

LAURENCE MAZAURIC

LE TRAITEMENT COGNITIF DE L'INFORMATION:
POUR UNE PERSPECTIVE DOCUMENTAIRE

Mémoire de DEA

Professeur: M.J.P. Metzger

DEA Sciences de l'Information et de la Communication
Université Jean Moulin Lyon III
1992-1993

LAURENCE MAZAURIC

**LE TRAITEMENT COGNITIF DE L'INFORMATION:
POUR UNE PERSPECTIVE DOCUMENTAIRE**

Mémoire de DEA

Professeur: M.J.P. Metzger

**DEA Sciences de l'Information et de la Communication
Université Jean Moulin Lyon III
1992-1993**

SOMMAIRE

INTRODUCTION	page 4
Présentation de la problématique	page 5
Présentation du mémoire	page 9
LES CONCEPTS DE CONNAISSANCE ET REPRESENTATIONS DES CONNAISSANCES	page 11
Introduction	page 12
Psychologie cognitive	page 14
La perception	page 14
La mémoire	page 16
Connaissances et représentations	page 17
Acquisition des connaissances	page 20
Le langage	page 24
La sémantique	page 24
La pragmatique	page 27
L'intelligence artificielle	page 29
Connaissance et représentations des connaissances	page 29
Raisonnement et IA	page 32
Conclusion	page 35
ETUDE DES PROCESSUS DU FONCTIONNEMENT COGNITIF ET DU TRAITEMENT DE L'INFORMATION	page 37
Introduction	page 38
Des cartes objectives et des paysages subjectifs	page 40
Eléments de didactiques	page 44
Dimensions pragmatiques de la relation d'information	page 48
L'apport de l'IA: le dialogue homme-machine	page 58
Modélisation cognitive	page 67
Conclusion	page 73

ETUDE DE CAS: LA BANQUE DE DONNEES BID
ET SES UTILISATEURS

page 75

BIBLIOGRAPHIE

page 92

I N T R O D U C T I O N

PRESENTATION DE LA PROBLEMATIQUE

Cette étude a pour objet le transfert de l'information au sein d'un service documentaire. Cette notion est au coeur de l'activité de tout centre de documentation, la raison même de son existence. La première préoccupation du documentaliste est de pouvoir apporter une réponse, sous forme d'informations ponctuelles ou de documents, aux demandes des utilisateurs.

Trois pôles sous-tendent cette activité documentaire:

- un ensemble structuré d'informations
- l'utilisateur
- le documentaliste

Intermédiaire entre un système informatif et l'utilisateur, l'action du documentaliste se déploie dans deux directions:

- connaissance approfondie du système informatif
- écoute des besoins en information de l'utilisateur.

En ce qui concerne le premier point, le documentaliste s'est donné les moyens techniques afin de dominer au mieux l'environnement informatif auquel il est confronté.

Une grande activité intellectuelle a été déployée pour déterminer les meilleures approches possibles afin de permettre un accès facilité aux domaines de la connaissance.

Ces travaux ont porté sur les fonds documentaires et en ont permis une meilleure gestion: analyse, indexation, classifications,...

La connaissance scientifique, structurée par des classifications complexes emboîtées les unes dans les autres, est sujette dans le même temps à des caractéristiques dynamiques. Chaque jour d'innombrables faits se produisent dont une infime partie sera publiée, diffusée et communiquée. La connaissance scientifique est le résultat d'une perpétuelle restructuration.

- (1) Dartois, C.: Approche de la relation
d'information,...
- (2) Chaumier, J.: Travail du documentaliste, p.8

Traditionnellement, l'information transmise doit répondre à ces critères:

- exhaustivité de l'information
- fiabilité de l'information
- nouveauté de l'information
- rapidité de l'information

L'objectif d'un service documentaire se définit à partir de l'étude des besoins de l'utilisateur. De nombreuses études ont été faites aussi dans ce domaine et des stratégies ont été développées pour une meilleure écoute des besoins, autant sur le plan relationnel afin de faciliter cette "relation d'information" (1) qu'à un niveau plus scientifique et systématique d'analyses techniques et psychologiques basées sur des interviews.

En termes de théorie des communications, l'objectif du documentaliste est donc de fournir la réponse la plus exhaustive (absence de silence) et la plus appropriée à la demande (absence de bruit) (2).

L'utilisateur se trouve confronté au besoin d'information dans le cadre de son activité, et pour des raisons multiples:

- mise à jour des connaissances et suivi du domaine dont il est spécialiste;
- informations ponctuelles en vue de la réalisation d'une tâche spécifique;
- informations ponctuelles ou acquisition d'une culture de base dans un domaine dont il n'est pas spécialiste.

Les études sur la relation utilisateur/service documentaire ont montré que celle-ci est loin d'être idéale. Le rapport souvent conflictuel que l'utilisateur entretient avec le centre de documentation souligne le décalage qui peut exister entre une politique d'offre d'un service et la manière dont celui-ci est perçu par le destinataire.

Ainsi, au sein des organisations qui se donnent les moyens d'accéder à l'information afin d'accroître leur connaissance de l'environnement et donc leur

implantation , le statut de la documentation est souvent mal défini. Le centre de documentation est perçu comme un service marginal, à disposition mais facultatif.

En effet, le service documentaire est rarement la source d'information privilégiée. Il arrive loin derrière la documentation personnelle, les collègues de travail, les anciens de l'université ou les experts rencontrés dans les congrès et avec lesquels on maintient des liens personnels.

A un moment où le documentaliste peut se sentir menacé dans son statut professionnel par de nouveaux agents économiques qui se veulent tous spécialistes de l'information, il ne serait peut-être pas inutile d'essayer de déterminer un terrain d'action spécifique au centre de documentation.

La spécificité d'un centre de documentation est d'être "le" service d'information d'une organisation. Pour cette dernière, la maintenance d'un tel service est un coût non négligeable. L'exigence d'un retour économique ne peut pas être sans conséquences sur la nature même de l'activité documentaire. Il ne s'agit pas ici d'affronter le problème particulièrement complexe du coût de l'information, mais de proposer quelques réflexions relatives à la nature du transfert de l'information.

Comme je l'ai déjà souligné, une organisation développe un service documentaire afin d'accroître sa capacité de connaissance. Cela est rendu possible par la possibilité qu'aura chacun des acteurs d'accroître sa propre connaissance dans le cadre des tâches qui lui sont dévolues.

C'est à ce niveau qu'entre en jeu le service documentaire: accroître la connaissance de l'utilisateur dans le cadre de ses fonctions spécifiques à l'entreprise.

Le service documentaire est au service de l'entreprise à travers le service à l'utilisateur, et non le contraire.

Cela signifie qu'il n'est pas suffisant de transmettre des documents ou des informations ponctuelles, même si ceux-ci sont pertinents par rapport à la demande.

L'entreprise a besoin que cette information soit systématiquement transformée en connaissance. Or, "information" est loin d'être synonyme de "connaissance".

(3) Brookes, B.C.: The foundations of Information Science

Le documentaliste pourrait être l'acteur privilégié de cette transformation.

Une nouvelle approche, rompant avec la tradition, tend à se focaliser sur la structure cognitive de l'utilisateur.

Des études ont montré comment la structure d'une recherche documentaire (interrogation, structure sémantique) est déterminée par la propre structure cognitive de la personne. La recherche documentaire ne doit pas prendre en compte seulement la représentation de la connaissance scientifique proposée par les sources d'information, mais aussi les représentations de la connaissance subjective.

La psychologie cognitive s'interroge sur comment la connaissance du monde est perçue, représentée, mémorisée, transformée et rendue accessible par notre structure cognitive.

La connaissance est structurée par des "modèles mentaux". Si l'information reçue appartient à un domaine pour lequel il existe déjà un modèle structuré, elle est assimilée et encodée directement à cette structure. Si l'information est inédite, son assimilation requiert la création d'un nouveau modèle et donc, une restructuration fondamentale.

Brookes (1976) a exprimé ainsi "l'équation fondamentale des sciences de l'information" (3):

$$I + K = K'$$

pour: I --> l'information
K --> la structure des connaissances
K' --> la structure modifiée

Le problème se pose alors de savoir comment s'assurer que l'information diffusée soit bien transformée en connaissance.

Une des réponses déjà apportées consiste à développer la notion de "conditionnement" de l'information afin de faciliter son assimilation, mais elle est loin d'offrir un moyen de contrôle systématique. Cette solution se base toujours sur l'adéquation demande de l'utilisateur/réponse du système informatif.

Une solution qui consisterait à jouer directement sur les structures cognitives de l'utilisateur pourrait permettre une possibilité de contrôle plus systématique.

L'attention du documentaliste devra se porter alors sur une autre adéquation: demande de l'utilisateur/ représentations des connaissances de l'utilisateur. Le passage de l'un à l'autre nécessite la mise en place de modèles: un modèle de l'utilisateur et un modèle de la représentation des connaissances.

PRESENTATION DU MEMOIRE

Ayant ainsi défini la relation qui s'établie dans un centre de documentation, je suis partie en quête de documents, limitant la recherche à la littérature relative au monde de la documentation.

J'ai pu constater très vite que le discours "documentaliste" s'étendait dans deux directions:

- appropriation d'éléments des discours des sciences cognitives et des sciences du langage afin de constituer une base théorique pour la recherche de l'information;
- adhésion totale au discours de l'intelligence artificielle qui a le mérite de clarifier et mettre en lumière les différentes étapes de la constitution d'un corpus de connaissances, et permettre ainsi une schématisation des fonctions documentaires.

L'ensemble de ces disciplines m'étant inconnu, je me suis lancée dans un premier moment à la découverte de ces disciplines, et navigant entre ces domaines, je me suis retrouvée noyée sous une masse de documents privilégiant les approches les plus diverses.

Afin d'éviter une dispersion chronique, je me suis donnée pour but dans la première partie de clarifier le concept qui se trouve à la base de ma problématique: le concept de connaissance et représentations des connaissances et leur modélisation, prenant en considération les disciplines qui se penchent sur ce domaine.

La deuxième partie intégrera ces notions dans l'étude des processus du fonctionnement cognitif et du traitement de l'information.

Cette étude se présente en trois parties:

- Ière partie: quelques définitions des concepts de connaissance et représentations des connaissances
- IIème partie: étude des processus du fonctionnement cognitif et du traitement de l'information.
- IIIème partie: présentation d'une étude de cas: la banque de données BID (Cour de Cassation Italienne).

1ère PARTIE

**LES CONCEPTS DE CONNAISSANCE
ET REPRESENTATIONS DES CONNAISSANCES**

(4) Morin, E.: La Méthode 3, p.11

INTRODUCTION

Le concept de connaissance est une notion complexe qui recouvre de nombreuses assertions. Ma première démarche a consisté à sonder les dictionnaires et encyclopédies pour essayer de faire le tour de la question.

Définitions

Le grand Dictionnaire Larousse propose quatre définitions du concept de connaissance:

- faculté de connaître en général;
- notion, idée de quelque chose;
- discernement, compréhension;
- expérience, habileté, instruction.

La notion de connaissance comporte en elle-même diversité et multiplicité. On ne saurait donc la réduire à une seule notion. Il faut plutôt concevoir en elle plusieurs modes ou niveaux.

Edgar Morin constate que toute connaissance comporte nécessairement:

- une compétence: aptitude à produire des connaissances;
- une activité cognitive qui nécessite un appareil cognitif, le cerveau;
- un savoir: résultant de ces activités. C'est dans le cadre de la culture qui lui a été transmise que l'esprit humain élabore et organise sa connaissance (4).

Ces présentations de la notion de connaissance mettent en évidence la nature multidimensionnelle de ce concept.

Les différentes disciplines scientifiques tentent de bien encadrer les diverses notions, se réservant chacune une part du concept général.

Le Grand Larousse Encyclopédique propose trois domaines mis ainsi en évidence:

- Nature de la connaissance: c'est le domaine de la psychologie cognitive. Il traite des notions liées au mode d'acquisition des idées, aux sensations, à la perception, à l'imagination, à la mémoire, aux concepts, aux jugements, au raisonnement, et aux sources d'erreurs possibles;
- Critères de la connaissance: c'est le domaine de la logique. Il traite des notions de vérité et d'erreurs. C'est la méthode des sciences, de la connaissance exacte;
- Valeur de la connaissance: c'est le domaine de la métaphysique.

L'Encyclopédie Universalis propose comme descripteur "Sociologie de la connaissance" qu'elle définit comme la science traitant du rapport entre connaissance et cadres sociaux.

C'est le domaine de la "nature de la connaissance" que je développerai ici.

Je me propose de présenter un tableau synthétique du sujet qui me préoccupe à travers les différentes approches de la psychologie cognitive, l'intelligence artificielle et les sciences du langage.

- (5) Approaches to Knowledge Representation, p.118
- (6) Traité de psychologie cognitive 1, p.3
- (7) Les Sciences cognitives, p.40
- (8) Vignaux,G.: Les sciences cognitives, p.188

PSYCHOLOGIE COGNITIVE

La psychologie cognitive, contrairement à la psychologie générale qui explique l'action humaine en termes de motivation et de désir, s'exprime en termes de processus mental et de représentations, présentés sous forme de modèles et de théories (5).

Nos représentations sont le résultat de nombreux processus que cette science s'applique à décrire.

La perception

La perception désigne l'ensemble des mécanismes et des processus par lesquels l'organisme prend connaissance du monde et de son environnement sur la base des informations élaborées par les sens. L'Ecole du Gestalt considère que les produits finis de la perception sont le reflet immédiat de la structure du monde.

La conception cognitive de la perception considère que le système perceptif est un système de traitement de l'information qui élabore des connaissances sur le monde environnant (6).

Les perceptions sont, en effet, des élaborations actives du cerveau, sélectionnant, traduisant et intégrant les sensations élémentaires fournies par les mécanismes de la transduction sensorielle au niveau des récepteurs périphériques (7).

Perception visuelle des formes

La reconnaissance d'un objet implique la médiation d'une représentation de l'objet: rôle des "prototypes" (8).

Il y a trois ordres de représentations cognitives:

- une représentation structurale: l'objet est reconnu dès que l'on a perçu un de ses traits caractéristiques, et non après une vision complète;

(9) Traité...1, p.66

(10) Traité...1, p.217

- une représentation sémantique: rapport que l'objet entretient avec un environnement donné;
- une représentation phonologique: dénomination de l'objet.

Ces représentations fonctionnent en parallèle, dans une interaction constante: on peut parler de "modèle de cascade" (Humphreys, Riddoch et Quilan) (9).

