E.N.S.S.I.B

ECOLE **N**ATIONALE **S**UPERIEURE

DES **S**CIENCES DE L'**I**NFORMATION

ET DES **B**IBLIOTHEQUES

UNIVERSITE

CLAUDE BERNARD

LYON I

DESS en INFORMATIQUE DOCUMENTAIRE

Rapport de recherche bibliographique

Les applications de l'intelligence artificielle à l'assistance ou l'automatisation du catalogage

Jérôme GUTMAN

Sous la direction de

Madame Danielle ROGER

E.N.S.S.I.B

Année 1995-1996

Les applications de l'intelligence artificielle à l'assistance ou l'automatisation du catalogage

Jérôme GUTMAN

RESUME Les techniques d'intelligence artificielle peuvent aider à automatiser le catalogage descriptif dans les bibliothèques et centres de documentation. Nous présentons les différentes applications concernées autour de trois catégories : les conseillers, les créateurs d'enregistrements et les systèmes totalement automatiques.

DESCRIPTEURS CATALOGAGE, BASE CONNAISSANCE, DESCRIPTION BIBLIOGRAPHIQUE, INTELLIGENCE ARTIFICIELLE, SYSTEME EXPERT

ABSTRACT Artificial intelligence methods can help to automate descriptive cataloguing in libraries and documentation centers. This work deals with various AI applications in three categories : advisors, record generators and automatic systems.

KEYWORDS CATALOGING, KNOWLEDGE BASE, BIBLIOGRAPHIC DESCRIPTION, ARTIFICIAL INTELLIGENCE, EXPERT

SYSTEM

Table des matieres

1ERE PARTIE : METHODE DE TRAVAIL

I DESCRIPTION 1
1. Le demandeur *1*
2. Le sujet *1*
3. La démarche générale *2*
4. Le choix des descripteurs *3*
5. Le choix des sources *4*
II RECHERCHES 6
1. Personnes ressources (entretiens et courriers électroniques) *6*
1.1. Contact direct 6
1.2. Courriers électroniques 6
1.2.1. Adresses personnelles 6
1.2.2. Les listes de discussion 7
1.3. Bilan 9
2. Recherche à la bibliothèque de l'ENSSIB *10*
2.1. L'OPAC 11
2.2. Les monographies 11
2.3. Les périodiques 12
3. Les sources électroniques *12*
3.1. Recherches sur CD-Rom 12
3.1.1. PASCAL 13
3.1.2. LISA 13
3.1.3. DOCTHESES 14
3.2. Recherche "en-ligne" 14
3.2.1. PANCATALOGUE 14
3.2.2. DIALOG 14
3.2.2.1. Présentation des bases 15
3.2.2.2. Stratégie d'interrogation 16
3.2.2.3. Résultats 20
3.3. Internet 23
III ELABORATION DU RAPPORT 24
1. Bibliographie *24*
1.1. Sélection et lecture des informations 24
1.2. Présentation des références 25
2. Synthèse *26*
2.1. Acquisition des documents 26
2.2. Plan et rédaction 26
3. Démarche *27*
4. Outils de recherche *27*
IV COUTS ET TEMPS DE TRAVAIL 28

1. Temps passé.....*28*
2. Coût.....*28*
CONCLUSION 29

2EME PARTIE : SYNTHESE DES DOCUMENTS.....30

INTRODUCTION 31
1. Présentation générale.....*31*
2. Les projets, prototypes et produits développés.....*33*
2.1. Les conseillers.....33
2.2. Les créateurs d'enregistrements.....35
2.3. Les systèmes experts automatiques.....36
2.3.1. Reconnaissance de documents (OCR, SGML, ...).....37
2.3.2. Contrôle d'autorité.....39
2.3.3. Nettoyage et dédoublonnage des catalogues.....39
2.3.4. Etudes théoriques.....39
CONCLUSION 41

3EME PARTIE : BIBLIOGRAPHIE.....42

1. Présentation.....*43*
2. Projets, prototypes et produits développés.....*52*
2.1. Conseillers.....53
HEADS.....53
CATALYST.....53
BIBLIOTHEQUE NATIONALE AMERICAINE DE MEDECINE.....53
MAPPER.....54
CONFER.....56
AACR2EXPERT.....56
2.2. Créateurs d'enregistrements.....57
ESSCAPE.....57
BIBLIOTHEQUE DU CONGRES (USA).....58
2.3. Systèmes experts automatiques.....59
2.3.1. Reconnaissance de documents (OCR, SGML, ...).....59
OCLC.....59
BIBLIOTHEQUE UNIVERSITAIRE DE TUBINGEN (ZURICH).....60
MORE.....60
LAURA.....61
CAROL.....62
RIDDLE.....62
GRAPHEIN.....64
2.3.2. Contrôle d'autorité.....66
QUALCAT.....66
2.3.3. Nettoyage et dédoublonnage des catalogues.....66
KARIN.....66
2.3.4. Etudes théoriques.....67
MOLTO et SVENONIUS.....68

RESEAUX NEURONAUX....68
JENG....69
MEADOR et WITTIG....71
ABRERA et SHAW....72

INDEX....73

1ère partie : Méthode de travail

I Description

Nous présentons ici la démarche de recherche qui, de la prise de connaissance du sujet à la quête d'informations, a conduit à établir une liste de références et leur synthèse. Enfin, nous apprécierons les moyens qui ont été nécessaires, puis rendront compte de quelques remarques qui ont émergées de cette expérience.

1. Le demandeur

Madame Danielle ROGER est conservateur de bibliothèque à l'ENSSIB. Elle a proposé aux étudiants du DESS "Informatique Documentaire" plusieurs sujets pour ce rapport. Je l'ai rencontré plusieurs fois lors de mon travail, au cours desquelles nous nous sommes entretenus dans l'ordre : première approche du sujet ; précisions et sources à parcourir ; interrogation DIALOG ; présentation d'une première bibliographie ; remise du rapport.

2. Le sujet

Il semble que des informations sur le domaine des systèmes experts en catalogage aient été demandées à madame Roger, qui a donc présenté ce sujet pour en savoir plus. L'intitulé de départ était celui-ci : "Les applications de l'intelligence artificielle à l'assistance ou l'automatisation du catalogage" ; avec les indications suivantes : "Techniques évoluées pour l'automatisation du catalogage : utilisation d'outils linguistiques, de systèmes experts ou d'aide à la décision. A titre indicatif, parmi les applications possibles, on pourra rechercher des exemples de reconnaissance des caractères et de la structure d'une notice pour des opérations de rétroconversion, d'extraction et de traitement d'informations provenant des pages de titre, de reconnaissance de la structure de sommaires de périodiques, etc..."

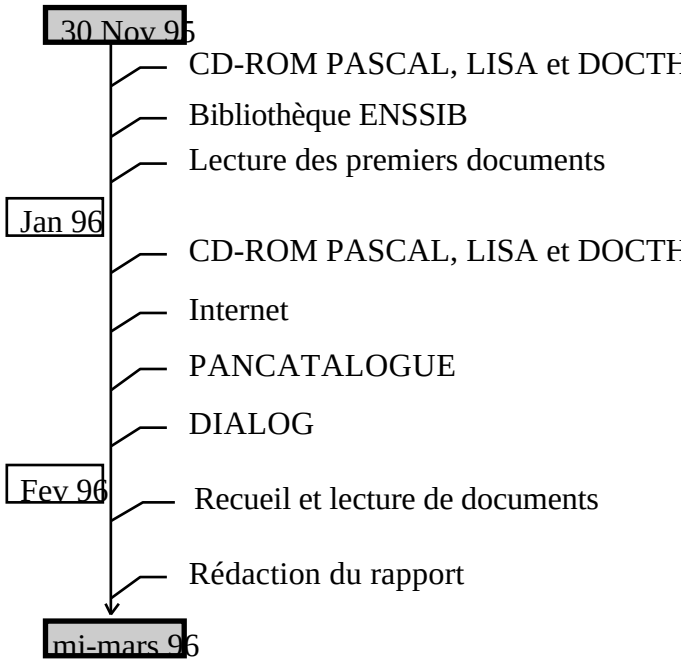
J'ai une formation en linguistique, avec des connaissances en informatique et sciences cognitives. J'ai choisi ce thème de travail, car il abordait des notions qui m'étaient familières : la linguistique, l'intelligence artificielle, l'informatique. La

connaissance du catalogage m'a été donnée pendant cette année de cours en DESS et s'est avérée suffisante pour traiter le sujet.

L'idée de départ était donc de rechercher des applications concrètes en catalogage, faisant intervenir l'intelligence artificielle. La période de recherche devait être assez récente, la date butoir de 1989 (jusqu'à aujourd'hui) a donc été choisie.

3. La démarche générale

Ce rapport est le fruit de trois mois et demi de travail (30 novembre 1995 - mi-mars 1996). Le point de départ de toutes les recherches a été la définition du sujet par le demandeur, puis sa précision au fur et à mesure de l'avancée des travaux.



Ceux-ci ont suivis une orientation "en entonnoir", du général au particulier : la réflexion sur les sources à consulter, vague au départ, s'est précisée lorsque de nouvelles données ont, au fil du temps, donné un sens plus précis à la démarche.

Après une première rencontre avec madame Roger, une base a été élaborée sur papier quant aux produits de recherches disponibles. Une première requête sur les CD-Rom de la base PASCAL ont permis de se faire une idée plus précise, ainsi que sur DOCTHESE et LISA. Une orientation plus précise du travail est apparue. J'ai ensuite parcouru certains rayons de la bibliothèque de l'ENSSIB. Arrivé fin décembre, une partie des documents a alors pu être rassemblé puis traité aux vacances de Noël. Ce premier mois peut être donc considéré comme une approche du sujet.

A la rentrée 1996, une nouvelle interrogation des CD-Rom Pascal, LISA et DOCTHESES a été traitée plus en profondeur. Le reste du mois a été consacré a un bref parcours sur Internet, l'accès au PANCATALOGUE et au serveur DIALOG. Quinze jours ont suivis pour regrouper les références, puis une durée équivalente pour leur lecture, et enfin quinze derniers jours pour la rédaction du rapport.

4. Le choix des descripteurs

Les descripteurs sont présentés ici selon leur forme d'entrée sur la base PASCAL[1].

Les premiers m'ont été fournis par madame Roger : *catalogage, intelligence artificielle, système expert.* Puis, plus tard, *base connaissance* (base de connaissance). Par la suite, d'autres mots-clés ont été trouvés dans les documents obtenus, les notices indexées des bases de données ou après réflexion sur les champs sémantiques susceptibles d'être impliqués dans ce domaine. La pertinence des descripteurs a pu être testée lors de l'interrogation des CD-Rom, en consultation gratuite à l'ENSSIB, avant de passer au serveur DIALOG, pour lequel de mauvais choix auraient ramené du bruit et donc des dépenses inutiles.

En fait, seul un petit nombre a été utilisé, et ce sont les premiers "ressentis" qui se sont avérés les plus adaptés. Les mots généraux (*bibliothèque, documentation, sciences de l'information, logiciel, informatique*) n'ont pas été d'un grand secours, surtout dans des bases spécialisées (*sciences de l'information* dans LISA !). Tous les termes possibles concernant la linguistique (*sémantique, morphologie, syntaxe,* etc...) n'ont fait que produire du bruit.

Enfin, l'équivalent du mot français *catalogage* s'écrit avec un *u* en anglais (*cataloguing*) et sans en américain (*cataloging*). J'ai commencé à utiliser ce mot avec un

---

[1] Voir le dictionnaire des CD-Rom Pascal ou la référence suivante : INIST. - Pascal Lexique [Sciences exactes et technologies. Sciences de la vie. Lexique français-anglais-espagnol]. - Vandoeuvre-lès-Nancy, ed. INIST, 1991. - ISBN 2904975638. - ENSSIB U025.49 CEN

opérateur de troncature après le *g* (*catalog?*), ce qui aurait regroupé les deux orthographes, mais l'ai abandonné d'après les conseils de madame Roger (provoquait trop de bruit : *catalogne*, *catalogue d'oeuvres d'art*, etc...). Du coup, toutes les références en anglais m'ont échappé (interrogation sans *u*), avant de m'en rendre compte et de refaire le travail : on ne m'y reprendra plus... !

5. Le choix des sources

La partie *Recherches* présente les sources qui ont été effectivement parcourues. Nous ne les aborderons donc pas ici.

La liste des chemins suivis a été un assemblage des conseils du demandeur, de mes propres connaissances en recherche bibliographique (la plupart issues des cours du DESS) et de l'observation des pratiques de mes collègues étudiants. Il me semble que le résultat obtenu est assez complet, au vu des divers recoupements entre les références.

Compte tenu du temps imparti, il n'a cependant pas été possible de consulter toutes les pistes pressenties au départ. Ce travail ne pouvant être exhaustif, d'autres "puits d'informations" auraient pu être sondés :

- autres périodiques spécialisés (comme Annual Review of Oclc Research), en dehors de ceux parcourus à l'ENSSIB (voir page 12). En particulier, le centre de documentation de l'école possède peut-être certaines ressources intéressantes.

- travail poussé sur Internet : par moteurs de recherches, listes de discussions spécialisées (PACS-L par exemple). Il est difficile de se retrouver dans cet amalgame d'informations, d'accéder aux plus pertinentes, c'est pourquoi le "réseau des réseaux" n'a pas été fouillé.

II Recherches

Nous présentons ici les différentes sources d'informations explorées et les stratégies d'informations.

1. Personnes ressources (entretiens et courriers électroniques)

En dehors du demandeur qui m'a fourni quelques références, je suis parfois allé chercher l'information "de vive voix" ou par courrier électronique, auprès de spécialistes.

