



**HAL**  
open science

**Statut de la terminologie et du travail linguistique dans  
le nouveau paradigme des sciences de l'information :  
exemple du thésaurus des “ domaines de formation ” du  
Conseil général du Nord**

Nicolas Tricoche

► **To cite this version:**

Nicolas Tricoche. Statut de la terminologie et du travail linguistique dans le nouveau paradigme des sciences de l'information : exemple du thésaurus des “ domaines de formation ” du Conseil général du Nord. Sciences de l'information et de la communication. 2000. dumas-01706562

**HAL Id: dumas-01706562**

**<https://dumas.ccsd.cnrs.fr/dumas-01706562>**

Submitted on 12 Feb 2018

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

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

**Nicolas TRICOCHÉ**

**MAITRISE EN  
SCIENCES DE L'INFORMATION ET DE LA DOCUMENTATION**

**RAPPORT DE STAGE :**

*STATUT DE LA TERMINOLOGIE & DU TRAVAIL LINGUISTIQUE DANS LE  
NOUVEAU PARADIGME DES SCIENCES DE L'INFORMATION :  
EXEMPLE DU THESAURUS DES « DOMAINES DE FORMATION »*

*Du  
CONSEIL GENERAL DU NORD*

**Stage effectué du 13 Juin 2000 au 18 Août 2000**

**Au  
CONSEIL GENERAL DU NORD**

*Sous la direction de  
Madame WIDAD MUSTAFA ELHADI, responsable universitaire  
Madame SALIHA KHATIR, responsable professionnel*

**Lille 3  
Université Charles De Gaulle  
UFR IDIST**

**05 octobre 2000**



Y-01\* 1770

*Je tiens vraiment à remercier l'ensemble du service « formation » pour son accueil, son aide précieuse mais surtout pour la chaleur humaine de ses membres, pour leur gentillesse et leur grande capacité de compréhension. Merci à tous...*

# SOMMAIRE

<i>INTRODUCTION</i>	P.1
1- LA DEMANDE LINGUISTIQUE DU CONSEIL GENERAL DU NORD : PRESENTATION DE LA MISSION & METHODOLOGIE MISE EN PLACE	P.3
1-1 FORMULATION & FORMALISATION DE LA TACHE	P.3
1-1-1 OBJECTIFS ET NATURE DE LA DEMANDE	P.3
1-1-2 LE THESAURUS DES « <i>DEMANDES DE FORMATION</i> » : UNE DEMANDE LINGUISTIQUE & TERMINOLOGIQUE	P.5
1-2 ETUDE DE L'ENVIRONNEMENT	P.8
1-2-1 FONCTION DU SERVICE « <i>FORMATION</i> »	P.8
1-2-2 FONCTIONNEMENT DU SERVICE	P.9
1-2-2-1 PRESENTATION GENERALE DU SERVICE	P.9
1-2-2-2 SCHEMA DE CIRCULATION DE L'INFORMATION & NATURE DU FONDS	P.10
1-3 APPROCHE ORIENTEE-USAGER	P.11
1-3-1 ANALYSE DE L'USAGE	P.12
1-3-2 ANALYSE DES BESOINS	P.13
2- ADAPTATION DES OUTILS LINGUISTIQUES & TERMINOLOGIQUES AUX NOUVELLES EXIGENCES INFORMATIONNELLE : NATURE & SPECIFICITES DU THESAURUS DES « <i>DOMAINES DE FORMATION</i> »	P.16
2-1 POLITIQUE D'INDEXATION	P.16
2-1-1 RESEAU CONCEPTUEL & DESCRIPTEURS	P.16
2-1-1-1 MODELISATION DES CONCEPTS : LE TERME	P.16
2-1-1-1-1 UNE DEMARCHE ONOMASIOLOGIQUE	P.16
2-1-1-1-2 UNE NECESSAIRE PRECOORDINATION	P.18
2-1-1-2 LES RELATIONS ENTRE CONCEPTS	P.19
2-1-1-2-1 LA RELATION & SON IMPORTANCE DANS LA CONSTITUTION DU SENS	P.19
2-1-1-2-2 MISE EN RESEAU & COGNITIVITE	P.21
2-2 LE THESAURUS DES « <i>DOMAINES DE FORMATION</i> » : UN THESAURUS « DOCUMENTAIRE » ?	P.22
2-2-1 APPROCHE TELEOLOGIQUE	P.22

2-2-1-1 RECADREMENT THEORIQUE	P.22
2-2-1-2 UN OUTIL « HORS NORME » ?	P.24
2-2-2 UNE NATURE HYBRIDE	P.25
2-2-2-1 UN THESAURUS TERMINOLOGIQUE ?	P.25
2-2-2-2 COLLABORATION ENTRE ELEMENTS LINGUISTIQUES & EXTRA-LINGUISTIQUES	P.27
<b>3- SYSTEME TERMINOLOGIQUE ET APPLICATIONS LINGUISTIQUES</b>	<b>P.29</b>
3-1 NOUVEAUX ENJEUX DE LA MAITRISE TERMINOLOGIQUE	P.29
3-1-1 DIMENSION COMMUNICATIONNELLE DE LA TERMINOLOGIE	P.29
3-1-2 IMPLICATION DE LA TERMINOLOGIE DANS L'ENTREPRISE	P.30
3-2 TERMINOLOGIE DANS LE CONTEXTE DES « NTIC »	P.32
3-2-1 LE PRADIGME USAGER	P.32
3-2-1-1 LA SUPPRESSION DES INTERMEDIAIRES	P.32
3-2-1-1-1 LES INTERFACES INFORMATIQUES	P.32
3-2-1-1-2 LA PERTE DE VITESSE DES LANGAGES D'INTERROGATION	P.34
3-2-1-1-3 LE TRAITEMENT AUTOMATIQUE DES LANGUES NATURELLES	P.35
3-3 INTERNET : BIBLIOTHEQUE VIRTUELLE	P.37
3-3-1 INTERNET OU LE CHAOS « NUMERIQUE »	P.37
3-3-2 INSUFFISANCE DES MOTEURS DE RECHERCHE TRADITIONNELS	P.38
3-3-3 DE NOUVELLES PISTES DE TRAVAIL...	P.39
<i>CONCLUSION</i>	<i>P.41</i>
<i>BIBLIOGRAPHIE</i>	<i>P.43</i>
<i>U.E 7 : CONDUITE D'UNE ACTION D'ETUDE EN SCIENCES   DE L'INFORMATION &amp; DE LA DOCUMENTATION</i>	<i>P.44</i>

## **ANNEXES**



## INTRODUCTION

Le 1<sup>er</sup> septembre 2000, la revue « *Le Monde Informatique* » choisit de consacrer un de ses articles à la présentation d'un méta-annuaire chargé de mettre en ordre toutes les données de type identificatrices, qui sont présentes sous la forme d'annuaires. Jean-Luc Rognon, le rédacteur de cette article, met en avant l'importance que revêt l'information dans le bon fonctionnement de toute entreprise (et, serions nous tenter de dire, de toute communauté). Mais cet exemple ne représente que la partie immergée de l'iceberg. En effet, à l'heure de la netéconomie et de l'avènement du multimédia, l'information apparaît comme une données vitale et à haute valeur stratégique. Ainsi, il est évident que dans ce contexte, le management de l'information (plus connu sous l'appellation de *Knowledge Management*) devient une composante incontournable de la productivité de cette nouvelle « *Société de l'information* ». La tâche de tout individu chargé de la gestion de l'information (encore détenue par les professionnels de l'« info-doc ») consiste donc en la transformation des données en vue de leurs intégrations dans un processus d'action et de décision. En d'autres termes, l'objectif de cette gestion est une création d'intelligence c'est à dire qu'il s'agit d'apporter à ces données une plus-value. C'est cette plus-value qui donnera à ces données le statut d' « Informations ». Il apparaît donc clairement, en filigrane, que ce qui prime ce n'est pas tant la possession de cette « information », que sa maîtrise.

Ainsi, le basculement dans cette « nouvelle » société de l'information (et donc de la mise en exergue du paradigme communicationnel en dépendant) renvoie à la problématique de la mise en ordre et de l'organisation des données : ce qui est vital n'est pas l'information à proprement parler, mais son bon ordonnancement. De la bibliothèque d'Alexandrie (et sans doute antérieurement) à l'heure du réseau, les objectifs n'ont finalement pas tant varié que cela : durant plusieurs millénaires, les hommes ont cherché (et donc inventé) les moyens les plus pertinents pour le repérage et l'organisation des informations. Mais quelles sont elles ? Les méthodes créées sont l'indexation, la classification et l'activité résumante ; les outils sont les

listes, les répertoires, les annuaires, les index, les dictionnaires, les encyclopédies. Toutes ces approches ont comme point commun, dans leur mode d'organisation, l'élément essentiel qu'est le langage. En effet, la langue et le langage (l'objectif n'est pas ici de débattre de leurs différences) représente le véhicule le plus courant de transmission d'informations et d'organisation. C'est pour cette raison que la langue est au centre des préoccupations des spécialistes des sciences de l'information. Notre objectif est donc ici de voir, à travers la description de la tâche qui nous a été confié par l'organisme le « Conseil Général du Nord », le statut du travail linguistique et de la terminologie (qui ont été au cœur de notre travail) dans ce nouveau paradigme des sciences de l'information. Pour cela, nous opérerons à une description de la nature de la mission et de la méthodologie mise en place, afin de soulever les nécessaires adaptations et les nouvelles orientations (dictées par les exigences informationnelles de cette nouvelle manière de penser) du travail linguistique et terminologique. Cette analyse nous permettra de voir, dans un troisième temps, les nouvelles pistes de travail et les nouvelles applications qui sont permises par les incessants progrès technologiques et techniques.

1- LA DEMANDE LINGUISTIQUE DU CONSEIL GENERAL DU NORD :  
PRESENTATION DE LA MISSION & METHODOLOGIE MISE EN PLACE

1-1 FORMULATION & FORMALISATION DE LA TACHE

1-1-1 OBJECTIFS & NATURE DE LA DEMANDE

Dès notre arrivée, la mission nous a été formulé de la sorte : il s'agissait d'établir un « cahier des charges » qui devait comprendre à la fois une évaluation de la manière de travailler et une proposition d'organisation. Ainsi, la tâche qui nous incombée se décomposait en deux pôles : le premier, purement critique, consistait à évaluer la gestion de la documentation opérée par le « service formation ». La deuxième partie avait comme finalité de réorganiser les ressources documentaires sous forme de thesaurus. L'ensemble du service partait donc avec le postulat que la manière de travailler était mauvaise (ou plutôt inexistante), puisque nécessitant une réorganisation de fonds. Ce fut de cette manière que se formula la demande dite « officielle ». Bien entendu, les membres du service mirent surtout l'accent sur le deuxième volet de la mission car c'était à ce niveau là que se situait la raison de notre présence et c'est à ce niveau que notre présence allait pouvoir être utile au service, c'est à dire dans cette optique de « réorganisation » (de leurs propres termes). Cependant, même si la nécessité du travail (mais aussi la quantité de travail) se situait au plan du thesaurus, la première partie (constituée par « l'audit ») était pour nous d'une grande importance. En effet du point de vue méthodologique, toute élaboration de produit documentaire doit, idéalement, être fondée sur une étude contextuelle, qui a comme finalité, tout en garantissant l'adéquation de l'outil avec ses futurs utilisateurs, d'assurer sa pertinence. Ainsi, ce nécessaire travail préparatoire d'analyse et d'étude était incarné dans cet « audit » et recevait de fait une caution officielle. Ce qui n'était qu'un travail de préparation devint un des volets de la mission, et en ce sens, devait faire l'objet de plus d'attention et de précision. De plus, le souhait du responsable de service de voir cet examen sous forme de « rapport » écrit, nous aida dans notre propre formalisation et modélisation des données.

Mais rentrons plus dans le détail afin de voir en quoi consistait précisément l'attente du CGN (entendre « Conseil Général du Nord »). La mission se situait dans le cadre du service « formation » et pour comprendre toute la problématique de la situation (et donc de notre travail), il faut opérer à quelques précisions quant au fonctionnement du service (qui sera faite de manière plus complète et rigoureuse dans un deuxième temps).

De manière schématique, l'organisation des formations se déroulent comme suit : un agent passe une demande de formation à son chef de service qui le valide et qui l'envoie au service « formation », les membres de ce service étudie cette demande pour voir si on peut dispenser cette formation c'est à dire si la formation existe. Pour cela, le chargé de formation regarde dans sa base de données si un organisme dispense ce stage. Comment se constitue cette base de données ? Les organismes de formation (qui sont de type privé ou public) démarchent le CGN en leur proposant la liste des stages (sous forme de « catalogues ») qu'ils assurent ; la base de données est donc constituée par les stages que chaque organisme dispense : cette liste globale, alliant à la fois les intitulés de stage et les différents organismes, représente la liste de formations « POTENTIELLES » que peut recevoir tout agent du CGN qui en fait la demande. C'est donc par recoupement entre la demande de l'agent et cette liste de formations que le chargé de formation voit si le souhait de l'agent pourra être satisfait ou non. Si l'agent demandeur souhaite une formation qui n'est pas présente dans la base de données, alors le chargé de formation ne peut valider la demande de formation puisque le type de formation demandée n'existe pas pour le chargé de formation (et donc pour l'ensemble des agents dépendant du CGN).

