu'utilisateur [Chaudiron 02a].

Bien sûr, il ne s'agit pas de soutenir la position selon laquelle la modélisation ignore ouvertement la question des utilisateurs mais elle part fréquemment du postulat qu'une situation modélisée et rationalisée puis informatisée sera plus aisément appréhendable par les utilisateurs même si ces derniers ne sont pas informés du ou des modèles élaborés.

Le sociologue Francis Pavé [Pavé 89] a montré à partir de cas concrets que ces présupposés relevaient d'une forme d'illusion informatique. Stéphane Chaudiron, dans la lignée des recherches anglo-saxonnes, montre en matière d'ingénierie documentaire liée à la recherche d'information que le paradigme système ne peut en aucun cas être privilégié ni se substituer au paradigme utilisateur [Chaudiron 02b].

Le monde des bases de données (relationnelles et objet) et des modèles conceptuels qu'il impose pour toute transposition de situations relevant du « monde réel », fournit, à ce titre, de nombreux exemples de cette rupture qui existe entre nécessité de la modélisation et exploitation effective de celle-ci par les utilisateurs [Planche 98].

Pourtant dès 1982, Y. Corson s'intéressant aux aspects psychologiques liés à l'interrogation des bases de données,

souligne l'importance de la vision conceptuelle [Corson 82] que peuvent avoir les opérateurs de l'organisation des bases de données et qui pourrait améliorer la performance de la recherche d'informations dès lors que les opérateurs ont à leur disposition les « *ossatures schématiques reproduisant les type de relations existant entre données* ».

Par ailleurs en ce qui concerne les catalogues, [Kulthau 04] montre que les études et les réalisations de dispositifs de recherche de type catalogues bibliographiques informatisés, sont contraints par la définition des besoins du système et qu'elles relèguent au second plan les considérations relatives aux usagers.

1.1 Modélisation des Systèmes d'information des bibliothèques

Cette question de la modélisation demeure entière, qu'il s'agisse de l'analyse de systèmes d'informations (SI) existant ou de la conception de nouveaux systèmes. Les systèmes d'information des bibliothèques ne dérogent pas à cette règle puisqu'ils sont depuis longtemps largement informatisés [Maisonneuve 03] [Lupovici 01]. L'informatisation a concerné principalement la chaîne de traitement des ouvrages (tâches professionnelles d'acquisition, de préparation, d'identification, de conservation, *etc.*) pour aboutir à proposer aux usagers la possibilité de consulter les catalogues des notices bibliographiques de façon électronique. Ce sont les OPAC (*On Line Public Access Catalog*) [Le Marec 90] qui permettent l'interrogation des données bibliographiques et qui sont, depuis la vague internet, consultables au sein - et hors - de la bibliothèque à partir de navigateurs Web.

Lévine et Pomerol constatait qu'en ce qui concerne le monde de l'entreprise la modélisation est difficilement dissociable de la gestion des connaissances puisque le modèle est la représentation des connaissances explicites de l'entreprise [Lévine 01]. Ce constat s'applique sans difficulté au monde des bibliothèques dont la raison d'être s'articule sur les points saillants de la conservation et de la diffusion des connaissances. La prééminence, la complexité, la lourdeur de la chaîne de traitement des ouvrages et les exigences de l'interopérabilité technique (protocole Z39.50) ont fortement orienté le processus de modélisation vers des questions manifestement plus proches du paradigme système (modèles conceptuel des données et des traitements).

1.2 Une organisation intellectuelle peu visible par les usagers

Ces impératifs de la gestion de la chaîne documentaire ont réduit les processus d'organisation intellectuelle et de cohérence globale du fonds et des collections, où interviennent l'expertise professionnelle des bibliothécaires (spécialités disciplinaires, connaissances des formations proposées par l'université) à quelques propriétés qui n'apparaissent que trop discrètement dans les OPAC.

C'est dans le cadre d'une action de recherches pluridisciplinaire (Sciences de l'Information et de la Communication, Informatique, Sciences de l'Éducation, Cartographie, Psychologie-Ergonomie) que nous avons cherché à vérifier l'hypothèse qu'un dispositif rénovant les fonctionnalités des OPAC pourrait améliorer les performances en Recherche d'Information (RI) des usagers principalement quand les recherches s'inscrivent dans un processus de repérages et de reconnaissances de champs et de frontières disciplinaires [Dinet 02] [Rice 01].

Il s'agit de rendre visible les objets conceptuels et leurs relations qui interviennent de façon significative dans l'organisation intellectuelle de la bibliothèque. On constate souvent que les représentations de la bibliothèque qu'ont les usagers sont erronées [Jolly 01]. Cette différence profonde entre la représentation qu'ils en ont et la réalité des missions de la bibliothèque universitaire et de recherche provient d'un référent inadapté.

En effet les usagers qui n'ont eu la possibilité de suivre dans leur cursus universitaire un cours de méthodologie documentaire [Fé0 98] gardent une représentation de la bibliothèque universitaire proche de celle qu'ils peuvent avoir du Centre de Documentation et d'Information (CDI) de leur période lycéenne ou bien encore construire à partir de leur pratique personnelle de la bibliothèque municipale.

Il arrive cependant que ces usagers aient pu suivre quelques heures de formations (dites formations aux usagers) mises en place par les Services Communs de Documentation (SCD) eux-mêmes. Ces formations de quelques heures subissent l'effet instrumental des dispositifs en libre accès proposés par les SCD. De fait, c'est la formation à la logique des outils de RI généralistes et spécialisés (moteurs de recherches, annuaires, encyclopédies, catalogues, bases de données) qui l'emporte lors des séances planifiées (2 x 3h).

Quelles que soient leurs qualités et l'implication du personnel du SCD, ces formations demeurent parcellaires (elles ne permettent pas de toucher tout le public de l'université) et insuffisantes pour aboutir chez l'utilisateur au développement de véritables habiletés pour maîtriser les dispositifs techniques. L'offre toujours plus importante d'outils informatiques et de bases de données destinés à la RI accapare le temps des formations, qui n'apportent alors que très indirectement un apprentissage de la cohérence globale du système d'information de la bibliothèque dans un objectif d'appropriation [Dinet 02].