1.1. Contact direct

Jean-Marc PROUST est conservateur à l'ENSSIB. Dans le cadre du DESS, il nous a donné plusieurs cours sur la recherche bibliographique, mais aussi le catalogage (format MARC). Ce dernier sujet m'a semblé (impression personnelle) lui tenir à coeur. C'est pourquoi je lui ai demandé des informations sur mon sujet. Il n'avait malheureusement rien à me fournir sur les systèmes experts.

1.2. Courriers électroniques

1.2.1. Adresses personnelles

Sur les conseils de madame Roger, j'ai pu obtenir les coordonnées électroniques des co-concepteurs de RIDDLE (voir référence [38] page 62), par l'intermédiaire d'une adresse de type web (Internet) sur leurs travaux. Voici le message que je leur ai envoyé :

Envoyé le :	24/01/1996
Destinataires :	Kees.van-t.Hoff@cwi.nl, F.A.Roos@cwi.nl, Wouter.Mettrop@cwi.nl, alanh@lcd.co.uk, RET@ib.rl.ac.uk
Sujet :	information about works similar to RIDDLE
Texte :	Dear Sir, Madam, I'm a french student in information science, at the ENSSIB school, France. I'm doing bibliographical research about applications in use of artificial intelligence for assisting or automating cataloguing tasks (descriptive cataloguing and not subjet indexing). I read about the RIDDLE project you have worked on, and I think that it deals with the same subject : expert system and cataloguing. Would it be possible for

	you to give me information about recent studies (after 1992), publications or projects in this area : bibliographical references, URL adresses on internet, etc... ? This information would be extremly helpful to me for my research. Thank you for your time, sincerely Jerome GUTMAN

En retour, je n'ai reçu qu'une réponse :

Reçu le :	25/01/1996
Expéditeur :	faroos@cwi.nl
Sujet :	Re : information about works similar to RIDDLE
Texte :	Dear Jerome Gutman, Thank you for showing interest in the RIDDLE project ! More information about the (finished) RIDDLE project (including public deliverables = consortium documents) can be found at : http://www.cwi.nl/cwi/projects/riddle.html If you give me your postal address, I can also send you a reprint of an article that appeared in the "Electronic Library" journal in 1995. Best regards, Frank A. Roos, CWI - Centre for Mathematics and Computer Science / Library, Amsterdam, The Netherlands

Même si cela part d'un bon sentiment, l'aide de monsieur Roos ne m'a pas été d'un grand secours : il ne m'a rien apporté sur d'autres projets (Pour promouvoir RIDDLE essentiellement ? Question mal formulée de ma part ?) et j'avais déjà lu l'article dont il m'a fait part. Il est cependant le seul à m'avoir répondu parmi ses collègues.

1.2.2. Les listes de discussion

Afin de scruter les messages qui auraient pu avoir attrait à mon sujet, ou pour envoyer une demande d'informations, je me suis abonné (gratuitement) à deux listes électroniques de discussions françaises spécialisées en sciences de l'information : biblio-fr (biblio-fr@univ-rennes1.fr) et adbs-info (adbs-info@univ-rennes1.fr). Seule la première est régie par un "modérateur" qui fait le tri des messages envoyés à la liste et les renvoie aux abonnés.

Je n'ai rien trouvé de pertinent dans les archives de biblio-fr (accessibles sur un serveur web). Ni dans les messages des deux listes reçus sur ma messagerie, tout au long de ces trois mois et demi, parmi des thèmes pourtant variés.

J'ai envoyé aux deux listes la question suivante :

Envoyé le :	13/12/1995 (biblio-fr), 24/01/1996 (adbs-info)
Destinataires :	biblio-fr@univ-rennes1.fr, adbs-info@univ-rennes1.fr
Sujet :	catalogage et intelligence artificielle
Texte :	Bonjour, Je suis étudiant en DESS "Informatique Documentaire" de l'ENSSIB et je recherche actuellement des documents sur le thème suivant : "les applications de l'intelligence artificielle à l'assistance ou l'automatisation du catalogage". Ce sujet tend à dresser une liste des applications récentes (depuis 1990) d'automatisation partielle ou totale du catalogage (catalogage descriptif et non indexation) à partir de documents primaires ou des notices papier par exemple. Il semble que de nombreuses études dans ce domaine ont été effectuées, qui ont abordé les thèmes suivants : systèmes experts, reconnaissance de caractères, opération de rétroconversion, outils linguistiques... Même si elles n'ont pas abouti a un produit utilisable, les problèmes qu'elles ont soulevés m'intéressent. Je lance donc un appel aux personnes de la liste qui pourraient me fournir des sources d'informations sur ce thème, quelles que soient ces références : monographie, périodique, personne à contacter, CD-Rom, site Internet... Cette recherche bibliographique se terminera autour du 15 Février 1996. Merci d'avance, Jerome GUTMAN

Voici les réponses reçues :

Reçu le :	24/01/1996
Expéditeur :	Francoise.Renzetti@imag.fr
Sujet :	Re : catalogage et intelligence artificielle
Texte :	peut-être ce qui concerne SGML serait un point de départ ? Francoise.Renzetti@imag.fr http://www-mediatheque.imag.fr/Mediatheque.IMAG IMAG, Institut d'Informatique et de Mathématiques Appliquées - Médiathèque

Une réponse un peu vague ! J'ai consulté l'adresse web accompagnant le message (fonds des centres de documentation de l'IMAG et l'INRIA, entre autres), où je n'ai rien vu d'intéressant.

Reçu le :	25/01/1996
Expéditeur :	Pascale.Laurent@irisa.fr
Sujet :	Re : catalogage et intelligence artificielle
Texte : (une partie n'a pas été notée :	Bonjour, je ne sais s'il s'agit vraiment d'intelligence artificielle mais voici l'url du colloque sur l'information élaborée : http://crrm.univ-rms.fr/ile-

voir les points d'interrogation)	rousse/prog.htm A suivre trois références de rapports qui sont disponibles sur la machine ftp de l'irisa (ftp.irisa.fr), dans le répertoire techn?? : 1993/PI-718.ps.gz Boucher, Paul?? : Compounds : an intelligent?? ; 1995/PI-928.ps.gz Danna, Frédéric?? : Modélisation de l'apprenant?? ; 1995/PI-974.ps.gz Calculability of the Semantics?? Pascale Laurent, responsable documentation irisa

L'indication du colloque m'avait déjà été donnée par un autre abonné : rien ne correspondait à mon sujet. Enfin, les trois rapports en question concernaient les mots composés anglais !

1.3. Bilan

Au bout du compte, aucun résultat supplémentaire n'a été amené par ces contacts. Une recherche et sélection plus poussée des personnes-ressources spécialistes aurait certainement pu être plus bénéfique : notamment dans d'autres pays, comme en témoigne l'internationalité des travaux mentionnés dans la synthèse.

De même, il se peut que mes questions n'aient pas été correctement amenées, précisées.

Enfin, mon opinion sur le courrier électronique est celle d'un moyen pratique, mais superficiel. J'ai bien conscience que, dans tout milieu professionnel, les gens ont d'autres choses à faire que prendre du temps pour répondre à un étudiant. Il m'a semblé en définitive que les personnes contactées n'ont pas suffisamment pris le temps d'analyser la question : à part une mauvaise formulation de la demande (cela est fort possible), comment expliquer la non-pertinence des informations fournies ? Peut-être est-ce aussi un problème de sujet, pas toujours aisé à appréhender. Tout cela pourrait expliquer le peu de réponses envoyées. Une autre remarque sur ce mode de communication est la désagréable impression de recevoir des "publicités camouflées" sur les produits développés par les expéditeurs : je pense particulièrement aux rapports de l'irisa et, dans une moindre mesure, aux informations sur RIDDLE... mais pas sur les autres systèmes experts !

Malgré les critiques qui reflètent ma déception en recevant ces (quelques) messages, je tiens à remercier les personnes qui auront fait l'effort de me répondre.

2. Recherche à la bibliothèque de l'ENSSIB

Le fonds de la bibliothèque de l'ENSSIB, spécialisée en sciences de l'information, est le mieux placé pour répondre à mes attentes. Le premier mois a été consacré à y rechercher des ouvrages et articles, et a permis une première approche de la méthode, plus traditionnelle.

Le nombre total de documents parcourus n'a pas été retenu, c'est pourquoi nous n'avons pas fait apparaître la pertinence (voir définition page 20). La couverture est fournie à titre indicatif.

Références trouvées à la bibliothèque (Documents)	Références de la bibliographie (Documents)	Couverture (%)
21	51	41,18

2.1. L'OPAC

Une des premières démarches a été de rechercher des ouvrages dans le catalogue des monographies, consultable sur les ordinateurs de la bibliothèque. Au cours du temps, les recherches par sujet ont donné lieu à quatre équations :

- catalog% ET intelligence artificielle
- catalog% ET système expert
- expert ET librarie
- expert ET bibliothèque

Seules les deux dernières ont amené quelques références. L'indexation de cet OPAC ne m'a pas semblé très profonde, en plus d'une interface peu conviviale, et ne m'a pas beaucoup aidé pour cette bibliographie.

2.2. Les monographies

Compte tenu de l'interrogation de l'OPAC, j'ai pensé qu'un rapide parcours de rayons choisis de la bibliothèque pourrait faire ressortir certaines monographies intéressantes. J'avoue qu'il s'agit là d'une méthode très aléatoire (livres de petits formats peu visibles, titre absent sur la tranche, etc...), peu scientifique (lecture de la table des matières seulement si le titre paraissait évocateur), laborieuse et archaïque (je sais que les catalogues existent !), mais elle s'est révélées assez fructueuse ! Le déchiffrage des rayons concernant l'intelligence artificielle, l'informatique documentaire, le catalogage, les aspects techniques des bibliothèques, etc... a été ainsi le seul moyen de trouver le manuel d'AACR2EXPERT (voir référence [28] page 56). Cela n'est bien évidemment pas faisable dans une bibliographie générale aux rayons immenses, et met de côté les ouvrages stockés ailleurs (dans le fonds par exemple).

2.3. Les périodiques

Comme il s'agit d'un fonds spécialisé, j'ai parcouru les classeurs de sommaires des périodiques (j'aurais pu, avec plus de temps, en sélectionner davantage). Voici les revues concernées (seule la date de début est mentionnée) :

- Archimag (1990)
- Cataloging and Classification Quarterly (1993)
- The Electronic Library (1994)
- Information Processing and Management (1991)
- Information Retrieval and Library Automation (1990
- Library Hi-Tech (vol. 10)

La lecture de toutes ces photocopies de sommaires amène vite à une saturation visuelle (et intellectuelle), car elle demande une certaine concentration, même si la tâche n'est pas d'une complexité exemplaire ! Encore une fois, il n'est pas parfait de se baser sur les tables des matières qui ne proposent souvent que les titres des articles : ceux qui

correspondaient au sujet, avec un titre peu évocateur, auront donc été écartés. Cependant l'article décrivant RIDDLE ([38] page 62) a pu être connu de cette manière.

3. Les sources électroniques

3.1. Recherches sur CD-Rom

Les bases PASCAL, LISA et DOCTHESES, ont été choisis parmi un ensemble de CD-Rom de l'ENSSIB. Nous ne développerons ici qu'une impression d'utilisation pour ces supports. En effet, elles ont aussi été consultées sur le serveur DIALOG, plus tard, et les résultats sont plus complets pour faire part de pourcentages . Quant à DOCTHESES, qui n'est pas sur ce serveur, aucun document n'a été trouvé, d'où l'inutilité d'une telle analyse. Cette recherche sur CD-Rom a principalement permis de tester la stratégie et de se rendre compte :

- des descripteurs à utiliser (termes en sciences de l'information inutiles, comme ceux en linguistique ; découverte des différents équivalents anglophones de catalogage)
- que l'interrogation par noms de systèmes experts, d'auteurs ou de revues n'amenait rien de plus par rapport à l'utilisation de mots-clés ci-dessus. Ceci appuie une constatation, à savoir qu'il est difficile de découvrir de nouvelles publications d'un auteur ou sur une application après avoir déjà trouvé une première référence.

3.1.1. PASCAL

Les neuf CD-Rom de 1990 à 1995 ont été interrogés. La base sur ce support n'est pas des plus maniables : les disques ne sont pas cumulatifs (deux par an) et demandent beaucoup de manipulations pour passer de l'un à l'autre. La stratégie peut être sauvegardée, imprimée, mais cela demande de passer par des étapes contraignantes et répétitives. L'interface n'est pas non plus spécialement agréable. Cependant le

fonctionnement par menus est assez intuitif (deux modes : assisté vs expert). Enfin, la base semble indexée en profondeur, ce qui facilite la recherche.

Les champs utilisés ont été DXF (mot composé français : pour *intelligence artificielle, système expert* et *base connaissance*) et LI (Titre + Résumé + mots-clés : pour *catalogage*).

3.1.2. LISA

Le CD-Rom LISA est cumulatif. L'interrogation se fait sans menu, mais en précisant les champs au clavier d'une manière simple. La combinaison des questions (champ *cs*, pour *combine set*) est assez pratique. L'indexation semble complète.

J'ai surtout utilisé les intitulés *kw* (mot-clé), ou l'interrogation en mot libre (recherche sur tous les champs, pour les noms d'applications par exemple). *Da* (date) et *la* (langue : français ou anglais) m'ont permis d'affiner les résultats.

3.1.3. DOCTHESES

Ce CD-Rom présente les thèses soutenues dans les universités françaises (depuis 1972 pour les lettres et les sciences). L'interface est sous Windows, très agréable : conviviale, intuitive. Cela ne m'a cependant pas servi à grand chose, car rien de pertinent n'en est sorti.

3.2. Recherche "en-ligne"

Il s'agit ici d'interrogation de serveurs situés dans des sites distants.

3.2.1. PANCATALOGUE

Le PANCATALOGUE est un catalogue des ouvrages de l'enseignement supérieur (universités). Il s'agit essentiellement de monographies, dans tous les domaines. A la rentrée 1996, la bibliothèque de l'ENSSIB proposait un accès à cette base, c'est pourquoi je me suis brièvement plongé dans le programme de consultation (assez barbare !). A tort, car je n'ai rien ramené.