Cependant, comme le service « formation » (composé d'une dizaine de personnes) est chargé, seul, de la gestion des formations des quelques 4500 agents territoriaux dépendant du CGN, la majorité de leur travail est constituée par les tâches administratives attenantes à ces formations (composition des groupes, paiement des organismes, réservation des salles de cours ect...). Les membres du service « formation » n'ont donc pas le loisir de dépouiller (et ne serait ce qu'ouvrir) les différents catalogues qu'ils reçoivent chaque jour. Ainsi, si un agent demande un

stage que les membres du service ne connaissent pas ou qu'ils n'ont jamais organisé, ils ne peuvent valider cette demande, alors que ce stage est peut être présent dans un des catalogues posés sur l'armoire ou sur une table.

C'est dans ce cadre que s'insère la mission qui nous a été confiée : comme les agents faisant partie du service n'ont matériellement pas le temps de dépouiller ces catalogues, c'est nous qui avons été chargé de la faire. Et comme, les formations se déclinent en termes de thèmes (« *informatique* », « *Word* », « *les tableaux sous Excell* ») et de manière hiérarchisée, il est possible de créer un thesaurus « des domaines de formation » venant palier cette carence informationnelle. L'objectif de la mission était donc de proposer un outil qui « recensait », en quelque sorte, l'ensemble des stages que proposait les différents organismes et que pouvait donc recevoir les agents territoriaux. La finalité de cet outil était bien entendu d'optimiser l'efficacité et le travail en démultipliant son champs d'action (ou l'étendue de son champs de formation).

#### 1-1-2 LE THESAURUS DES « DOMAINES DE FORMATION » : UNE DEMANDE LINGUISTIQUE & TERMINOLOGIQUE

La description de l'information, c'est à dire la représentation du contenu des documents, et l'accès au contenu des documents représentent des clés essentielles dans la bonne organisation des systèmes d'information. Ainsi, les outils de représentation des connaissances (tels que les thesaurus) semble alors être des éléments nécessaires pour le bon fonctionnement de ces systèmes d'information. En effet, la bonne maîtrise de l'information repose sur une compréhension sémantique et conceptuelle du domaine dans lequel s'inscrit la spécificité du service concerné par ce type de données. Chaque service (ou domaine) génère des problématiques qui lui sont propres et qu'il lui faut maîtriser. La première étape consiste alors à l'appropriation des différents concepts qui sont mis en avant dans le contexte précis. Et comme nous l'avons dit précédemment, des outils de description conceptuelle tels que les thesaurus semblent être une réponse pertinente.

Mais il convient de voir en quoi consiste précisément un thesaurus. Retenons d'abord la définition figurant dans la norme française ISO 2788 « *Principes directeurs pour l'établissement et le développement de thesaurus monolingues* » : « *Un thesaurus peut être défini selon sa fonction ou selon sa structure. Du point de vue de sa fonction, un thesaurus est un instrument de contrôle de la terminologie utilisée pour transposer en un langage plus strict (langage documentaire, langage d'information) le langage naturel employé dans les documents et par les indexeurs ou les utilisateurs. Du point de vue de sa structure, le thesaurus est un vocabulaire contrôlé, et dynamique, de termes ayant entre eux des relations sémantiques et génériques et qui s'applique à un domaine particulier de la connaissance* ». Nous pouvons donc déduire plusieurs choses de cette définition. La première est que le thesaurus est un outil linguistique pour l'indexation des documents. Selon Jacques Chaumier, l'indexation est « *l'opération qui consiste à caractériser un document à l'aide de représentation des concepts contenus dans un document, c'est à dire à transcrire en langage documentaire les concepts après les avoir extraits du document par une analyse* » (dans « le traitement linguistique de l'information »). Et nous verrons que cette fonction peut se révéler précieuse si notre thesaurus fait l'objet un jour d'une intégration informatique (voir pour de précision « 1-3-1 Analyse de l'usage »). En effet, en contrôlant le vocabulaire, le thesaurus a aussi comme fonction de palier les inconvénients du langage naturel : regroupement de signifiés différents sous la même forme signifiante et dispersion de l'information sous des termes plus ou moins proche sémantiquement parlant. Il permet ainsi de réduire le taux de silence et de bruit lors de la recherche d'information.

Mais ce qui est véritablement prégnant dans ce contexte précis est que notre thesaurus des « domaines de formation » est aussi un instrument à volonté et à objectif terminologique (et non pas seulement en terme de « normalisation » et de contrôle de la terminologie comme il est de rigueur dans les thesaurus documentaires traditionnels) . En ce sens, il devient un outil d'acquisition du vocabulaire et, à travers lui, un mode d'appropriation plus efficace du domaine au niveau conceptuel. Or, cet aspect est très important dans le contexte qui nous occupe aujourd'hui : une meilleure compréhension du domaine (au plan conceptuel et sémantique) permettrait à l'agent de moins avoir recours à des instruments d'aide. En effet, mettons nous en situation de l'agent chargé de formation qui reçoit une

demande de formation : dans un premier temps, le thème du stage demandé ne lui évoque rien. S'il n'a pas connaissance du domaine (en terme de « réseau sémantique » c'est à dire avec les notions de hiérarchisation et d'inclusion inhérentes à ce type de connaissances), il ne pourra pas donner suite à la demande. Par contre, s'il a une connaissance sémantique du domaine, il pourra opérer à certains recoupements et certains rapprochements qui lui permettront de pouvoir tout de même donner suite à cette demande. Par exemple, on lui demande une formation sur les « firewalls » et s'il ignore totalement ce dont il en retourne, il ne pourra pas satisfaire la demande. Mais par contre, s'il connaît un peu le domaine et qu'il sait (tout en étant sûr que cette formation n'est pas dispensée par le CGN) que c'est un programme de protection informatique sur Internet, il pourra tout de même proposer au demandeur un stage équivalent de « détection d'intrusions » ou une formation sur un programme comparable de « protection informatique ». La connaissance sémantique du domaine permet à l'individu de prendre du recul par rapport à la demande, cela lui apporte une plus grande flexibilité c'est à dire une plus grande capacité d'adaptation. De plus, en ayant une connaissance du domaine, il sait qu'une formation peut avoir de multiples appellations : ce stage pourrait être présent dans la base de données sous une autre appellation, et pourrait alors valider la demande. Mais le « non-connaissant » lui ne validerait pas le stage car ne connaîtrait que la dénomination première et ne serait pas capable de faire le lien entre les deux termes. Il agirait comme un automate et c'est en ce sens qu'on peut dire que l'appropriation du domaine permet de prendre du recul par rapport aux mots, et donc penser en termes de concepts. Ce ne serait plus une réflexion de surface mais une réflexion ontologique de fonds. D'un automate, il passe à un agent intelligent. Ce thesaurus n'a donc pas comme unique objectif de permettre une meilleure recherche d'information par la normalisation de son vocabulaire, mais a aussi, une visée terminologique.

Cette mise en situation permet de voir en quoi une approche linguistique d'analyse des contenus sémantiques et de leur mise en réseau (sous forme articulée et relationnelle de thesaurus) peut avoir une réelle influence sur l'activité professionnelle des membres du service « formation ». Dans ce cas précis, le

thesaurus des « domaines de formation » (en tant qu'une approche linguistique) a donc toute raison d'être et se révèle légitime.

## 1-2 ETUDE DE L'ENVIRONNEMENT

### 1-2-1 FONCTION DU SERVICE « FORMATION »

Dépendant de la « Direction des Ressources Humaines » (DRH), la fonction première du « service formation » consiste en la gestion des formations des quelques 4500 agents territoriaux. Il est donc chargé de répondre aux demandes en formation des différents acteurs le sollicitant (agents, chefs de service...), en montant des « plans de formation » et les calendriers de stages, internes au Conseil Général du Nord.

De fait, le caractère stratégique de la mission du « service formation » apparaît comme évidente. En effet, le rôle des agents de la fonction publique suit le schéma du service collectif c'est à dire qu'ils ont une mission de service à l'égard de la collectivité. Ainsi, l'exigence de la formation apparaît comme nécessaire pour l'exercice de leurs propres fonctions. Elle est le moyen de renforcer et d'optimiser la pertinence de l'intervention de chaque agent dans cette optique de réponse sociale. Comme exemple de cette volonté de coller le plus possible à la conjoncture, nous pourrions soulever le fait que l'informatique est une formation dite « institutionnelle », car exigée par la présidence du conseil général ; en effet, quoi de plus logique et de plus pertinent de privilégier la formation informatique quand la société s'apprête à basculer dans « une société de l'information » où l'informatique et les NTIC occuperont une place prépondérante ? La démarche de formation constitue alors un élément fondamental dans la gestion des ressources humaines car visant l'amélioration du service rendu à la collectivité.

Mais outre ce caractère purement social, nous pouvons aussi souligner que le Conseil Général (comme tout organisme publique) est une structure qui se particularise par une forte potentialité de mouvements de personnel (recrutement, concours, changements de service, évolution de carrière...). Ainsi, ce service se trouve de fait comme vecteur de cette problématique de mobilité du personnel.

Tant au niveau externe (répondre à cette exigence d'adaptation sociale de la compétence des agents) qu'au niveau institutionnel pur et interne ( cristallise le caractère particulièrement mobile du personnel d'une telle structure), le « service formation » semble occuper une place essentielle dans l'architecture de gestion du personnel du Conseil Général du Nord. En effet, ce service apparaît comme la condition de possibilité de l'adaptabilité des fonctionnaires à la tâche qui leur est confiée.

### 1-2-2 FONCTIONNEMENT DU SERVICE

#### 1-2-2-1 PRESENTATION GENERALE DU SERVICE

Le « service formation » auquel vient se greffer la cellule « emploi-jeunes » comptabilise une vingtaine de personnes. L'objectif des formations est de rafraîchir ou d'acquérir de nouvelles connaissances dictés par l'évolution des missions des différents services départementaux et des souhaits personnels.

Ainsi, une note de cadrage et d'orientation, qui émane du directeur général des services, fixe les objectifs généraux de formation et les différents services déclinent leurs objectifs autour de leurs projets. A partir de ces données et en fonction des conditions de réalisation, chaque direction conçoit des plans de formation qui sont ensuite communiqués au « service formation » qui les valide (ou pas), et qui organisent les calendriers des stages avec la collaboration des acteurs concernés (services, agents et organismes de formation).

A l'intérieur même du service, la répartition a été faite entre les agents selon les différents thèmes souhaités : ainsi, une équipe s'occupe du domaine social, une autre de la Bureautique et de l'informatique, ect...

1-2-2-2 SCHEMA DE CIRCULATION DE L'INFORMATION ET NATURE DU FONDS

De manière grossière, nous pouvons dire que les « flux informationnels » se caractérisent par un trajet « externe⇒ interne ». En effet, chargés de gérer les formations de près de 4500 agents, le Conseil Général constitue un gros marché pour les organismes de formation (la plupart du temps de type « privé »). Ainsi, dans une démarche purement commerciale, ce sont les organismes de formation qui démarchent le CGN, et qui lui envoient de manière continue de nouveaux intitulés de stages ou de nouveaux catalogues. L'arrivée de nouvelles informations est donc continue.

Ce sont par ces informations que leurs dispensent les organismes, que les agents chargés de la formation traitent les demandes en formation. En effet, ce sont ces informations et ces organismes qui leur donnent les matériaux pour former leurs plans de formation : quand un agent demande telle ou telle formation, le gestionnaire de formations recherche dans la base de données ou dans les catalogues (envoyés de manière complètement spontanée par les organismes de formation) si le stage qu'il demande est dispensé. S'il le stage se pratique alors il peut le mettre en place, et mettre en branle tout le processus administratif qui vient en périphérie de sa fonction et de son « utilité » première. Notons que ce travail administratif représente la charge la plus lourde pour l'agent chargé de la formation, alors que le cœur même de la mission de ces agents est de répondre à une demande (c'est à dire trouver un stage qui satisfasse le demandeur).

Nous avons pris le temps d'évoquer ces aspects de fonctionnement du service formation (qui ne sont étrangères à aucune des personnes y travaillant) pour montrer une chose centrale : l'information est un outil incontournable et essentiel de ce service ; par voie de conséquence, il ne faut pas mésestimer toute l'importance de la bonne gestion de cette information. Une gestion pertinente de celle-ci accroît le potentiel d'action et l'efficacité de ces agents chargés de formation. De même, l'efficacité du service renforce le rayonnement de leur action : un agent pertinemment formé est un agent véritablement efficace.

Cependant, il convient de préciser de quelle nature ce flux informationnel est composé et sur quel type de documents avons nous travaillé. Comme nous l'avons dit précédemment, les différents organismes envoient de manière spontanée leurs catalogues, présentant ainsi de manière détaillée les différentes formations qu'ils proposent. Tous les thèmes de stage sont donc précisés dans ces catalogues. Ainsi, les supports sur lesquels nous avons travaillé sont constitués par ces catalogues et brochures de présentation (Annexe 1) : tous les éléments permettant une indexation sûre et pertinente étaient réunis dans ces documents.

Du point de vue méthodologique nous nous sommes efforcés de travailler sur les sommaires de ces catalogues (et donc sur les intitulés que ces organismes donnaient à leurs formations) pour plusieurs raisons : la première est une raison pratique de gain de temps. En effet, la totalité des stages est contenue de manière synoptique sur quelques pages (au maximum) et l'ensemble pouvaient être connu en très peu de temps ; cela nous évitait une fastidieuse compulsation de fonds du document (si un intitulé ne nous semblait pas clair, nous pouvions toujours nous reporter à la description précise correspondante à ce stage et contenu à « l'intérieur » du document). La deuxième raison est que le découpage et le classement sémantique de ces formations apparaissaient de manière plus claire et explicite, et dans le contexte qui nous intéresse ces classifications étaient des auxiliaires précieux de compréhension des différents domaines (et sous domaines) que décrivaient le document.