1.3 Modélisation orientée vers la Recherche d'Information

Rouet *et al.* [Rouet 02] ainsi que Baccino [Baccino 04] ont montré que les tâches liées à la RI étaient à distinguer des tâches de lecture et de compréhension ou des tâches d'apprentissage.

La consultation d'un catalogue en ligne répond schématiquement pour un usager à la localisation d'un ouvrage recherché. Cette démarche, particulièrement vraie pour les

plus jeunes étudiants (1^{er} cycle) est très éloignée des pratiques des étudiants plus avancés notamment ceux dont la validation du diplôme passe par la rédaction d'un mémoire [Couzinet 96] [Polity 00].

La possibilité de pouvoir apprendre l'existence d'ouvrages dont ils ignoraient initialement la présence dans le fonds documentaire constitué (sérendipité) [Toms 00] et qui pourraient s'avérer constructif dans l'élaboration de leur travail universitaire repose pour une grande part sur la maîtrise de tâches relatives à la RI [Lefèvre 2000]. Au sein de la bibliothèque, ces activités propres à la RI demandent une connaissance importante des principes de structuration et d'organisation des savoirs.

La classification et les répertoires d'autorité matière, les logiques de distribution des ouvrages en salles de lecture et en magasin, la nature de l'information bibliographique interrogeable via le catalogue (champs et sous-champs UNIMARC) vont déterminer la capacité de compréhension de la cohérence globale de l'espace plurivoque de la bibliothèque et de l'exploiter activement.

Ce sont toutes les stratégies reliées à cette compréhension qui conduiront l'utilisateur à utiliser efficacement l'ensemble des services (PEB, expertise des professionnelles pour l'accès à des bases de données spécialisées, réseau de bibliothécaires) et des outils (catalogues électroniques en et hors la bibliothèque tels que le SUDOC ou ELECTRE, bases de données de dépouillement bibliographiques ou de texte intégral : FRANCIS, EUROPRESS, Persée, etc.) disponibles.

La réalité montre que les ressources acquises par les bibliothèques (traditionnelles et électroniques) sont peu voire pas pour certaines d'entre elles - utilisées par les usagers. A titre d'exemple le rapport d'activité 2004 publié par la bibliothèque de l'université Paris 8 montre que des 24 abonnements onéreux souscrits par la bibliothèque, le plus utilisé porte sur le bouquet de presse : « un succès important, côté public, du bouquet de presse Europresse (env. 20 500 connexions) ... une sous-utilisation frappante des bases de données de dépouillement des périodiques ».

Cette réalité fort dommageable en regard des efforts financiers supportés en direction de la documentation électronique montre de façon flagrante que saisir la cohérence et le sens de l'organisation des données est fondamental pour comprendre les données proposées à la consultation et se représenter mentalement leurs places dans la structure d'ensemble. Si cette représentation n'est pas claire, c'est tout le processus de RI qui est remis en cause [Baccino 04] [Ciaccia 05].

Les modèles conceptuels existent pourtant (FRBR, CIDOC) mais leurs orientations vers une approche données et le paradigme système les rendent peu utilisables par les usagers.

2 Modèles Conceptuels Muséographiques et bibliographiques

Les structures institutionnelles vouées à la conservation du patrimoine culturel (musées, bibliothèques, archives) se présentent comme des systèmes complexes, multidimensionnels, « de vrais objets de sens » qui s'articulent autour de notions relevant autant de l'informationnel (objets, données, méta-données) que du communicationnel (communautés, services aux usagers).

Le Modèle Conceptuel de Référence CIDOC [Gill 04] (cf. Fig. 1) est une ontologie de domaine orienté objet pour l'échange d'informations hétérogènes relatives au patrimoine culturel géré par les musées, les bibliothèques et les archives. Ce modèle est le résultat d'un travail de collaboration internationale de normalisation dirigée par ICOM/CIDOC, le comité international de la Documentation du Conseil International des Musées.

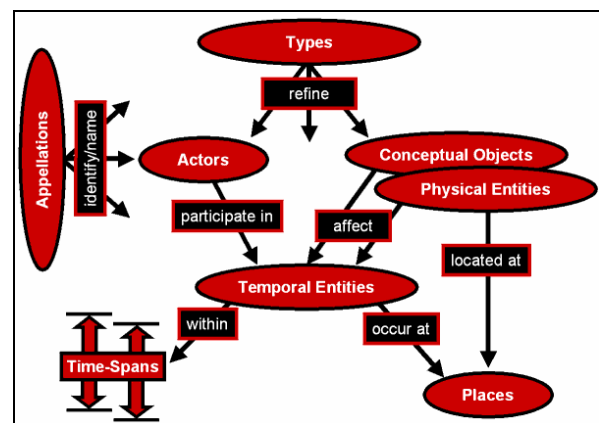


Fig. 1 : Vue du Modèle Conceptuel de Référence CIDOC CRM centré sur les événements [tiré de Gill 04]

A la suite de ces modèles, les bibliothèques, les archives et les musées ont élaboré des règles pragmatiques pour décrire leurs fonds et les rendre exploitables par les usagers. Ils ont créé des formats informatiques pour stocker et indexer ces descriptions sur machines. Or ces formats informatiques, reflets d'une modélisation évidente sont des modèles de données qui avantagent la position des systèmes (paradigme système) plus que celle des usagers (paradigme utilisateur). Ils représentent un effort d'explicitation des concepts et des relations qui s'arrête uniquement à la structure des données présentes dans les descriptions.

Le modèle CRM a été conçu à l'origine pour rendre compte de la sémantique des informations muséographiques, il est adaptable pour servir à expliciter la sémantique des informations bibliographiques contenues dans les notices au format MARC (*MAchine Readable Cataloguing*). Néanmoins, un modèle plus spécifiques aux bibliothèques existe : le FRBR (*Functioal Requirements for Bibliographic Records*).

C'est un modèle conceptuel de l'information bibliographique créé par l'IFLA [IFLA 98] qui distingue 4 niveaux dans l'objet que décrivent les catalogues : l'item, la manifestation, l'œuvre et l'expression.