3.2.2. DIALOG

Le serveur DIALOG regroupe des bases de données sur des thèmes très variés. Leur choix s'est fait en consultant le catalogue papier et en accord avec le demandeur Nous présentons, d'après les feuilles descriptives ("blue sheets"), celles interrogées avant de décrire la stratégie et d'analyser les résultats. La description est assez sommaire, car il ne nous a pas paru utile pour cette étude de rendre compte de tels détails.

3.2.2.1. Présentation des bases

Nom	n°	Domaine	Date début
LISA (Library & Information Science Abstracts)	61	Sciences de l'information et bibliothèques	1969
ISA (Information Science Abstracts)	202	Sciences de l'information et bibliothèques	1966
INSPEC	4	Physique Electronique Informatique	1983
NTIS	6	Recherches, développements, études d'organismes gouvernementaux américains (aussi étrangers) : NASA, HUT, CNRS...	1964
Dissertation Abstracts Online	35	thèses, mémoires des universités américaines (aussi britanniques)	1861
Pascal	144	Information Scientifique et Technique	1973
LLBA (Linguistics and Language Behavior Abstracts)	36	Linguistique	1973
LC MARC-BOOKS	426	Monographies de la Bibliothèque du Congrès (USA)	1968
Books in Print	470	Ouvrages en langue anglaise et disponibles dans le commerce, surtout aux USA	Courante
British Books in Print	430	Ouvrages en langue anglaise et disponibles dans le	Courante

		commerce, surtout en Grande-Bretagne	

Toutes les bases, sauf Pascal, sont en langue anglaise. Les notices contiennent toutes les informations bibliographiques nécessaires, ainsi que la plupart du temps un résumé et des mots-clés.

3.2.2.2. Stratégie d'interrogation

Nous avons regroupé les bases en trois groupes, afin de scinder la recherche (gain de temps et test sur le premier lot, avant reconnexion et essais sur les deux suivants) :

1) Sciences de l'information

   LISA

   ISA

2) Information Scientifique et Technique

   INSPEC

   NTIS

   Dissertation Abstracts Online

   Pascal

3) Fonds de monographies

   LLBA

   LC MARC-BOOKS

   Books in Print

   British Books in Print

Les questions qui vont suivre n'ont pas utilisées de champs précisés, sauf l'année de publication (PY) pour restreindre les réponses. La recherche s'est toujours faite sur les index de base ("Basic index") contenant, pour ce qui nous intéressait : titre, descripteur, résumé. Le problème de l'indexation en mots ou en phrases a été résolu par l'emploi de l'opérateur de proximité (W) ou ( ) seulement (avec "*mot1( )mot2*", la recherche se fait sur *mot2* suivant immédiatement *mot1*).

Voici les questions posées :

A) (CATALOG?ING *OR* BIBLIOGRAPHIC?) *AND* (ARTIFICIAL( ) INTELLIGENCE *OR* EXPERT( ) SYSTEM? *OR* KNOWLEDGE( )BASE?) *AND* PY>1992

Cette question a pu être établie grâce aux tests effectues sur les CD-Rom. La première parenthèse regroupe les termes pour atteindre le domaine du catalogage, commun (opérateur *AND*) avec celui de la seconde, l'intelligence artificielle, le tout restreint par une date de publication supérieure à 1992. Le point d'interrogation (opérateur de troncature) permettait de ramener les éventuels pluriels, les termes commençant par BIBLIOGRAPHIC et l'épineux problème de CATALO*GUI*NG/CATALO*GI*NG. Seule cette première recherche a été conservée pour la présentation des résultats (voir graphique et tableau plus bas, page 20). En effet, elle s'est révélée la plus pertinente et a été lancée sur les trois groupes de bases. Les deux questions qui suivent, n'ont rien ramené d'utile sur le lot 1 et n'ont donc pas été renouvelées pour les groupes 2 et 3. Cette première question a pris un certain temps (message « processing, processing, processing... ») sur l'ensemble 2 (bases quantitativement importantes). De plus, elle a ramené des notices sur des formes de catalogage particulières : images du ciel, images animales, etc... Au bout du compte, ces dernières ont été écartées (plutôt de l'indexation ou de la classification).

B) (AUTOCAT? OR ESSCAPE? OR MACCAT OR MITINET? OR DOBIS? OR CATTUTOR OR URICA OR HYPERCAT) AND PY<1993

J'ai souhaité tenter une interrogation par noms d'application déjà trouvés dans la littérature, pour savoir si quelque chose de nouveau m'avait échappé. Cela n'a rien donné de pertinent et a ramené beaucoup de bruit sur "DOBIS/LIBIS" que je recherchais mais qui, à la lecture des résumés, s'est révélé extérieur à cette recherche.

C) (MAPPER OR NAL OR RIDDLE OR APEX OR EXETER OR HEADS OR CATALYST) AND (CATALOG?ING OR EXPERT?) AND PY<1993

Ces autres intitulés de programmes ou lieux (NAL = National Agricultural Library ; EXETER = université de deux auteurs, voir page 35) pouvaient prendre plusieurs significations : HEADS = têtes ; APEX = pointe de la langue, etc... La deuxième parenthèse a donc permis de préciser le domaine d'application. Rien n'a été trouvé sur le groupe 1, ce qui ne m'a pas encouragé à la tester sur les deux autres.

Notons enfin que l'utilisation de la commande ID pour l'identification des doublons à l'intérieur de chaque groupe n'a rien éliminée, alors que l'affichage des notices (commande TYPE, en format complet 5) en a fait apparaître.

3.2.2.3. Résultats

Nous rappelons que ce bilan sur DIALOG concerne uniquement la question A.

Le tableau présente le nombre de documents trouvés dans chaque base, suivi de l'ensemble des documents retenus, puis les taux de pertinence et de couverture par rapport aux 52 documents papier de la bibliographie finale. La définition[2] des deux derniers nombres, seuls sur le graphique, est la suivante[3] :

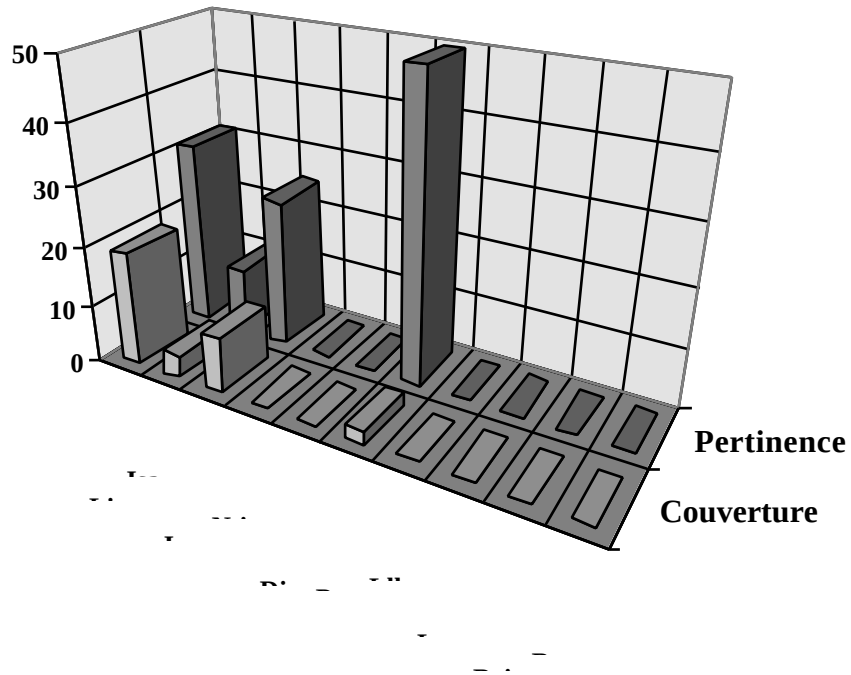
$$\text{PERTINENCE} = \frac{\text{nombre de documents pertinents fournis en réponse à une question}}{\text{nombre total de documents, pertinents et non pertinents, fournis en réponse à la même question}}$$

$$\text{COUVERTURE} = \frac{\text{nombre de documents pertinents sélectionnés par la base}}{\text{nombre total de documents de la bibliographie}}$$

[2] Adapté de VAN SLYPE (Georges). - Conception et gestion des systèmes documentaires. - Paris, Les éditions d'organisation, 1979, pp. 170-172. - ISBN 2708103245. - ENSSIB U651.8 VAN

[3] Le taux de couverture d'après VAN SLYPE s'applique sur l'ensemble des documents présents dans le monde. Nous l'avons repris ici en le comparant aux documents de la bibliographie : il s'agit donc de la couverture par rapport à la bibliographie finale.

Base	Question	Documents trouvés	Documents pertinents	Pertinence (%)	Couverture (%)
LISA	1	32	10	31,25	19,23
	2	13	0	0	0
	3	2	0	0	0
ISA	1	18	2	11,11	3,84
	2	4	0	0	0
	3	0	0	0	0
INSPEC	1	20	5	25	9,61
NTIS	1	1	0	0	0
Dissertation Abstracts Online	1	1	0	0	0
Pascal	1	2	1	50	1,92
LLBA	1	0	0	0	0
LC MARC-BOOKS	1	0	0	0	0
British Books in Print	1	0	0	0	0



Books in Print	1	0	0	0	0

Rappelons, sauf erreur de notre part, que les taux de bruit et de silence peuvent être obtenus en faisant respectivement les compléments à 100 % des pourcentages de pertinence et de couverture (c'est-à-dire : pertinence + bruit = ensemble des données de la base ; silence + couverture = ensemble des documents de la bibliographie)

Il apparaît clairement que LISA a été la source d'informations en-ligne la plus importante pour la recherche. INSPEC se trouve en seconde place, les autres (ISA et Pascal) donnant des résultats négligeables. Il ne faut pas non plus oublier de mettre en correspondance les 50 % de pertinence impressionnants de Pascal (calculés sur seulement deux documents !), avec le taux de couverture qui est, lui, très faible. Enfin, le taux de pertinence assez bas (environ 30 %) de LISA, qui a pourtant été la base la plus utile, amène à s'interroger sur une éventuelle trop grande généralité de la question A. Je n'ai pu cependant trouver d'autres descripteurs qui auraient correctement affiné les résultats sans écarter des références intéressantes (les noms d'auteurs, de revues, de projets, rappelons-le, n'ont rien donnés).

3.3. Internet

Comme il a déjà été précisé, Internet avait été désigné au départ comme une des dernières rubriques à consulter s'il restait du temps. Seul un début a pu être amorcé, au cours de quelques approches. A part le courrier électronique, voici certains points qui pourraient être approfondis :

- Consultation des archives de PACS-L, une liste de discussions sur l'informatique en bibliothèque. L'abonnement se fait à l'adresse PACS–L@UHUPVM1.UH.EDU

- Accès par moteurs de recherches (Lycos, Yahoo, Alta Vista, etc...) à des sites sur le catalogage. Notamment un guide qui regroupe les ressources Internet sur ce sujet, et qui semble intéressant. Son adresse web est : HTTP://WWW.LIB.UMICH.EDU/CHDOCS/LIBCAT/CATTOOLS.HTM

III Elaboration du rapport

La recherche et sélection des documents a permis de construire une bibliographie (3ème partie, page 42), puis d'en faire leur synthèse (2ème partie, page 30). Une fois ces deux étapes terminées, je me suis concentré sur l'exposition de ma démarche, présentée ici.

Il est difficile, mais cela doit s'acquérir avec l'expérience, de doser le temps consacré à chaque stade. Ces trois étapes demandent quasiment toutes autant de travail, mais de types différents.

1. Bibliographie

1.1. Sélection et lecture des informations

Pour sélectionner les documents à intégrer dans la bibliographie, mise à part une date de publication supérieure à 1989, je me suis d'abord basé sur les descripteurs et les mots-clés des notices trouvées. Un résumé complet me suffisait, s'il était plus vague cela nécessitait d'aller voir dans le document lui-même. Leur nombre important (bien plus que ceux de la bibliographie finale, évidemment) a posé des problèmes de démarche de lecture : il n'est pas toujours facile de se faire une idée précise au vue du titre, de la table des matières ou même d'un court résumé. Il a donc fallu souvent se plonger dans les chapitres. De plus, si l'on parle en général de lecture entre les lignes, ce n'est pas du tout facile lorsqu'il s'agit d'un texte en anglais, langue que je ne maîtrise pas totalement.

En conséquence, j'ai donc dû véritablement lire une grande partie des références, en écartant les données trop techniques ou les points en dehors du sujet. J'ai essayé pour chacune d'entre elles de faire un résumé.

Au niveau du contenu, un des plus grands problèmes de ce travail a été de faire le tri entre les programmes plus "classiques" (algorithmiques) et ceux "intelligents", lorsque cela n'était pas clairement exprimé par les auteurs. C'est pourquoi les cas qui sont

présentés dans la synthèse viennent d'écrits qui les mentionnent en tant qu'outils d'intelligence artificielle. En revanche, on trouvera dans la bibliographie certains projets qui ne sont pas repris dans la synthèse, quand nous n'étions pas certains du type de logique sous-jacente. Dans le doute (système expert ou programme "banal" ?), nous avons préféré les mentionner dans les résumés de certaines références plutôt que les écarter : nous laissons au demandeur ou aux personnes spécialistes du domaine de l'IA le soin d'écarter ces zones d'ombres. De même, il fut assez difficile de savoir tout au long des lectures si les systèmes experts ont produit des réalisations concrètes ou s'ils n'ont pas dépassé le stade de prototype : les articles qui les citent ne le précisent pas toujours et n'ont souvent aucune suite.

Au final, je peux dire que mon approche de la lecture, la prise d'informations, a été modifiée, en bien. Elle est devenue plus rapide, plus pertinent, plus ciblée : une expérience documentaire en quelque sorte.