### 1-3 L'APPROCHE ORIENTEE-USAGER

Pour Yves Le Coadic, le nouveau paradigme des Sciences de l'Information exige que l'utilisateur occupe une place prépondérante dans la réflexion sur les systèmes d'information. En effet, si le service d'information se veut en adéquation avec les besoins des usagers, il faut que ce dernier en soit un élément constitutif. Le changement de paradigme voit donc passer l'utilisateur au centre des préoccupations, venant remplacer le document. Cette réflexion sur l'utilisateur nécessite deux temps

consécutifs qui sont l'étude de l'usage (c'est à dire comment l'utilisateur utilisent le système ?) et une étude du besoin (comment les usagers se représentent la satisfaction de leur manque informationnel ?). L'objectif de ces deux approches complémentaires étant d'en dégager un profil indirect d'utilisateur en formalisant son besoin d'information.

Notons que pour cela, le questionnaire distribué aux différentes personnes revêt une grande importance (Annexe 2) et a constitué la base de notre analyse sur les usagers.

### 1-3-1 ANALYSE DE L'USAGE

Dans un premier temps, il convient de souligner que le service « formation » est déjà pourvue d'une classification sémantique qui est contenue dans la base informatique (le logiciel est « Paradox 3.1 » qui est un logiciel de gestion) et sur papier (Annexe 3). Cette arborescence assujettie à un code numérique est appelée « *Code Secteur* ». Dans un premier temps, il ne faut pas se méprendre quant à la finalité de cette classification. En effet, elle n'a pas été faite avec la volonté de rendre compte des domaines de formation dans leur exhaustivité et leur globalité. Elle a la fonction d'« élaguer » le terrain et de faciliter le travail et la recherche des agents chargés de la formation. Cependant, même si nous pouvons féliciter la démarche d'intégration d'une classification thématique, un seul regard suffit pour voir que cette classification est incohérente, et de ce fait, totalement obsolète (chose que nous avons démontré de manière précise dans l'audit). De plus, le logiciel « Paradox » permet une recherche par « code secteur » (qui revient somme toute à une recherche thématique). Or, il est apparu après dépouillement des différents questionnaires que ce mode de recherche n'était jamais utilisé. Cela pour plusieurs raisons. La première est que l'arborescence ne se déploie pas lors de l'interrogation c'est à dire que le système ne permet pas de voir les stages qui dépendent des inférieurs hiérarchiques. Par exemple, si on tape le code secteur de la catégorie « Bureautique-Informatique », n'apparaîtront que les stages qui ont été rangés sous cette catégorie précise c'est à dire que les stages attenants aux sous catégories telles que « Word » ou « Excell » ne sont pas accessibles par ce mode d'interrogation. La deuxième raison est constituée

par le fait que l'indexation est mauvaise et parfois même source d'erreurs (de l'avis même des agents interrogés). Ainsi, les « intitulés de stage » apparaissent par moment totalement obscurs et on ne peut pas parler dans ce cas de véritable « représentation conceptuelle de contenu », d'où la nécessité d'une refonte terminologique. La dernière raison (et non la moindre) est que du fait de l'activité des agents, il est totalement laissé à l'abandon. Cependant, même si ce système ne semble pas pertinent (pour des raisons d'incapacité technique de la part du logiciel), il a tout de même le mérite d'exister.

Ainsi, la recherche d'information, du fait des raisons de la mauvaise classification et d'incapacité technique de la part du logiciel, ne peut être efficace et pertinente par le système déjà existant. Ce constat implique que dès lors que l'on a analysé l'existant, nous nous tournions vers le deuxième aspect c'est à dire vers le besoin informationnel de l'utilisateur.

### 1-3-2 ANALYSE DES BESOINS

Après avoir vu l'analyse des usages, il convient donc de faire une analyse des besoins qui intervient comme explication du désintérêt des usagers vis à vis de la base (en pointant du doigt ce que la base ne « permet pas » de faire et en montrant, de ce fait, l'inadéquation de celle-ci avec leurs attentes). Il est donc essentielle, à ce niveau là de l'analyse, de laisser la parole de manière concrète aux usagers pour que ceux ci orientent l'élaboration de la « politique informationnelle » et documentaire. Car n'oublions pas que ce sont eux qui se révéleront les utilisateurs du système, et que toute cette approche usager tend vers une idée qui est l'adéquation du système avec leurs souhaits afin de garantir la satisfaction de leurs exigences c'est à dire de répondre de manière adéquate à leur besoin informationnel. Ici, le questionnaire est aussi un élément prépondérant de notre étude : les questions concourant à cette analyse du besoin sont les questions 6 et 7 (les questions précédentes étaient plus orientées vers l'analyse de l'usage). Pour ne pas interférer sur leurs réponses, nous avons privilégié l'approche « ouverte » du questionnement, où ils sont tout à fait

libres de leurs réponses et de leurs dires. Que nous apprennent alors ces réponses sur la représentation qu'ils se font de leur propre besoin informationnel ?

De manière évidente et unanime, la réponse qui est attendue à cette véritable carence est la conception d'un outil sémantiquement plus complet et représentatif de la complexité du domaine (le terme d'« exhaustivité » a été plusieurs fois mentionné). Ce qui est réclamé est, somme toute, la création d'un outil qui permette une recherche thématique plus pertinente et surtout, plus efficace. Notons que le terme de « thesaurus » a été mentionné par deux fois. La représentation idéale que chacun se fait d'un système véritablement adapté est un dispositif permettant une recherche thématique ; reliant ce thème (ou domaine de stage) aux différents organismes dispensant des stages du domaine recherché. Il apparaît comme évident que l'approche documentaire que fournit « Paradox » ne permet pas la mise en place d'un tel outil. Ce dépouillement permet de souligner un aspect qui n'avait pas été mis en avant lors de la présentation de la mission : l'organisation de l'information doit s'orienter autour de deux entités ontologiquement différentes, et non pas autour d'une seule comme il est traditionnellement question dans les thesaurus. En effet habituellement dans les thesaurus, le travail est axé sur un seul type d'élément qui est l'unité linguistique (c'est à dire autour des descripteurs et de l'extraction). Or, il apparaît ici que cet élément linguistique (les descripteurs, qui sont dans le cas présent des « intitulés de stage ») doit être articulé autour d'un élément extralinguistique qui est la dénomination de l'organisme dispensant ces stages. Sans l'existence de ce questionnaire, nous serions sans doute passé à côté de cet élément non linguistique, pourtant central dans le contexte de recherche d'information qui est ici le nôtre. Notre thesaurus, sans l'ajout de cette donnée, aurait très certainement été taxé d'obsolescence et réduit, de ce fait, notre travail et nos efforts à néant. C'est donc l'alliance de ces deux unités (l'une linguistique et l'autre pas) qui crée l'information : en combinant ces deux types de données et en les convertissant sous une forme répondant aux besoins de recherche des utilisateurs, nous intégrons une plus value d'organisation qui crée l'information. La notion d'organisation est donc un élément constitutif de l'information.

Cependant, nous devons souligner un point, au niveau méthodologique, qui vérifie les propos d'Yves Le Coadic sur l'analyse de ce besoin d'information. En effet pour lui, l'étude ce besoin est altérée par les difficultés et les retenues qu'ont les usagers à formuler et à exprimer leur propre besoin informationnel : il suffit pour cela de voir combien de questionnaires nous sont revenus sur le nombre distribué, et dans quels délais . Mais cela ne veut pas dire que ceux ci ne voulaient pas répondre, mais cela met juste en valeur le fait qu'on peut difficilement faire l'expertise (et donc apporter des éléments de réponses aux problèmes soulevés) d'un système (informationnel ou non) dans lequel on est trop investi et impliqué : il nous manque ce fameux « recul ». Et c'est pour tout le monde la même chose, qu'on le veuille ou non ! C'est pour cela qu'une audit doit toujours être pratiquée par une personne totalement étrangère à la structure.

2- ADAPTATION DES OUTILS LINGUISTIQUES & TERMINOLOGIQUES AUX NOUVELLES EXIGENCES INFORMATIONNELLE : NATURE ET SPECIFICITE DU THESAURUS DES « DOMAINES DE FORMATION »

2-1 POLITIQUE D'INDEXATION

Il semble important de souligner que toute politique d'indexation (entendre les choix faits au niveau de l'extraction - ou de la formation des termes - et de la mise en relation de ces termes) est rendue possible par l'analyse du besoin (qui a été faite précédemment). Toute détermination de l'orientation conceptuelle qui a été la notre ne prend sens qu'à travers l'analyse du besoin, qui considère l'utilisateur comme source et base de tout travail concernant la gestion de l'information et de la documentation.

2-1-1 RESEAU CONCEPTUEL & DESCRIPTEURS

2-1-1-1 MODELISATION DES CONCEPTS : LE TERME

2-1-1-1-1 UNE DEMARCHE ONOMASIOLOGIQUE

Bien entendu, chaque organisme ne propose pas des formations toutes originales. Bien que certains soient plus spécialisés dans un domaine bien précis (par exemple, l'organisme « Maquet » est spécialisé dans les formations informatiques de tout niveau : cela va du traitement de texte aux langages de programmation en passant par les réseaux), il arrive irrémédiablement qu'apparaissent des redondances. Ainsi, des stages rigoureusement identiques pouvaient être dispensés par de nombreux organismes. Chacun de ceux ci ayant leur propre logique de présentation, de formulation (sans oublier les aspects marketing dans la formulation des intitulés), le fait que le même stage pouvait prendre différentes appellations fut un phénomène très observé. Cela a représenté une réelle difficulté qui a complexifié notre travail d'indexation et d'extraction ; en effet, les dénominations pouvaient sembler fort éloignées alors qu'ils étaient en fait rigoureusement identiques. Prenons

un exemple simple pour étayer notre propos (un exemple complexe aurait tendance à altérer la bonne compréhension de notre argumentation et n'apporterait rien) : la formation « Word initiation » pouvait être décliné de multiples façons comme « Word débutant » ou « Word-niveau 1 ». Bien entendu, ici, la ressemblance est immédiatement appréhendée, mais ce ne fut pas toujours le cas.

Il semble donc clair que les différentes appellations renvoient à la même réalité, à un objet distinct et unique (bien qu'immatériel). En ce sens, nous pouvons donc parler de « concept » qui, pour Cabré, est « *un élément de pensée, une construction mentale, qui représente un objet individuel, matériel ou immatériel* ». Toute la difficulté de notre travail était donc à partir d'un concept unique et d'appellations plurielles, de déterminer une seule unité linguistique révélatrice du sens de ce dernier. Nous pouvons donc dire que nous avons suivi une démarche onomasiologique car nous partions du concept, et le travail consistait à déterminer sa dénomination la plus pertinente c'est à dire à former un terme qui circonscrivait correctement la notion (la démarche sémasiologique suit rigoureusement le chemin inverse).

Cela nous renvoie à toute l'importance que revêt l'indexation et la détermination des termes lors de l'élaboration de ce type d'outil linguistique. En effet, un des aspects du thesaurus (mais non pas le seul) est la normalisation du vocabulaire ; mais hormis en cas d'interrogation de base de données, en quoi cette normalisation est elle essentielle ? Un des objectifs de la normalisation du vocabulaire est de permettre, par son intermédiaire, une structuration de la pensée via cette structuration de la langue : elle permet un apprentissage univoque du domaine étudié. Et par l'univocité de ces unités linguistiques (ou « termes ») et donc par cette meilleure appréhension de la réalité, la communication est de fait facilitée. Cela nous permet de dire que le terme est tri-dimensionnel en ce qu'il allie des aspects linguistiques (bien entendu), cognitifs et communicationnels : « *Un terme est une unité qui désigne une réalité, qui s'exprime au moyen d'une forme qui sert à communiquer* » (Cabré, 1998). Ainsi, tout individu qui effectue un travail terminologique et linguistique doit bien prendre en compte ces trois dimensions du terme : tout terme doit correctement décrire la réalité qu'il représente par une forme

linguistique appréhendable, caractère de lisibilité qui rend alors possible sa communication.

#### 2-1-1-1-2 UNE NECESSAIRE PRECOORDINATION

La précoordination s'est révélée nécessaire pour trois raisons. La première est la raison évoquée plus haut c'est à dire la nécessité que le terme soit une description cohérente du concept duquel il constitue son « incarnation » linguistique. Cependant, la complexité du domaine fait que l'utilisation de termes uniques (entendre « unitermes ») ne fut pas toujours possible si le thesaurus voulait garantir une bonne représentation linguistique de ses concepts. Il fallut alors recourir à la « précoordination » et passer du stade de « l'uniterme » à celui de « pluritermes ». En effet, l'utilisation de termes précoordonnés permet le développement d'un vocabulaire plus spécifique qui favorise la précision de la description par l'unité linguistique. Cependant, la difficulté, qui a été la notre, était de ne pas tomber dans l'abus et de ne pas produire des descripteurs trop volumineux (le descripteur n'est pas nécessairement un « uniterme » mais peut aussi être un « pluriterme » car un descripteur est la représentation linguistique, nous pourrions presque dire « matériel », d'une notion donnée).