Ces modèles intéressants du point de vue des données ne mettent pas en évidence l'activité productrice de cohérence des professionnels de la bibliothèque. Cette cohérence s'appuie d'une part sur des règles précises de catalogage, de classification (CDU, Dewey) et de description sémantique unifiée des notices établie par le réseau des professionnels des bibliothèques (BnF, bibliothèques universitaires, bibliothèques de lecture publique et de recherche) et d'autre part sur la connaissance des secteurs disciplinaires et des formations dispensées par l'université qui héberge le SCD. (cf. Fig 2)

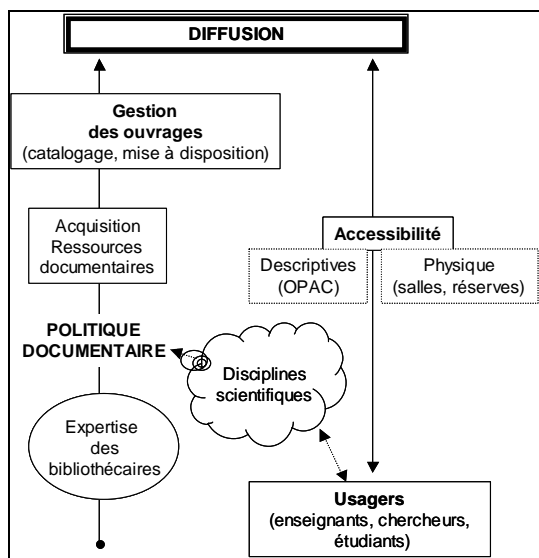


Fig. 2. Représentation générale de l'acquisition (politique documentaire) et de la diffusion (accessibilité) des ressources documentaires dans une bibliothèque.

3 Modèle conceptuel pour la RI dans un macro-système d'information

Les catalogues que les bibliothèques, musées et archives mettent à la disposition de leurs usagers sont pour l'essentiel des assemblages de meta-données descriptives qui répondent aux exigences du catalogage.

3.1 La bibliothèque : un macro-système d'informations, de connaissances et de communication inter-intra communautaire

La bibliothèque polarise différents objets (techniques, documentaires, cognitifs), différentes approches (bibliéconomiques, éducatives, institutionnelles, scientifiques) et différentes facettes (sociales, culturelles, économiques, prospectives). Cette attractivité conduit de nombreux systèmes à fonctionner ensemble, à subir une hiérarchie, un ordre, des conditions d'activation, des relations de dépendances, etc.

3.1.1 Un système d'informations

Essentiellement technique, il assure une tâche purement mécanique, sorte de courroie de distribution envoyant à certains sous-systèmes des données alimentées par les autres : commandes des ouvrages, livraisons, catalogage, prêts, traitements bibliéconomiques, etc.

Il est consultable par tout le personnel qualifié avec des finesses de visibilité et de mise à jour qui dépendent des fonctions assurées au sein de la bibliothèque.

Ce système d'information se décline en Système Informatisé de Documentation (SID) ou en Système Intégré de Gestion des Bibliothèques (SIGBD).

3.1.2 Un système de connaissances et de connaissances sur les connaissances

La bibliothèque n'est pas un lieu qui se caractérise par une production de matière première brute : les seules (et rares) données endogènes de ce type, sont des données techniques qui vont entrer dans un processus de gestion et alimenter le tableau de bord du directeur et des responsables de service et de collections.

Les matériaux intellectuels qu'elle propose à ses usagers ; alphabétiques, ouvrages de références, fondamentaux, méthodes, actes, essais, mémoires, etc. ont intrinsèquement une qualité intellectuelle ou informative et constituent des vecteurs de connaissances. Ainsi, la représentativité disciplinaire générique de l'université qui héberge la bibliothèque, les formations dispensées, les orientations de recherches des différents laboratoires, équipes et groupes de recherches, les propositions bibliographiques des enseignants, vont influencer la politique documentaire de l'établissement.

C'est toute l'expertise professionnelle des responsables d'acquisition corrélée à la connaissance du milieu universitaire ambiant qui va amener, au fil des acquisitions successives et de la constitution des collections, à établir la cohérence globale du fonds documentaire.

Des sélections sont alors indispensables, alimentées par des obligations autant d'ordre éducatif (tel ou tel secteur disciplinaire qui se développe en relation avec l'offre de

formations de l'Université) que scientifique (pôles d'excellence, partenariat scientifique, commissions du conseil scientifique, écoles doctorales). Au nom du choix, ce sont les connaissances professionnelles des bibliothécaires et la connaissance de leur lieu d'exercice qui sont finalement convoqués pour maintenir et faire évoluer qualitativement les collections.

3.1.3 Un système instrumental au service des usagers universitaires

Cet instrument intellectuel, conceptuel, technique et culturel représente, au-delà de sa mission d'archivage et de conservation, un lieu privilégié d'exercice du travail intellectuel où se construit et progresse la pensée de l'individu qui a accompli la démarche de s'approprier ce lieu de règles et d'organisation destiné à son usage [COULON 97]. La bibliothèque suggère de multiples approches et sollicitations. La diversité des démarches d'investigation informationnelle (qui inclut la démarche méthodologique) émanant des usagers et la multiplicité des ressources et services documentaires disponibles conduisent à des combinaisons infinies d'associations entre demandes d'information et objets documentaires.

3.1.4 Un système de communications issues de relations entre les communautés en présence.

La bibliothèque est un espace social [POLITY 00] de collaboration, d'échanges, de rencontres et de confrontations. Elle met en relation deux grandes communautés : les professionnels d'un côté, les usagers de l'autre qui introduisent de nombreuses variations.

Plus restreintes dans la communauté des bibliothécaires, les variations renvoient à des qualifications, responsabilités, profils, spécialités et compétences différentes qui dépassent le cadre des catégories professionnelles A, B et C.