1.2. Présentation des références

Les articles, ouvrages et adresses internet sont regroupés dans la même bibliographie. Ils sont suivis d'un résumé, pour la grande majorité (il a été jugé préférable de le placer à la suite des références, plutôt qu'à part, ce qui aurait amené de nombreux renvois désagréables pour le lecteur). Ils sont classés d'après le plan de la synthèse, puis par ordre chronologique de publication décroissant à l'intérieur de chaque catégorie. Si cela ne suffisait pas, nous les avons classés alphabétiquement sur le premier nom d'auteur, puis alphabétiquement sur le titre. Certains ouvrages qui traitent de plusieurs catégories sont placés dans leur thème dominant.

La description bibliographique s'inspire des normes françaises AFNOR Z44–005[4] (catalogage simplifié pour les bibliographies) et AFNOR Z44–060[5] (uniquement

[4] AFNOR. - Norme française Z44-005 [Références bibliographiques]. - In : AFNOR, Recueil des normes françaises 1986 : Documentation [présentation des publications, translittération, thesaurus et indexation]. - tome 1, 3ème ed.. - Paris, Afnor, 1986, pp. 138-148. - ENSSIB 025.324 AFN

chapitre sur les actes de colloque). Certaines références sont suivies de l'ISBN, de l'ISSN et/ou de la côte à la bibliothèque de l'ENSSIB. Les séparateurs et la présentation sont un choix personnel : l'essentiel a surtout été de s'assurer une cohérence dans la présentation.

2. Synthèse

2.1. Acquisition des documents

Une fois les documents sélectionnés, il a fallu s'en procurer une bonne partie lorsque le résumé ne suffisait pas ou si l'article paraissait très intéressant à lire pour la synthèse. Pour cela, le CD-Rom MYRIADE du CCN (Catalogue Collectif National : localisation des périodiques) a été utilisé. Cela n'a posé aucun problème.

La plupart des documents ont été trouvés à l'ENSSIB. Six ont transités par le service du PEB (Prêt Entre Bibliothèques). Trois ou quatre ont dû être photocopiés à la Bibliothèque Municipale de Lyon. Enfin, l'article de Rossella DINI (voir [1], page 43) a été copié à l'INFB (Institut National de Formation des Bibliothécaires, 69100 Villeurbanne).

2.2. Plan et rédaction

De nombreux articles suivent le plan que j'ai directement repris dans ma synthèse : 1. Systèmes conseillers ; 2. Créateurs d'enregistrements ; 3. Catalogage automatique et études théoriques. Il m'a paru bien adapté pour le récupérer tel quel.

Le regroupement des informations autour de cette structure s'est fait en reprenant chaque résumé de référence et en notant dans les catégories concernées ce qu'elle apportait. Ensuite, le rassemblement de toutes ces notes a donné lieu à la synthèse, en veillant à la présence des numéros de références de la bibliographie.

---

[5] AFNOR. - Norme française Z44-060 [Catalogue d'auteurs et anonymes]. - In : AFNOR, Recueil des normes françaises 1986 : Documentation [présentation des publications, translittération, thesaurus et indexation]. - tome 2. - Paris, Afnor, 1986, pp. 81-84. - ENSSIB 025.324 AFN

3. Démarche

Tout au long du travail j'ai pris des notes pour me souvenir, au moment de rédiger le rapport, des démarches effectuées. Chaque étape a ainsi été référencée dans un tableau avec sa durée et son coût éventuel. Les stratégies de recherche sur DIALOG (sauvegardées par captures d'écrans) et sur CD-Rom m'ont aussi été utiles pour cette réflexion rétrospective.

4. Outils de recherche

Il m'a paru très utile d'intégrer dans ce dossier des facilités pour le lecteur qui souhaiterait retrouver des informations. Nous avons donc choisi d'insérer, en plus de la classique table des matières, des renvois (« voir page... ») et un index regroupant les auteurs, les applications et certaines données de cette première partie.

Notons qu'une version informatique de tout rapport de recherche bibliographique sous une forme hypertexte (ou pourquoi pas hypermédia) semblerait très intéressante, mais demanderait tout une période de développement. En l'absence, nous espérons que ces outils auront apportés un "plus".

IV Coûts et temps de travail

1. Temps passé

Ces nombres ne sont pas exacts au chiffre près, mais ils permettent de se faire une idée. On pourrait ainsi considérer que, pour un salarié travaillant au même rythme pendant 39 heures par semaine, ce rapport aurait pris environ une période de trois semaines à temps plein (117 heures).

Type de travail	Durée (heure)
Rendez-vous avec le demandeur	2,5
Mises au points de la recherche	5
Bibliothèque ENSSIB	7,75
CD-Rom	13,25
PANCATALOGUE	1
DIALOG	0,75
Internet et courriers électroniques	2,5
Lectures et résumés	32,5
Ecriture du rapport	20,5
Tâches matérielles : photocopies, traitement de texte, impressions, reliures	40
TOTAL	124,75

L'on voit que, dans la recherche, la lecture des articles a été le plus prenant. Le chiffre concernant les tâches matérielles n'est pas très révélateur en tant qu'effort intellectuel.

2. Coût

Certaines évaluations sont signalées par une étoile "*". Il apparaît clairement que le plus cher a été le recours à des services extérieurs : PEB et DIALOG.

Descriptif	Prix unitaire (francs)	Quantité	Total (francs)
Photocopies	0,25	13	3,25
Photocopies	1	18	18
Photocopies	*0,15	250	37,5
PEB	29 la 1ère tranche de 10 pages	6	174
	14,5 par tranche supplémentaire	3	43,5
Interrogation DIALOG	*1 $ = 5 Frs	15,83	79,15
Transports (pour acquisition de documents	5,4	4	21,6
Toutes impressions (rapport final, notices LISA, captures d'écran, etc...)	*0,10	500	50
Reliures	20	5	100
		TOTAL (francs)	527

Conclusion

Ce travail a consisté en une véritable expérience dans le domaine de la recherche bibliographique : différentes sources, approche de supports variés (CD-Rom, DIALOG, Internet, etc...), sélection et traitement des documents, synthèse, normes de présentation du travail, etc... Les apports nouveaux ont surtout été une autre approche de la lecture (atteindre les informations pertinentes, rapidement), mais aussi une méthode de prise de notes stricte, qualitativement et quantitativement importante : savoir gérer les sources consultées, celles à parcourir, garder une trace de ses recherches, des descripteurs, etc... Au fur et à mesure, cette démarche a pris mieux forme, un début d'expérience a vu le jour et les différentes informations sont passées d'un ensemble assez flou à un ensemble bien structuré.

Enfin un dernier bénéfice, non négligeable pour mon avenir professionnel, est l'amélioration de la lecture d'articles en anglais, du moins pour ce domaine spécialisé.

2ème partie : Synthèse des documents

Introduction

Ce travail traite des programmes d'intelligence artificielle en catalogage descriptif, et non en catalogage de sujet (indexation). Il s'agit de systèmes experts que nous présenterons dans leur ensemble, pour nous pencher ensuite vers les applications concrètes réalisées, ainsi que certaines études théoriques. La bibliographie présentée plus loin reprend ce schéma d'organisation et les références sont mentionnées par leur numéro dans la troisième partie.

1. Présentation générale

Les systèmes experts, programmes d'informatique à base d'intelligence artificielle (IA), ont été créé pour effectuer des tâches, résoudre des problèmes au même niveau d'expertise qu'un spécialiste, professionnel, humain. De très nombreux écrits présentent l'IA et les systèmes experts ([1], [3], [11], [28] entre autres) en général.

Un tel programme se compose d'une base de connaissance (faits et règles : contient le savoir, l'expertise du domaine), mise en oeuvre par un moteur d'inférence (savoir procédural : fait fonctionner le tout) autour d'un problème particulier à résoudre, inscrit dans la mémoire de travail. [3] ajoute les caractéristiques suivantes : une interface avec l'utilisateur et une facilité d'explication de son raisonnement.

Il existe différentes conceptualisations du savoir dans un tel programme ([3]) : systèmes à base de règles, réseaux sémantiques, "frames" (trames), programmes orientés objet, règles de grammaire, raisonnement par incertitude, tableau noir, apprentissage par induction, apprentissage par réseaux neuronaux. Pour les développer, les concepteurs ont à leur disposition des langages de programmation adaptés (type PROLOG) ou des "shells" qui nécessitent peu de notions en informatique.

Dans le domaine des bibliothèques et des sciences de l'information, les systèmes experts ont fait l'objet des références [3], [5], [10], [17], [12] et [16] qui, pour la plupart, les présentent dans une visée professionnelle et pratique.

Pour le cas qui nous intéresse, le catalogage, on trouvera dans [1] une synthèse récente et assez complète. [6], plus ancien, présente ce qui a été publié en catalogage descriptif jusqu'en 1992 et aborde les systèmes experts. [2] étudie les opportunités pour le catalogage des périodiques.

D'autres articles nous font part de remarques "pratiques" sur les systèmes experts en catalogage : le domaine d'application doit être précis, bien limité, stable, fonctionnant d'une manière logique et non subjective ([2], [3], [21], [50]). Ils peuvent s'appliquer à toutes tâches demandant un niveau d'expertise, un jugement, un raisonnement de type heuristique (vs algorithmique). Les avantages sont importants : disponible 24 heures par jour, ils libèrent les experts de la fatigue, des tâches répétitives, leur permettent de s'atteler aux cas complexes en laissant les plus simples aux programmes. A une époque où les budgets sont restreints, de telles applications peuvent donc économiser les ressources humaines, mais aussi (c'est souvent le principal attrait) financières. Ils ne donnent cependant pas des résultats totalement parfaits. Pour certains ([3], [14]), les systèmes experts actuels ne sont donc capables que de conseiller, pas d'automatiser totalement et sans erreur un travail spécialisé. Pour développer de tels systèmes, beaucoup de ressources sont nécessaires ([2]) : une équipe motivée, du temps, de l'argent et le support de la hiérarchie et des futurs utilisateurs. Le plus compliqué, mais le plus crucial, est l'établissement de la base de connaissance, c'est-à-dire toute la conceptualisation du savoir, de la pratique des experts humains. [5] se penche ainsi sur cette étude ("knowledge engineering"), qui amène souvent à repenser ses propres procédures de travail ([2], [21]). De plus, une base de connaissance doit être réévaluée constamment, puisque l'expertise évolue ([3]). Lors du développement, il est important selon [21] de créer rapidement un prototype utilisable. Il apparaît cependant pour la plupart que de tels systèmes soient encore trop chers à implanter dans des bibliothèques ([5]).

2. Les projets, prototypes et produits développés

Nous présentons ici tous les systèmes experts en catalogage dont les auteurs font part, à l'état d'ébauche, d'étude ou déjà réalisés. Enfin, nous décrivons les travaux théoriques parallèles.

Une classification à trois niveaux est souvent communément adoptée : les conseillers, les créateurs d'enregistrements et les systèmes automatiques. Le niveau de compétence s'accroît bien sûr d'une catégorie à l'autre, ainsi que le degré de complexité nécessaire pour développer de tels programmes. A l'intérieur de chaque classe, les travaux sont présentés par ordre chronologique.

2.1. Les conseillers

Les systèmes experts conseillers aident les catalogueurs en panne d'inspiration, lorsqu'un doute survient par exemple pour la compréhension des règles ou le catalogage d'un support inhabituel. Ils sont souvent à vocation pédagogiques ([7]), pour permettre aux catalogueurs novices d'atteindre le niveau... expert !

L'utilisateur doit parcourir une série de questions, de menus, le programme suivant pas à pas une démarche catalographique guidée par ses réponses. Au bout du compte, une solution du type "la règle numéro tant d'AACR2 s'applique" ou "tel nom d'auteur doit être l'entrée principale" est proposée. Cela reste donc de l'ordre de l'incitation et n'est pas suivi d'une intégration de notice dans un système automatisé.

HEADS fut ainsi développé par Black, Hargreaves et Mayes en 1985, à la bibliothèque du Polytechnique de Teesside ([8], [19]).

Eyre (1986), pour sa part, s'est soucié du chapitre 22 des règles de catalogage AACR2 au Polytechnique de Londres-Nord ([8]).

APEX (Access Point Expert) de Piotr Murasik (université de Gdansk, Pologne) a dû traiter des points d'accès, comme son nom l'indique, mais peu d'informations ont été trouvées sur ce système ancien([8]).

Gibb et Sharif (1988) ont ensuite construit CATALYST à l'université de Strathclyde. On pourrait parler de CAO (Catalogage Assisté par Ordinateur) pour ce conseiller : il est cependant considéré par d'autres comme un programme classique de type algorithmique, plutôt qu'un véritable système expert de nature heuristique.

Le projet du système expert assistant en catalogage de la Bibliothèque Nationale Américaine de Médecine (1989-1991) consiste en un conseiller semi-automatique pour la sélection des noms de personnes([21]). Il a été testé sur les noms à préfixes et utilise une comparaison avec les entrées de fichiers d'autorité. Trop coûteux, il n'a pas donné de suite.

MAPPER (1990) apparaît dans de nombreuses références ([3], [6], [8], [18], [19], [22], [23], [24], [25], [26], [50]). Les premiers travaux de Zorana Ercegovac ont donnés lieu à une thèse à l'université de Californie. Il s'agit d'un système conseiller en catalogage de matériel cartographique, développé dans un environnement hypertexte. Il intègre la formalisation des règles AACR2, ainsi que le savoir-faire de catalogueurs professionnels. Il pose des questions, conseille, explique, fournit les champs bibliographiques (responsabilités intellectuelles, titre et publication). Il s'agit pour certains de la première approche positive des systèmes experts en catalogage, montrant sa faisabilité, puisqu'il s'appuie sur un domaine bien circonscrit : la cartographie. L'application a donné lieu à plusieurs rapports, un test comparatif parmi deux groupes d'étudiants en catalogage et une période d'essai à l'OCLC et la National Science Foundation (USA).