De plus, l'utilisateur ne se révélant pas un spécialiste de la documentation, nous avons utilisé la précoordination qui se rapproche du langage naturel pour garantir une meilleure lisibilité (et donc compréhension), et assurer une meilleure convivialité lors de l'utilisation de ce thesaurus. En effet, les termes précoordonnés sont beaucoup plus proches du langage naturel que ne le sont les unitermes (qui s'apparentent beaucoup plus au langage contrôlé) et il faut reconnaître que l'utilisation du langage naturel semblait plus adéquate que l'utilisation d'un langage rigoureusement contrôlé. En effet, cet utilisateur formulera plus sa requête sous la forme « *Méthodologie de projet* », que sous la forme « *Méthodologie + Projet* » qui s'identifie, elle, plus avec les langages documentaires traditionnels (et nous pourrions multiplier les exemples). C'est pour cette raison que nous n'avons pas eu de scrupules à utiliser cette forme de termes qui s'imposait (comme nous l'avons évoqué

précédemment), de plus, pour une meilleure représentation linguistique des concepts mis en jeu.

Enfin, la dernière raison de cette précoordination est à rapprocher de la précédente et concerne la faible familiarité de ce type d'outils documentaire par les utilisateurs (ne m'a t-on pas parler de « liste » ou de « listing » en parlant du thesaurus ?). En effet, nous avons volontairement alourdi les descripteurs pour que ceux ci ne soient pas sources d'erreur pour l'utilisateur alors que la structure de « réseau » de la présentation hiérarchique systématique (avec la conception de relation *hiérarchique*, c'est à dire de dépendance de l'inférieur par rapport au supérieur) aurait permis que l'on explicite moins les termes. Afin d'être plus clair, prenons un exemple significatif: dans la facette « *Informatique* », le TS3 « METHODE D'ELABORATION DE BASE DE DONNEES » aurait pu prendre la forme « *METHODE D'ELABORATION* », selon sa relation hiérarchique d'infériorité sémantique par rapport à son TS2 « BASE DE DONNEES ». Mais pour l'utilisateur néophyte de ce type d'outil documentaire, il était plus parlant de préciser « METHODE D'ELABORATION DE BASE DE DONNEES ».

Ainsi, la formation des termes (suivant une démarche et une logique nécessairement onomasiologique de par les sources et le domaine) a elle aussi obéi au principe de l'approche orientée-usager qui est prégnante dans le nouveau paradigme des Sciences de l'information.

## 2-1-1-2 LES RELATIONS ENTRE LES CONCEPTS

### 2-1-1-2-1 LA RELATION & SON IMPORTANCE DANS LA CONSTITUTION DU SENS

Un thesaurus est un outil de description conceptuelle d'un domaine. En ce sens, nous pouvons dire qu'il est une retranscription de la réalité ; réalité qui depuis Platon et Aristote se base sur une conception taxinomique. En effet, celle ci se décline en termes de classes (inclusion) et donc en termes de hiérarchie : toute chose (ou tout concept ou encore, en termes platoniciens, toute « Idée ») ayant la particularité

d'appartenir à une classe. Ainsi, chaque concept fait sens dans cette organisation hiérarchique. Toute classification repose sur cette même notion de découpage taxinomique : un terme ne reçoit son sens que par rapport à son appartenance à un réseau sémantique c'est à dire par rapport à la relation que celui ci entretient avec ses supérieurs et inférieurs sémantiques. Ainsi, un terme, de par sa nature, ne peut donc être un élément isolé. De plus, c'est en s'inscrivant dans un domaine spécifique que l'unité linguistique reçoit le statut de « termes » : tout mot n'est donc pas nécessairement un « terme ». Il apparaît alors que le mode de présentation systématique hiérarchique n'est pas un mode de présentation comme les autres : il est une description de la réalité (sémantique) d'un domaine, cette description formant alors le « champs conceptuel » de ce domaine. Une description des relations hiérarchiques constitutives du thesaurus est donc un élément incontournable dans le travail de description ontologique de celui ci. Il faut donc savoir, pour comprendre ce que sont les thesaurus, qu'ils se caractérisent par une description conceptuelle à deux niveaux : une description conceptuelle par le « terme » propre, mais aussi par sa mise en relation avec les autres termes ; c'est à dire que le sens d'un terme naît à la fois de sa « composition » linguistique, mais aussi dans sa relation avec les autres termes. La mise en réseau d'une unité linguistique est donc constitutive de son sens.

Le thesaurus se compose essentiellement par la relation hyponyme/hyperonyme, plus souvent appelée la relation « générique/ spécifique ». Bien entendu, il peut apparaître d'autres types de relation, mais notre objectif n'est pas de décrire de manière exhaustive le mode de fonctionnement relationnel du thesaurus, mais de voir quelle peut être une relation pour comprendre que celles-ci sont véritablement constitutives du sens du terme. Hyponyme et hyperonyme définissent les relations d'infériorité et de supériorité sémantique que deux termes entretiennent entre eux. Ainsi, on peut dire que *chien* est hyponyme d'*animal* et qu'*animal* est un hyperonyme de *chien*. D'une manière peut être moins technique et moins formelle, nous pouvons dire que la relation « générique/spécifique » équivaut à la relation « est un type de » (uniquement dans le sens hyponymique) : *chien* « est un type » d'*animal*. De même , l' « Euro » ou les « 35 heures » constituent des types de « mutation conjoncturelle » ( l' « Euro » ou les « 35 heures » sont deux hyponymes

de « **mutation conjoncturelle** »). Un autre exemple, la « **maltraitance** » constitue un hyponyme de « **problème familial** » c'est à dire que la « **maltraitance** » est un type de « **problème familial** », ect...

#### 2-1-1-2-2 MISE EN RESEAU & COGNITIVITE

Si l'on veut faire une représentation conceptuelle d'un domaine, il faut donc sans conteste l'organiser sous forme de réseau sémantique. Nous avons dit dans la partie précédente que l'expertise du sens d'un terme doit tenir compte de sa « manifestation » linguistique, mais aussi de la relation qui l'unit avec les autres termes. Ainsi, cela signifie que la relation nous éclaire sur le sens du concept, ou en d'autres termes, la notion nous est compréhensible par le biais de cette mise en réseau. Ces relations notionnelles, puisqu'elles nous renseignent sur le sens des termes, devient alors un vecteur essentiel de la compréhension de l'utilisateur. Par extension, nous pouvons dire que si le concepteur veut que l'utilisateur ait une véritable appréhension et compréhension du domaine (qui est exprimé par le thesaurus), il faut qu'il lie les notions entre elles. En cela, nous pouvons alors suivre Felber quand il nous dit : « *Un domaine (ou une sous section de domaine) n'est accessible mentalement que si le champs notionnel est structuré, c'est à dire s'il constitue ce que l'on appelle un système de notions. Dans cet ensemble, chaque notion révèle ses rapports avec les autres notions.* » (Felber, *manuel de terminologie*, 1987).

Cela nous permet de justifier un choix qui nous semble important dans l'élaboration de notre thesaurus : c'est la présence de la liste hiérarchique systématique dans notre thesaurus, qui pourrait sembler ne pas être pertinente dans le contexte de recherche d'informations qui nous intéresse ici. En effet, il faut pour cela commencer par rappeler ce contexte de recherche informationnelle : l'agent chargé de formation reçoit une demande de formation, il cherche alors ce stage pour voir si un organisme le dispense (et donc s'il existe dans notre thesaurus). Qu'elle va être sa démarche ? il va se faire une liste de mots-clés qu'il va considérer comme différentes « entrées-matière » éventuelles et va donc tout naturellement se reporter à la présentation alphabétique structurée permutée (générale ou particulière selon sa

dimension de connaissance sur le domaine) qui permet une telle recherche alphabétique. A priori, la présence de la présentation hiérarchique n'a donc pas lieu d'être. Mais ce n'est qu'une impression car seule cette liste permet à l'utilisateur d'élever son degré de connaissance (par son apport cognitif certain) : or, cette progression apparaît comme un moyen d'accroître son efficacité et d'optimiser son travail, c'est à dire d'être plus à même de remplir correctement la mission qui est la sienne. Mais il peut aussi rencontrer des situations concrètes qui lui imposeraient de consulter cette liste. En effet, si la demande de formation qu'il reçoit lui apparaît obscure, il peut consulter le mode systématique de présentation pour bien circonscrire la notion extraite de cet intitulé de stage. N'oublions pas que cette mise en réseau renseigne l'utilisateur sur le sens du terme (et donc du domaine) : elle semble donc essentielle pour une représentation conceptuelle complète et cohérente. En effet, quand la manifestation linguistique du concept ne permet pas une bonne identification du sens (comme ici dans notre exemple) il faut faire appel à son contexte conceptuel, c'est à dire à son réseau sémantique.

## 2-2 LE THESAURUS DES « DOMAINES DE FORMATION » : UN THESAURUS « DOCUMENTAIRE » ?

Pour déterminer la nature du thesaurus, après avoir vu son organisation conceptuelle (c'est à dire comment il représente et détermine les concepts, comme nous venons de le voir), il faut se pencher sur sa finalité : que vise-t-il ? Quelle est son utilité première ? Mais pour permettre une bonne circonscription de ses objectifs, il faut opérer d'abord à un recadrement fonctionnel théorique, et voir comment il se situe par rapport aux traditionnels « thesaurus documentaires ».

### 2-2-1 APPROCHE TELEOLOGIQUE

#### 2-2-1-1 RECADREMENT THEORIQUE

- **AFNOR 1981** : « *Un thesaurus est une liste d'autorité organisée de descripteurs et de non descripteurs obéissant à des règles terminologiques propres et reliés entre eux par des relations sémantiques (hiérarchiques, associatives, ou*

*d'équivalence). Cette liste sert à traduire en un langage artificiel dépourvu d'ambiguïté des notions exprimées en langage naturel. »*

- *« Une liste structurée de concepts, destinés à représenter de manière univoque le contenu des documents et des questions dans un système documentaire déterminé, et à assister l'utilisateur dans l'indexation des documents et des questions. » (Van Slype, 1987 dans Hudon, 1994)*

Que nous apprennent ces définitions sur la finalité des thesaurus, traditionnellement « documentaires » ? Il est donc utilisé pour le contrôle et la structuration du vocabulaire, la cohérence de l'indexation et pour une meilleure performance lors du repérage (c'est à dire réduire le plus possible le bruit et le silence). Le thesaurus représente alors un moyen de garantir une meilleure indexation (cohérence, prévisibilité, rappel et pertinence) qui est, rappelons le, une représentation conceptuelle du contenu des documents et qui vise le repérage de l'information contenu dans ce même document - N'oublions pas, de plus, que les thesaurus se destinent en priorité au documentaliste chargé de l'indexation -. Si l'on veut alors accéder à cette information, il faut se reporter au document. C'est donc un outil de récupération de l'information via le retour vers le document : il constitue une source documentaire (et informationnelle) secondaire au même titre que les notices bibliographiques. Il ne permet donc pas un accès direct à l'information mais un moyen de récupérer plus sûrement cette information.

Or, avec le changement de paradigme, les objectifs ont changé : on ne vise plus l'élaboration d'outils qui fournissent les moyens de récupérer l'information, mais la récupération et l'accès direct à cette information. On veut apporter la réponse à la question de l'utilisateur, alors qu'avant on lui apportait les moyens d'y répondre par lui-même. L'analyse de la finalité de notre thesaurus viendra nous montrer où il se situe par rapport à ces deux pôles : outil d'accès direct à l'information ou outil de récupération, et donc d'accès indirect à l'information ?

### 2-1-1-2 UN OUTIL « HORS NORME » ?

En guise d'introduction, il convient de souligner que nous entendons le terme « hors norme » en son sens premier, c'est à dire comme se soustrayant à toute forme de normes.

Comme nous le précisions plus haut, pour que l'information puisse éclore, il fallait nécessairement que soient articulés deux entités distinctes : l'intitulé de stage (élément thématique et sémantique) et le nom de l'organisme auquel il était rattaché. Sans cette collaboration, l'outil aurait été inutile à ses éventuels usagers pour lesquels ce travail a été réalisé. Ainsi, l'exigence était de pouvoir rattacher la formation à sa structure. Ce qui a été demandé est donc un produit qui avait comme caractéristique première la mise en forme et l'organisation de l'information, par le balancement entre ces deux pôles: de la seule consultation du thesaurus devait naître l'information. Il devait être une réalisation auto-suffisante, en ce que sa seule compulsation devait permettre de mener à bien toutes les étapes successives de la recherche d'information. Ce document constitue donc une source primaire, et se devait d'être un accès direct à l'information. L'indexation, qui a donc était faite sur la base des catalogues, n'avait pas alors comme objectif de renvoyer les demandeurs d'informations vers eux (dans un deuxième temps), mais, a contrario, de leur permettre d'éviter ce retour aux documents.

Le thesaurus ne constitue donc pas un outil de repérage de l'information, mais de création d'informations ; et cela par son unique utilisation. Or, il a été dit précédemment que les thesaurus documentaires traditionnels étaient en quelque sorte des instances de renvois qui avaient comme objectif de faciliter la localisation de l'information (à travers la localisation du document). Il est donc manifeste que le thesaurus des « domaines de formation » n'est pas un thesaurus documentaire : il ne renvoie pas son utilisateur à une quelconque autre source mais est un système complet. Il n'est documentaire que de par sa source. Nous pouvons donc dire que la finalité fait de cet outil un instrument qui se situe hors des normes préétablies. Il n'a

pas la même utilité qu'un thesaurus documentaire classique, ou, en termes de tous les jours « *il ne sert pas à la même chose* ».