Chez les usagers, la multiplicité des profils est plus importante encore car elle ne peut se limiter à la seule qualité réglementaire des personnes les plus couramment autorisées à accéder à cet espace situé dans l'enceinte universitaire : étudiants, stagiaires de la formation permanente, personnels administratifs et techniques, enseignants et chargés de cours, chercheurs et enseignants-chercheurs. Ainsi, la figure de l'utilisateur ne peut être dissociée de sa démarche informationnelle. De l'étudiant en première année de 1er cycle à celui en dernière année de doctorat, le tout factorisé par le nombre de formations diplômantes et de secteurs disciplinaires, il est aisé, sans rajouter les profils des membres du corps enseignant et de la recherche, de comprendre que la communauté des usagers est difficile à saisir et à circonscrire en terme d'identité.

3.2 Un modèle conceptuel en couche : classes, notices et éléments conceptuels RAMEAU

Les opérations de classification et de description sémantique à partir d'un langage documentaire et d'un vocabulaire contrôlé organisés ou non en thésaurus, la distribution d'un ouvrage dans une ou plusieurs classes thématiques constituent de la part des professionnels de ces lieux, de véritables actions meta-cognitives visant à homogénéiser et organiser les collections et le fonds documentaire d'un point de vue intellectuel [Svenonius 00] [Hunter 00] [Rowley 92].

L'organisation des salles de lecture et la distribution des secteurs disciplinaires au sein de chaque salle constituent de véritables connaissances de références introduites par les conservateurs et bibliothécaires afin de donner aux usagers tous les moyens de saisir les logiques d'organisation des collections et des proximités disciplinaires.

Le modèle que nous présentons (Fig. 3) rassemble sur 3 niveaux les objets qui sont porteurs de la visibilité de cette méta-connaissance d'organisation.

L'organisation de ce modèle en couches s'appuie et unifie les différents éléments utilisés par les bibliothécaires et/ou les usagers de façon distincte et à l'occasion de situations d'activités particulières.

En effet, la couche « Classe » correspond pratiquement aux exigences professionnelles du métier de bibliothécaire en conformité avec la politique documentaire du SCD. La classification adoptée pour les ouvrages disposés en libre accès (puisque les ouvrages en réserve sont uniquement affectés d'un numéro d'inventaire) constitue un repère intellectuel pour les responsables des acquisitions. Cette organisation intellectuelle globale leur permet de se représenter les secteurs disciplinaires couverts par les acquisitions et de mieux apprécier les efforts d'acquisition à porter vers tel ou tel autre secteur moins bien pourvu ou nécessitant une actualisation importante du fonds.

La couche « notice » représente l'activité de catalogage et de traitement pour chacun des ouvrages mis à disposition des usagers au sein de la bibliothèque. La notice est un lieu de convergence des activités des professionnels de la bibliothèque et des usagers. L'existence, le traitement et la disponibilité d'un ouvrage sont intimement liés à l'existence de la notice créée et actualisée par les bibliothécaires (acquisition, catalogage, traitement technique, prêt). La notice représente également l'objet de la recherche bibliographique informatisée. En effet, toutes les requêtes d'interrogation adressées à l'OPAC portent sur les champs de la notice bibliographique.

Enfin la couche « RAMEAU » renvoie à une activité de collaboration des professionnels des bibliothèques dont l'objectif est d'établir un langage d'indexation capable de décrire en quelques termes (les autorités-matières) le contenu d'un ouvrage. L'activité des professionnels au sein

du réseau RAMEAU est bien spécifique et ne recouvre pas les activités de ces mêmes professionnels au sein de leur service de rattachement.

La couche supérieure « Classe » signale les grandes catégories scientifiques relatives à la classification adoptée (CDU ou Dewey) et aménagée en fonction des spécialités développées par les établissements de l'enseignement supérieur. Chacune de ces classes regroupe plusieurs ouvrages qui sont reconnus par les acquéreurs comme conformes aux propriétés de la catégorie où ils figurent. Le sens des flèches du graphique partant de chacun des nœuds de la couche « Classe » peut être interprété comme « subdivision composée de ».

Ce sont les notices bibliographiques qui sont représentées dans la couche intermédiaire « entités ». La notice et les informations descriptives qui la composent (champs et sous-champs UNIMARC) apportent tous les éléments facilitant l'indexation et la recherche de l'ouvrage. Elles rassemblent des méta-données destinées aux usagers (Titres, auteurs, éditeurs, format, *etc.*) et aux bibliothécaires (zone SIBIL).

Les méta-données descriptives portent sur l'ouvrage en tant qu'entité et sur ses instances (les exemplaires). Les exemplaires sont dotés de propriétés spécifiques comme le statut (empruntable, exclu du prêt, en traitement, *etc.*) et la classe de rattachement. C'est par l'intermédiaire des exemplaires que la notice se trouve rattachée à une, voire plusieurs catégories scientifiques.

La notice donne l'identité bibliographique de l'ouvrage dans la bibliothèque. A un ouvrage donné (au sens bibliographique du terme), il existe au moins un exemplaire, mais il est fréquent que plusieurs exemplaires soient acquis. Le nombre d'exemplaires pour un ouvrage donné va dépendre de son prix, de son type (monographie, dictionnaire, encyclopédie, *etc.*), de la notoriété de l'ouvrage, du nombre d'étudiants ou de formations qui le demandent mais également des conditions de disponibilités (prêt à courte durée, consultable sur place, empruntable, magasin). Ces informations d'exemplaires sont appelées données d'exemplaire et constituent les instances de la notice bibliographique. Ce sont ces instances qui seront affectées d'un indice (héritée de la classification en cours dans le SCD) dès l'instant où les exemplaires sont mis à disposition en salle de lecture (pour consultation sur place ou pour prêt). Il peut arriver que les acquéreurs choisissent d'acheter plusieurs exemplaires d'un même ouvrage moins pour des questions de popularité de l'ouvrage qu'en raison de la difficulté à l'affecter à un et un seul secteur disciplinaire. Par exemple, à la bibliothèque de l'université Paris 8, l'ouvrage « Internet en entreprise (2001, ouvrage coordonné par Pierre-Jean Benghozi, Patrice Flichy et Alain d'Iribarne) » existe en deux exemplaires. Disponibles en salle de lecture, chacun d'eux est affecté d'une cote spécifique : respectivement 659.23 INT et 070.3 INT. De fait, les deux exemplaires se trouvent rattachés à des classes

génériques de la CDU bien différentes (6 pour les Sciences appliquées et 0 pour les Généralités).