CONFER (Zainab Awang Ngah, 1991) conseille sur le catalogage des actes de colloques d'après AACR2 ([27]).

Enfin, [28] regroupe le manuel et la disquette du programme AACR2EXPERT (Smith, Poulter et Shaw, 1993), la seconde édition d'un système expert en catalogage qui choisit les points d'accès et leurs formes d'après les chapitres 21 à 25 d'AACR2. Disponible dans le commerce, il est constitué de cinq fichiers : un pour le shell (moteur d'inférence) et quatre pour la base de connaissance, modifiable. Ce conseiller pédagogique

ne crée aucun enregistrement, mais propose des conclusions et explique les mécanismes qui l'y ont conduit. Un chapitre dans [1] lui est consacré, qui critique l'inadéquation de sa base de connaissance, fondée sur des règles AACR2 jugées ambiguës. De plus, d'après Dini, les algorithmes du système ne feraient pas appel à l'intelligence artificielle.

2.2. Les créateurs d'enregistrements

Conseiller, c'est bien. Conseiller et produire une notice bibliographique en final, d'après les indications du catalogueur, c'est mieux. C'est le rôle des créateurs d'enregistrements qui, pour la plupart, fournissent après vérification de l'utilisateur une fiche catalographique informatique aux normes voulues (MARC, SGML, etc...).

Davies et James ([8], [11], [15], [18]) ont ainsi travaillé sur un projet d'automatisation de la seconde partie d'AACR2 (université d'Exeter, 1983-1984).

Le plus connu de cette catégorie est sans doute le système expert de Hjerppe et Olander (1983, 1985, 1989), baptisé ESSCAPE : Expert Systems for Simple Choice of Access Points for Entries ([8], [11], [15], [29], [50]). Développé en Suède à l'université de Linköping pour une bibliothèque universitaire (LIBLAB), ce sous-système en catalogage du projet HYPERCATALOG s'est centré sur le choix des points d'accès en identifiant les entrées principales et secondaires. Il a montré qu'un système expert était réalisable, mais pas assez économique pour une bibliothèque. Lors des différentes études, il est apparu de plus un besoin de révision des règles AACR2. Incluant des capacités pédagogiques, le programme a donné lieu à deux systèmes experts, basés sur des shells différents : ESSCAPE/EMYCIN, qui a nécessité une réécriture des règles de catalogage, et ESSCAPE/Expert-Trees, une application capable d'apprendre à partir d'exemples.

Notons en dernier lieu un système expert particulier, capable de créer des fichiers d'autorité. L'article [30] se base ici sur un contact personnel de l'auteur à la Bibliothèque Américaine du Congrès pour ce projet en cours en 1995.

2.3. Les systèmes experts automatiques

L'idéal est enfin de posséder un système de catalogage entièrement automatique, capable de repérer les informations bibliographiques de n'importe quel support et de les codifier, sans intervention humaine, dans un système informatique (OPAC) pour de futures recherches. Une fois intégrées, d'autres applications intelligentes pourront se charger de mettre à jour les informations dans l'ordinateur, à savoir la concordance des données avec des listes d'autorité, puis le nettoyage et la suppression d'enregistrements doublons. Enfin, à la base de tels projets, on trouve de nombreuses études théoriques.

Il est à noter qu'il est difficile de véritablement différencier les créateurs d'enregistrements des systèmes automatiques, puisque ceux-ci créent aussi des enregistrements. Le premier type nécessite toujours un minimum d'intervention humaine. Le second ne devrait pas en avoir besoin, mais il est souvent nécessaire d'indiquer un minimum de structure pour la reconnaissance de documents. Le classement des produits parmi ces deux catégories s'est donc effectué d'après ce qu'en ont écrit les différents auteurs.

2.3.1. Reconnaissance de documents (OCR, SGML, ...)

Les systèmes présentés jusqu'ici sont la plupart du temps anciens. A l'époque (avant 1990), un programme totalement automatique n'était pas concevable, comme en témoignent de nombreuses références, ceci souvent à cause des imperfections de la reconnaissance optique de caractères (en anglais : "OCR"). Avec les progrès techniques dans ce domaine, les projets d'acquisition de documents, de conversion rétrospective de notices papiers vers un support informatique se sont répandus. Ils se heurtent aux problèmes de structure de documents, d'extraction des informations bibliographiques et création de notices en formats normés (ISBD, MARC, SGML,...) etc...

Weibel, Oskins et Vizine-Goetz ont mené un projet d'automatisation du catalogage d'après des pages de titre ([3], [11], [18], [31], [32]) à l'OCLC, en 1990. Les pages numérisées devaient être entrées selon les règles AACR2 dans un système de catalogage automatisé. Seuls 70 % des éléments ont été correctement reconnus

(problèmes de l'OCR). Weibel a effectué, seul, des recherches plus poussées : identification des composants bibliographiques, encodage SGML, conversion vers un format MARC.

[18] cite un autre projet de l'OCLC nommé ADAPT (il pourrait s'agir du même...), ainsi que celui du Centre de Recherche National Allemand en Informatique.

Dans [8], le successeur d'un programme nommé AUTOCAT (université technique de Darmstadt, Allemagne), développé par Rauth et Schnellbach, est mentionné. Il s'agit d'AUTOCAT2, en cours de réalisation lorsque l'auteur écrivait (1992). Il se penche sur le catalogage d'articles de journaux scientifiques.

Un système expert de la bibliothèque universitaire de Tubingen (Zurich, 1992) est référencé par [33]. Il est chargé de la conversion rétrospective du catalogue.

Le prototype MORE ("Marc Optical REcognition") est voué à la même tâche à la Bibliothèque Nationale de Belgique. Construit de décembre 1992 à décembre 1992, il structure les documents numérisés dans un code SGML ou ODA et traduit les informations vers un format UNIMARC ([34]).

Dans les bibliothèques universitaires espagnoles, le projet LAURA (1992-1994) a donné lieu à un système expert récupérant de la numérisation d'ouvrages les informations pour des champs ISBD ou un autre format ([35]). Il s'intègre dans un programme de la CEE, nommé (ou par ?) VERBA LOGICA (voir [36]).

En 1993, CAROL ([37]) est du même type, resté visiblement à l'état de prototype.

RIDDLE (Rapid Information Display and Dissemination in a Library Environment, 1993-1994) est un prototype développé aussi sous l'égide de la CEE ([38], [39], [40]) pour le catalogage automatique de journaux scientifiques. Les documents passent une phase d'"océrisation", le résultat subit une transcription en SGML pour être intégré dans un OPAC. L'article [38], récent, présente deux liens vers

des sites web (voir [41] et [42] page 62, le premier étant la correction d'une indication fausse dans [38]).

Pour terminer, le système GRAPHEIN ([43], [44]), centré sur l'analyse et la reconnaissance de documents (ARD) en général, a été testé pour la conversion rétrospective de notices ISBD papier de la Bibliothèque Nationale de Belgique (transcription par OCR vers SGML puis UNIMARC).

2.3.2. Contrôle d'autorité

Lors de la création de notices bibliographiques, il faut pouvoir veiller à la normalisation des entrées sous une forme d'autorité unique. Applicable dans un même catalogue, ce principe l'est encore plus lorsqu'un stock de notices de différentes origines est créé pour un partage du travail catalographique.

L'article [45] mentionne un système expert du projet QUALCAT (1994) chargé de centraliser le contrôle d'autorité entre plusieurs bases de données bibliographiques.

2.3.3. Nettoyage et dédoublonnage des catalogues

Une fois le catalogue alimenté, il faut pouvoir le nettoyer des "impuretés" (champs vides, manquants, mal remplis, etc...), voire supprimer après détection les éventuels doublons.

C'est le rôle du système KARIN de la bibliothèque universitaire de Karlsruhe en Allemagne ([46]).

De même, dans le projet QUALCAT cité plus haut ([45]), il existe un programme capable de détecter les doublons et de sélectionner le meilleur candidat pour la conservation d'un exemplaire unique.

2.3.4. Etudes théoriques

Les systèmes experts automatiques s'accompagnent aussi de travaux plus théoriques qu'il nous a paru utile de présenter. Ils traitent de l'expertise catalographique,

de la reconnaissance d'indices visuels, des règles nécessaires pour un système expert, etc...

Ainsi Molto et Svenonius (1988-1990) ont élaboré deux algorithmes de reconnaissance automatique des noms (personnels et de raisons sociales) d'après les pages de titre ([11], [18], [47], [48], [50]).

Dans un tout autre genre, [4] et [49] propose l'implémentation des techniques de réseaux neuronaux dans les programmes de catalogage.

Jeng, de son côté, est un auteur très productif ([11], [18], [50], [51], [52]) et s'est penchée sur l'expertise catalographique. En 1990 elle a tenté de reformuler les règles AACR2 pour un éventuel système expert : celles-ci doivent, selon elle, prendre la forme de paires logiques "Condition/action", spécifiant le domaine et le degré d'application. Plus tard, elle a étudié les indices bibliographiques visuels et linguistiques contenus dans 203 pages de titre. A la Bibliothèque Nationale d'Agriculture des Etats-Unis, elle s'est intéressée à un modèle général de l'expertise catalographique en trois dimensions : tâches, capacités, niveau de compétence.

Meador et Wittig (1991) ont tenté de trier les règles les plus utiles à un système expert ([53]). Dans un même registre, Abrera et Shaw (1992) ont montré que seul un noyau de ces règles était véritablement utilisé ([54]).

Enfin, au sujet des règles de catalogage anglo-américaines AACR2, une désapprobation générale s'est faite sentir dans ces études, mais aussi dans l'ensemble des autres références. Ces règles ne sont pas appliquées par tous et elles ont l'inconvénient d'être parfois ambiguës : leur interprétation fait appel à du subjectif, des connaissances personnelles (appréhension du monde réel) et publiques (politique de catalogage locale) variables, ce qui est contraire à leur conception. Au fondement de nombreuses bases de connaissance, elles devraient pour de nombreux auteurs être révisées et, qui plus est, ne sont peut-être plus adaptées au catalogage des supports de demain.

Conclusion

Comme attendu dès la définition du sujet, les projets conduisant à des applications concrètes ne sont pas très nombreux dans la littérature. Nous avons pu constater que les systèmes les plus récents se tournent vers un catalogage entièrement automatique, alors que ceux plus anciens ne dépassaient pas le stade de création d'enregistrements. Les progrès de l'OCR y sont pour beaucoup.

Pour terminer, soulignons de nouveau une remarque importante. Il a été très difficile, à la seule lumière des articles et de mes connaissances en intelligence artificielle, de déterminer si un système appartenait ou non à ce domaine ou s'il consistait en un programme algorithmique classique (en particulier les projets d'assistance en catalogage, comme l'EAO, le classement d'images scientifiques, le catalogage de supports particuliers, etc...). Ceci explique la présence dans la bibliographie de noms qui n'ont pas été cités dans ces paragraphes.

3ème partie : Bibliographie

Pour les explications de classement, résumé et normes, se reporter à la page 25.

1. Présentation

[1] DINI (Rossella). - Sistemi esperti e catalogazione [italien - Expert systems and cataloguing]. - *Bollettino AIB*, 35, 2, Juin 1995, pp. 159-210. - INFB

L'ouvrage le plus récent et le plus complet (mais en langue italienne) sur les systèmes experts en catalogage. Dans une première partie, l'auteur les définit : base de connaissance, moteur d'inférence, base de faits, règles de production, direction du raisonnement (chaînage avant vs arrière). La seconde partie est consacrée à l'étude d'AACR2EXPERT dans son édition de 1993 : fonctionnement, critiques. Dini dresse ensuite une synthèse des systèmes experts en catalogage : historique (les origines), les conseillers (AACR2EXPERT, projet du Polytechnique de Londres-Nord, HEADS, CATALYST, MACCAT, APEX, MAPPER), les créateurs d'enregistrement (travaux de James et Davies, ESSCAPE), les études théoriques pour un catalogage entièrement automatique (Svenonius et Molto, Clarke et Cronin, Davies, Helga Schwarz, Hjerppe et Olander, Jeng, projets de l'OCLC et de la Bibliothèque Nationale Américaine de Médecine). Enfin une réflexion prospective est menée, puis la bibliographie est présentée, intéressante car récente. L'article est suivi d'un complément (Expert systems and cataloguing, pp. 209-210), qui n'est autre qu'un résumé, en langue anglaise, de l'article précédent.

[2] WEISS (PJ). - Getting the expert into the system : expert systems and cataloging. - *Serials Librarian*, 25, 3/4, 1995, pp. 235-241

Les rôles et opportunités des systèmes experts pour les catalogueurs de périodiques : article explicatif général. Les critères de faisabilité (tâches susceptibles d'être ainsi formalisées), domaines, supports humains et financiers. Donne quatre secteurs d'applications possibles : AACR2, classification, détection des doublons et "shelflisting" (ordre des champs ?).

[3] HAWKES (CP). - Expert systems in technical services and collection management. - *Information Technology and Libraries*, 13, 3, Sept 94, pp. 203-212

Avec la crise, il a fallu repenser le fonctionnement des bibliothèques, en terme de management, avec moins de personnel et de crédits. Les systèmes à base de connaissance peuvent être utiles à leurs services techniques pour développer un personnel plus autonome. L'auteur décrit les utilisations possibles des systèmes experts, les domaines appropriés pour les développer, ses composantes, le problème de construction de la base de connaissance, les différents types de moteurs d'inférences (à base de règles, réseaux sémantiques, frames, orientés objet, etc...) et les outils de développement (langage de programmation vs shell). Il fait part de certaines applications en bibliothèque, dont quelques unes en catalogage : le projet OCLC de Weibel et al., MAPPER, les études de Jeng sur l'identification du titre propre, SHELFPRO (classification et classement), CATALYST et CATTUTOR. Il conclut sur les solutions envisageables actuellement : des systèmes experts conseillers gérant les cas communs (80 %) en catalogage, ce qui permettrait aux experts humains de se pencher sur les 20 % restants, plus complexes.