Perception du langage parlé

Il n'existe pas de véritable modèle de la perception de la parole, mais des cadres théoriques ont été élaborés:

- théorie motrice de la perception de la parole: hypothèse selon laquelle la perception et la production des sons de parole ont les mêmes mécanismes (Liberman et Mattlingly);
- théorie d'analyse par synthèse: hypothèse d'une médiatisation d'un niveau de représentation abstrait entre la perception et la compréhension (Halle et Stevens; Zue);
- théorie des détecteurs de propriétés: le processus perceptif est envisagé sous la forme d'une séquence d'étapes de traitement liées à la mise en oeuvre de détecteurs spécifiques (Studdert, Kennedy);
- modèle d'intégration des traits: appariement de propriétés acoustiques indépendantes (Massaro) (10).

- (11) Les Sciences cognitives, p.58
- (12) Traité de psychologie cognitive 2, p.4
- (13) Traité...2, p.11

La mémoire

Traditionnellement, la mémoire est considérée comme une simple forme de la cognition, spécialisée dans la représentation du passé (Piaget).

Les sciences cognitives proposent un autre postulat: la mémoire est probablement la forme même de la cognition.

En effet, les processus mnésiques ne déterminent pas seulement les possibilités d'accès aux représentations des événements du passé, mais les propriétés des mémoires transitoires et permanentes déterminent aussi notre présent psychologique et ce qui sera encodé, représenté et stocké dans notre environnement immédiat (11).

La mémoire produit de façon permanente des schémas hiérarchisés et emboîtés qui façonnent nos anticipations: elle permet l'apprentissage (12).

Le paradigme général reconnu comporte trois phases:

- une phase d'encodage et de stockage: l'information d'origine perceptive est transformée en une représentation mentale;
- une phase de rétention;
- une phase de réactivation de ces représentations mentales.

Une distinction a été établie entre mémoire transitoire et mémoire permanente:

- mémoire transitoire: système de mémoire dont la fonction est de maintenir les informations à un niveau d'évocabilité pendant une période de temps restreinte;
- mémoire permanente: système dont le but est d'organiser et de stocker les informations pour les utiliser ultérieurement (13).

Les procédés mnémoniques jouent un rôle central dans tous les traitements cognitifs de l'information. La mémoire n'est pas seulement une forme particulière de connaissance, elle est sans doute la forme même de la connaissance (Tiberghien).

- (14) Traité...2, p.23
- (15) Richard,J.F.: Les Activités mentales, p.10
- (16) Richard,J,F,: les Activités mentales, p.10
- (17) Traité...2, p.36
- (18) Richard,J.F.: Les Activités mentales, p.11

Connaissances et représentations

Ces notions sont centrales en psychologie cognitive.

Les représentations sont des structures au caractère transitoire. Elles sont élaborées dans le cadre de tâches et résultent des activités de construction d'interprétations (compréhension).

Elles disparaissent généralement après la réalisation de la tâche pour laquelle elles ont été construites. Sous certaines conditions, certaines composantes de ces représentations peuvent être stockées en mémoire: on parle alors d'acquisition ou construction de connaissances (Richard) (14).

Les représentations prennent en compte l'ensemble des éléments de la situation et de la tâche: elles sont donc très particularisées, occasionnelles et précaires par nature. Il suffit que la situation change ou qu'un élément non remarqué de la situation soit pris en compte pour que la représentation soit modifiée (15).

Les connaissances sont des structures stabilisées en mémoire à long terme. Elles constituent le savoir de base pour l'action et pour la compréhension des messages et des situations. Elles ne sont pas entièrement dépendantes de la tâche à réaliser. Tant qu'elles n'ont pas été modifiées, elles sont supposées se maintenir sous la même forme (16).

La différence entre connaissance et représentation tient dans le fait que la première structure a besoin d'être activée pour être efficiente, tandis que la seconde structure l'est immédiatement (17). Toutes les informations en mémoire à long terme ne sont pas disponibles: une faible partie de celles-ci seulement le sont, celles qui ont un niveau d'activation suffisant ou qui font l'objet d'une recherche en mémoire couronnée de succès (18).

(19) Traité...2, p.45

Typologie des connaissances et représentations

On distingue trois formes de représentation, correspondant aux trois formes classiques d'appréhension de la réalité:

- les représentations imagées expriment les structures spatiales caractéristiques de la perception visuelle;
- les représentations liées à l'action reposent sur la sensori-motricité, expression privilégiée des structures temporelles.
- les représentations propositionnelles sont basées sur le modèle prédicatif, caractéristique du langage, à la base de sa fonction majeure de communication et de transmission d'information.

On distingue deux formes de structures sémantiques:

- la notion de réseau relationnel: ce réseau sémantique est utilisé pour représenter les relations d'inclusion entre les concepts (Collins et Quillian);
- la notion de schéma (frame) pour décrire l'organisation des connaissances en mémoire et comment ces connaissances sont utilisées pour comprendre, mémoriser, faire des inférences (Minsky).

On peut considérer quatre formes de connaissances:

- les connaissances concernant les catégories naturelles: s'opposant à la conception aristotélicienne de catégorisation conceptuelle, cette hypothèse affirme le caractère privilégié de certaines représentations-types (Rosch) (19).

Trois concepts clés sont ici à introduire pour illustrer les structures sémantiques:

(20) Les Sciences cognitives, p.68

(21) Traité...2, p.46

(22) Traité...2, p.58

- le concept de niveau de base traduit l'existence d'un niveau privilégié dans l'organisation hiérarchique de référence qu'est la taxonomie: correspond au niveau d'abstraction des représentations sémantiques;
- le concept de typicalité traduit le fait que les individus s'accordent à considérer certains exemplaires des catégories plus représentatifs que d'autres. Les exemples typiques sont appris les premiers, sont plus rapidement actualisés et mis en oeuvre. Ce concept suppose la création de normes;
- le concept de prototypes montrent que les représentations des différents exemplaires d'une catégorie ne sont pas identiques, mais se distribuent selon leur degré de représentativité par rapport à un exemplaire particulièrement typique ou représentatif: le prototype (20).
- Les connaissances concernant les procédures: cette conception s'inscrit dans le cadre de l'opposition entre savoir et savoir-faire (Winograd).
 - la connaissance déclarative: éloignée de l'action concrète;
 - la connaissance procédurale: proche de l'action concrète, elle spécifie des structures de contrôle directement utilisables dans la réalisation de l'action.

Ces deux notions ne sont pas nécessairement liées entr'elles (21).

- Les connaissances concernant les conceptions du monde physique: les chercheurs ne peuvent que constater la persistance d'une "physique naïve", de réactions intuitives (Viennot, McClosey) (22);

(23) Morin,E.: La Méthode 3, p.107

- les connaissances concernant les conceptions de l'environnement technique: modélisation d'une connaissance opérationnelle qui permet d'utiliser une machine donnée (Carroll, Rouse et Morris).

La représentation est une synthèse cognitive dotée des qualités de globalité, cohérence, constance et stabilité.

Elle est connaissante, connaissable, analysable, descriptive par un esprit-sujet qui, de plus, en échangeant ses informations et descriptions peut mieux objectiver et enrichir sa perception et ainsi, vérifier sa connaissance du monde extérieur.

C'est une configuration mentale formant image. Cette image, projetée sur le monde extérieur, s'identifie totalement à ce monde: processus que l'auteur appelle "boucle perceptive" (Morin) (23).

Acquisition des connaissances

Dans une interaction constante avec son environnement et des situations particulières, le sujet est amené à construire de nouvelles représentations et donc à acquérir de nouvelles connaissances et modifier les connaissances acquises précédemment.

La compréhension

Comprendre, c'est construire une représentation: c'est à dire élaborer une interprétation.

La compréhension a trois composantes:

- la situation
- les connaissances
- la tâche

- (24) Traité...2, p.79
- (25) Morin,E.: La Méthode 3, p.144
- (26) Traité...2, p.92
- (27) Morin,E.: La Méthode 3, p.59

Nous pouvons mettre en évidence deux types de compréhension:

- la compréhension à visée épistémique: les informations sont intégrées dans les structures mnésiques en vue d'une récupération ultérieure (Denhière);
- la compréhension à visée pragmatique: finalisée à l'action, elle consiste à construire une représentation spécifique, orientée par la recherche d'une solution (Richard) (24).

La compréhension est un processus de connaissance qui appréhende tout ce dont nous pouvons nous faire une représentation concrète ou que nous pouvons saisir de façon immédiate par analogie.

Ce processus implique une projection et une identification, et donc, subjectivité et affectivité (Morin) (25).

L'apprentissage

L'apprentissage est un processus de modification des connaissances, ou du comportement, toujours au cours d'interactions entre le sujet et son environnement (George) (26).

Deux formes principales d'apprentissage sont considérées:

- l'apprentissage par l'action
- l'apprentissage par le tutorat

Apprendre signifie acquérir des savoir-faire, faire acquisition de savoir.

C'est une conjonction de la connaissance et de la découverte, l'union du connu et de l'inconnu (Morin) (27).

Cette théorie est une alternative à l'innéisme qui considère que l'on n'apprend que ce que l'on connaissait déjà; et à l'acquisitionnisme qui considère que seule l'expérience est source de connaissances.

- (28) Les Sciences cognitives, p.29
- (29) Traité...2, p.107
- (30) Les Sciences cognitives, p.31
- (31) Traité..., p.120
- (32) Les Sciences cognitives, p.33

Le raisonnement

L'étude du raisonnement relève traditionnellement de la logique pour ce qui concerne les propriétés formelles des opérations permettant de passer de prémisses à des conclusions par des inférences justifiées, et de la psychologie pour les conditions mentales, comportementales, voire sociales, dans lesquelles s'effectuent ces activités.

On peut distinguer deux grands niveaux d'analyse psychologique du raisonnement:

- la macrostructure: il s'agit des "raisonnements longs" mis en jeu dans les situations d'acquisition de connaissances et de certains apprentissages, d'argumentation, de démonstration, de résolution de problème et de conduite de dispositifs complexes. Les concepts pertinents pour l'analyse de ces situations sont ceux de buts et de stratégie;
- la microstructure: il s'agit des raisonnements courts qui forment la maille des premiers (28).

Nous pouvons mettre en évidence cinq modes de raisonnement:

- le raisonnement déductif: il s'agit d'un raisonnement inférentiel dont le but est d'obtenir une conclusion sûrement vraie à partir de prémisses posées comme vraies (Politzer) (29);
- le raisonnement par abduction: permet de déterminer les causes susceptibles d'expliquer un fait, en remontant des conclusions aux hypothèses (Demolombe, Farinas del Cerro) (30);
- le raisonnement inductif: approche probabiliste qui consiste à obtenir une conclusion générale, à partir de prémisses de portée particulière (Politzer) (31);
- le raisonnement analogique: mode qui utilise une situation déjà analysée pour expliquer une situation nouvelle (Bourelly, Chouraqui) (32);

(33) Les Sciences cognitives, p.36

(34) Morin,E.: La Méthode 3, p.48

- le raisonnement heuristique: finalisé par la tâche, dans une situation de résolution de problèmes (Richard, Kayser) (33).

Conclusion

Connaître, c'est principalement computer.
La connaissance est nécessairement:

- une traduction en systèmes signes/symboles;
- une construction de systèmes cognitifs;
- une solution à un problème: l'adéquation construction/traduction de la réalité qu'il s'agit de connaître.

La connaissance ne reflète pas directement le réel, mais elle le traduit et le reconstruit en une autre réalité: c'est le "computo cellulaire" (Morin) (34).

(35) Les Sciences cognitives, p.8

(36) Problèmes de psycholinguistique, p.19

LE LANGAGE

L'approche pluridisciplinaire du postulat cognitif renouvelle l'étude du langage.

Le but de la linguistique est de caractériser ce que les sujets connaissent à propos de leur langue par le biais de descriptions aussi précises que possible de la structure formelle des différentes composantes de la grammaire.

La grammaire d'une langue est donc censée refléter le savoir linguistique partagé par les locuteurs de cette langue.

La psycholinguistique rend compte de l'acquisition et de l'utilisation effective de ce savoir dans le cadre d'une modélisation de type traitement de l'information.

Les traitements psycholinguistiques sont conçus comme des processus de transformation conduisant d'une forme de représentation mentale à une autre forme de représentation mentale (35).

La sémantique (Le Ny) (36)

Les processus cognitifs qui sont mis en jeu entre l'entrée et la sortie de l'information sont constitués par un ensemble complexe et interactif d'opérations cognitives.

On considère en général que chaque opération cognitive consiste schématiquement dans les trois moments suivants:

- l'opération saisit, en entrée, de l'information organisée sous une certaine forme;
- elle transforme cette information;
- elle fournit en sortie de l'information organisée sous une nouvelle forme.

Trois notions supplémentaires s'ajoutent à ce schéma:

- les opérations de traitement de l'information ont localement un caractère séquentiel: la forme de sortie d'une opération est généralement la forme de sortie d'une autre opération;
- une opération peut prendre en entrée plusieurs sources d'information: c'est la convergence et le traitement conjoint des informations venues à la fois de la source nouvelle d'information et de la mémoire qui assure la compréhension;
- il existe des formes naturelles d'organisation de l'information qui sous-tendent le traitement sémantique.

Les grandes phases successives du traitement de l'information généralement admises sont au nombre de six:

- les activités perceptives qui ont pour aboutissement l'identification des unités signifiantes et le repérage de certaines caractéristiques locales ou relationnelles;
- la construction au fur et à mesure de la signification des phrases, et sa conservation transitoire dans la mémoire de travail;
- la mise en mémoire à long terme de cette signification;
- la conservation en mémoire à long terme de l'information sémantique stockée. Cette conservation n'est pas exempte de modifications;
- la recherche et le recouvrement de l'information sémantique qui a lieu dans toute situation impliquant la réactivation de l'information. D'importantes différences selon les situations existent durant cette phase;
- la production des énoncés qui constituent le protocole de rappel. La façon dont l'information sémantique recouverte chez le sujet est exprimée, est

également soumise au jeu d'un certain nombre de facteurs de situation.

L'analyse du traitement de l'information en psychologie du langage repose sur l'hypothèse d'une organisation "propositionnelle" de l'information. Dans les représentations humaines la forme propositionnelle est la forme dominante.

Selon cette hypothèse, les premières opérations de traitement transforment l'information de telle manière que, lors de la sortie de ces opérations, l'information se trouve pour l'essentiel sous forme propositionnelle: les opérations de traitement ultérieures maintiennent cette forme et s'effectuent à partir d'elle.

Les notions auxquelles on fait appel sont issues de la logique: "proposition", "prédicat", "calcul" (non numérique):

Proposition:

- la plus petite unité de discours à laquelle puisse s'appliquer une valeur de vérité, vrai ou faux (point de vue logique);
- la plus petite unité sémantique intégrée susceptible d'être traitée ou mémorisée (point de vue psychologique).