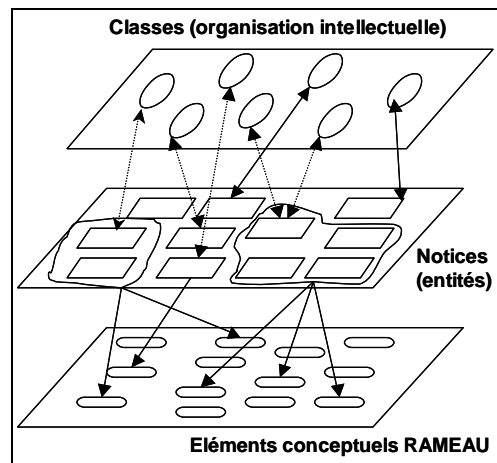


Fig. 3 : Modèle d'organisation intellectuelle des collections dans une bibliothèque : classes, éléments conceptuels RAMEAU et notices

La dernière couche (éléments conceptuels RAMEAU) contient, dans des notices d'autorité reliées provenant du répertoire national RAMEAU (Répertoire d'Autorité-Matière Encyclopédique Alphabétique et Unifié) (cf. <http://rameau.bnf.fr>), le vocabulaire et les indications qui permettent de construire les vedettes-matières dans un fichier bibliographique.

Il est important de rappeler que l'expression « Vedette-matière » se définit comme un terme polysémique qui désigne tantôt les vedettes contenues dans la liste d'autorité-matière, tantôt la vedette (ou la combinaison de vedettes) établie par le catalogueur pour servir de point d'accès à une notice bibliographique (pour lever cette ambiguïté, on emploie les termes de : vedettes, pour celles qui figurent dans la liste d'autorité et vedettes-matière, pour celles qui figurent dans le catalogue bibliographique).

Le répertoire évolue sur la base des propositions faites par le réseau des utilisateurs professionnels. La liste d'autorité est complétée par le guide d'indexation qui en assure la lecture cohérente et le bon usage. A la différence d'un thésaurus, la liste d'autorité encyclopédique n'est pas constituée *a priori* mais au fur et à mesure des besoins d'indexation et évolue sur la base des propositions faites par le réseau de ses utilisateurs.

Une notice peut-être caractérisée par plusieurs vedettes-matières. Une vedette-matière peut être rattachée à plusieurs notices.

3.3 Un modèle utilisable par l'utilisateur

Par rapport à une implication intellectuelle immédiate que l'utilisateur développe à l'endroit de la bibliothèque poussée par les travaux de natures différentes que lui imposent les exigences de son parcours universitaire, les réseaux liés à

la classification et au vocabulaire RAMEAU s'avèrent selon nous les plus adaptés pour pouvoir exploiter le plus exhaustivement possible le fonds documentaire en limitant la désorientation et la surcharge cognitive.

Comme l'illustre la figure 3, les deux réseaux organisent par des logiques bien distinctes le fonds documentaire. Celui de la classification utilise un regroupement thématique pré-établi qui créera une relation entre l'exemplaire, l'univers intellectuel et l'espace de la bibliothèque (les exemplaires sont généralement placés en salles de lecture).

Le second, à l'inverse, s'affranchira des exemplaires pour signer sémantiquement la notice d'un document. Cette information offre alors la possibilité de croisement dans des réserves bibliographiques mutualisées.

A partir de ce modèle, nous avons développé un dispositif expérimental qui tout en utilisant les mêmes métadonnées, rénove le fonctionnement des OPAC traditionnels [Papy 04]. Il se compose de 3 modules : l'interrogation, l'exploration de la classification et la localisation à partir d'une cartographie interactive. Cette approche complémentaire à la recherche documentaire repose sur l'intégralité du fonds bibliographiques du SCD de l'université Paris 8 (360 000 notices bibliographiques).

3.4 Des difficultés d'interrogation des OPAC traditionnels

Les difficultés inhérentes à une recherche sur catalogue [YEE 98] se manifestent :

- d'une part par le fait que la recherche d'information est un processus complexe. De nombreuses étapes par lesquelles passent les usagers durant leur recherche demeurent non décrites.
- d'autre part parce que rechercher est un processus itératif, dans lequel les essais et les échecs sont suivis d'autres essais, il est donc difficile de définir le début et la fin d'un processus de recherche.

La recherche sur catalogue est une négociation entre l'utilisateur et le système, durant laquelle l'utilisateur apprend comment fonctionne le système (il s'en fait une représentation mentale) et intègre ce nouveau paramètre dans les recherches futures.

Du point de vue du dispositif de recherche, les difficultés peuvent surgir des facteurs suivants :

- Le *design* de l'interface du dispositif, le type d'index choisi par les usagers, le type d'indexation disponible, l'affichage des fichiers, le type de fonctions disponibles pour la recherche et l'affichage des enregistrements.
- La taille et la nature du fonds sur lequel porte la recherche
- Les connaissances de l'utilisateur et ses capacités d'analyse et d'observation.
- Les pratiques de catalogage qui déterminent la forme des enregistrements.

Du point de vue des usagers :

- ils n'ont ni la connaissance de l'univers bibliographique dans lequel ils cherchent ni de sa structuration. L'étude menée à la BPI [DUJOL 85] en vue de déterminer si le lecteur établit un lien entre son parcours physique (le spatial) et intellectuel (le signifié), révèle que le rôle de la classification n'est pas perçu car le libellé de la cote est essentiellement identifié comme indicateur du rayon.

- ils n'ont pas conscience de la source des problèmes qu'ils rencontrent lors de la consultation du catalogue, ils sont dans l'incapacité de les signaler (sur ce point on peut noter le problème de l'interrogation sur « Titre ». L'utilisateur interroge un champ mais la réalité est tout autre puisque la notion de titre correspond à plusieurs champs UNIMARC).