[4] HOLTHOFF (Timothy N). - Library expert systems. - In : Encyclopedia of Library and Information Science. - ed. Allen Kent, 53, 1994, pp. 203-209. - ISBN 0824720539. - ENSSIB 020 ENC

Les systèmes experts sont présentés : composantes, la base de connaissance, les outils de développement. Plusieurs applications pour l'administration sont proposées (gestion budgétaire, management du personnel, planification), puis pour les services techniques de bibliothèque (catalogage, classification, collection). En catalogage on trouvera des références à des systèmes plutôt anciens : ESSCAPE, MAPPER, MITINET/MARC. L'auteur conclut sur les systèmes experts dans les services publics et les possibilités offertes par les réseaux neuronaux.

[5] POULTER (Alan), MORRIS (Anne) et DOW (Julie). - LIS professionals as knowledge engineers. - In : SCHAMBER (Linda) et al., Annual Review of Information Science and Technology. - ed. Williams Martha E. [pour ASIS], 29, 1994, pp. 305-350. - ISBN 0938734911. - ISSN 00664200. - ENSSIB 020ANN

Article sur l'expertise et son étude en bibliothèque, à travers la formalisation du savoir ("knowledge engineering") : historique, buts, étendue, représentation des connaissances, outils de développement, interface, aspects commerciaux. En catalogage on trouvera peu d'exemples.

[6] KNUTSON (G). - The year's work in descriptive cataloging, 1992. - *Library Resources and Technical Services,* 37, 3, Juil 93, pp. 261-275

Synthèse de ce qui fut édité en 1992 sur le catalogage descriptif. Une partie sur les systèmes experts et l'intelligence artificielle.

[7] FEINMAN (VJ). - Computers and library instruction : expert systems. - *Computers in Libraries,* 13, 3, Mars 93, pp. 53-55

Passe en revue les systèmes experts à vocation pédagogique en bibliothèque.

[8] DAVIES (Roy). - Expert systems and cataloguing. - In : MORRIS (Anne), The application of expert systems in libraries and information centres. - Londres, ed. Bowker-Saur, 1992. - pp. 133-166. - ISBN 0862912768. - ENSSIB

Historique des systèmes experts en catalogage et études théoriques des années 80 : Hjerppe et Olander, Schwarz, Davies, Jeng, Svenonius et al. Classement ensuite des systèmes experts en trois types. Tout d'abord les conseillers : Eyre, HEADS, CATALYST, MACCAT, APEX. Puis les créateurs d'enregistrements : université d'Exeter, ESSCAPE/EMYCIN et ESSCAPE/Expert-Trees. Enfin, des études se font pour créer des systèmes entièrement automatiques, notamment sur la reconnaissance des informations bibliographiques d'après les pages de titre : Schwarz, Davies, Jeng, Svenonius et al., projets de l'OCLC (Weibel, Oskins et Vizine-Goetz). Actuellement une

application, AUTOCAT2, est en cours de développement à l'université technique de Darmstadt (Allemagne ; Rauth et Schnellbach, 1987), pour le catalogage automatique de journaux. L'auteur conclut sur le catalogage automatique, le contrôle des doublons, une possible approche connexionniste et l'étude des erreurs. L'avenir est dans le partage travail/travail complexe vs travail du système expert/travail répétitif et automatisable, ainsi que dans la révision des codes de catalogage AACR2.

[9] WECKERT (John) et MCDONALD (Craig). - Artificial intelligence, knowledge systems and the future library. - *Library Hi Tech*, 10, 1/2, 1992, pp. 7-13

Discute des applications possibles de l'intelligence artificielle dans les futures bibliothèques : catalogage, recherche en-ligne, classification, indexation et sélection de documents. Prévoit une bibliothèque du futur.

[10] MORRIS (A). - Expert systems for library and information services. A review. - *Information Processing and Management*, 27, 6, 1991, pp. 713-724

Listage de 22 prototypes de systèmes experts en bibliothèque et sciences de l'information : recherche d'informations, catalogage, résumé, bibliographie, indexation et classification. Ils pourraient libérer les professionnels des tâches ingrates.

[11] ANDERSON Barbara. - Expert systems for cataloguing : will they accomplish tomorrow the cataloguing of today ? . - *Cataloging and Classification Quarterly*, 11, 2, 1990, pp. 33-48

La complexité et le coût de la tâche catalographique ont amenés l'idée de son automatisation. Les principes du système expert sont présentés : parties constituantes, importances de la base de connaissance, nombre de règles en général. Vient ensuite les applications en catalogage : MITINET/MARC, MAPPER, ESSCAPE, projet Exeter de Davies et James. Puis les études théoriques : Jeng, Weibel et al., Svenonius. Conclut sur un constat d'impossibilité actuelle de systèmes totalement automatiques, l'intervention humaine en catalogage, le contrôle d'autorité, l'interprétation automatique des pages de

titre, le coût trop élevé pour les bibliothèques, les exceptions aux règles de catalogage, les nouveaux types de documents (« nous développons des technologies de demain basées sur les pratiques d'aujourd'hui ! »), le manque d'intérêt d'applications en sciences de l'information chez les chercheurs en intelligence artificielle et le besoin impératif d'impliquer les bibliothécaires dans de tels développements.

[12] BREWIS (W L-E). - Application of artificial intelligence in information work. - *Mousaion* (????),8, 2, 1990, pp. 65-81

Examine les systèmes experts en sciences de l'information (recherches et prototypes) : classification, catalogage, recherche d'informations et bibliographie.

[13] CHANG (R). - Developing a cataloging expert system. - *Ill. Libr.*, 72, 8, 1990, pp. 592-596

Développement de MITINET/MARC, un système demandant des informations à l'utilisateur et produisant des enregistrements au format MARC. Conclut à l'impossibilité d'un système expert pour cause de difficultés d'interprétation des règles AACR2.

[14] FENLY (Charles). - Technical services processes as models for assessing expert system suitability and benefits. - In : Artificial intelligence and expert systems. - ed. Lancaster, 1990, pp. 50-66

Fenly détermine une grille pour déterminer de l'utilité d'un système expert dans les services techniques. Il reste pessimiste pour le catalogage descriptif. Le contrôle d'autorité pourrait être un champ d'application, mais une analyse des coûts doit être envisagée.

[15] JACQUESSON (A). - L'intelligence artificielle. - *ARBIDO-R*, 5, 1, 1990, pp. 18-27

L'intelligence artificielle : applications documentaires et bibliothéconomiques. Historique de l'IA, représentation des connaissances (réseaux sémantiques, frames, règles de production), systèmes experts en sciences de l'information. Suivi de quelques applications documentaires et bibliothécaires (Davies et James, ESSCAPE), médicales et juridiques. Enfin, les perspectives.

[16] SCHUEGRAF (Ernst J). - A survey of expert systems in library and information science. - *Canadian Journal of Information Science*, 15, 3, Sept 90, pp. 42-57

Examine les systèmes experts en sciences de l'information (recherches et prototypes) : classification, catalogage, recherche d'informations, bibliographie, indexation, archivage et sélection d'articles.

[17] WECKERT (John) et COOPER (Clare). - Artificial intelligence, expert systems and librarianship : a review of the literature. - *Australian Library Review*, 7, 4, Nov 90, pp. 281-299

Passe en revue la littérature de l'intelligence artificielle et des systèmes experts, particulièrement dans les champs bibliothécaires et documentaires : applications en architecture, criminologie, agriculture, management, bibliographie, recherche d'informations, catalogage, classification et indexation.

[18] WEIBEL (Stuart). - Automated cataloging : implications for libraries and patrons. - In : Artificial intelligence and expert systems. - ed. Lancaster, 1990, pp. 67-80

Les bénéfices du catalogage automatisé dans les bibliothèques : pour un meilleur travail, un meilleur service. Présentation de travaux : Davies et James, Schwarz, Svenonius et Molto, Jeng, Ercegovac, Weibel et al. Parmi d'autres projets plus récents, on trouvera ADAPT, en cours (1990) au bureau de recherche de l'OCLC : système de conversion automatique de documents papier (numérisation) en enregistrement SGML lisibles sur machine (segmentation d'après des statistiques, traitement OCR, récupération des informations bibliographiques). Un autre du même type est proposé : "Prototype system for conversion of paper documents to SGML structured documents", au centre de recherche national allemand en informatique. Aucune référence sur ces derniers n'est donnée.

2. Projets, prototypes et produits développés

2.1. Conseillers

HEADS

[19] Expertni systemy v knihovnickych a informacnich sluzbach [tchèque - Expert systems in library and information work]. - I'92 Casopis [colloque]. - 34, 9, 1992, pp. 217-218

Les applications des systèmes experts en sciences de l'information et bibliothèques : recherche en ligne, catalogage (HEADS, CATALYST, MAPPER), résumé, services bibliographiques, classification et indexation.

CATALYST

[20] GIBB (Forbes) et SHARIF (Carolyn). - CATALYST : an expert assistant for cataloging. - *Program*, 22, Janv 1988, pp. 62

Conseiller pédagogue en catalogage.

BIBLIOTHEQUE NATIONALE AMERICAINE DE MEDECINE

[21] WEISS (PJ). - The expert cataloging assistant project at the National Library of Medecine. - *Information Technology and Libraries,* 13, 4, Déc. 1994, pp. 267-271

Dans le cadre d'un projet plus général (impliquant aussi la Bibliothèque du Congrès et la Bibliothèque Nationale d'Agriculture aux USA), la Bibliothèque Nationale de Médecine a essayé, à partir de Juillet 1989, de développer un système expert en catalogage pour la sélection des noms de personnes en rapport aux contrôles d'autorité. L'application assiste le catalogueur. Elle compare le nom entré aux fichiers d'autorité de la bibliothèque, télédécharge l'enregistrement approprié ou, s'il n'existe pas, recommande au catalogueur le nom à prendre selon les règles AACR2. Aidée par un consultant extérieur, l'équipe de travail a testé un prototype sur le cas particulier des noms à

préfixes (chapitre 22.5D1 AACR2). Cet article décrit les objectifs du projet, la sélection du matériel et du logiciel de conception, la formalisation du savoir d'experts humains, les problèmes rencontrés face aux règles anglo-américaines, l'expérimentation du projet et le coût de l'étude. Finalement, il fut abandonné après deux ans, car il demandait trop de moyens. Malgré cela, il souleva pour la Bibliothèque de Médecine des points intéressants : processus susceptibles d'être implémentés dans un système expert, besoin de réécriture des règles traditionnelles pour une automatisation.

MAPPER

[22] ERCEGOVAC (Zorana). - A multiple observation approach in knowledge acquisition for expert systems : a case study. - ***Journal of the American Society for Information Science*****, 43, 7, Août 1992, pp. 506-517**

Décrit comment le savoir personnel des catalogueurs de cartes et le savoir publié sur ce sujet ont été intégrés dans le système MAPPER.

[23] ERCEGOVAC (Zorana) et BORKO (Harold). - Design and implementation of an experimental cataloguing advisor : Mapper. - ***Information processing and management,*** **28, 2, 1992, pp. 241-257,**

Décrit un conseiller expérimental, MAPPER, pour assister le catalogage descriptif de cartes américaines. Etude de l'interface et de la formalisation du savoir des experts humains pour implémentation dans la base du système. Développé sous Hypercard (Apple).

[24] ERCEGOVAC (Zorana) et BORKO (Harold). - Performance evaluation of Mapper. - ***Information processing and management,*** **28, 2, 1992, pp. 259-268**

Evaluation sur deux ans de MAPPER à l'OCLC et la NSF (National Science Foundation), conseiller en catalogage de cartes US. A montré que des étudiants novices dans le domaine donnent de meilleurs résultats avec que sans.

[25] ERCEGOVAC (Zorana). - Design, implementation, and evaluation of an experimental cataloguing advisor - Mapper. - *Geography and Map Division Bulletin*, 163, Mars 1991, pp. 2-29

Premier rapport sur MAPPER, développé à l'initiative de l'OCLC et la NSF (National Science Foundation) dans le cadre du "Library and Information Science Research Grant Program". Conception, formalisation du savoir humain et tests avec des novices.

[26] ERCEGOVAC (Zorana). - Research on knowledge-based descriptive cataloging of considerable importance cartographic publications : an experimental advice-giving system, MAPPER (cataloging, HYPERCARD). - Phd Thesis [Thèse]. - UCLA, Los Angeles [USA], 1990

Thèse à l'origine de MAPPER

CONFER

[27] NGAH (Zainab Awang). - Using CRYSTAL : an expert system shell to create CONFER, a guide for cataloguers in determining the correct form of main headings for conference documents. - *Kelal Abadi*, 10, 3, sept 91, pp. 1-9

Le système expert CONFER, shell conseiller, est un guide pour les novices décrivant des actes de colloques. Règles appropriées (AACR2) et format pour les entrées principales. Présentation des travaux sur l'étude des règles pour ce système, crées d'après l'amalgame d'AACR2 et des conseils d'un enseignant en catalogage.