La proposition est analysable en prédicats et arguments.

Prédicat:

- en logique aristotélicienne, attribut d'un mot, d'une proposition;
- utilisé en logique par référence aux valeurs de vérité. C'est une fonction qui fait correspondre à tout exemple particulier l'une des deux valeurs vrai ou faux.

(37) Problèmes de psycholinguistique, p.88

(38) Problèmes de psycholinguistique, p.113

La pragmatique (Costermans, Hupet)(37)

Traditionnellement, la pragmatique est l'étude des relations entre les signes et leurs usagers (Pierce; Morris).

L'acte d'énonciation est au coeur de cette problématique: il ne s'agit pas seulement d'informer, mais il s'agit aussi de tenir compte de ce que l'interlocuteur sait déjà, de signaler ce que l'on croit que l'interlocuteur sait, croit ou suppose.

Les sujets n'ont pas seulement à faire à des énoncés, mais ils sont impliqués personnellement, comme locuteurs ou auditeurs, dans une situation d'énonciation: mise en fonctionnement de la langue par un acte individuel d'utilisation (Benveniste).

La pragmatique vise à l'étude des trois aspects de l'acte d'énonciation:

- l'acte en tant que tel: l'énoncé signifie d'abord "qu'il signifie" avant même de signifier "ce qu'il signifie" (Austin; Searle).
Il s'agit d'un processus d'interaction sociale;
- les situations où il se réalise: analyse des fonctions de représentation et de communication;
- les instruments de son accomplissement:
 - les marqueurs illocutoires;
 - les marqueurs d'attitude.

Les actes de langage (38)

La langue est un instrument de communication, c'est à dire un moyen par lequel les sujets agissent les uns sur les autres (Caron).

Le fait de parler est une façon d'agir et particulièrement d'agir sur les autres.

Deux types d'énoncés sont retenus:

- les énoncés constatifs: décrivent des faits étrangers à l'acte d'énonciation;
- les énoncés performatifs: décrivent une action présente du locuteur et leur énonciation a pour fonction d'accomplir cette action (Austin).

Tout énoncé, dès lors qu'il est placé dans une situation énonciative, constitue un acte qui modifie le contrat qui régit les rapports entre les interlocuteurs.

Trois notions sont mises en évidence:

- le locutionnaire: désigne l'activité cognitive et phonatoire impliquée par le fait de parler, et nécessaire pour rendre la parole possible;
- le perlocutionnaire: désigne l'ensemble des conséquences qui découlent de cette activité locutionnaire;
- l'illocutionnaire: ce que l'on fait par le fait même de l'énonciation (Austin).

Le langage est au centre de ma problématique. C'est essentiellement à travers lui que les représentations et les processus cognitifs seront connaissables.

(39) Les Sciences cognitives, p.61

(40) Haton, J.P.: L'Intelligence artificielle, p.39

L'INTELLIGENCE ARTIFICIELLE

L'intelligence artificielle est une science nettement pluridisciplinaire.

On peut dégager deux directions à son activité:

- étude des mécanismes de l'intelligence afin de concevoir des machines performantes pour la réalisation de tâches spécifiques;
- éclairer les mécanismes cognitifs en développant des modèles de raisonnement et des programmes qui permettent l'expérimentation.

Elle contribue par sa démarche à l'étude de l'intelligence humaine.

L'objet central de l'intelligence artificielle concerne la manipulation de symboles, appelés "connaissances".

Il s'agit de les représenter sur ordinateur, de mécaniser des raisonnements les manipulants et de les acquérir, éventuellement, par apprentissage.

Il existe de nombreuses représentations des connaissances qui ont des pouvoirs d'abstraction variés et sont associées à diverses formes de raisonnement (39).

Connaissance et représentations des connaissances

On entend par connaissance toutes les formes du savoir de l'homme:

- objets, faits et événements;
- concepts et relations entre les concepts;
- heuristiques et stratégies de savoir-faire;
- procédures de raisonnement (40).

Un système d'intelligence artificielle utilise des connaissances sur le domaine dans lequel il est opératif. Il doit aussi pouvoir manipuler les

(41) La Recherche en informatique, p.71

(42) Haton, J.P.: L'Intelligence artificielle, p.40

connaissances dont il dispose afin d'expliquer ce qu'il a trouvé, d'apprendre en comprenant les raisons de ses succès et de ses échecs.

Cela demande de travailler à deux niveaux:

- niveau de base: résolution du problème;
- niveau "meta": observation du déroulement de la résolution du problème (41).

Le problème de la représentation des connaissances est celui de leur transcription sous une forme symbolique qui puisse être exploitée par un système de raisonnement.

Les représentations logiques

Issues des développements théoriques dans le domaine de la logique formelle (Newell, Shaw, Simon), elles concernent surtout la logique mathématique:

- la logique des propositions;
- la logique des prédicats du premier ordre;

La logique formelle est parfaitement adaptée au raisonnement exact disposant de données complètes. Mais elle ne fournit pas de moyens d'organisation de l'ensemble des connaissances manipulées.

De nouveaux formalismes ont été étudiés en vue d'étendre le champ d'utilisation de la logique:

- les logiques multivaluées: permettent de rendre compte de l'ambiguïté ou du flou dans l'expression des connaissances;
- les logiques modales: expriment la notion intuitive d'implication;
- les logiques non monotones: permettent de traduire les raisonnements par défaut (42);

Les réseaux sémantiques

Un réseau est une représentation graphique composée de noeuds reliés par des segments.

(43) Approaches to Knowledge Representation, p.45

Les noeuds représentent les concepts, les segments représentent les relations de nature sémantique (43).

Raisonnement revient à calquer un réseau représentant le problème posé sur un réseau de connaissances par un mécanisme de mise en correspondance structurelle.

Les réseaux sémantiques sont utilisés dans les systèmes de compréhension du langage naturel pour représenter des connaissances traduisant des associations d'idées entre concepts et des taxonomies de connaissances où les liens entre les concepts sont de nature hiérarchique.

Les règles de production

Les règles sont des morceaux de connaissance indépendants qui ne se réfèrent pas les uns aux autres. Chaque règle contient dans sa partie prémisses les conditions de son application.

Elles constituent un moyen simple et naturel de traduire la connaissance heuristique mis en jeu dans un raisonnement de type conditionnel.

Les représentations procédurales

Ensemble des procédures algorithmiques, souvent associées au mode de représentation déclaratif.

Les représentations par objets structurés

La nécessité d'organiser de façon structurée l'ensemble des connaissances disponibles a donné naissance à la notion de frames et scénarios (Minsky, Schank).

Ces formalismes permettent la restructuration des objets conceptuels manipulés dans un certain type d'activité.

- Frame (ou prototype): regroupe de façon structurée l'ensemble des connaissances relatives à un objet physique, un concept, une situation-type.

Il se compose d'un ensemble d'attributs correspondants à autant de notions associées au concept. Chaque attribut est décrit selon différentes facettes.

(44) Haton, J.P.: L'Intelligence artificielle, p.51

(45) Les Sciences cognitives, p.83

- Scénario (= script): repose sur le même principe de représentation mais il est relatif à des séquences typiques et non plus à des descriptions statiques d'objet.

Les représentations par objet permettent:

- structuration et modélisation d'un univers d'application;
- rassemblement de l'ensemble des éléments de connaissance relatifs à un concept;
- couplage des aspects déclaratif et procéduraux;
- hiérarchisation des connaissances et des raisonnements (44).

Des approches différentes se développent, en particulier le courant connexionniste avec le neurocalcul, inspiré par les neurosciences.

Cette modélisation considère le neurone comme système élémentaire de traitement de l'information. Les unités fonctionnelles du cerveau peuvent alors être modélisées comme des populations de tels neurones schématisés, reliés par des connexions transmettant en la modulant, une information quantitative (45).

Raisonnements et intelligence artificielle

Le raisonnement dans les systèmes à bases de connaissances est multiforme.

Un dénominateur commun est le concept d'hypothèse, parcelle de connaissance nouvelle, créée par un mécanisme de raisonnement pour rendre compte d'un phénomène observé dans le cadre d'un univers d'application.

Un raisonnement peut être défini comme un enchaînement d'énoncés ou de représentations symboliques, conduit en fonction d'un but. Un tel enchaînement n'est pas linéaire et nécessite des retours en arrière, présents dans la quasi totalité des systèmes d'intelligence artificielle tout comme dans la démarche humaine.

Cette définition générale du raisonnement met en évidence l'étroite imbrication entre raisonnement et

(46) La Recherche en informatique, p.64

(47) La Recherche en informatique, p.64

connaissances: aucun mécanisme ou modèle de raisonnement ne peut être séparé des connaissances sur lesquelles il opère (46).

Un mode de raisonnement associe deux notions imbriquées:

- la structure des données pour représenter l'information;
- la méthode associée d'exploitation de cette information, ou de raisonnement;

Le mécanisme de raisonnement permet de découvrir dynamiquement de nouvelles informations et connaissances sur le problème traité.

- Le raisonnement formel: fondé sur la manipulation syntaxique de structures symboliques à l'aide de règles, dans le cadre d'une sémantique.
- le raisonnement procédural: toutes les connaissances, la façon de les utiliser et la conduite du raisonnement sont entièrement figées sous formes d'algorithmes.
- le raisonnement par analogie: très naturel et efficace chez l'homme, mais difficile à mettre en oeuvre car peu connu.
Une solution est le raisonnement par cas: résoudre un problème nouveau en se fondant sur la solution d'un problème similaire.
- le raisonnement par généralisation et abstraction: lié à l'apprentissage par induction.
Ce raisonnement peut utiliser le mécanisme d'inférence par héritage et se rattache directement au raisonnement par classification.

Le problème central de la mécanisation d'un raisonnement est de parvenir à mener une suite d'opérations sur des structures symboliques, tout en préservant la véracité et la cohérence des conclusions déduites (47).

(48) La Recherche en informatique, p.70

(49) Lévy,P.: Les Technologies de l'intelligence,
p.175

L'apprentissage

Le mot apprentissage a trois sens en intelligence artificielle:

- acquisition des connaissances par l'homme: il concerne l'enseignement assisté par ordinateur qui s'occupe de la création et de l'utilisation de modèles cognitifs d'un apprenant humain, dans le but de l'aider à apprendre;
- acquisition des connaissances par la machine: l'homme aide la machine à acquérir des connaissances. Ce domaine est basé sur les interfaces conviviales permettant à l'expert d'exprimer sa connaissance dans un formalisme compréhensible par la machine;
- auto-acquisition des connaissances de la machine: corrections progressives des données (48).

Conclusion

L'intelligence artificielle est encore loin de pouvoir rendre compte de la grande complexité de l'activité cognitive humaine.

Cependant, le travail de modélisation qu'elle a entrepris des structures cognitives de l'homme permet de comprendre comment des pouvoirs d'abstraction et de raisonnement formel se sont développés au sein de notre espèce (49).

- (50) Le Moigne, J.L.: La Modélisation des systèmes complexes, p.4
- (51) Le Moigne, J.L.: La Modélisation des systèmes complexes, p.11

CONCLUSION

Cette présentation de l'analyse des structures cognitives de l'homme proposées par les sciences cognitives, de la perception au formalisme du raisonnement, s'est voulue dans un premier temps descriptive. Ceci afin de mettre en lumière et clarifier les différentes démarches et leurs postulats.

C'est à l'ensemble de cette activité scientifique que la documentation fait appel afin d'affronter le problème du transfert de l'information.

Une préoccupation commune est au centre de toute cette activité:

La notion de modélisation

Modéliser signifie élaborer et concevoir des modèles: constructions symboliques susceptibles de rendre intelligible un phénomène perçu complexe, et d'amplifier le raisonnement de l'acteur projetant une intervention délibérée au sein du phénomène (50).

Deux façons d'aborder un système complexe pour le comprendre et donc lui donner du sens:

- le simplifier afin de découvrir son intelligibilité et lui donner une explication.
Mais en simplifiant il y a risque de le mutiler et donc d'empêcher son intelligibilité (Bachelard).
- le modéliser pour construire son intelligibilité et donc sa compréhension (Valéry) (51).

Méthode de modélisation

- Méthode du "modèle précédent": raisonnement par analogie;

- (52) Le Moigne, J.L.: La Modélisation des systèmes complexes, p.36
- (53) Le Moigne, J.L.: La Modélisation des systèmes complexes, p.38
- (54) Le Moigne, J.L.: La Modélisation des systèmes complexes, p.65
- (55) Le Moigne, J.L.: La Modélisation des systèmes complexes, p.15

- méthode analytique: hypothético-déductive, basée sur la logique formelle.

Trois axiomes:

- d'identité
- de non contradiction
- du tiers-exclu

- méthode systémique: axiomatico-inférentielle, basée sur la logique conjonctive.

Trois axiomes:

- d'opérationnalité téléologique ou de synchronicité
- d'irréversibilité téléologique ou de diachronicité
- d'inséparabilité ou de récursivité (52).

Pour représenter un système complexe, on le représente comme et par un système, système assez général et stable pour permettre de rendre compte de tous les types de complexités à considérer: s'établie par la construction de la forme canonique du système général (53).

La notion de récursivité postule que l'action de modéliser n'est pas neutre et que la représentation du phénomène s'établie dans une interaction entre le système modélisé et le système modélisant (54).

Modéliser, c'est donc identifier et formuler des problèmes en construisant des énoncés, et chercher de résoudre ces problèmes en raisonnant par simulation.

Modèle: représentation intelligible artificielles, symbolique, des situations dans lesquelles nous intervenons (55).

IIème PARTIE

**ETUDE DES PROCESSUS DU FONCTIONNEMENT COGNITIF ET DU
TRAITEMENT DE L'INFORMATION**

(56) Richard, J.F.: Les Activités mentales, p.11

INTRODUCTION

Cette étude a pour objectif de décrire les différents éléments qui constituent le système cognitif et leurs relations, en tant que "architecture cognitive" (56), c'est à dire l'ensemble des fonctions qui permettent la construction des représentations et de la connaissance.

Je m'attacherai à décrire ces fonctions dans leur ensemble, c'est à dire comme un système intégré de traitement de l'information (Simon). Ce qui correspond davantage à la relation qui s'établit dans un centre de documentation que le découpage en secteurs proposé par les manuels de psychologie cognitive (perception, mémoire, compréhension,...).

En effet, la construction des représentations est le résultat de l'interaction de toutes ces fonctions; c'est donc cette interaction que le documentaliste doit pouvoir analyser et modéliser, et non chacune des fonctions en soi.