4 Interrogation, exploration et visualisation des données bibliographiques pertinentes

L'intérêt de mettre en œuvre les principes d'exploration des classifications avaient déjà été soulevé, notamment à l'IFLA où des réflexions en vue de faire de la Dewey la colonne vertébrale d'une indexation ont mené au développement par l'OCLC (*Online Computer Library Center*) d'un navigateur hypertextuel DDC (Dewey Decimal Classification) en tant que dispositif de recherche avec accès par sujets (cf. l'OCLC DeweyBrowser : <http://ddcresearch.oclc.org/ebooks/fileServer>).

Avec le *Visual...Catalog* (cf. Fig. 4) [Papy 05], au-delà même de l'expression technique du dispositif, nous avons tenté de rendre explicite les différentes parties du modèle qui accompagne toute démarche de recherche et de traitement documentaire au sein de la bibliothèque.

Plutôt que de donner une vision tronquée des nécessités méthodologiques inhérentes à toute recherche d'informations, au risque même de « saturer cognitivement » l'utilisateur, nous avons résolument maintenu à la disposition immédiate de lisibilité/visibilité de l'utilisateur, les objets du modèle.

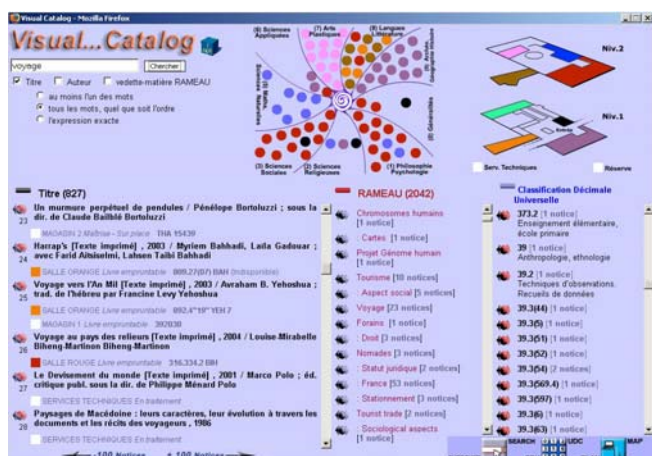


Fig. 4 : *Le Visual...Catalog, Interface d'interrogation basé sur le modèle classes, notices et éléments conceptuels RAMEAU*
(<http://visualcatalog.univ-paris8.fr>)

Le dispositif propose un mode de navigation transversale qui s'articule en trois niveaux et cinq pôles de recherche consultables simultanément (un champ de saisie, trois listes et deux représentations graphiques) :

1. Au premier niveau (recherche initiale) il s'agit du champ de saisie renseigné d'un ou plusieurs termes (mots clés) génériques ou spécifiques. Les options de la requête s'articulent entre trois champs (titre, auteur, vedette-matière) pouvant être croisés avec trois traitements dont les booléens et l'expression exacte. La validation va distribuer les résultats dans chacune des trois listes.
2. Au deuxième niveau de recherche s'élabore une relation transitive entre les trois listes : Titres, RAMEAU et CDU. Trois modes de navigation sont utilisables pour la recherche de références :
 - a) La sélection d'un titre de la liste permet d'accéder à ses sujets associés et sa classe.
 - b) La sélection d'un sujet propose l'affichage d'un ou plusieurs titres et leurs classes respectives ou communes.
 - c) La sélection d'une classe active d'autres sujets et la ou les notices bibliographiques en correspondance.
3. Au troisième niveau et parallèlement aux deux premiers, deux synoptiques graphiques renseignent sur la représentativité disciplinaire des ouvrages résultant de l'interrogation et leur localisation par salles de lecture.

Ainsi, à l'issue d'une interrogation que l'utilisateur – utilise – adresse à l'un ou à l'autre des trois champs titre, auteur ou vedette-matière RAMEAU, il obtient cinq groupes d'informations interdépendantes ; deux représentations graphiques animées et trois listes textuelles :

- une liste des titres des ouvrages répondant à la requête

- une liste de vedettes-matières RAMEAU cumulative (extraction des notices associées aux ouvrages)
- une liste des classifications (CDU) concernées par la requête (et déduite de la cote des ouvrages)
- une carte synoptique de localisation des ouvrages dans la bibliothèque

- une synthèse graphique métaphorique illustrant les secteurs disciplinaires concernés par les résultats de la recherche et le degré de spécialisation. Dans cette distribution aléatoire des cercles issue de l'interrogation, seuls sont significatifs le nombre, la taille et la couleur des cercles répartis dans les secteurs d'une, de plusieurs ou de toutes les classes génériques concernées. Chaque cercle correspond à une subdivision contenant elle-même une ou plusieurs notices, d'où la variation de diamètre.

Les 3 listes textuelles (TITRE, RAMEAU et CDU) sont interdépendantes et reliées « une à deux » la sélection d'un élément de l'une des 3 listes régénère dynamiquement les deux autres, soit :

- sur une notice; en proposant dans les listes RAMEAU et CDU, les vedettes-matières associées à la notice sélectionnée et la/les subdivision(s) de la CDU concernée(s) par la/les exemplaire(s).
- sur une vedette-matière RAMEAU; en proposant les notices communes à la vedette-matière sélectionnée, et transitivement les subdivisions de la CDU concernées par les exemplaires associés aux notices.
- sur une subdivision de la CDU; en repérant dans la liste TITRE, les notices associées à la subdivision sélectionnée, et transitivement, dans la liste « RAMEAU », les vedettes-matières associées à ces notices.

Conforme aux principes de transitivité inclus dans le modèle, l'utilisateur peut obtenir précisément pour chaque ouvrage, chaque vedette-matière RAMEAU ou chaque classe CDU, une sorte de « focus sur contexte » dynamique et sélectif permettant d'associer d'une part à chaque ouvrage une liste de termes (les vedettes-matières RAMEAU) plus explicites que le titre seul, et d'autre part le secteur intellectuel (la subdivision de la CDU) dans lequel l'ouvrage a été affecté par le catalogueur.

La connaissance des vedettes-matières permettra de resituer l'ouvrage dans une famille d'ouvrages décrits avec des termes « contrôlés » au sein d'une classe de la CDU. La sélection d'un titre, d'une vedette-matière ou d'une classe CDU ne rend visible que les éléments en correspondance dans les deux autres listes.