AACR2EXPERT

[28] SMITH (David), POULTER (Alan) et SHAW (Malcolm). - Using the new AACR2, an expert systems approach to choice of access points. - 2ème ed. [1ère ed. 1980]. - Londres, ed. Library Association Publishing, 1993. - coll. Student Textbook. - ISBN 1856040860. - livré avec une disquette du programme. - ENSSIB 025.32 SMI

Nouvelle édition par la "Library Association" d'un ancien ouvrage (Using AACR2 : a step by step algorithm approach, 1980). Elle se compose d'un manuel explicatif et d'une disquette contenant un système expert en catalogage, premier de ce type disponible dans le commerce. AACR2EXPERT est un conseiller pédagogique sur l'application des règles AACR2 (Il concerne les points d'entrée des catalogues : choix et formes). Le logiciel ne crée pas d'enregistrement, mais propose la description des mécanismes qui l'ont amené aux conclusions du type « la règle n° 4 de ce chapitre s'applique » : choix des questions posées à l'utilisateur, règles appliquées selon ses réponses, commentaires sur la règle finale, etc... Le manuel décrit les deux algorithmes du système sous la forme d'organigrammes "si...alors", après avoir brièvement présenté les systèmes experts et l'historique d'AACR2. La disquette contient un fichier du shell (moteur d'inférence) et quatre autres sur la base de connaissance, paramétrable et complétable : points d'accès (chapitre 21 AACR2), titres uniformes (chap. 25), noms de raison sociale (23-24), noms de personnes (22).

2.2. Créateurs d'enregistrements

ESSCAPE

[29] HJERPPE (R) et OLANDER (B). - Cataloging and expert systems : AACR2 as a knowledge base. - *Journal of the American Society for Information Science,* 40, 1, 1989, pp. 27-44

Etude de deux systèmes experts utilisés pour le catalogage d'une bibliothèque suédoise. Leur tâche principale était de choisir les points d'accès en identifiant les entrées principales et secondaires. Les résultats du projet ESSCAPE mis en place en 1983 ainsi que les processus du catalogage et les règles AACR2 adoptées dans le programme dont discutés. Aperçu des résultats et réflexion sur l'ensemble des règles. Il apparaît que les systèmes experts actuels de catalogage peuvent fournir des notices correctes et être utiles dans un environnement non traditionnel, mais s'avéreraient moins intéressants dans une bibliothèque.

BIBLIOTHEQUE DU CONGRES (USA)

[30] REYNOLDS (RR). - Tools for a new age : an overview. - *Serials Librarian*, 25, 3/4, 1995, pp. 223-233

Les projets pour le catalogage de périodiques dans le futur. Chapitre sur SGML. Exposition d'un projet (pas de références) : "Electronic cataloging in publication (CIP) at LC". Etude en cours pour produire des enregistrements CIP à partir d'ouvrages reçus sous forme électronique. Fonctionne ainsi : un éditeur envoie une version informatique via FTP (ou version papier qui serait scannée), possédant un minimum d'encodage SGML ou HTML. En utilisant un éditeur de texte, le catalogueur récupère la page du titre du manuscrit, ajoute certaines ponctuations ISBD et sélectionne chaque type bibliographique à la souris, en cliquant alors sur le menu approprié à cet élément. Il est aussi possible d'ajouter automatiquement des notes à la future notice. Une fois que toutes les données ont été mises en valeur, le programme les formate et les labelle en un enregistrement descriptif prêt à être vérifié par un spécialiste humain. Un autre projet, d'après un contact personnel avec Gary Strawn (bibliothécaire chargé des contrôles d'autorité à l'Université du Nord-Ouest, USA ; pas de références) fait part d'études en cours à la Bibliothèque du Congrès pour l'assistance à la création de fichiers d'autorité : le programme utilise le système NOTIS en inspectant les enregistrements bibliographiques, extrayant les champs sous contrôle et en les comparant aux fichiers d'autorité en listant les nouvelles entrées nécessaires. Sont présentés aussi les outils de mise à niveau du

savoir des catalogueurs de la Bibliothèque du Congrès : le CD-Rom "Catalogers Desktop", en hypertexte, contient les LCRI ("Library of Congress Rule Interpretations"), des manuels d'indexation et de classification, les formats MARC, etc... Expose aussi l'attente d'une version électronique d'AACR2 codée en SGML. Enfin d'autres rubriques sont proposées, comme possibilité d'un avenir catalographique : internet, l'hypertexte, les catalogues et bases de données communs, les nouvelles technologies, la réalité virtuelle.

2.3. Systèmes experts automatiques

2.3.1. Reconnaissance de documents (OCR, SGML, ...)

OCLC

[31] WEIBEL (Stuart), VIZINE-GOETZ (Diane) et OSKINS (William). - Automating descriptive cataloging. - In : Expert Systems in Libraries. - Norwood, N.J., ed. Aluri Rao et Riggs Donald E, 1990, p.123, 125-126 et 134

Un projet de l'OCLC : automatisation de capture d'informations bibliographiques d'après les pages de titre.

[32] WEIBEL (S), VIZINE-GOETZ (D) et OSKINS (W). - Automated title page cataloging : a feasibility study. - *Information Processing and Management,* 25, 1989, pp. 187-203

BIBLIOTHEQUE UNIVERSITAIRE DE TUBINGEN (ZURICH)

[33] LEONHARD (J-F). - Managing projects for retrospective catalogue conversion in the federal republic of Germany. - *International Cataloguing and Bibliographic Control*, 21, 3, Juil/Sept 92, pp. 38-41

Projet de conversion rétrospective de catalogage, dont un système expert de structuration de données bibliographiques à partir de numérisations (Bibliothèque universitaire de Tubingen, Zurich).

MORE

[34] SANDERPIJPEN (W), LUPOVICI (C), TOUZE (J) et BELAID (A). - MARC Optical Recognition (MORE). - Communication of the European Communities, DG XIII, Dec 92/Dec 93

Etude de faisabilité d'un système OCR de conversion rétrospective de catalogue de bibliothèque sous forme papier (MORE = "MARC Optical Recognition"). Plusieurs études : rôle des dictionnaires, analyse des structures de données, intégration des outils optiques. Développement d'un prototype et intégration dans l'environnement de bibliothèque : test en conditions réelles à la Bibliothèque Royale Albert Premier (Belgique).

LAURA

[35] ORTIZ-REPISO (V) et RIOS (Y). - Automated cataloguing and retrospective conversion in the university libraries of Spain. - *Online and CD-Rom Review*, 18, 3, Juin 94, pp. 157-168

Les méthodes de conversion rétrospective de catalogues papiers dans les bibliothèques universitaires espagnoles. Apparition des systèmes experts en 1993. Le programme d'intelligence artificielle LAURA est utilisé par la faculté Complutense (Madrid) et le conseil de recherche scientifique. Le mécanisme est le suivant : l'enregistrement est transféré par OCR ou traitement de texte vers un code ASCII, puis codifié par des méthodes d'intelligence artificielle pour former un enregistrement ISBD (les champs sont repérés par la présence de mots spécifiques, leur position, le contenu sémantique, la ponctuation, etc...). On crée ensuite un modèle standard qui pourra être transféré automatiquement vers un format particulier (MARC). La démarche s'est révélée fructueuse, les seuls problèmes étant les "coquilles" graphiques sur le document ou après "océrisation", qui ont affecté les décisions du système. Il est prévu une codification d'enregistrements autres que bibliographiques. LAURA fait partie d'un

projet de la CEE nommé Verba Lógica : project for retrospective conversion of bibliographic records using OCR+AI techniques, 1992.

[36] VERBA LOGICA. - Project for retrospective conversion of bibliographic records using OCR+AI techniques. - 1992

Projet de la CEE (LAURA), cité par [35].

CAROL

[37] SCHMIDT (J) et PUTZ (W). - Knowledge acquisition and representation for document structure recognition : the CAROL project. - In : proceedings of 9th IEEE Conference on Artificial Intelligence for Applications Conference Sponsor (9, 1-5 Mars 1993, Orlando, FL [USA]). - Los Alamitos, CA [USA], ed. IEEE Comput. Soc. Press, 1993. - pp. 177-181. - ISBN 0818638400

Description d'un système de reconnaissance automatique de documents papiers. Ils sont scannés et passés à l'OCR. Les informations obtenues servent comme entrée de règles d'interprétation, basées sur un document type et générées par un module d'apprentissage. Ces règles doivent permettre d'intégrer le document reconnu dans un catalogue informatisé de bibliothèque. Le prototype CAROL était encore testé par des bibliothécaires à l'époque de l'article.

RIDDLE

[38] HARRISON (AD), ROOS (FA) et THOMAS (RE). - (Semi) automatic capturing of bibliographic information from journal contents for inclusion in online library catalogues : the RIDDLE project. - *Electronic Library*, 31, 1, Fev. 95, pp. 15-20

Décrit le projet RIDDLE ("Rapid Information Display and Dissemination in a Library Environment"), initié par le programme des bibliothèques de l'Union Européenne ("Telematics Research and Development Programme"). Etude de faisabilité (Fev. 1993 -

Dec. 1994) de numérisation de périodiques scientifiques, puis extraction des informations bibliographiques des articles et chargement dans un OPAC, d'une manière la plus automatique possible. Ne s'applique pas aux journaux électroniques. Décrit l'utilisation de l'OCR, l'étiquetage des données obtenues en SGML et le chargement dans les catalogues de bibliothèques. Insiste sur le problème d'identification des périodiques durant la numérisation. Termine par une étude des coûts pour acquérir ce système. Un prototype a été testé dans la bibliothèque d'un des centres de recherche du projet (CWI, Hollande). Propose deux liens vers les pages web des deux groupes impliqués, dont un n'est plus à jour (voir [41] plus loin).

[39] LOOK (Huge). - Rapid information display and dissemination in a library environment (RIDDLE). - In *: First reported to CURRENT RESEARCH 1993* [Comission of the European Communities], 1993

Premier rapport de la CEE sur le projet RIDDLE. Sont impliqués : RAL Oxfordshire (Grande-Bretagne) et CWI (Amsterdam, Hollande). Prévoit le développement d'un prototype.

[40] EUROPEAN LIBRARY PLAN RESULTS. - In : *Aslib Information*, 20, 7/8, Juill/Août 1992, pp. 282-283

Les différents projets reçus en réponse à l'appel d'offres de la CEE ("European Library Plan"), dont RIDDLE.

[41] HTTP://WWW.CIS.RL.AC.UK/PROJ/RIDDLE.HTML [adresse web internet]

Propose la description du projet RIDDLE, les chercheurs impliqués et des versions disponibles du logiciel (par FTP). Coordonnée exacte d'un site mentionné comme erroné dans l'article principal sur RIDDLE ([38]) : http://web.inf.rl.ac.uk/proj/riddle.html

[42] HTTP://WWW.CWI.NL:80/CWI/PROJECTS/RIDDLE.HTML [adresse web internet]

Même type de contenu que [41].

GRAPHEIN

[43] BELAID (Abdel). - Analyse et reconnaissance de documents. - In : INRIA, Le traitement électronique du document [Cours INRIA , 3-7 Octobre 1994, Aix-en-Provence]. - ed. ADBS, 1994. - ISBN 2901046762. - ISSN 11602376. - ENSSIB 025.04 INS

Présentation de l'ARD, l'analyse et la reconnaissance de documents (non spécifique aux sciences de l'information) : convertit l'information papier pour qu'elle soit manipulable informatiquement. Description des obstacles (support physique, dualité physique/logique, aléas de la segmentation), types de structure (physique et logique, les objets et leurs relations), niveaux de structure (information principale/secondaire, corps du document/illustration/mise en page, typographie/géométrie, macro et microstructure), normes (SGML, ODA/ODIF), et des principales composantes d'un système d'ARD (acquisition et pré–traitements, segmentation physique, étiquetage logique, reconnaissance de la structure). Le but est le développement d'un système générique pouvant s'adapter facilement et efficacement à tout type de document : c'est le projet GRAPHEIN, qui suit la technique du tableau noir ("blackboard"). Plusieurs tests ont été effectués, dont la reconnaissance des notices bibliographiques (ISBD) papier en vue de la conversion rétrospective du catalogue de la Bibliothèque Royale Albert Ier (Belgique). L'auteur décrit leur structure (problèmes typographiques, lexicaux, syntaxiques), leur modélisation (modèle physico-logique, attributs des objets), les actions effectuées (pré et post-analyse), et la stratégie de reconnaissance ("océrisation", balisage SGML et présentation au système). Le résultat est un enregistrement en UNIMARC. Des essais ont été faits sur 400 notices environ, donnant un taux de satisfaction de 85 %, pour un temps de 20 secondes par notice. En conclusion, l'auteur déclare les progrès de l'ARD insuffisants : la solution se trouve à un haut niveau de processus cognitif, de lecture et de compréhension du sens.

[44] CHENEVOY Patrick. - Reconnaissance structurelle de documents imprimés : études et réalisations. - Thèse sous la direction de BELAID (Abdel). - Institut National Polytechnique de Lorraine, 1992

Thèse sous la direction d'Abdel Belaïd (voir [43]) : le système GRAPHEIN et la conversion rétrospective de notices ISBD à la Bibliothèque Royale de Belgique.

2.3.2. Contrôle d'autorité

QUALCAT

[45] AYRES (FH), CULLEN (J), GIERL (C), HUGGILL (JAW), RIDLEY (MJ) et TORSUN (LS). - Qualcat : automation of quality control in cataloguing. - In : *British Library Research and Development Department Report,* source 6068 [d'après le CD-Rom LISA], 1994, 112 p.

Le projet QUALCAT sur la mise en commun de bases de données. Inclut le développement d'un système expert capable de détecter les meilleurs enregistrements parmi plusieurs doublons, ainsi qu'un autre pour centraliser les fichiers d'autorités.

2.3.3. Nettoyage et dédoublonnage des catalogues

KARIN

[46] REICHART (M) et MONNICH (MW). - Dublettenkontrolle in bibliographischen Datenbanken [allemand - The control of duplicates in bibliographic databases]. - *Bibliothek Forschung und Praxis,* 18, 2, 1994, pp. 193-216

Le problème des doublons dans les bases de données. Décrit l'implémentation d'un système expert développé à la bibliothèque universitaire de Karlsruhe (Allemagne) pour le système de catalogage KARIN.