La "relation d'information" dans un centre de documentation

On retiendra ici comme définition générale que l'information est un processus de communication de la connaissance.

Un transfert d'information suppose:

- un contenu: la signification;
- un processus: l'établissement d'une relation.

La notion de signification fait intervenir trois niveaux:

- l'utilisateur: son savoir, élaboré au cours des années d'études et de recherche, intègre toute nouvelle connaissance dans une structure logique.

(57) Marino, J.B.: Utilisation de la théorie
mathématique de la
communication, p.38

La signification du contenu d'une connaissance correspond à tout un réseau complexe de liens avec la structure logique. On peut parler d'assimilation;

- le documentaliste: le savoir est compris comme un assemblage de symboles compatibles avec les langages documentaires. Ce savoir pourrait être représenté par une géographie personnelle de champs sémantiques symbolisés par des clés où tout savoir de toute discipline peut trouver une place, une fois reconditionné. Les clés de langages documentaires sont autant d'unités distinctes véhiculant la signification. On peut parler de rangement intelligent;
- le système d'information: l'accès à l'information nécessite l'expression d'un savoir intelligible pour le système. On peut parler de rangement commandé(57).

La compréhension et l'acquisition de la connaissance se ramènent donc à l'établissement d'une correspondance.

Cela suppose nécessairement la compatibilité des trois acteurs du processus de signification:

- le documentaliste et le système d'information doivent être en symbiose. Nous considérerons ici que l'assimilation des représentations du système par le documentaliste est acquise.
- le documentaliste et l'utilisateur gagnent à connaître chacun les règles du jeu de l'autre afin de progresser dans une interaction constante.

Je ne m'intéresserai ici à la problématique largement débattue de l'adéquation entre la demande et l'indexation des documents que dans la mesure où ces travaux mettent en lumière les structures cognitives qui sous-tendent les représentations.

Je m'attacherai plutôt à essayer de mettre en lumière les moyens à disposition du documentaliste afin d'accéder aux représentations de l'utilisateur.

DES CARTES OBJECTIVES ET DES PAYSAGES SUBJECTIFS

Ce chapitre est directement issu de la lecture d'une série d'articles de Bertram C. Brookes (1980) dans lesquels l'auteur analyse l'enjeu qui se trouve à la base des sciences de l'information.

Le postulat ontologique de Popper

L'hypothèse ontologique de Popper se présente sous la forme de trois mondes :

- le monde 1: le monde physique des objets, l'univers;
- le monde 2: le monde de la connaissance subjective et des représentations;
- le monde 3: le monde de la connaissance objective, c'est à dire le savoir de l'humanité rendu tel grâce à son accessibilité et à sa diffusion.

C'est par son enregistrement que le savoir humain atteint un degré d'universalité et d'objectivité, refusé à la connaissance subjective.

Ces trois mondes évoluent indépendamment mais dans une interaction permanente.

Le documentaliste travaille à recueillir et organiser les objets du monde 3 afin de permettre leur utilisation.

Une orientation théorique consiste à décrire et expliquer ces interactions afin d'aider à l'organisation de la connaissance plutôt que des documents, pour une utilisation plus effective.

La notion de dimension

Les trois mondes évoluent dans des dimensions différentes, chacune de ces dimensions étant elle-même susceptible d'évoluer dans des espaces multiples.

L'ensemble de ces trois mondes évolue dans des espaces régies par les représentations mentales. Les représentations des mondes 1 et 3 sont liées à une époque et une culture et donc sont admises et assimilées par tous.

Mais le monde 2 est plus complexe. Composé des représentations subjectives, il n'existe pas un espace spécifique à ce monde, mais une multitude d'espaces. Chaque représentation est en elle-même une dimension, et une dimension unique.

Représentations objectives et représentations subjectives

La question alors se pose de savoir s'il est possible d'accéder à une "science du monde 2".

Il est possible de montrer que les représentations ont des caractéristiques communes:

- une structure cognitive;
- des intérêts communs.

Ces intérêts communs pourraient être les objets du monde 3: l'étude des interactions entre le monde 2 et 3, telles qu'on les retrouve dans le monde 3.

L'observation de la façon dont la connaissance est enregistrée et se développe sur une période donnée dans un domaine spécifique pourrait être un moyen d'étude.

Une représentation subjective ne devient objective que dans la mesure où elle est exprimée et enregistrée dans le monde 3. Elle devient ainsi accessible à d'autres.

Mais en réalité, cette connaissance enregistrée est loin d'être immédiate: elle est perçue à travers d'autres représentations qui la filtrent et qui à leur tour se trouvent remises en cause par cette connaissance.

- (58) Brookes, B.C.: The foundations of Information Science: part 1, p.131
- (59) Brookes, B.C.: The foundations of Information Science: part II, p.209

En se référant à l'équation vue plus haut :

$$K(S) + \Delta I = K(S + \Delta S)$$

on peut dire qu'une donnée reçue doit être subjectivement interprétée par la structure cognitive pour pouvoir devenir information.

L'absorption d'information dans une structure cognitive ne provoque pas de simples additions mais des réajustements qui impliquent de nouvelles relations entre les concepts ou de nouveaux concepts.

On peut donc conclure que l'organisation de la connaissance scientifique est un reflet de la façon dont sont construites les représentations (58).

Un processus d'objectivisation

Chaque représentation étant unique, un effort social constant est nécessaire afin de permettre une entente et un discours scientifique compréhensible.

La croissance de la connaissance est un processus d'objectivisation des représentations subjectives face aux stimuli que l'on reçoit (59).

La science ne se développe pas par simple accumulation de faits, mais par la création d'un ensemble de théories qui agissent réciproquement les unes sur les autres.

Quand une nouvelle théorie atteint un consensus scientifique, de nombreux faits peuvent être réinterprétés dans les termes de la nouvelle théorie.

Le développement de la science est donc un processus qui implique à la fois accroissement et restructuration.

Accès à la connaissance

Les systèmes informatifs qui proposent des références bibliographiques présentent un côté d'inefficacité dans la mesure où chaque document doit être lu, même ceux qui ne répondent pas au besoin d'information spécifié.

(60) Brookes, B.C.: The foundations of Information
Science: part III, p.269

Mais ils offrent un avantage important: l'information extraite est directement associée à un document, ce qui permet à l'utilisateur de situer l'information dans son contexte.

Ainsi, toute information doit être présentée conditionnée par la structure cognitive qui l'a produite, de façon à ce que l'utilisateur puisse juger de sa validité et de son autorité.

Les cartes de distribution de l'activité scientifique dans un domaine donné proposées par la bibliométrie sont un moyen de présenter l'information dans un contexte précis et donc de faciliter l'interprétation de cette information dès sa réception.

Prendre en considération le fait que l'information (partie de la connaissance) et la connaissance (structures cohérentes d'informations) existent seulement dans une dimension cognitive implique une prise en considération majeure de l'individualité et une tentative d'associer la recherche cognitive au domaine de la documentation (60).

(61) Astolfi, J.P.: La Didactique des sciences,
p.10

ELEMENTS DE DIDACTIQUE

La didactique, bien que liée au domaine de l'enseignement, peut être fort utile ici, dans son approche des situations d'appropriation de savoir: études des représentations des élèves, des modes de raisonnement (61).

Les concepts de la didactique

L'enseignement est confronté à l'importance des pré-savoirs structurés pour l'assimilation scientifique:

- variété infinie des représentations possibles;
- lenteur des transformations cognitives réelles au-delà des apprentissages et évaluations à court terme;
- la coexistence de plusieurs systèmes parallèles d'interprétation, différemment mobilisés selon les contextes et les situations.

Le concept de représentation

la "cartographie" des représentations

On recueille par des questionnaires, entretiens, dessins, etc., les représentations d'individus sur un concept et un niveau donnés. On détermine ainsi une sorte de "géographie pré-notionnelle".

Mais les modalités de questionnement présupposent de manière trop forte le caractère stable et invariant des représentations:

- les représentations sont d'abord des stratégies cognitives en réponse à un problème. Les réponses obtenues doivent donc toujours être rapportées à leur contexte de production;

(62) Astolfi, J.P.: La Didactique des sciences,
p.31

- Une réponse donnée à une question est toujours simultanément une réponse à l'interrogateur. Elle cherche inévitablement à se situer par rapport à des attentes supposés de celui-ci, et par rapport à une image de soi que l'on souhaite donner;
- toute interprétation intègre les cadres conceptuels de l'observateur: le repérage et la sélection d'un élément qualifié de représentation implique une hypothèse de représentation.

Origines des représentations

- une orientation psycho-génétique: relative au développement de l'intelligence de l'enfant selon Piaget;
- une orientation historique: correspondance entre des représentations actuelles et l'état de la science à une période donnée;
- une orientation sociologique: réappropriation sociale des concepts scientifiques, selon Moscovici;
- une orientation psychanalytique: caractère surdéterminé de certaines représentations relevant du travail de l'inconscient (62).

Prises en compte didactique des représentations

Il s'agit d'explorer toutes les voies qui permettent de prendre appui sur les représentations des sujets afin de favoriser leur dépassement.

L'enjeu de l'enseignement est de s'assurer que les nouvelles connaissances soient bien assimilées. Les erreurs peuvent naître d'une mauvaise assimilation, mais aussi d'une incapacité de la structure cognitive à encoder correctement les informations.

L'enseignant a plusieurs moyens à sa disposition pour pallier à cela:

(63) Capper, J.: Cognitive science could unlock the
Secrets of how
we learn, p.38

- procéder lentement à chaque nouveau concept;
- relier chaque nouvelle information au bagage précédemment acquis.

L'importance d'une participation active du sujet à été soulevée pour mettre en lumière la notion de métacognition: c'est à dire les stratégies utilisées par le sujet pour guider sa pensée et sa compréhension.

Cela implique une auto-évaluation de ses propres capacités, connaissances, inclinations et motivations. Cela implique aussi une évaluation du degré de capacité nécessaire à l'accomplissement d'une tâche.

La liste des stratégies de métacognition est longue et les activités qui impliquent une métacognition sont nombreuses:

- focaliser son attention sur la tâche;
- s'assurer que la tâche à effectuer est clairement comprise avant de poursuivre;
- découper la tâche à effectuer en de petites unités;
- séparer les informations pertinentes des informations non pertinentes;
- relier les informations nouvelles à ce que l'on sait déjà;
- sélectionner et ordonner les actions retenues;
- focaliser et évaluer les progrès; revenir dessus si c'est nécessaire (63).

Une médiation documentaire

La fonction du documentaliste ne consiste pas à transmettre un savoir mais plutôt à créer les conditions d'une construction autonome de ce savoir par le sujet lui-même.

On peut définir la médiation documentaire comme une didactique de la médiation qui conduit à un savoir

(64) Alava: Pour une didactique de la
documentation,
p.14

construit au travers d'un média et favorise la construction d'un savoir sur ce média (64). Le sujet doit donc à la fois construire son savoir et construire les compétences pour accéder à celui-ci, c'est à dire analyser la structure et la mise en forme de la connaissance.

Le documentaliste construit sa didactique dans deux directions:

- interaction avec le savoir général requis par les fonctions du sujet dans une organisation;
- interaction avec les représentations du sujet face à une demande spécifique.

Une confrontation croisée est nécessaire entre ces différents pôles afin de permettre une évaluation des actions entreprises.

La prise de connaissance du savoir préexistant du sujet se fait par l'intermédiaire du langage, toujours au centre du concept de communication.

DIMENSIONS PRAGMATIQUES DE LA RELATION D'INFORMATION

L'accès aux représentations des connaissances de l'utilisateur ne peut se faire qu'à travers le langage.

Une demande d'information se fait par écrit ou à l'oral. Le rôle du documentaliste est d'aller au delà des mots pour remonter à la structure cognitive de l'utilisateur et à ses représentations.

Cela ne peut se faire que par l'analyse de l'acte d'énonciation produit dans une situation de communication.

Les actes de discours

Extraction des connaissances

Nous ne disposons pas à chaque instant de tout le contenu de nos connaissances. Cela sous-entend une opposition entre le contenu immédiat de la conscience et l'ensemble des connaissances fixées par notre cerveau.

L'introspection nous renseigne sur la nécessité fréquente d'un effort pour faire parvenir à la conscience des données de connaissance manifestement acquises.

Stimulations d'interfaces et données cognitives élémentaires ne sont pas immédiatement utilisables. Elles doivent auparavant être confrontées les unes aux autres et être intégrées dans les systèmes de connaissances. Bien que la confrontation des données cognitives soit en partie spontanée, cela est insuffisant pour construire un système cognitif très élaboré.

Ainsi s'explique l'importance du discours, moyen privilégié de mises en formes des données cognitives élémentaires venant des stimulations d'interface.

Le discours, intérieur ou extérieur, est déterminant dans l'élaboration des systèmes cognitifs, autorisant une étendue et une complexité que le maniement direct des impressions perceptives ne permettrait pas.

Le discours joue également un rôle essentiel dans la transmission cognitive.

Connaissance totale et connaissance transmise ne se superposent pas intégralement. La transmission a ses propres exigences:

- les émotions, sentiments et impressions sont intransmissibles;
- toute connaissance s'inscrit dans une culture.

Le discours est un moyen commode d'organisation et de transmission de données qui ont une existence propre, non verbale.

Se pose alors la question de la nature des relations entre discours et connaissances et l'analyse révèle que ces relations sont complexes et conflictuelles.

Indispensable à l'optimisation de la connaissance, le discours est aussi l'occasion de biais dans la mise en forme des systèmes cognitifs comme dans leur transmission.

Le concept et son signifiant apparaissent comme des fonctions algébriques, c'est à dire comme des entités associant des relations et des constantes, mais aussi des variables pouvant prendre une infinité de valeurs:

- le mot isolé a une signification imprécise. Cette signification ne s'enrichit qu'à partir de contextes;
- un message verbal supposé complet acquiert un sens plus précis dans le contexte cognitif global de l'orateur ou de l'auditeur.

Le discours n'est pas fondement de la connaissance (Platon), il serait plutôt un ensemble de mécanismes pouvant traiter les données d'interfaces et constituant un véritable système autonome (Andreewsky).

Il est pourvu de lois internes qui lui fournissent une cohérence propre.

Chaque mot présente un accroissement de signification du fait de tous les autres mots du message.

Le sens du discours est la forme d'équilibre qui supprime les contradictions entre les différents composants.

(65) Systémique et cognition, p.84

Mais ce sens est toujours recommencé car l'introduction de n'importe quelle donnée extérieure nouvelle suffit à rompre un équilibre précaire.