5 Conclusion et perspectives

Le *Visual...Catalog* se présente comme un amplificateur visuel des structurations sous-jacentes de la bibliothèque et favorise les relations constructives entre la libre associativité des usagers et les documents disponibles.

Il révèle l'organisation globale résultant des missions des bibliothèques en matière de conservation et de diffusion

des connaissances. Placé à la périphérie du cœur des activités intellectuelles complexes associées à la transmission, la construction, l'acquisition et l'organisation de connaissances, le dispositif tend à suggérer à l'utilisateur la nécessité d'acquérir les habiletés « méta-documentaires » dépassant les démarches documentaires individuelles des étudiants et enseignants-chercheurs, afin d'utiliser ce lieu de connaissances à la hauteur des ressources documentaires, techniques et humaines imbriquées et reliées, qu'il met à leur disposition.

Ce dispositif expérimental fonctionne depuis octobre 2004 à la bibliothèque de l'université Paris 8 et exploite les mêmes données bibliographiques (actualisées quotidiennement) que l'OPAC propriétaire (Absys).

Le Visual...Catalog ne propose pas à l'affichage autant d'informations que les notices bibliographiques de l'OPAC propriétaire. Par ailleurs, sa performance technique est moins bonne que l'OPAC propriétaire puisque le volume de données qu'il fait transiter entre le serveur et le poste client est beaucoup plus important.

Les sessions d'interrogation longues (entre 45 minutes et 4 heures, en local ou en distant) que nous avons pu constater grâce aux croisements de données des fichiers-journaux du serveur Web et des requêtes des usagers représentent selon nous un indicateur pertinent d'une transformation du comportement de RI de l'utilisateur. Il n'est malheureusement pas possible de confronter ces données avec celles de l'OPAC propriétaire, les données d'enregistrement que celui-ci fournit étant beaucoup moins fines.

A partir de ces sessions de consultations longues, rien ne permet cependant de confirmer l'hypothèse d'une transformation significative dans la façon dont l'utilisateur envisage la cohérence globale de la bibliothèque, rien ne permet de l'infirmier non plus.

A ce jour, le module d'interrogation a fait l'objet de 2 types d'évaluation. La première a porté sur un questionnaire de 24 questions (23 questions fermées, 1 question ouverte) remplies entre octobre 2004 et février 2005 par 82 usagers (sur le site même de la bibliothèque après l'utilisation du module d'interrogation et pendant des séances de cours). De ce questionnaire, il ressort que les éléments que notre modèle met en évidence et qui sont instrumentés dans l'interface sont, soit globalement inconnus des usagers (CDU, RAMEAU) soit connus ("connaissance intellectuelle") mais inexploités en pratique.

A cette première évaluation, s'est ajoutée 2 analyses ergonomiques, la première visant à confronter le dispositif d'interrogation aux critères ergonomiques et la seconde cherchant à appréhender au travers de l'observation de 10 sujets soumis à un protocole expérimental sur des situations de recherche documentaires, les types d'usages des outils de recherche bibliographiques (recherche thématique, recherche ciblée d'une référence, d'un article ou d'un ouvrage, recherche ciblée des publications d'un auteur sur une période donnée).

Cette deuxième étude a porté sur la qualité perçue d'une part de l'OPAC habituel de la bibliothèque (Absys) et d'autre part du Visual...Catalog.

La première analyse avait pour objectif de détecter les défauts du dispositif par rapport aux critères ergonomiques, défauts pouvant être sources de difficultés d'utilisation pour les usagers.

Synthétiquement, cette analyse a révélé que malgré le non respect de plusieurs critères ergonomiques (densité informationnelle, signifiante des codes et des dénominations, homogénéité, compatibilité avec les schémas d'utilisation), le dispositif d'interrogation restituait les informations les plus importantes en 2 ou 3 étapes (au maximum) à partir de la page d'accueil et qu'il semblait éviter la surcharge cognitive dans l'exploitation des résultats de la requête.

L'analyse relative à la qualité perçue entre les deux dispositifs d'interrogation, établie notamment à partir des verbalisations des usagers, a confirmé que le Visual...Catalog répondait à une situation de recherche documentaire que ne pouvaient pas couvrir le catalogue en ligne. De ce point de vue, les éléments mis en correspondance dans le modèle et instrumentés par l'interface sont bien au service de l'activité distincte des sujets, activité portée par la recherche thématique.

Remerciements

Ce travail de recherches s'effectue avec le SCD de l'université Paris 8 et le groupe de recherche C3U dans le cadre d'une convention de recherche.

Références

- [ACSIOME 90] ACSIOME, Modélisation dans la conception des systèmes d'information, Masson, 1990
- [BACCINO 04] Baccino T., Colombi T., « Exploration visuelle et navigation dans les hypertextes : quelles stratégies ? », Ergo'IA 2004, Biarritz
- [BALPE et al. 96] Balpe J-P., Lelu, A., Saleh, I., Papy, F., « Techniques avancées pour l'hypertexte », Editions Hermès, Paris, 1996.
- [CHAUDIRON 02a] Chaudiron, S., « L'évaluation des systèmes de traitement de l'information textuelle : vers un changement de paradigme », Mémoire d'Habilitation à Diriger les Recherches, Univ. Paris X, Nanterre, Nov. 2002.
- [CHAUDIRON 02b] Chaudiron S., Ihadjadene M., « Quelle place pour l'utilisateur dans l'évaluation des SRI ? », actes du colloque "Recherches récentes en Sciences de l'information", 21-22 mars 2002, Toulouse, Adbs-Éditions, V. Couzinet et G. Régimbeau (dir), Paris : ADBS éd., 2002, pp 211- 230
- [CIACCIA 05] Ciaccia, A., Martins, D., « Recherche d'informations sur le web : Etude de l'influence de facteurs liés à l'interface, à l'utilisateur et à la tâche » in "Alternatives en sciences cognitives, enjeux et débats", RSTI série RIA, Volume 19, Hermès Lavoisier, 2005