Nous pourrions conclure ce point par une petite remarque méthodologique. La notion de « rapidité » est un élément prégnant dans l'évaluation d'un mode de recherche d'information, nous avons essayé de mettre en pratique cette idée dans l'élaboration de notre thesaurus. En effet, il est de coutume dans tout thesaurus documentaire de n'y inclure qu'une seule liste alphabétique structurée permutée ; ainsi, selon le nombre de descripteurs, cette liste peut s'avérer volumineuse. Ce volume peut de fait se révéler un obstacle à une consultation rapide et donc optimale. Nous nous sommes efforcés donc de produire deux types distincts de listes permutées : la première, classique, reprend tous l'ensemble des descripteurs et non descripteurs du domaine (utile pour un usager qui a une faible connaissance du domaine dans laquelle la demande se situe). La deuxième liste permutée constituée se divise en thèmes et a, donc, comme avantage de réduire considérablement le volume de la liste ; ce qui rend la recherche moins fastidieuse et plus rapide. L'ajout de cette originalité n'a donc pas comme unique avantage de lutter contre la polysémie et le caractère parfois approximatif de certains descripteurs, mais d'accélérer l'exécution de la recherche.

### 2-2-2- UNE NATURE HYBRIDE

#### 2-2-2-1 UN THESAURUS TERMINOLOGIQUE ?

De l'avis des spécialistes des sciences de l'information, la terminologie se rapprochent de la formation de langages documentaires tels que les thesaurus. En effet, **une terminologie** se définit comme « *l'ensemble des termes d'un domaine donné* » (Cabré, 1998), notion de « spécificité » qui est aussi présente dans les multiples définitions qui ont été donné des thesaurus documentaires. De plus, la terminologie se définit comme l'ensemble des principes et des fondements conceptuels qui régissent l'étude des termes. Cela nous renvoie à toute la problématique linguistique de la formation des thesaurus documentaire lors de l'indexation des termes et de la formation des descripteurs. De plus par nature, la terminologie est forcément

impliquée dans la formation des thesaurus car l'expression de n'importe quel concept se fait à l'aide des termes. En ce sens, la terminologie est au service de la documentation. Mais cessons là la liste des rapprochements qui se révélerait relativement longues. Pour Louise Larivière, les terminologies et les thesaurus documentaires ont suffisamment de points de concordances et sont suffisamment proches pour qu'on puisse créer un répertoire commun c'est à dire un outil « *qui servirait tant à l'indexation de documents, qu'à la consultation des données terminologiques* ». Ainsi, Louise Larivière formalise une idée qui est présente dans les esprits depuis plusieurs années : mêler la spécificité des terminologies avec celle des thesaurus. Au départ, sans doute était ce pour gagner du temps et produire un instrument plus complet, mais depuis l'idée a fait son chemin. En effet, il se peut que par le biais de l'informatisation, les usagers auront tendance à exiger des outils nettement plus souples et plus clairs que ce langage documentaire, qui peut paraître obscur aux utilisateur néophyte. Cette idée, soyons en sûr, est le fruit de la demande de ces usagers qui réclament des outils plus proches de la réalité qu'ils côtoient et qu'ils pourront consulter sans l'aide de l'intermédiaire qu'est le documentaliste. Remarquons, en guise d'appartée, que cette notion d'autonomie de la recherche est prégnant dans le nouveau paradigme informationnel.

Cependant, notre thesaurus des « domaines formations » concorde t il avec cette conception ? Au départ, les langages documentaires étaient utilisés, à travers les classifications, pour notamment l'organisation physique (le classement) des documents dans les bibliothèques. A l'heure actuelle, il a plus comme finalité d'organiser les différents savoirs de l'homme, organisation qui passe encore par le truchement du classement physique. Ainsi, le thesaurus documentaire (et l'ensemble des langages et des pratiques documentaires) renvoie à l'idée du document : répétons le, il s'agit d'un outil de repérage de l'information par le biais du retour au document. Or, nous avons aussi dit que ce détour au document n'était pas la finalité de notre thesaurus (qui se veut un système autonome et clos), cet outil se veut « organisateur du savoir universel » (tout comme les thesaurus traditionnels), mais uniquement de manière conceptuelle...c'est à dire à la manière d'une terminologie. Vu sous cet angle, notre thesaurus s'apparente plus à une terminologie qu'à un

langage documentaire. Louise Larivière proposait l'idée de la création d'un thesaurus à consonances terminologiques, nous nous avons élaboré une terminologie avec des résonances de thesaurus (par le mode de présentation, de mise en relation des concepts...). Ainsi, nous avons bien élaboré un outil mêlant approche terminologique et documentaire, donc bien un « thesaurus terminologique », mais dans un sens différent de celui entendu par Louise Larivière.

[Nous développerons l'aspect communicationnel inhérent à la terminologie dans une prochaine partie et, verrons dans quelle mesure notre thesaurus participe à ce critère téléologique de la terminologie.]

#### 2-2-2-2 COLLABORATION ENTRE ELEMENTS LINGUISTIQUES ET EXTRA-LINGUISTIQUES

Nous avons déjà évoqué le fait que c'était de la collaboration de ces deux types d'éléments que naissait l'information. Mais il convient de voir comment elle se formalise dans le thesaurus des « domaines de formation ». Ce langage documentaire (?) s'oriente autour d'une approche linguistique d'où ressortent les descripteurs, qui reste la matière première de cet outil. Ces unités linguistiques occupent une fonction d'organisation du domaine décrit. En effet, il s'agit ici d'un agencement conceptuel de ce même domaine, véritable vecteur d'appréhension de celui-ci par l'utilisateur (voir la partie 2-1-1-2-2).

A ces indicateurs linguistiques, sont adjoints un nombre de codes numériques qui renvoient à une liste située en annexe du thesaurus. A chacun de ces chiffres correspond un organisme bien précis. Ainsi, si nous prenons par exemple, le terme (ou la formation) « **Maintenance informatique** », nous pouvons voir que ce descripteur est accompagné des codes « 7-116-158-451 » ; en se reportant au récapitulatif des organismes de formation, nous pouvons voir que la formation « **Maintenance informatique** » est dispensé par *ACTIF, CESI, CNFPT, PACEI* qui sont tous des organismes de formation. Par ce biais, nous pouvons répondre à toute question du type « *quel organisme propose le stage... ?* » ou « *est-ce qu'un organisme propose le stage... ?* ». Ce sont ces deux types de questions qui ralentissaient beaucoup

les chargés de formation, ne disposant pas d'outil permettant la liaison de ces deux entités entre elles : la réponse à cette question ne pouvait passer que par le biais de leurs mémoires. Bien entendu, nous avons appliqué cette liaison à chaque mode de présentation, pour que l'utilisateur quelque soit son mode d'entrée débouche sur une information complète (Annexe 4). Cependant, nous sommes obligés de regretter le peu de temps que nous avons disposé à l'élaboration de ce thesaurus : pour rendre le système le plus complet possible, nous aurions aimé y rajouter d'autres données (comme une fiche signalétique des différents organismes, par exemple).

Cette collaboration ne représente pas uniquement la condition de possibilité de l'émergence de l'information, mais constitue une réelle originalité par rapport aux thesaurus traditionnels, qui ne renferment pas d'articulation entre deux types de données distinctes (ou très rarement). Cette exigence était avant tout dictée par un besoin informationnel précis et spécifique qui ne pouvait se satisfaire qu'en prenant certaines libertés vis à vis des normes d'élaboration de thesaurus.

### CONCLUSION - TRANSITION

*Les exigences informationnelles particulières ont fait que le thesaurus des « domaines de formation » a du subir des adaptations par rapport aux normes d'élaboration des thesaurus encore en vigueur. Ce qui au départ était perçu comme un outil purement linguistique d'indexation de documents et de repérage de l'information, a du s'adapter au contexte de recherche d'information. En effet, c'est toujours l'exigence contextuelle d'information qui donne les directions à suivre pour l'élaboration d'un tel outil (et non l'inverse), car il est issu de cette exigence, il intervient et est crée pour venir palier les insuffisances des systèmes et outils déjà existants. Devant les flux exponentiels d'information et l'avènement des nouveaux moyens de communication, l'importance du travail linguistique d'organisation des connaissances est exacerbée : il reste le meilleur moyen de communication et de structuration de l'information. Cependant, la montée en puissance des systèmes informatiques de gestion des connaissances implique deux variations. La première est que tous ces systèmes langagiers doivent s'adapter aux nouveaux supports et à l'écriture sur écran. La deuxième est que les progrès technologiques, non contents de permettre un accès plus rapide à l'information, permettent à l'homme de penser de nouveaux outils de*

*traitement des informations, et ainsi, développer de nouvelles pistes de travail et de nouvelles applications à ces approches linguistiques et terminologiques.*

### 3- SYSTEME TERMINOLOGIQUE & APPLICATIONS LINGUISTIQUES

#### 3-1 NOUVEAUX ENJEUX DE LA MAITRISE TERMINOLOGIQUE

##### 3-1-1 DIMENSION COMMUNICATIONNELLE DE LA TERMINOLOGIE

Personne n'ignore, notamment avec le tapage médiatique autour d'Internet, que nous sommes entrés dans une nouvelle société de communication. Les nouveaux moyens de communication et de transmission d'information (les fameux NTIC) permettent une circulation plus rapide des messages, et donnent l'impression de ce fait d'une optimisation de la communication. La mode est maintenant à la communication...La nouvelle quête du Graal consiste à chercher tous les moyens possibles et imaginables pour améliorer sa communication. Tout a fait logiquement, on a donc vu se développer, notamment dans l'industrie, de nouveaux outils terminologiques : les bases de connaissances terminologiques (les BCT). Mais il convient de revenir sur la dimension communicationnelle de la terminologie afin de bien comprendre de quelle manière elle peut s'inscrire dans ce nouveau paradigme du « hourra - communicationnel ».

Comme nous l'avons évoqué de nombreuses fois, le terme est une unité linguistique servant à désigner une notion (ou concept) qui lui préexiste. Or, Les problèmes rencontrés lors de la formation des termes (ou de l'indexation) n'est pas nécessairement le problème de savoir si le signe renvoie de manière cohérente à son concept. En effet, la problématique inhérente à toute indexation constitue deux faces distinct du même aspect : un véritable terme doit à la fois correctement circonscrire la notion, mais en même temps être compréhensible pour celui qui sera amené à manier ce terme. Tout le paradigme orienté-usager repose sur ce problème et vise, à l'aide d'une meilleure connaissance de l'usager, à répondre au critère de compréhensibilité du terme et du système, en plus général. Il faut donc en quelque sorte qu'il y ait un échange possible entre deux usagers utilisateurs du système

linguistique, c'est à dire qu'il faut qu'il y ait communication possible. En effet, les termes sont avant tout des unités mises au service des locuteurs, et donc leur fonction première est de servir la communication et la circulation d'une information. Tout travail de traitement linguistique et terminologique vise donc à l'amélioration de la communication en axant son travail sur sa composante première : le langage (ou plutôt les langages). De plus, remarquons à titre révélateur, que La bonne communication d'une connaissance (ou d'une science) passe par un travail important sur sa terminologie, c'est notamment ce que Maryvonne Holzem affirme dans son habilitation au doctorat (1997). Il n'est donc pas étonnant de voir autant d'ouvrages consacrés à la vulgarisation des sciences qui insistent sur le rôle et l'importance de leurs terminologies pour leur bonne inscription dans la sphère sociale, inscription qui reste la condition de leur pérennité.

De même, toute notre entreprise vise cette facilitation de la communication entre individus, Que ce soit pour la collaboration entre les agents du service « formation », ou idéalement, entre les agents demandeurs de formation et ces mêmes chargés de formation. Si tous les services (entendre toutes les directions de service qui transmettent les demandes de leurs propres agents ) venaient à formaliser leurs requêtes de formation à l'aide du thesaurus des « domaines de formation », les agents chargés de formation verraient leur tâche grandement simplifiée. C'est dans cette optique intra et inter services que nous nous sommes efforcés d'élaborer notre thesaurus, et c'est, en outre, pour cette raison que la formation des descripteurs a été au centre de toutes nos préoccupations.

### 3-1-2 IMPLICATION DE LA TERMINOLOGIE DANS L'ENTREPRISE

L'information étant une donnée centrale dans l'entreprise, sa gestion (le *Knowledge Management*) et sa communication se révèle de fait comme des vecteurs de progression et de productivité de leur inscription dans l'économie de marché. Ainsi, l'investissement dans la construction de produits terminologique peut être un atout dans le développement des affaires. En effet, l'élaboration de bonnes bases de connaissances permettra, par exemple, que la « rédaction » d'un produit

commercialisé (c'est à dire toute la dimension marketing de mise en valeur du produit) soit plus précise et mieux ciblé vers le groupe d'acheteurs ou d'utilisateurs potentiels visé par le produit. Même si les consommateurs sont de mieux en mieux formés, ils continuent de n'acheter que ce qu'ils comprennent c'est à dire ce qui se sentent capables d'utiliser à bon escient.