Lorsqu'un discours qui traduit une activité mentale originale est transmis à un interlocuteur, ni les concepts ni les étiquettes qui les désignent n'ont obligatoirement une signification immédiate et identique pour cet interlocuteur.

Il peut pressentir un contenu cognitif mais ne peut le définir comme l'orateur puisqu'il n'a pas lui-même effectué les démarches d'abstraction.

L'auditeur est donc contraint d'intégrer le message verbal présenté dans son système cognitif, avec les risques d'erreur que cela présente (65).

Compréhension du discours

C'est à l'auditeur d'identifier l'acte que le locuteur vise à accomplir par son énonciation.

Les actes de langages peuvent être rapportés à quatre classes principales (Austin, Searle):

- les régulations de l'échange (Jakobson): ensemble d'énoncés qui manifestent simplement qu'une interaction sociale s'instaure;
- les échanges d'informations: importance de faire une distinction entre questions ouvertes et questions fermées (Clark):
 - question fermée: la réponse est confirmation ou dénégation;
 - question ouverte: la réponse peut prendre des formes déclaratives variées (commentaire, récit, exposé argumentatif);
- les engagements: requête, autorisation, interdiction, qui correspondent à une promesse ou un refus;
- les appréciations: interactions entre le locuteur et le partenaire.

Une offre d'informations ne répond pas nécessairement à une demande explicite. On peut croire cependant,

qu'une demande au moins implicite est généralement attribuée par le locuteur à son partenaire (Hupet, Costermans).

Les déclarations sont souvent assorties de modalisateurs exprimant le degré de certitude que le locuteur attribue à son propre énoncé.

Une autre catégorie, celle des "transgressions", peut donner naissance à des interactions particulières:

- poser une question dont la réponse est déjà connue afin de contrôler l'état de connaissance du sujet;
- considérer comme mensonge une déclaration que l'on croit fausse;

Divers procédés lexicaux et syntaxiques permettent au locuteur de présenter certaines informations comme supposées connues et d'autres supposées nouvelles pour le destinataire.

De son côté celui-ci se révèle capable d'exploiter ces indications et de réserver à ces deux types d'informations des traitements spécifiques.

Pour assurer une bonne communication, le locuteur doit construire son discours selon deux principes:

- l'information présentée comme connue doit pouvoir être univoquement repérée et recouverte en mémoire par l'auditeur;
- l'information présentée comme nouvelle ne doit pas lui être déjà connue et doit pouvoir lui être efficacement intégrée (Clark).

Une fois identifiée, l'information connue est utilisée comme une "adresse" dans la structure sémantique en mémoire, à laquelle l'information nouvelle est encodée.

Les processus d'intégration de l'information peut être influencé par l'ordre des phrases (Hupet; Le Bouédec).

Le langage est ainsi fait et utilisé qu'il préside à un transfert de représentations alors même que les énoncés grâce auxquels s'opère ce transfert n'encode qu'une partie seulement de ces représentations (Greenfield).

Ainsi, pour déterminer ce qu'il énonce et la façon dont il l'énonce, le locuteur prend en compte un

(66) Problèmes de psycholinguistique, p.163

(67) Problèmes de psycholinguistique, p.116

certains nombres de paramètres caractérisant la situation énonciative.

Le discours dépend des présuppositions ou ensemble de conjonctures que le locuteur fait à propos du destinataire.

Le succès de l'échange reposera pour une bonne part sur la justesse de ces conjonctures et par conséquence, sur la coopération que le destinataire offrira.

Cette coopération prend un relief particulier dans la conversation, puisque les partenaires y sont à la fois locuteurs et auditeurs (66).

La coopération dans le discours

Quatre principes de coopération ont été précisés (Grice):

- quantité: donner autant d'informations qu'il est requis, mais pas davantage;
- qualité: n'asserter que ce que l'on croit vrai;
- relation: donner des informations qui soient en relation avec le thème traité;
- modalité: s'exprimer sans ambiguïté.

C'est au destinataire de juger de la quantité et de la qualité des informations reçues: lui seul peut savoir si ces impératifs sont rencontrés.

Comment le locuteur saurait-il ce qu'il en est? : par un échange constant avec le destinataire qui permettra d'acquérir une connaissance mutuelle et d'établir un "common ground" (Clark, Carlson), c'est à dire l'ensemble des connaissances, croyances et suppositions que les interlocuteurs partagent au moment de l'échange d'informations.

Pour fonder strictement un terrain commun, il faudrait vérifier une série infinie de conditions.

Cette connaissance mutuelle est une représentation mentale inférée à partir de certains types de données grâce à certaines démarches heuristiques (Clark; Marshall) (67).

(68) Problèmes de psycholinguistique, p.263

Approche différentielle du langage

L'approche différentielle s'applique à décrire les différences individuelles de façon intelligible et proposer des modèles pouvant donner un sens aux constats différentiels.

Les travaux dans ce domaine, encore peu battu, montrent que les différences en matière de langage sont stables dans le temps. Les variations observées dans le degré de stabilité seraient dues à des fluctuations temporelles propres à l'aspect étudié ou à l'imperfection des moyens de description utilisés (Malrieu).

Ils ont aussi montré que ces différences sont cohérentes.

L'analyse factorielle permet d'exploiter de façon plus systématique les cohérences interindividuelles observables en conduites langagières.

La cohérence des différences individuelles observées sur un ensemble d'épreuves verbales suggère qu'il y a un sens à considérer le langage comme une fonction intégrée, unitaire, dont on peut se demander quelles sont les modalités d'intervention dans le fonctionnement cognitif (68).

Deux approches peuvent être envisagées:

Le premier courant met en évidence l'existence de deux modalités de représentation: l'une utilisant le langage, l'autre des images indépendantes du langage. Des études ont montré l'emploi préférentiel de l'une ou l'autre de ces représentations selon les sujets (Paivio). Une série de tests sur l'emploi de procédures verbales ou visuelles de mémorisation et de compréhension ont montré comment les sujets se répartissent sur un axe verbal-iconique.

Les modèles issus de ces études permettent d'établir les temps de compréhension et de vérification employés en fonction du sujet et du matériel présenté (MacLeod; Hunt; Mathews).

Ces études suggèrent que dans les tâches cognitives un certain degré de vicariance peut s'établir entre processus verbaux et processus non verbaux de codage et traitement de l'information. Des différences entre sujets sont mises en évidence de façon cohérente et donc prédictible quant à la facilité avec laquelle sont évoqués les uns et les autres.

(69) Problèmes de psycholinguistique, p.271

Un second courant s'attache à mettre en évidence deux usages différents des propriétés du langage (Reushlin):

- usage pragmatique des propriétés sémantiques du langage: moyen de coder de l'information sur les référents ou de recevoir sur eux une information codée.
- usage formel des propriétés syntaxiques du langage: moyen d'évoquer des informations sur les référents, mais aussi de connaître certaines relations entre ces informations sans utilisation nécessaire de connaissances extérieures au message.

Les différences individuelles quant à l'usage préférentiel de procédés cognitifs verbaux ou non verbaux, quant à l'habitude d'utiliser les propriétés du langage de façon partielle ou complète, constituent des variations vraisemblables à cette source commune de variation se manifestant dans tout usage du langage (69).

Les études sur le rapport entre langage et mémoire sémantique montre que la vitesse à laquelle un sujet peut retrouver dans sa mémoire sémantique la signification attachée à un stimulus paraît être un facteur important des performances linguistiques globales dont le sujet est capable (Hunt; Davidson; Lansman).

L'acquisition des connaissances

Rôle des connaissances antérieures

Comme nous l'avons déjà vu dans la première partie, les connaissances en mémoire constituent un cadre assimilateur pour les informations nouvelles.

Dans une perspective de réseau sémantique (Anderson), l'acquisition de connaissances se fait par création de nouveaux arcs ou par création de nouveaux noeuds dans le réseau.

L'approche basée sur la notion de niveaux de traitement se base sur l'hypothèse qu'une information traitée à un niveau sémantique est mieux retenue

(70) Richard, J.F.: les Activités mentales, p.171

(71) Le Ny, J.F.: Science cognitive et
compréhension du langage,
p.202

qu'une information traitée à un niveau superficiel (phonétique ou orthographique).

Cette approche insiste sur le rôle des processus cognitifs qui consistent à relier l'information donnée à des informations en mémoire et qui sont d'autant plus importants que l'on est plus expert dans le domaine.

Cette approche met l'accent sur l'idée que l'accès en mémoire n'est pas automatique et peut en particulier requérir des inférences dont la mise en oeuvre dépend de la représentation que le sujet se fait de la situation (Craik; Lockart).

La théorie de l'activation (Anderson) admet que l'utilisation d'un schéma dépend de son accessibilité et que cette dernière est fonction de la fréquence et de la récence des activations dont il a fait l'objet. Ainsi, la mémorisation d'une information nouvelle congruente avec ce schéma sera facilitée.

Parallèlement, le temps de reconnaissance d'une information sera une fonction croissante du nombre d'informations correspondant à d'autres arcs partant du même noeud activé.

On peut construire ainsi une courbe à maximum (70).

Les modèles d'activations reposent sur trois idées fondamentales (71):

- celle du niveau d'activation d'une unité cognitive à un instant donné;
- celle d'activation proprement dite, c'est à dire de modification du niveau d'activation;
- celle de structure cognitive qui véhicule les activations.

Activation des connaissances

Une connaissance peut être examinée sous deux aspects:

- une dimension généralité-spécificité: elle peut être exprimée sous une forme particularisée ou sous une forme générale;
- une dimension description formelle-utilisation: elle peut être présentée sous ses aspects relationnels (concepts en jeu et leurs relations) ou sous

(72) Richard, J, F, : les Activités mentales, p.174

(73) Richard, J, F, : les Activités mentales, p.305

l'angle de l'utilisation qui peut en être faite dans le cadre de la réalisation d'objectifs d'action dans des contextes spécifiques.

La signification d'un concept englobe les situations auxquelles il s'applique et ce sont ces dernières qui lui donnent du sens (Vergnaud), mais dans le même temps, ces situations donnent du sens au concept parce qu'elles servent à le construire.

La fonction des énoncés généraux est d'établir les connexions nécessaires avec les concepts existants et d'aider à la généralisation en fournissant une expression qui donne un support au changement des constantes en variables.

Ces deux fonctions ne peuvent être remplies que si les énoncés généraux ont été précédés par des particularisations (Vezin) (72).

Contrôle des représentations

Dans le cadre de la réalisation d'une tâche ou de la résolution d'un problème, un résultat positif est le meilleur moyen d'apprécier la justesse de la représentation et son adaptation au contexte.

La résolution d'un problème implique que le sujet prenne conscience d'avoir épuisé l'espace de recherche. Ce sentiment est produit par le fait que le sujet est capable de relier la situation présente à son expérience passée.

Agir sur les hypothèses de l'expérience entraîne une modification des représentations de l'expérience.

Cela implique donc un changement dans les savoir-faire et un changement dans les significations.

L'expérience passée n'agit pas de façon automatique, il faut une orientation de pensée pour organiser l'expérience passée, pour sélectionner ce qui est utile pour la situation présente (Ellen; Weisberg; Alba).

On peut distinguer trois moyens d'évaluation des résultats (Richard) (73):

- évaluation de l'écart au but: appliquée dans le cas de problèmes numériques quand la recherche de la solution se fait en formulant des hypothèses;

- évaluation de l'adéquation de la solution trouvée: vérifier que la solution soit compatible avec les données du problème;
- reconnaissance de situations critiques et réflexion sur l'action: la confrontation à une situation d'incident ou d'impasse est le signe d'une erreur de procédure. Une réflexion sur l'action peut remettre en cause les représentations à la base de l'action.

L'APPORT DE L'INTELLIGENCE ARTIFICIELLE: LE DIALOGUE HOMME-MACHINE

L'action de l'intelligence artificielle dans le domaine des sciences de l'information est fondamentale.

Les questions relatives à un rôle actif des systèmes informatiques dans le processus de transfert de l'information a trouvé un terrain fertile dans ce domaine.

Objectifs du dialogue homme-machine

Les objectifs de l'intelligence artificielle peuvent s'exprimer ainsi:

- acquisition des connaissances;
- représentation des connaissances;
- élaboration de connaissances.

L'objectif théorique qui sous-tend l'ensemble des buts est la simulation parfaite des structures cognitives de l'homme.

Dans son processus de modélisation ce domaine se heurte à des problèmes spécifiques: compréhension du langage, représentation réciproque du système et de l'utilisateur, incidence de la situation sur l'échange, exigences pratiques de l'ergonomie.

Capacité de dialogue d'un système

- l'objectivité: les interventions du système utilisent une base de connaissances qui représente un domaine sémantique déterminé; établie un rapport entre cette communication et un domaine sémantiquement présenté;
- la textualité: le système doit maintenir la connexité entre ses interventions et celles de l'utilisateur, ainsi que la cohésion sémantique; elle assure la communication homme-machine;

(74) Rastier, F.: Sémantique et recherches
cognitives,
p.169

- l'apprentissage: le système doit pouvoir intégrer au moins temporairement dans sa base de connaissance des données contenues dans les propos de l'utilisateur; il établie une action homme-machine;
- le questionnement: le système doit pouvoir requérir les données, qui lui manque pour l'élaboration d'une réponse; il établie une action machine-homme, suivie éventuellement de sa converse;
- la rectification: le système peut suggérer des rectifications si l'utilisateur fait des erreurs en formulant sa requête; elle établie une action machine-homme;
- l'explicitation: le système doit pouvoir explicité sa réponse si l'utilisateur n'est pas satisfait. elleétablie une action machine-machine, suivie d'une action machine-homme.

Un système de dialogue évolué doit pouvoir ainsi formuler trois sortes de propos:

- sur ses propres connaissances;
- sur les propos de l'utilisateur;
- sur ses propres propos;

C'est dans le caractère réflexif de ces trois sortes de propos que résident ses facultés d'auto-adaptation et par suite, l'intérêt de la recherche cognitive (74).

Représentations de l'utilisateur et la machine

Le dialogue homme-machine reste très superficiel, dans la mesure où la machine n'a pas la souplesse d'une interaction interhumaine.

Il manque au système de nombreux éléments comme prendre l'initiative d'ouvrir ou clore le dialogue, donner des exemples, se contredire, mentir, négocier ou comprendre les sens multiples.

Faiblement socialisée, la communication homme-machine est fort simplifiée sous ce rapport, car les stratégies discursives n'y ont pas d'enjeu social direct.

Il n'existe pas d'utilisateur type, mais chaque application a ses utilisateurs potentiels, lesquels n'ont pas toujours de représentations précises du système utilisé.

Il devient alors nécessaire que le système parvienne à se représenter son utilisateur du moment afin d'accroître son efficacité, et d'autre part, qu'il se représente à lui en devenant aussi explicite que possible.