2.3.4. Etudes théoriques

MOLTO et SVENONIUS

[47] MOLTO (Mavis) et SVENONIUS (Elaine). - Automatic recognition of title page names. - *Information Processing and Management,* 27, 1, 1991, pp. 83-95

Travaux sur des algorithmes de reconnaissance automatique de noms (à distinguer d'autres chaînes de caractères) sur les pages de titre anglais,. Test de deux algorithmes, l'un sur les noms de personnes, l'autre sur les noms de raison sociale (comparaison de noms avec fichiers d'autorité et identification des marqueurs des noms sur la page). Les résultats sont différents selon les données (publiques ou universitaires), mais encourageants.

[48] MOLTO (Mavis) et SVENONIUS (Elaine). - Automatic derivation of name access points in cataloging. - *Journal of the American Society for Information Science,* 41, 4, Juin 90, pp. 254-263

Critères d'extraction d'informations d'après les pages de titre. De telles données sont-elles suffisantes pour générer automatiquement les points d'accès ?

RESEAUX NEURONAUX

[49] DOSZKOCS (TA). - Neural networks in libraries : the potential of a new information technology. - In : NIT'91 (2 Dec 1991, Budapest). - ed. Ching-Chih Che, 1991, pp. 27-33

Les réseaux neuronaux artificiels sont nés de la complexité de programmer des tâches "intelligentes" avec un modèle algorithmique classique (Von Neumann). Ils sont mieux adaptés aux tâches cognitives : association d'information, stockage, rappel, reconnaissance de formes, raisonnement et apprentissage. L'article décrit les potentialités de tels réseaux pour les bibliothèques : catégorisation dynamique d'informations, généralisation, classification, extraction de traits et apprentissage. Des applications existantes ou possibles sont présentées : catalogage, indexation et recherche en-ligne.

JENG

[50] JENG (Ling Hwey) et WEISS (KB). - Modeling cataloguing expertise : a feasibility study. - *Information Processing and Management,* 30, 1, Jan/Fev 94, pp. 119-129

Les auteurs passent d'abord en revue les systèmes experts en catalogage existants : rôles (Davies, Weibel et al., etc...), présentation des travaux (Davies et James, Hjerppe et Olander, MAPPER) dont seul le plus récent, MAPPER, confirme qu'une telle application est faisable, puisqu'il se limite à un domaine délimité (la cartographie). Décrit les principales difficultés d'un système expert : problème pas assez précis, manque d'analyse et de stratégie de résolution de problème, inadéquation des règles AACR2, etc... En conclut que l'étude doit se concentrer sur la base de connaissance : l'expertise catalographique. Jeng a déjà ouvert la voie dans son étude de description des indices visuels et linguistiques de pages de titre (schéma mental, heuristique). De même, Svenonius et Molto se sont concentrés sur l'utilisation des noms sur les pages de couverture en tant que points d'accès (fichiers d'autorité, indices typographiques). Jeng, enfin, a étudié les règles AACR2 qui sont, pour elle, le fondement de toute base de connaissance : elles doivent cependant d'abord être étudiées et fractionnées en paires logiques "condition/action" avant toute implémentation. Jeng et Weiss se penchent ensuite sur la définition de l'expertise d'après plusieurs auteurs : haut degré d'habileté dans un domaine, facilité d'action face à de nouvelles situations, mémoire fortement structurée, apprentissage par l'expérience, connaissance de ce qui peut être faux, etc... Pour déterminer son étendue, les auteurs ont conduit une large étude de faisabilité à la Bibliothèque Nationale Américaine d'Agriculture (Beltsville, Maryland), dans le département de catalogage : quelles tâches nécessitent de l'expertise, comment classifier cette dernière et qu'est-ce qui distingue un débutant d'un expert ? Seul le catalogage descriptif est concerné. Un modèle général d'expertise catalographique est apparu, en trois dimensions : tâches du catalogueur (recherche de bases de données bibliographiques ; détermination des points d'accès ; interprétation des données, règles et produits bibliographiques ; identification de problèmes ; gestion de la performance), capacités

nécessaires pour les accomplir (technicité : informatique, veille technologique... ; maîtrise d'outils professionnels : AACR2, MARC, LCRI... ; capacités de communication : échange de savoirs ; spécialisation dans un sujet ou une langue : acquis par l'expérience ou dès l'origine) et niveau de compétence indispensable pour effectuer une tâche. Jeng et Weiss passent ensuite en revue la relativité de l'expertise, le savoir public (politiques locales, produits de la bibliothèque...) vs privé, le transfert de savoir et le haut degré de stratégie cognitive nécessaire à un expert. Il apparaît en conclusion qu'une telle étude a été positive.

[51] JENG (Ling Hwey). - Knowledge representation of the visual image of the title page. - ***Journal of the American Society for Information Science*****, 42, 2, , Mars 91 pp. 99-109**

Rapport d'après l'étude de 203 pages de titre. Les indices visuels (éléments, séparations entre ces éléments) et linguistiques sont analysés pour un catalogage descriptif. Une structure est développée pour intégration dans un système expert ou un code catalographique.

[52] JENG (Ling Hwey). - The structure of a knowledge base for cataloguing. - ***Information Processing and Management,*** **27, 1, 1991, pp. 97-110**

Analyse des schémas généraux et du niveau d'application des règles AACR2 concernant le titre et les mentions de responsabilité. Tente de répondre à deux questions : qu'est-ce-qu'une règle AACR2 ? Quelles sont les domaines d'application ? Il y a deux types de règles : celles suivies par du texte et celles qui renvoient à d'autres. Les systèmes experts ont jusqu'à maintenant utilisés le premier type. Cette étude montre que la plupart des règles suivies de texte sont, en fait, constituées de plusieurs paires "condition/action", où chaque paire est une "sous-règle". Il apparaît qu'actuellement le degré d'application des règles est limité. Elles sont prévues pour cinq domaines principaux : définition, description, organisation, source d'identification de l'information et transcription. Les conclusions de cette enquête suggèrent qu'une base de connaissance doit être construite avec de telles paires et non des numéros de règles. Leur structure doit

contenir le domaine d'application, l'interprétation procédurale, le degré d'application et les conditions et actions spécifiées d'après le texte AACR2.

MEADOR et WITTIG

[53] MEADOR (Roy) et WITTIG (Glenn R). - AACR2 rules used in assigning access points for books in two subjects : implications for automatic cataloguing experts systems. - *Library Resources and Technical Services,* 35, 2, Avril 91, pp. 135-140

Etude théorique des règles AACR2 employées pour cataloguer deux groupes de livres : en chimie et en économie. Comparaison des résultats (un petit ensemble de règles du chapitre 21 est la plupart du temps suffisant) et implications pour la construction d'un système expert automatique.

ABRERA et SHAW

[54] ABRERA (Josefa) et SHAW (Debra). - Frequency of use of cataloguing rules in a practice collection. - *Library Resources and Technical Services,* 36, 2, Avril 92, pp. 149-161

Etude théorique des règles AACR2 utilisées le plus fréquemment, fondée sur une collection de 716 livres à cataloguer. 28,4 % des règles nécessaires au total ont été le plus souvent utilisées (choix du nom, 22.1A et 22.1B ; entrée sous le nom de famille, 22.5A1). Il apparaît, en comparaison avec d'autres études, qu'un tel noyau de règles "utiles" pourrait être incorporé dans un système expert

INDEX

—A—

AACR2 28, 29, 30, 31, 33, 34, 36, 39, 40, 42, 44, 45, 46, 54, 55, 56
AACR2EXPERT 29, 36, 45
Abrera 34, 56
ADAPT 31, 41
Adbs-info 6, 7
AFNOR 20
Anderson 39
APEX 15, 16, 28, 36, 38
ARBIDO-R 40
Archimag 10
ARD 32, 51
Aslib Information 50
AUTOCAT2 15, 31, 32, 38
Ayres 52

—B—

Base de connaissance 26, 27, 29, 36, 37, 39, 45, 54, 55
Belaid 48, 51
Biblio-fr 6, 7
Bibliothek Forschung und Praxis 52
Bibliothèque de l'ENSSIB 9
Bibliothèque Américaine du Congrès 13, 30, 42, 46
Bibliothèque Nationale Américaine d'Agriculture 34, 42, 54
Bibliothèque Nationale Américaine de Médecine 29, 36
Bibliothèque Nationale de Belgique 32, 51
Black 28
Bollettino AIB 36
Books in Print 13, 14, 17
Brewis 40
British Books in Print 13, 14, 17
British Library Research and Development Department Report 52
Bruit 3, 4, 15, 18

—C—

Canadian Journal of Information Science 41
CAROL 32, 49
Cataloging and Classification Quarterly 10
CATALYST 15, 28, 36, 37, 38, 42
CCN 21
CD-Rom 2, 3, 7, 11, 12, 15, 21, 22, 23, 24, 46, 52
CEE 32, 48, 49, 50
Centre de Recherche Centre de Recherche National Allemand en Informatique 31
Chang 40
Chenevoy 52
Computers in Libraries 38
CONFER 29, 44
Cooper 41
Couverture 9, 16, 17, 18, 54
Cullen 52

—D—

Davies 30, 38
Demandeur 1, 2, 4, 5, 12, 20, 23 *Voir* Roger
Descripteurs 3, 11, 18, 19, 24
DIALOG 1, 3, 11, 12, 16, 21, 23, 24
Dini 36
Dissertation Abstracts Online 13, 14, 17
DOCTHESES 2, 3, 11, 12
Doszkocs 53
Dow 38

—E—

EAO 34
Electronic Library 49
ENSSIB 1, 2, 3, 4, 5, 7, 9, 11, 12, 21
Ercegovac 29, 41, 43, 44
ESSCAPE 15, 30, 36, 37, 38, 39, 40, 45
European Library Plan Results 50
Exeter 30, 38, 39
Eyre 28, 38

—F—

Feinman 38
Fenly 40
First reported to CURRENT RESEARCH 1993 50

—G—

Gdansk 28
Geography and Map Division Bulletin 44
Gibb 28, 42
Gierl 52
GRAPHEIN 32, 51, 52

—H—

Hargreaves 28
Harrison 49
Hawkes 37
HEADS 15, 16, 28, 36, 38, 42
Hjerppe 30, 45
Holthoff 37
HTTP 50
Huggill 52
HYPERCATALOG 30

—I—

I'92 42
IA 20, 26, 40
Ill. Libr. 40
INFB 21

Information Processing and Management 10, 39, 43, 47, 53, 54, 55
Information Retrieval and Library Automation 10
Information Technology and Libraries 37, 42
INRIA 51
INSPEC 13, 14, 17, 18
intelligence artificielle 40 *Voir* IA
International Cataloguing and Bibliographic Control 47
INTERNET 3, 4, 5, 7, 18, 20, 23, 24, 32, 47, 50
ISA 13, 14, 17, 18
ISBD 31, 32, 46, 48, 51, 52

—J—

Jacquesson 40
James 30
Jeng 33, 36, 37, 38, 39, 41, 54, 55
Journal of the American Society for Information Science 43, 45, 53, 55

—K—

KARIN 33, 52
Kelal Abadi 44
Knutson 38

—L—

LAURA 32, 48, 49
LC MARC-BOOKS 13, 14, 17
Leonhard 47
Library Hi-Tech 10, 39
Library Resources and Technical Services 38, 56
Linguistics and Language Behavior Abstracts *Voir* LLBA
Linköping 30
LISA 2, 3, 11, 12, 13, 14, 17, 18, 23, 52
LLBA 13, 14, 17
Look 50
Lupovici 48

—M—

MAPPER 15, 29, 36, 37, 39, 42, 43, 44, 54
Marc 5, 13, 14, 17, 30, 31, 37, 39, 40, 46, 48, 54
Mayes 28
Mcdonald 39
Meador 34, 56
Molto 33, 36, 41, 53, 54
Monnich 52
MORE 32, 48
Morris 38, 39
Mots-clés 3, 11, 13, 19
Mousaion 40
Murasik 28
MYRIADE 21

—N—

Ngah 44
NSF 29, 43, 44
NTIS 13, 14, 17

—O—

OCLC 29, 31, 36, 38, 41, 43, 44, 47
OCR 31, 32, 34, 41, 47, 48, 49
ODA 32, 51
Olander 30, 45
Online and CD-Rom Review 48
OPAC 9, 10, 30, 32
Ortiz-Repiso 48
Oskins 31, 47

—P—

PACS-L 4, 18
PANCATALOGUE 3, 12, 23
PASCAL 2, 3, 11, 13, 14, 17, 18
PEB 21, 23
Pertinence 3, 8, 9, 16, 18
Poulter 29, 38, 45
Program 42
Putz 49

—Q—

QUALCAT 33, 52

—R—

Rauth 31
Règles 26, 28, 29, 30, 31, 33, 34, 36, 37, 39, 40, 42, 43, 44, 45, 46, 49, 54, 55, 56
Reichart 52
Réseaux neuronaux 26, 33, 37, 53
Reynolds 46
RIDDLE 5, 6, 8, 10, 15, 32, 49, 50
Ridley 52
Rios 48
Roger 1, 2, 3, 5
Roos 49

—S—

Sanderpijpen 48
Schmidt 49
Schuegraf 41
Serials Librarian 36, 46
SGML 7, 30, 31, 32
Sharif 28, 42
Shaw 29, 34, 45, 56
Shells 26, 29, 30, 37, 44, 45
Silence 18
Smith 29, 45
Strathclyde 28
Svenonius 33, 36, 38, 39, 41, 53, 54
Systèmes experts 1, 7, 8, 11, 20, 26, 27, 28, 29, 30, 33, 36, 37, 38, 39, 40, 41, 42, 45, 48, 54, 55 *Voir* IA

—T—

Teesside 28
The Electronic Library 10
Thomas 49
Torsun 52
Touze 48
Tubingen 32, 47

—U—

UNIMARC 32, 51

—V—

VERBA LOGICA 32, 49
Vizine-Goetz 31, 47

—W—

Weckert 39, 41
Weibel 31, 41, 47
Weiss 36, 42, 54
Wittig 34, 56