Cette première partie considérait la dimension externe et commerciale de l'information (c'est à dire sa valeur stratégique). Il convient de voir maintenant que la terminologie peut aussi se révéler un facteur d'optimisation et de facilitation de sa production. John Humbley, nous fait, pour étayer cette conception, la description d'une entreprise qui a intégré la terminologie au sein de sa production :

*« Au symposium d'INFOTERM de 1998 il a été question du rôle que joue le rôle de la terminologie dans la documentation des entreprises. Un témoignage qui a retenu toute notre attention concerne le cas d'une entreprise autrichienne de taille moyenne engagée dans la production industrielle. Depuis quelques années, cette entreprise intègre la terminologie au sein de la production, dès la conception d'un produit, puis dans toutes les étapes de sa production. Cette démarche permet de désigner les mêmes choses par les mêmes noms et facilite grandement la rédaction de toutes sortes de documents : spécifications techniques, matériels publicitaires, description de produit, manuel d'utilisation, catalogues de pièces détachées, notices de programmation, ect. Non seulement l'entreprise fait des économies en réalisant dès le départ des documents efficaces, elle prépare le terrain pour la traduction des documents dans la langue de le ses clients. L'outil qui permet ce suivi terminologique est une base de données tout à fait classique : l'originalité est son intégration à toutes les étapes du processus productif. »*

Nous voyons donc ici toute la pertinence de cette démarche et de son inscription dans le processus de production. En fondant une approche toute terminologique, cette entreprise autrichienne a compris une chose essentielle dans sa gestion informationnelle : c'est que la terminologie (par son approche de description de domaines de spécialité) permet le retour vers l'information utile c'est à dire le retour vers une conception de l'information appliquée et opérationnelle. En ce sens, la

gestion de l'information n'est plus considérée comme une simple fonction de stockage, mais comme une démarche pragmatique qui doit s'appliquer (et s'impliquer) dans le contexte qui l'a vu naître, et permettre ainsi une certaine dynamisation de la production.

### 3-2 TERMINOLOGIE & LINGUISTIQUE DANS LE CONTEXTE DES « NTIC »

En guise d'introduction, nous pourrions citer Jean-Claude Corbeil qui était cité dans l'ouvrage de Maria-thérèse Cabré :

*« L'information s'est glissée au cœur de l'avenir, comme matière première stratégique. Elle est servie par une technologie informatique qui en favorise le stockage, le traitement et la diffusion au moyen de logiciels de plus en plus diversifiés et performants qui fonctionnent tous grâce à des interfaces en langue naturelle. La terminologie est partout présente dans cet univers. Dans le vocabulaire des interfaces. Dans les systèmes de classement de l'information, qui utilisent le plus souvent un corpus de mots clés pour indexer les documents. Dans les systèmes de repérage de l'information, soit à partir des termes des textes, soit à partir des mots clés d'indexation. Dans les systèmes experts, où la réponse aux questions formulées par l'utilisateur exige un traitement préalable des informations spécialisée engrangée dans le système pour faciliter le passage entre les termes de la question et ceux des textes qui y répondent le mieux. Tous ceux et celles qui façonnent ces nouveaux accès à l'information sont des créateurs et consommateurs de terminologie. »*

#### 3-2-1 LA PARADIGME USAGER

##### 3-2-1-1 LA SUPPRESSION DES INTERMEDIAIRES

##### 3-2-1-1-1 LES INTERFACES INFORMATIQUES

Selon le Petit Larousse, les interfaces informatiques se définissent des « frontières(s) conventionnelle(s) entre deux systèmes ou deux unités, permettant des échanges d'information » c'est à dire que c'est un système qui permet la communication entre deux entités l'homme et le système de traitement des informations. De leur compatibilité dépendent la facilité d'interrogation, mais

surtout, la réussite de la recherche d'information et la satisfaction du besoin d'information. Tout interface doit donc se baser sur une modélisation des usagers qui seront amenés à utiliser le système. Toute élaboration d'interface informatique doit alors se moduler par rapport au profil cognitif du public vers lequel il se destine, et suivre une méthodologie de mise en forme des différentes fonctionnalités et des modes d'interrogation qui soit basée sur cette approche contextuelle. Ainsi, une interface est dite « flexible » quand elle se base sur cette approche orientée-usager, et rend possible l'utilisation de cette interface par plusieurs types d'utilisateurs (Yves Le Coadic). Par contre quand les différents systèmes d'information ont été créés, ils ont intégré des interfaces qui étaient destinés aux spécialistes de l'information-documentation, et toute recherche devait se faire par l'intermédiaire de ce spécialiste. Cependant, le nouveau paradigme des sciences de l'information (lequel intègre le « paradigme usager ») tend à prendre le contre coup de cette conception élitiste de l'interrogation d'un système d'information. La volonté qui habite aujourd'hui les professionnels de l'ingénierie de la connaissance est de permettre une interrogation de ces systèmes par tout le monde (et donc de supprimer l'existence de ces intermédiaires), et donc de créer des interfaces plus « conviviales » (terme connu et maintes fois répété pour qualifier la tendance des interfaces actuelles)

Ce qui est mis en jeu lors de l'interrogation du système (et donc par le passage par ces interfaces) est la réussite de la tâche que l'utilisateur se fixe. Ainsi, l'interface détermine les conditions du travail de tout interrogateur. L'évaluation de ces conditions de travail est l'objet d'étude de l'« ergonomie linguistique » qui a comme objectif d'« *étudier ces conditions de travail afin de mieux spécifier les attentes ou les besoins des utilisateurs et d'améliorer les produits.* » (Yolla Polity). La présence du terme « linguistique », dans l'appellation de cette nouvelle discipline de recherche, est expliqué par le fait que le langage tient un rôle important dans la relation de l'utilisateur au système d'information. En effet, tout n'est pas simplement l'objet de disposition spatiale de représentation des tâches possibles sous la forme d'icônes.

Cependant, l'élément linguistique qui détermine le plus la consultation des systèmes (et donc détermine la satisfaction ou non du besoin d'information) reste, bien entendu, le langage d'interrogation...

### 3-2-1-1-2 LA PERTE DE VITESSE DES LANGAGES D'INTERROGATION

Avec l'augmentation exponentielle des flux d'information, chaque individu est beaucoup plus enclin que précédemment à la consultation de base de connaissances : ces outils deviennent des médias incontournables dans la gestion de l'information (et pas seulement dans le stockage). De fait, le besoin d'adapter ces systèmes d'information à ses éventuels usagers devient essentiel : l'interrogation ne peut plus passer par l'intermédiaire des documentalistes qui doivent eux aussi s'organiser devant cette montée en puissance de la quantité des informations à traiter. On a donc assisté depuis plusieurs années à un décalage dans les modes d'évaluation de ces systèmes d'information : la pertinence de la réponse proposée par le système n'est plus le seul critère, existe maintenant l'exigence de l'adéquation et de l'adaptation de ces outils aux différents utilisateurs potentiels. C'est ce que l'on appelle le « paradigme usager » qui considère ici « *l'utilisateur, son comportement, son intention comme élément central* » (Widad Mustafa Elhadi). Pour ce faire, on cherche à simplifier les modes d'interrogation de ces bases de connaissances, et notamment son primat essentiel constitué par les langages d'interrogation. En effet, antérieurement, ces langages étaient d'une grande complexité car destinés avant tout à des professionnels. Les systèmes d'informations dits « traditionnels » possédaient des interfaces à « langage de commande » qui exigeaient de la part de son utilisateur un grand niveau de compétences : il devait notamment maîtriser une syntaxe élaborée (troncature, opérateurs booléens...).

Cependant, bien que ces systèmes n'étaient pas flexibles, le fort contrôle et la forte normalisation de ces modes d'interrogation en faisaient, dans les mains d'un professionnel, un outil tout de même efficace. Mais étant donné que dans le contexte actuel, chacun est susceptible d'interroger des bases de connaissances, les recherches d'information ne pouvaient s'opérer par le biais de ces langages d'interrogation. S'est présenté alors une alternative : soit on formait tous les usagers à ces modes d'interrogation, soit on adaptait l'outil à ses utilisateurs potentiels c'est à dire qu'on permettait une interrogation dans un mode moins contrôlé, c'est à dire dans un mode

d'interrogation plus libre. Le choix fut immédiatement fait. Mais comme nous l'avons dit précédemment, la normalisation et la rigueur de ces interrogations en faisait un instrument précieux car tout problème de recherche d'information est un problème d'appariement entre les termes de la requête et ceux qui ont servi à décrire les informations contenues dans le système. Comment alors garantir une réponse pertinente par le biais de mode d'interrogation libre, c'est à dire comment satisfaire ce critère d'appariement, condition sine qua none de l'adéquation de la réponse avec la question ? C'est l'objectif que se fixe les professionnels du traitement automatique des langues naturelles (TALN), qui veulent rendre possible une interrogation en langue naturelle, sans pour cela perdre la dimension de la pertinence de la réponse.

### 3-2-1-1-3 LE TRAITEMENT AUTOMATIQUE DES LANGUES NATURELLES

Si nous avons décidé de placer cette partie à ce niveau, c'est pour bien marquer le fait que le traitement automatique est né de cette problématique usager et que sa complexité théorique révèle paradoxalement une volonté de simplification des tâches des utilisateurs. Cette approche concoure à l'idée de permettre à tout utilisateur une certaine autonomie de recherche, tout en le satisfaisant quant à la réponse qu'il obtient.

D'emblée, nous devons faire une distinction. En effet, Le traitement automatique des langues naturelles comporte deux volets qui n'ont pas les mêmes objectifs. Le premier consiste dans l'indexation automatique (ou extraction automatique des unités linguistiques) et le deuxième vise à l'interrogation de bases de connaissances en langues naturelles. Cette interrogation consiste, pour le système, à un calcul de proximité sémantique entre une requête et un ensemble de textes : il s'agit donc d'un principe de reformulation. Mais nous devons préciser que ces deux systèmes se basent sur les mêmes outils qui sont des analyseurs sémantiques et linguistiques et qui sont, en d'autres termes, des cellules interprétatives. En effet, le traitement linguistique permet « *d'aider au choix dans la base, d'interpréter les questions, de les traduire dans le langage d'interrogation du serveur, il prend quelques fois en compte les problèmes de reformulation qui permettent l'assimilation des mots utilisés dans la*

*question à ceux utilisés pour décrire les mêmes choses dans les textes* » (Christian Fluhr). Il convient maintenant de voir comment procède ces analyseurs et de dégager les différentes étapes dans l'élaboration de ce nouveau type de produits. Cependant, il faut mettre en garde sur la finalité de cette présentation. En effet, celle-ci consiste à voir en quoi ce nouveau travail sur la linguistique vient enrichir les outils déjà existants, c'est à dire de quelle manière ces nouvelles applications relèvent d'un nouveau statut de la linguistique dans la société de l'information. Cet exposé ne consiste donc pas en une présentation exhaustive de la technicité de ces procédés informatiques de traitement linguistique.

Au départ, le traitement des données textuelles était basé sur des méthodes statistiques. Il reposait sur la simple occurrence des mots dans l'ensemble du corpus. On a depuis intégré d'autres éléments statistiques comme la fréquence dans l'ensemble du corpus, la régularité de la répartition. Ces méthodes visaient à « pondérer » les termes dans leur capacité à décrire les documents du point de vue conceptuel. Par la suite, on y a ajouté des traitements morphologiques sur les radicaux (la morphologie est l'étude des formes des mots – flexion et dérivation – par opposition à l'étude des fonctions ou syntaxes). Cependant, ces méthodes de « comptage » n'étaient pas rigoureuses car la polysémie est très présente et tout dépendait donc du sens que l'on mettait derrière la déclinaison particulière d'un terme. On a donc cherché à résoudre ces problèmes contextuels par l'intégration d'instruments sémantiques : par exemple, les thésaurus avaient comme objectif d'effectuer des regroupements sémantiques. Pour pallier aux insuffisances de ces méthodes statistiques (même si on lui a assigné des éléments linguistiques comme l'approche morphologique), on a appliqué des connaissances syntaxiques qui avaient comme objectif de détecter les syntagmes nominaux (le syntagme est l'unité syntaxique élémentaire). La multiplication de ces approches avait comme conséquences de résoudre des problèmes dans le traitement linguistique automatique comme notamment celui de la synonymie (c'est l'approche qui est adoptée par SPIRIT qui utilise les mêmes protocoles que le moteur de recherche ALTAVISTA). A l'heure actuelle, on essaie d'enrichir encore ces outils car l'ingénierie linguistique et l'ingénierie de la connaissance sont des domaines de recherche où la

quête du progrès est constitutive. Par exemple, on renforce le traitement par l'intégration de thesaurus de plus en plus complets et de plus en plus conséquents. L'objectif de tout ce travail est de rendre les systèmes de traitement de plus en plus indépendants, c'est pour cela que l'on parle aujourd'hui d' « Intelligence Artificielle » (ou IA).

### 3-3 INTERNET : BIBLIOTHEQUE VIRTUELLE ?

#### 3-3-1 INTERNET OU LE CHAOS « NUMERIQUE »

L'avènement de la société de l'information et des nouvelles technologies de l'information et de la communication nous annonce une nouvelle ère médiatique. Internet s'inscrit pleinement dans cette conception : le réseau permettrait à tout le monde d'accéder au savoir (à tout le savoir). Cependant, tout le monde connaît la difficulté à trouver de l'information pertinente, aspect essentiel dans toute démarche de recherche informationnelle. En effet, comment se repérer dans une « banque de données » que certains spécialistes jugent comme contenant près de 40 milliards de sites ? les moteurs de recherche tentent, tant bien que mal, de satisfaire certaines requêtes, mais ils se trouvent rapidement dépasser par la masse d'informations à gérer... Une métaphore très présente dans cet univers est la comparaison du réseau avec le monde de la bibliothèque : certains nous affirment, convaincus, que ce nouveau mode d'accès au savoir constitue une véritable bibliothèque « virtuelle » où tout le savoir serait enfermé et accessible à tout à chacun, c'est à dire « *la mise à disposition du monde dans son exhaustivité* » (Yves Jeanneret). Ainsi, l'homme semble être parvenu, plusieurs millénaires après, au but que s'était fixé la bibliothèque d'Alexandrie : posséder et capitaliser le savoir universel !