La pertinence est au coeur de cette communication. L'utilisateur n'adresse au système que des requêtes pertinentes, ou du moins dont il peut raisonnablement attendre une réponse.

Pour ne pas perdre de temps, il se montre coopératif et restreint le champ du dialogue au domaine sémantique représenté dans le système, pour autant qu'il le connaisse.

Pour satisfaire pleinement l'utilisateur, en s'adaptant à ses particularités, le système doit construire une représentation:

- construire un modèle préconçu de l'utilisateur, universel;
- classer les utilisateurs dans une typologie préconçue:
 - demander à l'utilisateur de se présenter lui-même: peu apprécié;
 - classer l'utilisateur d'après ses premiers propos: discutable.
- construire une représentation ad hoc pour chaque utilisateur: le système plus prometteur consiste à construire une représentation de l'utilisateur qui puisse s'enrichir et évoluer à chacune de ses interventions.
Cela suppose que le système construise un historique du dialogue au fur et à mesure de son déroulement.

D'un point de vue sémantique, les inférences et ajustements que le système opère en posant des

(75) Rastier, F.: Sémantique et recherches
cognitives,
p.172

questions coopératives à l'utilisateur ou en répondant à ses demandes d'explications ont pour effet d'établir certains types d'isotopies entre les propos des deux interlocuteurs et renforcer par là la cohésion textuelle (75).

Ainsi, la pertinence dialogique locale d'une intervention est définie par le type et le degré des isotopies qui la lient à celles qui précèdent. Paradigmatiquement, elle est déterminée au sein du réseau sémantique par la position respective des contenus manifestés dans les interventions des deux interlocuteurs.

La représentation de l'utilisateur permet de situer à quel niveau de profondeur du réseau sémantique les contenus indexés sur l'isotopie doivent être sélectionnés.

La place dans l'arbre des actions réalisées par le système lors de la recherche indique, ce que l'on peut appeler, son niveau de connaissance.

Cela s'établit en fonction de ce que l'utilisateur exprime et des connaissances dont dispose le système.

Les systèmes ne prennent pas en compte les intentions de l'utilisateur: chaque intervention est une demande quelle que soit sa structure.

Les systèmes peuvent prendre en considération les "plans" de l'utilisateur, c'est à dire passer d'un champ à l'autre, mais ils ne construisent pas de représentations de ces plans (Schank; Abelson).

Une nouvelle approche permettrait de prendre en compte les plans de l'utilisateur, et les actes de langages qui en dérivent. Il faudrait pour cela une théorie de l'intention qui soit implantable (Cohen; Allen; Perrault).

Les représentations de l'utilisateur par la machine tendent à une simulation aussi fidèle que possible. Mais cette relation est faussée car il procède d'une conception unilatérale du dialogue.

Elle néglige la représentation de la machine par l'utilisateur, ce qui fausse les conditions d'une simulation.

Des études ont mis en lumière certains faits: quand l'utilisateur s'adresse à la machine, et non à un autre homme, il respecte mieux les contraintes syntaxiques, il respecte mieux les repérages spacio-temporels, il va dans le sens de la concision.

- (76) Rastier, F.: Sémantique et recherches
cognitives,
p.174
- (77) Bryce, L.A.: Cognitive Research in Information
Science, p.7
- (78) Bryce, L.A.: Cognitive Research in Information
Science, p.7
- (79) Bryce, L.A.: Cognitive Research in Information
Science, p.7
- (80) Bryce, L.A.: Cognitive Research in Information
Science, p.10

Ainsi, les difficultés d'interprétation sont moindres dans ce cas là.

En pratique, l'utilisateur sait quand il s'adresse à un homme ou à une machine, et il s'adapte en conséquence (Luzzati) (76).

Systemes documentaires et interfaces

Les sciences cognitives distinguent le savoir, le savoir-faire et les représentations visuelles ou verbales. Un utilisateur serait ainsi connaissable à travers ses représentations du monde, le sujet de ses recherches, le savoir qu'il cherche à compléter et le système documentaire qu'il utilise (Sutcliffe) (77).

L'introduction de modèles de l'utilisateur dans le système permettrait à l'utilisateur de sélectionner le modèle qui semble convenir à son profil de connaissance (Daniels) (78).

La possibilité qu'aurait le système d'informer l'utilisateur de son fonctionnement, permettrait à celui-ci de s'adapter au système (Gilbert) (79).

L'idéal serait une interface qui soit le reflet des représentations de la connaissance de l'utilisateur et de la manière dont elles sont structurées (McDonald, Schaneveld) (80).

Un problème de communication se pose quand les représentations de l'utilisateur sont différentes de celles du système. En particulier, la description du domaine faite par le thésaurus ou le réseau sémantique peut être totalement étrangère à la structuration des connaissances de l'utilisateur. Le système idéal permettrait à l'utilisateur de proposer au système ses propres relations sémantiques et de les encoder au thésaurus.

En théorie, il serait possible d'indexer une banque de données par un nombre important de thésaurus alternatifs, suggérés par différents groupes d'utilisateurs.

Le thésaurus peut servir comme simple modèle statique d'un domaine de connaissance, tandis que des systèmes plus perfectionnés de réseaux sémantiques peuvent amener l'utilisateur à préciser son champ de recherche, à suggérer de nouveaux pôles d'intérêt et à

- (81) Bryce, L.A.: Cognitive Research in Information Science, p.13
- (82) Les Modèles mentaux, p.166

étendre le sujet de recherche à différents domaines (Piekara, Strube) (81).

Les systèmes experts s'appliquent à établir une correspondance entre le modèle de l'expert, le dialogue avec l'utilisateur et le logiciel de résolution de problème.

Les hypertextes permettent à l'utilisateur de naviguer dans la base de données en suivant librement les associations de concepts.

L'idée centrale de cette théorie de l'interaction est de considérer qu'élaborer ou comprendre une métaphore ou une analogie implique un changement de représentation de l'objet de la description.

Les représentations de la machine et de l'homme ont chacune une représentation stable et conventionnelle en mémoire à long terme: les représentations-types (LeNy). Ces représentations n'ont pas forcément une ressemblance apparente.

En juxtaposant ces deux représentations au moment du traitement, une représentation occurence est créée, correspondant à la représentation qui sous-tend la demande: la représentation du système est interprétée dans les termes de la recherche.

Ainsi, la création de ressemblance consiste à créer une nouvelle représentation de l'objet (82).

La notion sous-jacente à cette activité d'interaction est le concept de feedback: c'est à dire, donner la possibilité à l'utilisateur aussi bien qu'à la machine d'avoir une part active dans l'élaboration des stratégies de recherche et leur réévaluation tout au long de l'interrogation.

L'hypermédia et les processus d'apprentissage

Les systèmes experts s'appliquent à la représentation de la connaissance.

La connaissance est apportée par l'expert.

Les didacticiels multimédia se focalise sur la présentation de la connaissance.

La connaissance est apportée aussi par des experts en didactique.

(83) Ipermedia, non paginé

Les technologies éducatives qui se situent à un niveau de méta-apprentissage, s'appuient sur différents domaines pour structurer son discours:

- la psychologie: apporte les modèles d'apprentissage basés sur l'analyse cognitive et comportementale;
- la didactique: apporte les modèles d'enseignement basés sur l'organisation des contenus, les stratégies du processus d'enseignement et les situations qui s'établissent autour de ce processus;
- l'informatique: apporte une méta-connaissance par l'utilisation applicative qui en est fait.

Le paradigme de l'intelligence artificielle ne prend pas en compte l'élément créatif et les sentiments de l'utilisateur.

Ces notions sont au contraire au centre de l'activité d'enseignement.

Du point de vue d'un méta-apprentissage il est important de pouvoir transmettre la connaissance selon deux perspectives:

- sémantique: multimédialité et interactivité;
- syntaxique: réseaux conceptuels et structuration des concepts.

Les conditions efficaces d'un processus d'apprentissage peuvent être définies ainsi (Gagné) (83):

- attirer et contrôler l'attention du sujet: l'attention est provoquée par un stimulus extérieur au sujet. Il faut pour cela que le système enregistre une configuration de l'état d'esprit du sujet pour provoquer un stimulus adéquat;
- informer le sujet des objectifs: le sujet est informé, souvent verbalement, de la connaissance qui doit lui être transmise;
- activation des acquis nécessaires: activer dans la mémoire à court terme du sujet les

(84) Ipermedia, non paginé

connaissances qu'il possède déjà et qui sont indispensables pour la compréhension de la situation en cours;

- proposer les stimulus plus adéquats au processus d'apprentissage;
- guider le sujet: le sujet est accompagné dans son apprentissage par des suggestions, associations et aides, dans la ligne conforme aux objectifs de l'enseignement;
- mise en pratique immédiate par le sujet des connaissances à peine acquises;
- permettre une rétroaction afin que le sujet puisse apprécier son activité;
- provoquer l'assimilation des connaissances par une généralisation des concepts appris;
- permettre une nouvelle utilisation différée dans le temps afin que le sujet puisse juger de son acquis.

L'instrument technologique est inféré sur la base d'une analyse du message.

Les critères d'évaluation sont l'efficacité du système et la motivation du sujet (84).

Ainsi, quatre niveaux d'actions peuvent être définis dans le contexte de l'hypermédia:

- de communication: caractérise la nature du rapport homme-machine;
- morphologique: définit les stratégies d'organisation des informations et donc, les modalités opératives d'accès et d'orientation;
- cognitif: définit les possibilités de correspondance entre les modalités de structuration logique du savoir et les modalités de perception analogique et d'élaboration de réseaux sémantiques;
- didactique: qualifie les stratégies de l'apprentissage en prenant en compte l'autonomie du sujet dans l'action face au contenu du système.

(85) Ipermedia, non paginé

Le problème fondamental n'est pas comment concevoir et proposer l'information, mais comment la concevoir ensemble: c'est à dire que la communication homme-machine ne permette pas l'échange d'informations, mais l'échange d'un discours (85).

MODELISATION COGNITIVE

La modélisation consiste à passer d'une description des processus cognitifs faite dans le langage de la théorie psychologique à une expression dans un langage formalisé qui permet le calcul ou la simulation.

Cela implique de disposer de formalismes adaptés qui n'appauvrissent pas ce que l'on veut exprimer. Cela implique aussi d'avoir explicité la description des processus psychologiques à un niveau de précision suffisant pour que cette description soit complète. L'objectif est que les comportements simulés doivent correspondre aux comportements observés.

La modélisation dans le domaine du traitement de l'information ne peut pas être uniquement formelle. L'introduction de la sémantique dans le formalisme permet d'exprimer les concepts psychologiques et de traduire les représentations que l'on se fait du fonctionnement cognitif dans une situation donnée pour un sujet donné: systèmes de productions, schémas, réseaux sémantiques,...

Les modèles mathématiques sont peu adaptés pour exprimer des contenus de représentation (Richard).

la modélisation statistique

La modélisation statistique repose sur l'option méthodologique selon laquelle le niveau d'analyse pertinent n'est pas celui du détail du protocole, mais celui d'un regroupement des données fait selon des critères choisis pour leur pertinence en regard des processus à décrire.

Les modèles développés sont des modèles mathématiques qui font intervenir des paramètres. Les données pour estimer les paramètres et tester le modèle sont des statistiques calculées sur des groupes de sujets: cela sous-entend que les différences entre les sujets ne sont pas pertinentes.

Un modèle stochastique peut être vu comme un modèle dérivé d'un modèle de traitement de l'information, obtenu en regroupant des états du modèle, des stimulus

(86) Richard, J, F, : les Activités mentales, p.313

que l'on considère comme équivalents, en supposant implicitement les sujets comme équivalents. A chacun de ces regroupements correspond un paramètre (Richard).

Un modèle statistique est justifié si l'on ne peut tester le modèle sans faire de regroupements de données: c'est le cas quand les informations recueillies ne sont pas assez précises pour identifier les états du modèle.

C'est le modèle de traitement complet qui donne son sens au modèle statistique et qui oriente vers le type de données pertinentes pour tester le modèle (86).

Ce traitement implique que pour trouver des régularités il faut se placer à un niveau plus général que le détail des données.

La modélisation des protocoles individuels

Cette modélisation correspond davantage aux exigences d'explicitation de la problématique cognitive.

Il s'agit de donner une interprétation du protocole qui explique par quels processus sont engendrés les différents événements qui le constituent: actions et verbalisations.

Cela suppose que l'on considère comme pertinents tous les événements qui sont intervenus, et qu'on ne se contente pas d'un codage en classes; la prise en compte du déroulement temporel comme donnée à expliquer.

On distingue trois composantes dans les modèles, correspondant à l'architecture cognitive:

- l'état de la représentation et des connaissances sur la tâche: ce niveau concerne les informations qui interviennent dans le choix de l'action ou l'élaboration du contenu de la réponse:
 - informations ponctuelles sur l'état de la situation;
 - connaissances sur les possibilités d'action;
 - objectifs momentanés d'actions;

- règles générales d'action;
- connaissances générales mises en oeuvre dans la situation.
- les processus qui modifient l'état de la représentation ou des connaissances sur la situation:
 - mécanismes de construction de l'interprétation du problème ou de la tâche;
 - règles de production des buts qui engendrent des buts intermédiaires à partir du but général de la tâche et de l'état de la situation;
 - règles d'apprentissage qui permettent de construire de nouvelles connaissances sur la situation, et règles d'évolution qui permettent de changer l'état du système.
- les processus de décision et de contrôle de l'activité:
 - règles de sélection des actions, qui procèdent au choix de l'action à effectuer, compte tenu des informations qui constituent l'état de la représentation;
 - règles d'évaluation des résultats de l'action par confrontation de l'état de la situation et des buts poursuivis;
 - règles d'anticipation permettant de prendre en compte le résultat d'actions envisagées mais non effectuées;
 - règles d'exploration des éléments de la situation;
 - règles qui assurent la gestion des processus et explicitent dans

(87) Richard, J, F, : les Activités mentales, p.317

quelles conditions ils sont mis en oeuvre.

Ce cadre présente l'ensemble des composantes à partir desquelles on peut décrire un modèle (87).

Elaboration d'un protocole individuel

La façon d'analyser un protocole dépend du type de données dont il est composé et de sa structure.

Nature du protocole

- entièrement verbal: ensemble de réponses à des questions.
Mise en évidence de stratégies de raisonnement et de décision;
- exclusivement lié à l'action: ensemble de comportements correspondant à des tâches différentes.
Mise en évidence de raisonnement et de décision;
- les cas mixtes: enrichissement interactif de ces deux modes d'appréhension.
Suite ordonnée d'énoncés ou de comportements observés dans une même tâche et dont la chronologie est une information essentielle.
Ils permettent l'étude des phénomènes d'adaptation et d'apprentissage.