- [CORSON 82] Corson Y., « Aspects psychologiques liés à l'interrogation d'une base de données », Rapport de recherche de l'INRIA – Rocquencourt, 78 pages - Avril 1982, <ftp://ftp.inria.fr/INRIA/publication/publi-pdf/RR/RR-0126.pdf>
- [COULON 97] Coulon A., « Le métier d'étudiant. L'entrée dans la vie universitaire », Paris, PUF, 1997.
- [COUZINET 96] Couzinet V., Bouzon A., Normand R., « Les doctorants livrés à la recherche documentaire », *BBF*, Paris, T. 41, n°6, 1996, p.54-59
- [DINET 02] Dinet, J., Rouet, J-F., « La recherche d'information : processus cognitifs, facteurs de difficultés et dimensions de l'expertise », in *IHM et recherche d'informations*, C. Paganelli (dir), *Traité STI*, Paris, Hermes, 2002, pp 133-161.
- [DUJOL 85] Dujol A., « Le clair et l'obscur : perception et usage de la classification par le public de la BPI », Paris : Bibliothèque publique d'information, Centre Georges Pompidou, *Collection Etudes et recherches*, 1985.
- [FEO 98] Féo A., « L'enseignement de méthodologie documentaire à l'Université Paris 8. Un accompagnement bien tempéré ». *Documentaliste-Sciences de l'Information*, 1998, vol. 35, n°3
- [GILL 04] Gill T., « Building semantic bridges between museums, libraries and archives : the CIDOC Conceptual Reference Model », *First Monday*, vol. 9, n°5, may 2004,
- [HUNTER 00] Hunter, E., « Do we still need classification ? », in *The Future of Classification*, edited by R. Marcella and A. Maltby (eds), Gower Publishing, Vermont, USA, 2000, pages 1-17.
- [IFLA 98] IFLA Study Group on the functional requirements for bibliographic records, Final Report, Munich, K. G. Saur, 1998, <http://www.ifla.org/VII/s13/frbr/frbr.pdf>
- [JOLLY 01] Jolly C., « Bibliothèques universitaires. Regard sur les changements », *BBF* 2001 - Paris, t. 46, n° 6
- [KUHLETHAU 04] Kuhlthau C.C., « Seeking meaning: a process approach to library and information services », 2nd. ed. Westport, CT: Libraries Unlimited, 2004
- [LE COADIC 97] Le Coadic Y., « Usages et usagers de l'information », *ADBS*, Nathan, 1997, 128 pages
- [LEFEVRE 00] Lefèvre P., « La recherche d'informations. Du texte intégral au thésaurus », *Hermès*, 2000, 253 pages.
- [LE MAREC 90] Le Marec Joëlle, « Dialogue ou Labyrinthe : la consultation des catalogues informatisés par les usagers », *Etudes et Recherches*, BPI Centre George Pompidou, 1990.
- [LEVINE 01] P. Levine, J-C., Pomerol, « gestion des connaissances et modélisation d'entreprise reposant sur la sémantique des contrats », p. 121-148, *Management des connaissances. Modèles d'entreprises et applications*, M. Zacklad et M. Grundstein (dir), *Hermès*, 2001
- [LUPOVICI 01] Lupovici C., « Du catalogue informatisé à la navigation dans l'information numérique », *Revue de la BNF*, n°9, 2001, pp. 80- 83
- [MAISONNEUVE 03] Maisonneuve M., « Du catalogue de la bibliothèque aux ressources du Web », Paris, *ADBS*, 2003, 148 pages.
- [MEINARDIER 98] Meinardier J-P., « Ingénierie et intégration des systèmes », *Hermès*, Paris, 543 pages, 1998.
- [MOUNIER 2002] Mounier E., « Systèmes documentaires et systèmes de gestion de bibliothèques : place et rôle de l'opérateur professionnel », in *IHM et recherche d'informations*, C. Paganelli (dir), *Traité STI*, Paris, *Hermès*, 2002, pp 103-132
- [PAPY 05] Papy F., Chauvin S., « Pour une approche visuelle et ergonomique dans la recherche et l'exploration d'informations au sein d'un OPAC de SCD. L'exemple du Visual...Catalog », "Bibliothèques Numériques", *Hermès Lavoisier*, *Traité IC2, série Management et gestion des STICS*, 2005
- [PAPY 04] Papy F., Folcher V., Sidir M., Cerratto Pargman T., (2004) « E-Learning et technologies pour la coopération : inadéquations artefactuelles et logiques des activités instrumentées », *ERGO-IA*, 17-19 novembre 2004
- [PAVE 89] Pavé F., « L'illusion informaticienne », *L'Harmattan*, 1989
- [PLANCHE 88] Planche, R., « Maîtriser la modélisation conceptuelle », *Masson*, 1988
- [POLITY 00] Polity Y., « L'évolution des paradigmes dans le domaine de la recherche d'information », *Communication au groupe de travail "Théories et Pratiques scientifiques" (TPS) de la SFSIC*, le 3 mars 2000.
- [RAO 96] Rao, R., « Quand l'information parle à nos yeux », *La recherche*, n°285, Mars 1996
- [RICE 01] Rice R.E., Mc Creadie M., Chang S-J L., « Accessing and Browsing: Information and Communication », *The MIT Press*, London, England, 2001.
- [ROWLEY 92] Rowley E. J., « Organizing knowledge », 2nd Edition reprinted by Gower Publishing, England, 1995.
- [SALVAN 1972] Salvan, P. « Esquisse de l'évolution des systèmes de classification », *Ecole Nationale Supérieure des Bibliothèques*, Paris, 1972
- [SVENONIUS 00] Svenonius E., « Information Organization, in *The intellectual Foundation of Information Organization* », *MIT Press*, 2000
- [TOMS 00] Toms E., « Serendipitous Information Retrieval ». in *Proceedings of the First DELOS Network of Excellence, Workshop on Information Seeking, Searching and Querying in Digital Libraries*, Zurich, Switzerland: European Research, 2000.
- [YEE 98] Yee M.M., Shatford Layne S., « Improving Online Public Access Catalogs », *London : American library association*, 1998.