Mais qu'est ce qui caractérise justement une bibliothèque ? Toute conception bibliothéconomique repose sur le principe simple de classement et d'organisation du savoir, visant ainsi un accès plus aisé à l'information. En effet, une bibliothèque n'en serait plus une si elle n'était qu'un amas de livres. L'organisation de l'information semble donc être la condition d'accès à l'information. En ce sens, Internet est il une bibliothèque ? Bien que certains annuaires thématiques (tels que Yahoo !) essaient de

donner une présentation ordonnée du réseau, personne n'ignore que, justement, le réseau ne reste qu'un amas de sites. Personne ne peut soutenir que l'organisation est l'apanage d'Internet ! Or par un principe syllogistique de base, Internet ne peut donc être une bibliothèque. Le réseau sera devenu une réelle source d'informations quand celui-ci aura subi un travail purement bibliothéconomique d'organisation. En ce sens, la linguistique et la terminologie (se définissant comme des activités de dénomination et d'organisation du savoir) ont certainement un rôle à jouer dans ce désordre numérique. Manifestement pour des futurs professionnels de l'information documentation, Internet représente un terrain de travail inépuisable. Mais est ce que le succès de ce nouveau média ne repose t-il pas justement sur cette obscurité qui favorise le cheminement et le plaisir de la découverte ? Quand on voit les désaffections des bibliothèques peut on vraiment souhaiter qu'Internet en devienne une ?

### 3-3-2 INSUFFISANCE DES MOTEURS DE RECHERCHE TRADITIONNELS

Cependant, il apparaît que la structuration des informations sur Internet via la terminologie et la linguistique et la terminologie est une cause perdue d'avance. En effet, Vouloir indexer tous les sites (Près de 40 milliards comme nous l'avons évoqué précédemment) consiste en un travail trop titanesque pour qu'il puisse être effectué dans l'état actuel des choses. De plus, l'arrivée d'une quantité exponentielle de pages sur la toile par jour et de manière totalement incontrôlée (et incontrôlable) rend l'entreprise totalement impensable. Il semble que, dans ce constat, les moteurs de recherche soient, dans l'idée, un palliatif intéressant. Mais il ne faut se tromper sur les capacités des moteurs : tant au niveau quantitatif, qu'au niveau qualitatif, ils ne semblent pas qu'ils soient aussi performants que tout à chacun serait en droit d'espérer. En effet, le « bruit » (sélection de documents non pertinents) et le « silence » (absence de documents sélectionnés pourtant présents sur la base) semblent être incontournables tant la quantité et la complexité des informations est grande.

Mais pourquoi ces systèmes apparaissent ils si peu performants ? Une des raisons consiste dans le fonctionnement même des moteurs de recherche : chacun des moteurs est muni d'un « moteur d'indexation » qui effectue des opérations traditionnellement confiés au documentaliste (extraction, indexation...) de manière automatique. Cependant, ceux-ci indexent la totalité des mots, y compris les mots qui n'ont aucun intérêt à la recherche (ce que l'on appelle les « mots vides » : articles, adverbes, ect...). Les solutions envisagées (telle que la création d'une liste de mots vides qui ne ferait pas l'objet de l'extraction) n'ont jamais réussi à résoudre tous les problèmes : le problème de la synonymie et équivalents sémantiques, ainsi que ceux de la polysémie n'ont pas encore trouvé de solutions réellement satisfaisantes tant les problèmes sont nombreux.

Il existe une autre technique (celle là non textuelle) qui se développe depuis plusieurs années, c'est la technique des « métatags » (appelés aussi « métadonnées »). Celle ci consiste à établir des balises linguistiques permettant d'insérer dans la page des données (ou mots clés) qui ne servent qu'à l'indexation. Cette solution semblait à l'origine séduisante mais, laissés à la charge des concepteurs de sites, ceux ci avaient tendance à attirer le maximum de visiteurs sur leur site en employant des termes qui ne représentaient pas le contenu du site, mais avaient la simple fonction d'orienter les recherches des utilisateurs. Cette démarche qui vise à être sélectionné en employant des moyens purement attractifs et artificiels est appelé le « Spamming ».

Ainsi, les spécialistes remettent de plus en plus ce mode de recherche textuel (pour des raisons qui ont été cités précédemment), bien que les progrès soient, reconnaissons le tout de même, bien réels.

### 3-3-3 DE NOUVELLES PISTES DE TRAVAIL...

Un des moteurs les plus connus (et sans doute un des plus performants) a été élaboré à partir d'un autre mode d'indexation des pages. En effet, Altavista n'effectue pas l'indexation page par page, mais à l'aide des liens hypertextes dont un grand nombre de pages est pourvu. Ainsi, Altavista indexe chacun des liens et, comme le dit Jean Marc Font, « déroule le fil d'Ariane de sites en sites ».

Poursuivant cette idée, les membres du projet « *Clever* » se basent aussi sur l'étude de ces liens (qu'ils appellent « hyperliens ») pour créer une sorte de réseau sémantique entre ces pages. En effet pour eux, chaque lien est révélateur d'une proximité sémantique, entre ces deux pages mises en relation, qu'il convient d'exploiter. Il s'agit ensuite de travailler sur l'évaluation de la pertinence pour pouvoir intégrer le critère de pondération.

En outre, afin de répondre le mieux possible à ces exigences de pertinence, ils utilisent en priorité les sites d'autorité car la validité de leurs liens s'impose d'elle même. Cette méthode élimine donc les sites à informations douteuses ou fausses (la validité de l'information constitue un grand problème). Cependant, c'est un moyen qui ne semble pas respecter le principe premier d'Internet qui est constitué par l'idée de démocratie...

## CONCLUSION

Les outils documentaires sont en proie à une double révolution. La première, rendue possible grâce aux progrès technologiques, est constituée par une adaptation nécessaire aux outils informatiques et à l'écriture sur écran. La deuxième, prenant en compte la nouvelle valeur stratégique de ce que l'on appelle « information », tend à vouloir adapter ces outils au changement de statut de sa matière première (l'information) et, donc, de coller aux nouvelles exigences informationnelles. Les normes traditionnelles « éclatent »... La démarche usager représente dans ce contexte le plus sûr moyen d'avoir une totale adéquation entre le système d'information et ces exigences (dictées par la nouvelle société de l'information mais surtout par les utilisateurs). Changement de fonds et de forme qui concourent à une idée simple : revenir à une information utile et opérationnelle. On constate alors un glissement dans la fonctionnalité des centres de documentation : la constitution de la mémoire de l'entreprise (le stockage) n'est plus leur priorité. La terminologie et le travail linguistique, en tant que disciplines pratiques de structuration et de traitement de l'information visant à un meilleur accès à l'information, ont donc été convoqué de droit et de fait dans la constitution de cette nouvelle société de l'information. En effet, leur capacité à donner une plus value informationnelle aux données a rapidement été utilisé par les spécialistes de l'information qui ont vite compris quels pouvaient être leurs apports. C'est dans ce contexte que le thesaurus de « domaines de formation » a été élaboré : outil profondément hybride qui a mélangé des approches linguistiques, terminologiques, cognitives ; qui a fait collaborer et cohabiter des éléments de natures différentes...et a de cette manière donc malmené volontairement les normes et les traditions documentaires. Toute sa conception reposait sur la satisfaction des exigences et des besoins informationnels des destinataires du dit thesaurus. Le thesaurus des « domaines de formation », ainsi que les autres thesaurus actuels, n'est pas un simple outil d'indexation, comme il l'aurait été dans le passé. C'est un instrument autonome et suffisant de recherche d'information. De même, les thesaurus en général voit s'opérer un glissement de

leurs fonctionnalités : ceux ci peuvent jouer le rôle d'interface entre l'utilisateur et le système d'information (mais ici, cela ne pouvait être le cas car il n'existe pas réellement de système de gestion de l'information), ou encore être intégrés dans les programmes de traitement automatique du langage... bref, les thesaurus actuels ne sont plus des thesaurus documentaires. Ainsi dans le nouveau paradigme des sciences de l'information, la terminologie et la linguistique ont subi des changements d'orientation de travail mais pas de profonds changements de nature. En effet, la langue reste toujours le vecteur essentiel de transmission et d'acquisition du savoir, et en ce sens, les changements de paradigmes ne pourront jamais changer cet état de fait : les applications linguistiques et terminologiques restent des données constitutives dans toute pensée sur l'information, quelque en soit son statut ou son inscription sociale. En faisant du traitement de l'information, la clé de voûte de son fonctionnement, la société de l'information donne au support de l'information c'est à dire la langue elle même, une importance nouvelle tant sur le plan de sa maîtrise que de son traitement par ordinateur.

## **BIBLIOGRAPHIE**

Marjanovic, Mirjana. *Le monde des thesaurus documentaires*.  
[www.Ceveil.Qc.ca/traitem/marjanom/def.htm](http://www.Ceveil.Qc.ca/traitem/marjanom/def.htm)

Mustafa Elhadi, Widad. *Penser les nouvelles fonctionnalités du thesaurus*.  
 Terminologies nouvelles, 2000, n°20. (A paraître)

Bertrand-Gastaldy, Suzanne. *L'évolution de la gestion de l'information documentaire sous l'impulsion des nouvelles technologies*. Terminogramme, mars 1990, n°55, pp 25-31.

Les membres du projet Clever. *Recherche intelligente sur l'Internet*. Pour la science, août 1999, n°262, pp 52-58.

Polity, Yolla. *Evaluation des modes de recherche en langage naturel*. Documentaliste-Sciences de l'information, 1994, Vol.31, n°3, pp 136-141.

Larivière, Louise. *Vers un produit unifié en terminologie et en documentation : le thesaurus terminologique*. Méta, 1989, Vol. 34, n°3, pp 457-468.

Le Coadic, Yves. *Le besoin d'information*. ADBS éditions. Paris : 1998.

Fluhr, Christian. *Le traitement du langage naturel dans la recherche d'informations documentaires*. In *Interfaces intelligentes dans l'information scientifique et technique*. Sous la direction de Christian Bornes. INRIA, Mai 1992, pp 103-130.

Chaumier, Jacques. *Le traitement linguistique de l'information*. Entreprise moderne d'édition. Paris : 1988.

Holzem, Maryvonne. *Apport des recherches en terminologie à la communication des sciences*. Thèse de doctorat, Université de Rouen, 1997.

Campenhoudt, M.V. *cours de terminologie*.  
[www.refer.org/termisti/termisti.htm](http://www.refer.org/termisti/termisti.htm)

Cabré, Maria Thérèse. *La terminologie : théorie, méthode et applications*. Traduit du catalan par Monique Cormier et John Humbley. Armand Colin. Paris : 1998.

Humbley, John. *Terminologie et documentation*. Acte de la conférence sur la coopération dans le domaine de la terminologie en Europe.  
[www.unilat.org/dtil/aet/HUMBLET.htm](http://www.unilat.org/dtil/aet/HUMBLET.htm)

U.E. 7 : *CONDUITE D'UNE ACTION D'ETUDE EN SCIENCES DE  
L'INFORMATION ET DE LA DOCUMENTATION*

***LES ENJEUX DE LA CONSTRUCTION DE L'USAGER***

Le discours ambiant est unanime : visiblement, nous serions entrés dans une nouvelle ère, celle de la communication et du partage universel du savoir. Cette entrée dans une nouvelle « société de l'information et du savoir pour tous » est essentiellement due aux progrès des technologies de la communication (les fameuses NTIC). Ainsi, la glorification de ces médias les présente comme les seuls responsables de la facilitation de la communication : l'élément le plus révélateur est la hausse considérable de vente d'ordinateur, chacun voulant participer à la construction de ce nouveau monde. On considère que l'outil seul est garant de cet état de fait, que c'est lui seul qui rend possible cette communication en mettant les individus en relation les uns avec les autres. Cette pensée technique est donc un élément central dans cette rhétorique communicationnelle, et représente somme toute un formidable retour en arrière dans la réflexion socio-culturelle sur les systèmes d'information et de communication. Une fois de plus, le message véhiculé considère que tous les problèmes seront résolus avec les progrès technologiques. De même, quand l'informatique est entrée dans la sphère de la documentation, on l'a présenté comme la solution idéale de stockage, de traitement et de recherche d'information. La hausse des quantités d'informations rend maintenant chaque personne susceptible de manipuler ces outils, mais les problèmes sont vite apparus et l'on a tôt remarqué que l'outil, seul, ne permettait pas tout, notamment quand on ne sait pas l'utiliser ! C'est à ce moment que l'on a découvert l'« Usager » et que les spécialistes ont compris qu'une pensée orientée uniquement vers l'objet technique ne permettait pas de satisfaire tous les éventuels utilisateurs : un outil reste un outil . La pensée technique pris peu à peu en compte la dimension de l'utilisateur, pour finalement en faire une donnée centrale et essentielle. Actuellement, on considère que toute pensée technique exige au préalable une étude de l'utilisateur. A travers l'exemple de notre mission, nous allons donc voir comment peut s'effectuer cette modélisation de

l'usager, et toute la légitimité épistémologique et fonctionnelle qui fonde une telle orientation. Cependant, il nous faut préciser le choix du bien fondé de cette problématique. En effet, nous avons maintes fois évoqué ce « paradigme usager » dans notre précédent exposé, et l'avons toujours considéré comme central, mais nous en faisons un postulat sans réel procédé démonstratif : notre objectif est donc ici de montrer quelles sont ses assises épistémologiques et méthodologiques qui anime cette approche « usager ».