Structure du protocole

La structure du protocole est cruciale pour la définition de l'unité d'analyse.

Dans le cas d'un protocole verbal ou procédural, c'est le changement de question ou de situation qui délimite les unités.

Dans le cas d'une tâche continue, il faut des critères supplémentaires pour délimiter les épisodes:

- interventions définies à l'avance de l'expérimentateur dans certaines conditions;
- prise en considération des pauses et temps longs qui marquent une discontinuité dans la chaîne parlée ou comportementale;
- mise en évidence de comportements caractéristiques,...

La pertinence d'un critère dépend des hypothèses qui sont à la base de l'analyse.
La segmentation dépend donc d'options théoriques qui conditionnent la suite de l'analyse et ses résultats.

La définition d'épisodes permet des analyses statistiques préalables à l'analyse des protocoles individuels: cela permet de mettre en évidence des régularités et fait apparaître des phénomènes qui orientent la recherche des interprétations.

Interprétation du protocole

Cette interprétation consiste à repérer dans la suite des observables la trace de processus généraux qu'on doit retrouver dans d'autres situations et chez d'autres sujets.

Le protocole correspond à la partie observable du processus: pour reconstruire le processus dans son intégralité, il faut inférer des événements internes. L'analyse d'un protocole utilise donc les hypothèses provenant des connaissances que l'on possède sur les processus cognitifs.
Plus on émet d'hypothèses plus l'interprétation du protocole sera satisfaisante.

Le postulat de base est que tout comportement de recherche de solution est cohérent par le fait qu'ils sont déterminés par l'état de la structure cognitive.

Le risque majeur est de ne dégager dans l'analyse d'un protocole que ce qu'on y a mis.
Pour éviter cela, il faut se donner des contraintes:

- cohérence de l'interprétation tout au long du protocole;

(88) Richard, J, F, : les Activités mentales, p.361

- longueur du protocole et niveau d'observation pertinents.
- utilisation de données précédemment acquises.

Dans l'analyse d'un protocole, on s'efforce de remonter des actions aux objectifs, des objectifs aux interprétations qui ont permis de les construire, des interprétations aux connaissances qui les ont engendrées (88).

CONCLUSION

L'accès au monde 2 de Popper signifie pour le documentaliste prendre en compte l'ensemble des éléments qui constituent le processus de fonctionnement de la structure cognitive.

Vouloir agir directement sur la structure cognitive de l'utilisateur signifie devoir prendre en compte de manière exhaustive tous les éléments qui entrent en jeu dans la construction des représentations et leur utilisation.

Si cet objectif peut être atteint dans le contexte de la documentation, il ne pourra s'appliquer en temps réel que sous la forme d'un modèle.

Mais parler de modèle implique une analyse particulièrement fine de la situation d'échange qui s'établit entre le documentaliste et l'utilisateur . De nombreuses données entre en jeu, que l'exposé de ce mémoire a tenté de mettre en lumière.

De la réception de l'utilisateur au transfert de l'information

L'entretien avec l'utilisateur marque ainsi le début d'un long travail d'enquête:

- expliciter les termes de la demande;
- prendre connaissance du contexte opératif de la demande. Cela peut signifier aussi interroger l'organisation ou les collègues;
- prendre connaissance des connaissances de l'utilisateur. Cela signifie expliciter les termes du discours, extraire les connaissances non activées dans la situation présente;
- reconstruire les représentations de l'utilisateur relatives à sa tâche, à ses connaissances;

- recherche des documents dont le contenu serait susceptible d'être immédiatement assimilable par la structure cognitive de l'utilisateur;
- diffusion de l'information sous une forme sensée faciliter son assimilation;
- évaluer cette assimilation: par de nouvelles représentations ou par l'exécution de la tâche;

Cet ensemble d'opérations qui certainement est loin d'être complet soulèvent des interrogations.

La place à accorder à l'influence de l'observateur dans cette fonction d'interprétation n'est pas négligeable.

En effet, il s'agit pour le documentaliste de construire une représentation des représentations de l'utilisateur.

L'emboîtement de ces niveaux d'interprétation implique une situation difficile à gérer, surtout si on y ajoute les propres représentations du documentaliste: représentation du domaine et représentation des représentations de l'utilisateur.

Un moyen pour sortir de ce dilemme serait de s'assurer que le modèle proposé permette aussi l'analyse du fonctionnement cognitif du documentaliste dans son interaction avec l'application du modèle.

Des problèmes d'autre nature doivent être affrontés dans le cadre d'une situation réelle:

- implications matérielles d'un tel projet;
- acceptation par les documentalistes d'une nouvelle dimension de leur profession; et acceptation des entreprises;
- acceptation et coopération de l'utilisateur pour une action qui peut être ressentie comme une ingérence.

Ces dernières réflexions sont le fruit de discussions informelles menées avec des documentalistes et des usagers de systèmes d'information.

IIIème PARTIE

**ETUDE DE CAS: LA BANQUE DE DONNEES BID ET SES
UTILISATEURS**

INTRODUCTION

Cette étude s'est développée dans deux directions:

- l'étude d'un service d'information et des représentations sémantiques qui sous-tendent son organisation.
Mon choix s'est porté sur la banque de données BID dont je fais la présentation plus loin, active dans le domaine des sciences juridiques.
- l'étude du comportement des utilisateurs de cette banque de données .

Je me suis penchée sur la possibilité d'établir une correspondance entre leurs représentations sémantiques et celles du système.

Je remercie ici le centre de recherche et les chercheurs-réalisateurs de BID, pour leur grande disponibilité et leur intérêt pour l'objet de la recherche. L'accès libre à l'interrogation en ligne et à une documentation importante a facilité le déroulement de la recherche.

Je remercie aussi les personnes qui se sont prêtées à l'enquête et ont pris de leur temps pour se pencher avec moi sur la structuration du domaine dans lequel elles travaillent.

Les résultats obtenus ont permis de faire quelques constatations sur les représentations et confirmer ainsi mon propos.

Mais d'autres aspects ont émergé, non négligeables pour éclairer mon propos sous des aspects insolites, mais qui ont rendu parfois cette recherche difficile.

Hypothèses

- Deux structures cognitives produisent nécessairement deux représentations différentes.

- l'accès à ces représentations changent le déroulement de la recherche.

Démarche méthodologique

- choix des acteurs:

- un système informatif spécialisé dans un domaine, crée dans l'optique de répondre aux besoins d'une population bien définie;
- une population d'utilisateurs répondant aux critères suivants: spécialiste d'une des branches du droit procédant lui même à l'interrogation de BID; documentaliste, spécialisé ou non, opérant une recherche pour le compte d'un utilisateur.

- les méthodes:

1) méthodes non-réactives:

- initiation à la terminologie du droit;
- interrogation de BID et utilisation du thésaurus;
- représentations graphiques par les utilisateurs de leurs représentations conceptuelles;
- observations de recherches documentaires et d'entretiens entre documentaliste et utilisateur.

2) méthodes réactives:

- entretiens semi directifs avec les acteurs, comportants des questions fermées pour les données de présentation générale et des questions ouvertes pour l'explicitation des concepts.

Cette étude a été menée sur un nombre très limité d'acteurs, dans un laps de temps très court. Les résultats obtenus n'ont de signification que dans le cadre de l'étude et n'offre qu'un aperçu superficiel de la problématique de ce mémoire.

**BID: BIBLIOGRAFIA INTERNAZIONALE D'INFORMATICA E
DIRITTO**

La banque de données est produite par l'Istituto per la Documentazione Giuridica, centre de recherche du Conseil National de la Recherche (Florence).

L'IDG a pour objectif l'étude et l'utilisation dans le domaine juridique des techniques et méthodologies propres à l'informatique, la logique et l'intelligence artificielle.

Une des activités principales de l'IDG est la réalisation et l'actualisation de banques de données d'intérêt juridique.

Ces banques de données sont distribuées sur le réseau national et accessibles sur le système Italgiure de la Cour Suprême de Cassation ou le système informatif de la Chambre des Députés.

Le système informatif de la Cour de Cassation a pour mission de fournir aux magistrats, avocats et à tout citoyen toute l'information nécessaire à la connaissance et à l'application du droit.

Vingt fichiers informatiques sont accessibles dans les domaines de la législation, jurisprudence et doctrine juridique

BID:

- contenu: banque de données bibliographique à caractère international dans le domaine de l'informatique juridique et du droit en informatique;
- dimension: environ 30 000 documents;
- origine des données: environ 300 périodiques spécialisés italiens et étrangers, monographies, banques de données, littérature grise;
- période couverte: depuis 1969;
- unité documentaire: une notice bibliographique accompagnée d'un résumé, un code de classification, des mots-clés;

(89) Famelli, E.: Documentation in Legal
Informatics, ..., p.197

- distributeur: l'IDG et la Cour de Cassation;
- logiciel de recherche: Find;
- langue: l'Anglais.

Structure de BID (89)

BID est une bibliographie spécialisée dans le domaine particulier de l'informatique juridique, branche récente du droit, en cours de structuration, ce qui suppose une évolution constante du contenu de la banque de données.

Le domaine couvert est considérable: les notions obligatoires communes à tous les documents sont celles d'informatique et de droit. Cela implique que tous les domaines d'application sont pris en compte, ainsi que les analyses épistémologiques: cybernétique, philosophie, logique, mathématiques, électronique, sciences de l'information et de la documentation; et les implications de ces technologies au niveau social, politique, économique ou éducatif.

La grande variété du public considéré ajoute à la complexité et à l'hétérogénéité apparente de BID: hommes de loi et experts, théoriciens et praticiens, experts d'informatique juridique, politiciens, informaticiens, et tous ceux qui sont intéressés d'une façon ou d'une autre à l'application des nouvelles technologies au droit.

BID est sélective au niveau du contenu des documents, mais non par rapport au niveau scientifique de la publication et à la forme de sa présentation.

BID est analytique: sa mission n'est pas seulement d'attirer l'attention sur les nouvelles publications, mais aussi d'apporter des informations sur leur contenu. La notice est ainsi accompagnée d'un résumé et de plusieurs mots-clés.

BID est systématique: chaque notice reçoit un indice qui indique son contenu, issu d'une classification décimale, et des mots-clés, issus du thésaurus.

L'analyse des documents se fait en langage naturel.

Le choix a été fait d'utiliser deux systèmes d'indexation simultanément: la classification et les mots-clés, afin de permettre des recherches parallèles.

La classification permet d'avoir une vision générale de l'ensemble des domaines couverts et de la façon dont ils sont hiérarchiquement organisés.

La classification est par nature provisoire puisqu'elle doit s'adapter à l'évolution du domaine.

L'index des mots-clés présentent la liste alphabétique des descripteurs utilisés par le documentaliste pour l'analyse du contenu.

Un index systématique qui intègre les descripteurs à la fois du thésaurus et de la classification permet une recherche complètement articulée.

BID se compose ainsi de trois fichiers:

- un index des mots-clés (KWOC): liste alphabétique des descripteurs. Sous chaque descripteur se trouve les titres des publications indexés, ainsi que le code alphanumérique correspondant à la classification, le numéro d'enregistrement du document et sa typologie;
- le fichier systématique et analytique dans lequel se trouve la description des documents;
- l'index alphabétique des auteurs qui présente l'ensemble des publications de cet auteur.

La recherche documentaire dans BID se fait à la fois en utilisant le langage naturel (le système recherche alors dans les champs du titre et de l'abstract) et le langage contrôlé pour le choix des descripteurs.

Le système est composé de six zones d'interrogation:

- la typologie des documents;
- la cote issue de la classification;
- l'auteur;
- le titre, le résumé et les descripteurs;

(90) Informatica e ordinamento giuridico, p.63

(91) Biagini, P.M.: Metodologia di normalizzazione,..., p.307

- la description bibliographique;
- numéro d'enregistrement des notice.

Modalités de recherche du système (90)

Recherche par unité lexicale

Possibilité d'accéder à un document par reconnaissance morphologique du signe linguistique: la forme extensive des mots provoque la prise en compte aussi des formes de déclinaison et conjugaison. les troncatures étendent la recherche aux formes dérivés d'un ensemble de caractères.

Recherche par unité syntaxique

Utilisation de syntagmes pour définir le contexte, à travers des marqueurs syntaxiques.

Recherche par unité sémantique

Cette recherche par analogie et similitude de signification. Elle tend à établir les relations qui existent entre les concepts et les formes de représentation de ces concepts: se base sur l'analyse des sèmes, regroupés dans un fichier inversé.

Recherche par unité logique

Utilisation des opérateurs booléens pour affiner et faciliter la recherche. La recherche se déplace du niveau linguistique à un niveau logique d'organisation des objets en classe.

Normalisation du langage juridique (91)

La première distinction à faire quand on parle de langage juridique s'établit entre :

- le langage juridique: c'est à dire le langage par lequel le juriste annonce une règle juridique;

- le langage des juristes: c'est à dire le langage à travers lequel les juristes expriment et interprètent les normes.

Ce dernier langage peut être considéré comme un méta-langage dans la mesure où il contient toutes les expressions du langage juridique, et en plus toute la richesse du langage naturel, à travers lequel le juriste exprime son raisonnement juridique.

Un système juridique peut être considéré sous deux aspects:

- un ensemble d'expressions linguistiques: énoncés juridiques reliés entre eux et organisés selon des critères empiriques et arbitraires;
- un ensemble de concepts basé sur une structure logique.

Pour arriver à une cohérence et une exhaustivité complète de l'analyse du corpus normatif juridique, il est nécessaire d'arriver à partir du langage technique du législateur à une normalisation entendue comme traduction du langage naturel dans lequel sont exprimées les normes en une structure universelle d'après des règles d'analyse rigides.

L'imprécision du langage juridique génère une ambiguïté sémantique due à la signification des mots qu'il est difficile d'éliminer sans porter atteinte à l'élasticité nécessaire pour l'application pratique des normes aux cas réels.

THES/BID: un thésaurus pour l'informatique juridique

THES/BID est un thésaurus structuré de descripteurs (mots-clés ou locutions) élaboré en Anglais pour l'utilisation de BID, aussi bien pour l'indexation que pour l'interrogation.

Il s'agit de la première tentative d'organisation des concepts relevant du domaine de l'informatique juridique.

Le thésaurus a été conçu à l'aide d'un logiciel opérant des procédures automatiques de contrôle et d'édition.

Sélection des descripteurs

Le matériel lexical de base utilisé pour la préparation du thésaurus est issu de la terminologie de la classification, des mots et locutions utilisés dans les titres et abstracts des notices bibliographiques, ainsi qu'aux terminologies propres à d'autres secteurs entrant en ligne de compte.

Ainsi, une base lexicale de 2000 descripteurs a été obtenue, sélectionnés selon les critères suivants:

- contenu conceptuel;
- importance du concept dans le domaine;
- fréquence d'utilisation;
- prévision de son acceptation de la part de l'utilisateur;
- possibilité d'intégration avec les autres termes;
- exclusion des similitudes de sens avec renvois aux termes exclus;
- établissement des listes des noms géographiques ou noms propres.

Une fois sélectionnés, les descripteurs ont été formalisés:

- choix des variantes graphiques des mots: préférence accordée à la forme anglaise des mots, par opposition à la forme américaine;
- choix des formes grammaticales et des marqueurs;

Classification et structuration

Un numéro d'identification a été accordé à chaque descripteur afin de faciliter la structuration et le contrôle.

Les descripteurs sont regroupés par classes, relatives à la classification: cela a permis de faciliter l'établissement des rapports sémantiques entre les termes.

La structuration du réseau sémantique se base sur les trois relations classiques en documentation:

- les relations d'équivalence (USE; USE FOR): renvoient des non-descripteurs aux descripteurs;
- les relations hiérarchiques (BT; NT): expriment les rapports de généralité et spécificité entre les concepts;
- les relations associatives (RT): indiquent les analogies qui peuvent exister entre deux descripteurs.

L'accès à ce réseau sémantique est facilité par un ensemble de fichiers qui permettent des entrées multiples:

- le "KWOC index of permitted terms": permet de remonter des mots simples aux locutions dont ils font partie:

ex.--> ABSTRACT ABSTRACT
 ABSTRACT BULLETINS
 ABSTRACT JOURNALS

- la liste alphabétique structurée des descripteurs:

ex.--> ABSTRACT BT ABSTRACTING
 NT ABSTRACT BULLETINS
 RT SUMMARY

- liste hiérarchique des descripteurs: à chaque descripteur correspond le terme générique de la classe à laquelle il appartient:

ex.-->ABSTRACT see DOCUMENTATION

- liste des classes qui composent le thésaurus;

- liste des termes génériques accompagnés du nombre de noeuds qui dépendent d'eux:

ex.-->ABSTRACTING(7)

- représentation graphique de la liste hiérarchique: maintient la forme d'une liste, les relations hiérarchiques étant signalées par des points:

ex.-->DOCUMENTATION
 . AUTOMATED DOCUMENTATION
 . BIBLIOGRAPHIES
 .. LEGAL BIBLIOGRAPHIES
 .DATA BANKS
 ..ACCESS TO DATA
 ...ACCESS TIME
 ...FREE ACCESS

- Plan de la classification: permet de relier le contenu d'une notice à un système organisé de connaissance:

EX.-->0.1 LEGAL INFORMATICS IN GENERAL
 0.1.1 THEORY
 0.1.2 SYSTEMATICS

- Liste des descripteurs par classe: rapport entre le thésaurus et la classification:

ex.-->0.1 LEGAL INFORMATICS IN GENERAL
 1) LEGAL INFORMATICS
 2) JURIMETRICS

Le thésaurus permet une grande souplesse pour l'élaboration des stratégies de recherche: Il est cependant fortement lié au cadre conceptuel de la classification qui l'accompagne.

Cette combinaison des deux approches permet à l'utilisateur de vérifier que les mots-clés qu'il utilise correspondent bien dans le système aux concepts qu'il entend couvrir, et aux champs du domaine qui l'intéresse.

Cet emploi de la classification replace l'utilisateur dans une dimension qui lui est plus familière: celle du cadre théorique des domaines juridiques.

(92) THES/BID, . . ., p.155

Particularités de THES/BID:

- construit a posteriori;
- composé de descripteurs simples et de descripteurs composés qui couvrent des concepts complexes et articulés.

Le choix de la langue n'est pas sans poser de problèmes: la terminologie dans une langue couvre les concepts pris en considération dans les systèmes juridiques des pays qui parlent cette langue.

Le passage de l'Italien à l'Anglais, ou de l'Anglais à l'Italien, implique une adaptation linguistique et une réévaluation des concepts qui peut être source d'enrichissement ou de limitation.

Ainsi, la compréhension profonde des valeurs qu'implique l'usage d'un descripteur en Anglais demande une connaissance approfondie du système anglais.

Dans la réalité d'une recherche documentaire ces critères ne sont pas pris en compte (92).

BID ET SES UTILISATEURS

L'enquête a porté sur 10 personnes:

- 3 documentalistes;
- 7 utilisateurs.

Les documentalistes: spécialistes du domaine par une longue pratique de l'interrogation des banques de données.

L'un d'entre eux travaille en bibliothèque universitaire, les deux autres en entreprises.

L'enquête s'est déroulée principalement par entretiens sur le mode de la conversation, dans trois directions:

- représentations du système, de l'utilisateur, de sa mission;
- entretiens avec les utilisateurs;
- stratégies de recherche.

Les utilisateurs: juristes, spécialistes de domaines différents.

La population est composée de deux étudiants, trois chercheurs, deux avocats: les avocats sont les seuls à utiliser la banque de données, les autres utilisent les services des documentalistes.

L'enquête s'est déroulée dans un premier temps sur le mode de l'observation:

- stratégies de recherche;
- entretiens avec le documentaliste;
- représentations de réseaux sémantiques.

Je me suis rendue compte tout de suite que mes préoccupations n'étaient pas à l'ordre du jour, ni chez les documentalistes, et encore moins chez l'utilisateur.

La communication avec les documentalistes n'a pas été difficile dans la mesure où nous partageons le même milieu professionnel et donc, la même terminologie.

J'ai commencé l'enquête en assistant à une première série d'entretiens avec les utilisateurs sans révéler le but de cette observation.

Une seconde série d'observations a été faite après avoir mené de longues discussions avec les documentalistes et les utilisateurs séparément.

Les entretiens ont porté sur les représentations de leurs propres fonds, de BID, du domaine juridique, des utilisateurs.

Les observations des stratégies de recherches ont porté sur une moyenne de deux recherches par documentalistes. Ces recherches ont pris du temps puisqu'elles ont été précédées d'une explicitation de la stratégies et de ce que la personne attendait du système.

Le dialogue avec les utilisateurs a été plus difficile dans la mesure où la notion de représentation ne fait pas partie du paysage quotidien. Il m'a fallu longuement expliquer ce que j'attendais, sans obtenir beaucoup de résultats, ce qui est en soi assez révélateur.

Les entretiens ont porté sur la connaissance du domaine et du système informatif, la représentation du contenu des réponses attendues, le documentaliste.

Les observations se sont concentrées sur l'explicitation des demandes, les stratégies de recherche.

Ma tentative d'obtenir une représentation sémantique de quelques concepts s'est soldée par un échec: les utilisateurs n'utilisant pas les banques de données ont produit dans le meilleur des cas un lexique, les autres ont laissé au thésaurus le soin de s'exprimer à leur place.

Ce qui est aussi assez révélateur.

Résultats

La première constatation à faire consiste à reconnaître que les résultats ont été peu brillants. Cette étude peut être considérée comme une première étape dans l'approche de cette problématique.

L'enquête a consisté essentiellement à affronter les réticences des deux groupes de la population concernée.

La présentation du problème a trouvé beaucoup d'attention, et même de l'approbation.

Mais un refus généralisé s'est établi quand il s'est agit d'aller au delà des pratiques courantes quotidiennes.

Les deux populations ont sans doute exprimé ainsi une peur confuse de remettre en question des pratiques qui ont fait leur preuve et qui assurent une vitesse de croisière satisfaisante.

J'ai alors passé beaucoup de temps à expliquer que les pratiques d'information retrieval et d'entretiens avec l'utilisateurs n'étaient absolument pas remis en question, loin de là. Mais qu'il s'agissait de faire un retour sur soi même et exprimer ses propres représentations de l'environnement.

Ceci a pu m'amener à faire quelques constatations:

- projection inconsciente de la part du chercheur de ses propres représentations sur les représentations du système ou de l'interlocuteur.

La formulation des concepts est immédiatement intégrée à la structure cognitive qui la reçoit puisqu'il n'y a pas d'explicitation verbales, mais au contraire une acceptation systématique et passive.

- les hésitations dans la représentation de réseaux sémantiques pourraient montrer que les représentations sont loin d'être systématiquement connaissables par le sujet lui même.

Un dialogue interactif constructif ne peut l'être que si les deux parties ont conscience de l'enjeu et évoluent dans une même sphère d'approche cognitive.

Ainsi, ces constatations peuvent porter à deux conclusions:

- nécessité de prendre conscience de l'existence d'une projection des représentations des acteurs à chaque niveau de la relation d'information.

- prise de conscience de tous les acteurs face à l'importance de l'explicitation de ces représentations.

B I B L I O G R A P H I E

Approaches to Knowledge Representation: An Introduction / Edited by G.A.Ringland and D.A.Duce.
Taunton: Research Studies Press, 1989.
260 p.; 23 cm . - (Knowledge- Based and Expert Systems Series)
ISBN 0-86380-080-7

ASTOLFI, J.P., DEVELAY, M. / La Didactique des sciences
Paris: PUF, 1991
125 p.; 18 cm . - (Que-sais-je?; 2448)

BIAGINI, P.M., NATALI, F.S., GHELLI, D.T. / Metodologia di normalizzazione del linguaggio giuridico IN:
Informatica e diritto, 2, anno X, magg.-ag. 1984

BROOKES, B.C. / The foundations of Information Science:
part I: Philosophical Aspects , pp. 125-133
part II: Quantitative Aspects 1, pp. 209-221
part III: Quantitative Aspects 2, pp. 269-275
IN:
Journal of Information Science 2 (1980)

BRYCE, L.A. / Cognitive Research in Information Science: Implications for Design IN:
Information Science and Technologie, vol.26, 1991
pp. 3-37

CAPPER, J. / Cognitive Science could unlock the Secrets of how we learn IN:
The Executive Educator, June 1984
pp. 38-39

CHAUMIER, J. / Travail et méthodes du documentaliste: connaissance du problème
Paris: Ed. ESF, 1984
145 p.; 25 cm .- (Formation permanente en sciences humaines)
ISBN 2-7111-0247-5

DARTOIS, C. / Approche de la relation d'information dans un centre de documentation IN:
Documentaliste, vol.28, N°2, mars-avril 1991
pp. 63-66

FAMELI, E., NANNUCCI, R., DI GIORGI, R.M. / Documentation in Legal Informatics and the "International Bibliography on Computers and Law" IN:
Informatica e lavoro, 3, anno IX, sett.-dic.1983
pp. 183-239

HATON, J.P., HATON, M.-C. / L'Intelligence artificielle
Paris: PUF, 1990.
127 p.; 18 cm . - (Que-sais-je ?; 2444)
ISBN 2-13-043164-X

Informatica e ordinamento giuridico
Milano: Giuffrè, 1988

Ipermedia: nuovi strumenti per la didattica: atti ,
Torino, 30-31 oct.1991 / CSI Piemonte

LE MOIGNE, J.L. / La Modélisation des systèmes complexes.
Paris: Dunod, 1990

LE NY, J.F. / Science cognitive et compréhension du langage
Paris: PUF, 1989
249 p.; 21 cm . - (Le psychologue)
ISBN 2-13-042048-6

LEVY, P. / Les Technologies de l'intelligence: l'avenir de la pensée à l'ère de l'informatique
Paris: La Découverte, 1990
233 p.; 22 cm . - (Sciences et société)
ISBN 2-7071-1964-4

MARINO, J.B. / Utilisation de la théorie mathématique de la communication en sciences de l'information
Mémoire: Ecole des hautes études en Sciences sociales, 1982

Les Modèles mentaux: approche cognitive des représentations
Paris: Masson, 1993
183 p.; 25 cm . - (Sciences cognitives)
ISBN 2-225-84023-5

MORIN, E. / La Méthode 3: la connaissance de la connaissance: anthropologie de la connaissance
Paris: Seuil, 1992
243 pp.; 18 cm . - (Points.Essais; 236)
ISBN 2-02-014440-9

Problèmes de psycholinguistiques / textes rassemblés
par J.A. Rondal et J.P.Thibaut
Bruxelles: P.Mardaga, 1987
557 p.; 22 cm . - (Psychologie et sciences humaines; 165)
ISBN 2-87009-299-7

RASTIER, F. / Sémantique et recherches cognitives
Paris: PUF, 1991
262 p.; 22 cm . - (Formes sémantiques)
ISBN 2-13-043825-?

La Recherche en informatique: le Courrier du CNRS,
N°80, fév. 1993

RICHARD, J.F. / Les Activités mentales: comprendre,
raisonner, trouver des solutions
Paris: A.Colin, 1990
434 p.; 23 cm . - (Psychologie)
ISBN 2-200-31257-1

SARACEVIC, T., KANTOR, P. / A Study of Information
Seeking and Retrieving
part I : Background and Methodology p. 161-176
part II : Users, Questions and Effectiveness
p.177-196
part III: Searchers, Searches and Overlap
p. 197-216

IN:
Journal of the American Society for Information
Science, 39(3), 1988

Les Sciences cognitives: Le courrier du CNRS N°79,
nov.1992

SIMON, H.-A. / Sciences des systèmes sciences de
l'artificiel
Paris: Dunod, cop.1991
229 p.; 24 cm . - (Afcet systèmes)
ISBN 2-04-019815-6

Systémique et cognition / sous la dir. de E.Andreewsky
Paris: Dunod, 1991
194 p.; 25 cm . - (Afcet systèmes)
ISBN 2-10-000302-X

THES/BID: un "thesaurus" per l'informatica giuridica e
il diritto dell'informatica, redatto con l'ausilio
dell'elaboratore elettronico IN:
Informatica e diritto, 2, anno IX, magg.-ag.1983

Traité de psychologie cognitive 1: perception, action, langage / C.Bonnet, R.Ghiglione, J.F.Richard
Paris: Dunod, 1989
266 p.; 24 cm . - (Sciences humaines)
ISBN 2-04-018752-9

Traité de psychologie cognitive 2: le traitement de l'information symbolique / J.F.Richard, C.Bonnet, R.Ghiglione
Paris: Dunod, 1991
289 p.; 24 cm . - (Sciences humaines)
ISBN 2-04-018754-5

VIGNAUX, G. / Les Sciences cognitives: une introduction
Paris: La Découverte, 1992
359 p.; 22 cm . - (Textes à l'appui. Sciences cognitives)
ISBN 2-7071-2082-0

BIBLIOTHEQUE DE L'ENSSIB



801288A