### 1- LES MODES DE CONSTRUCTION DE L'USAGER

Le premier mode d'étude de l'usager est présentée comme un mode « armé » c'est à dire sous la forme de questionnaires ou de tout autre d'interrogation directe (entretiens). Cette démarche consiste à donner la parole de manière explicite à l'individu interrogé. Par la réponse à des questions « fermées » (c'est à dire que l'individu interrogé a le choix entre des réponses prédéterminées qu'il n'a plus qu'à cocher), l'examineur pose les questions qu'il juge comme essentielles à l'élaboration du produit dont il a la charge. Cela lui permet d'acquérir des données rigoureusement formalisées sur lesquelles il pourra se reposer dans son travail futur. Ce type de données est donc pourvu d'une certaine scientificité car il n'a pas à faire appel à son intuition qui reste malgré tout incontournable dans tout travail d'expertise. Le deuxième genre de questions est constitué par les questions « ouvertes » (ou à réponses libres). Celles ci permettent à l'individu de préciser explicitement quelles sont ses attentes : elles sont donc importantes à l'examineur pour la représentation qu'il peut se faire des orientations souhaitées par les futurs utilisateurs du système, et pour une deuxième reformulation de la mission qu'on lui confie. Cette reformulation est importante car elle est un autre point de vue c'est à dire qu'on lui présente le problème avec un autre angle visée. Elle est donc riche d'enseignements pour la représentation qu'il peut se faire de sa mission. Il convient donc de faire quand on entreprend une telle démarche de mêler les deux approches pour couvrir le maximum de champs et palier les insuffisances d'un mode de questionnement par l'autre (et vice versa). En effet, les questions fermées se révèlent

pertinentes mais leur principale insuffisance est qu'elles sont justement closes et donc ne laissant pas de marge d'expression, alors que tout questionnaire vise à cette expression de l'individu interrogé. Mais, cette faiblesse est comblée par le type « ouvert » du questionnement qui laisse la parole de manière complètement libre. Ainsi, nous avons fait le choix lors de notre mission de faire appel à cette méthode du questionnaire pour une ébauche du profil de l'utilisateur futur de notre thesaurus.

Cependant, cette méthode, bien que prépondérante dans la construction de l'utilisateur, pose plusieurs problèmes. Le premier est qu'en posant les questions, l'expert (au sens de « celui qui fait l'expertise ») oriente le débat dans un sens qui n'est peut être pas le plus légitime et le plus parlant. En effet d'une certaine manière, nous pouvons dire qu'en utilisant ce procédé, il cherche à cautionner de manière officielle les représentations qu'il s'était faites dans un premier temps : il cherche juste à valider ses premières impressions par l'intermédiaire des individus interrogés. Il faut donc voir dans quelle mesure l'expert ne connaît pas les réponses aux questions qu'ils posent. Bien sûr, la plupart du temps celui qui pratique cette étude n'a pas réellement conscience de cet aspect de son questionnement. Le deuxième problème est évoqué par Yves Le Coadic dans son « Besoin d'information » (1998). En effet, pour lui, l'utilisateur n'a pas suffisamment de recul pour exprimer une attente véritablement rigoureuse. Chose paradoxale, celui qui revendique un besoin ne parvient pas à l'exprimer de manière rigoureuse. Le Coadic va plus loin en nous disant que le besoin exprimé lors de cette interrogation ne correspond pas au besoin réel de l'individu interrogé.

Mais ce caractère flou peut être corrigé par une approche complémentaire, mais qui n'est pas formalisable de manière rigoureuse... et c'est précisément là que réside sa principale force. Cette démarche (qui n'en est pas réellement une) est dépendante de la communication interpersonnelle. En effet, pour recueillir le maximum de données, le questionnaire ne suffit pas : la discussion est très importante. Nous insistons bien sur le terme de « discussion » (cela n'est pas un entretien) : le tout consiste juste à laisser parler les futurs utilisateurs de leur perception et de leur souhait, mais tout en se garantissant de ne pas trop diriger la conversation, ni d'influer sur

le contenu des propos. Yves Le Coadic encore prétend que les usagers émettent quelques réticences à formuler leurs besoins informationnels, ce qui est tout à fait logique car il y a beaucoup d'éléments psychologiques qui rentrent en jeu quand on sait que nos réponses vont être analysées. Inconsciemment, il y a une démarche de défense qui se met en place. Le moyen de contourner cette difficulté méthodologique consiste donc à attendre que la discussion s'oriente dans cette direction. On pourrait opposer à ce mode de pensée qu'il ne présente pas le moindre caractère de rigueur. C'est à moitié vrai, et à moitié faux. Il faut comprendre dans quel contexte votre arrivée se présente : vous êtes là pour résoudre le problème des utilisateurs d'un système, chacun souhaite donc que votre travail soit efficace (car il doit logiquement se ressentir sur le travail de tout le monde). Les personnes à coup sûr viennent vous voir, mais à la différence du questionnaire, ce sont eux qui font la démarche d'en parler et, psychologiquement, cela compte énormément. On peut parfois apprendre beaucoup plus autour d'un café, qu'autour d'un questionnaire ! Bien sûr, il faut que le contexte soit en votre faveur : si votre présence n'est pas réellement souhaitée, le contact ne passera pas et il rend obsolète à l'avance cette démarche. En ce qui nous concerne, nous avons travaillé dans un contexte on ne peut plus idéal, tant et bien que nous avons déjà à l'esprit des données qui se sont révélées centrales. Nous n'exagérons pas quand nous affirmons que notre étude des usagers aurait pu quasiment se passer de questionnaires, tant la communication fut importante et riche.

## 2- DETERMINATION DU PROFIL COGNITIF :

Comme nous l'avons vu précédemment, l'étude de l'utilisateur à comme de pouvoir ainsi adapter l'outil réalisé à celui-ci. Mais de quelle nature est ce profil ? Pour cela, il faut comprendre la relation qui existe entre les usagers et les outils déjà présents. Traditionnellement, on présente le rapport « homme-système » comme une communication entre deux instances. Cependant, comme nous le montre Le Coadic, il s'agit plus d'un entretien forcé que d'une réelle communication ( « *c'est la machine qui a le dessus dans les systèmes humains/machines* », Habermas). Pour que le système

puisse fournir une réponse à la question de l'homme (car le contexte qui est le notre est un contexte de recherche d'information), il faut qu'il approprie le mode de fonctionnement de l'outil c'est à dire qu'il faut qu'il construise du sens à partir de ses connaissances préexistantes sur le rôle de ce même outil. Il s'agit donc d'un rapport d'appropriation du produit (la recherche d'information ne se fait pas toujours à partir d'un instrument informatique, comme c'est ici le cas) qui passe par l'intermédiaire d'une représentation de sa finalité et de son fonctionnement. C'est donc un principe cognitif qui relie ces deux instances, en ce sens, qu'une recherche d'information impose que nous sachions à la fois ce que l'on cherche mais que surtout, nous sachions comment nous le cherchons.

Ce qui est donc déterminé lors de l'élaboration du profil de l'utilisateur, ce sont ces deux choses (qui sont comme nous venons de le voir des éléments cognitifs) : la compétence dans le domaine (ou la compréhension qu'il peut faire de sa démarche et des instances convoquées lors de cette recherche) et le but qui anime cette recherche (ou ce qu'il cherche mais cet aspect étant déjà évoqué précédemment, nous ne l'évoquerons pas ici). Par exemple, dans l'élaboration de notre thésaurus des « domaines de formation », nous avons fait des choix qui reposaient sur ces deux principes : dans un premier temps, nous avons fait le choix de la précoordination pour rendre l'outil plus compréhensible c'est à dire que lors de son interrogation, il sera plus à l'aise si la modalité des réponses se fait selon la structure du langage naturel et non pas en termes d'équations booléennes en unitermes. Seuls les documentalistes ont été formés à penser de cette manière. Il faut donc qu'il appréhende la manière dont se déroule l'opération qu'il est en train d'effectuer : s'il ne l'appréhende pas, il ne sera pas capable de l'effectuer à bon escient et correctement. Il ne sera donc pas capable d'utiliser l'outil. Par cette approche d'étude des compétences, il faut faire le recensement des éléments de quotidienneté de l'utilisateur pour pouvoir les intégrer à l'outil. En d'autres termes, il faut essayer de présenter un système imprégné du quotidien des futurs usagers pour que ceux-ci ne doivent pas faire l'objet d'une éventuelle formation et qu'ils l'utilisent pratiquement immédiatement de manière productive. Quand on dit donc que l'outil doit s'adapter à l'utilisateur, on entend donc qu'il doit déjà faire partie du quotidien de celui-ci avant

même leur premier contact : c'est grâce à l'étude des structures cognitives (la démarche qu'il prend en situation de recherche et les éléments qu'il utilise) que nous sommes donc capables de définir ce profil cognitif de l'utilisateur. Il s'agit donc d'une sorte d'appariement de fonctionnement intellectuel entre un informé (l'homme) et un informant (le système d'information) qui est en jeu dans cet étude de l'utilisateur. Cet appariement constitue la condition d'une possible communication entre ces deux instances, ou, une possible prise d'information de l'homme par l'outil.

A cause de l'avènement de la société du savoir, il est aujourd'hui plus que jamais d'avoir rapidement accès à l'information en tout genre. Ainsi, les modes d'accès se multiplient, mais gardent tous présents à l'esprit le caractère central de l'utilisateur. En effet, pour s'assurer que les nouveaux moyens de communication et d'information permettent cet accès au savoir, plutôt que ne l'empêche, il convient de produire des outils réellement performants. Cette performance se mesure notamment par l'adaptation de ces moyens d'accès au contexte et à l'utilisateur. Cette démarche « usager » reste le plus sûr moyen, par une étude systématique de ses structures cognitives et des buts qu'ils expriment lors de sa recherche, de créer une adéquation entre l'utilisateur et le système employé. En effet, il faut permettre à l'utilisateur de se reconnaître à travers l'outil pour deux raisons : la première est que de ce fait il sera apte à l'utiliser et la deuxième est que de cette manière il ne le rejettera pas. Un profil cognitif semble donc être une étude préliminaire incontournable pour créer ce rapport de proximité et de quotidienneté entre ces deux entités. Ce qui est en jeu ici est donc bien la maîtrise des outils permettant l'accès à l'information ou, en d'autres termes, l'accès même à cette information. Or la maîtrise de cette information reste une habileté essentielle dans le contexte qui nous intéresse ici, et en ce sens, il faut que tout à chacun soit capable de la maîtriser : c'est pour cela que les approches « usagers » restent prégnantes dans la visée démocratique de cette nouvelle société du savoir.

## *ANNEXES :*

- Annexe 1 : sommaire de catalogue*
- Annexe 2 : questionnaire*
- Annexe 3 : retranscription des « codes secteurs »*
- Annexe 4 : collaboration descripteurs-codes numériques*

# *ANNEXE 1*



# CALENDRIER



Office  
International  
de l'Eau

CNFRUE

	MAI	JUIN	JUIL	AOU	SEP	OCT	NOV	DEC
--	-----	------	------	-----	-----	-----	-----	-----

## LES FORMATIONS TECHNIQUES

### ANALYSE DES EAUX

#### LES BASES

Notions de base sur la chimie de l'eau - 1<sup>er</sup> Degré (A.01)

#### EAU POTABLE

Analyses élémentaires d'une eau potable - 2<sup>ème</sup> Degré (A.02)

#### EAUX USEES

Analyse des eaux usées pour l'autosurveillance (A.04)

Analyse des eaux en laboratoire - 3<sup>ème</sup> Degré (A.05)

#### QUALITE

L'assurance qualité en analyse des eaux (A.08)

#### PISCINES

Traiter et contrôler les eaux de piscines (A.09)

#### BACTERIOLOGIE

Analyses élémentaires relatives à la bactériologie des eaux

1<sup>ère</sup> Partie (A.06)

2<sup>ème</sup> Partie (A.07)

Contrôle microbiologique : nouvelles méthodes d'évaluation des flores bactériennes (A.10)

Toxicité - Ecotoxicité (A.11)

### PRODUCTION D'EAU POTABLE

Exploiter les usines de production d'eau potable  
Initiation au traitement des eaux (B.01)

Exploiter les usines de production d'eau potable  
Perfectionnement 1<sup>er</sup> Niveau (B.04)

Exploiter les usines de production d'eau potable  
Perfectionnement 2<sup>ème</sup> Niveau (B.05)

Désinfection des eaux au chlore et au dioxyde de chlore (B.02)

L'ozone dans le traitement de l'eau - Principe et application (B.13)

Les réactifs en traitement d'eau potable (B.14)

Lutter contre les algues et leurs nuisances (B.15)

Les techniques de filtration (B.19)

Correction de la minéralisation (B.18)

Qualité et goûts de l'eau (B.17)

4,5	6995	28--3	10/14	18/22				
4,5	6995							
3,5	5530		25/28	26/29			13/16	
3	5390	20/23						
3	6250		3/6			2/5		
3	4730	6/9				2/4		
3,5	5530				20/23			
3	4730							
2	3150			11/14				21/23
2,5	3990				10/12			
4,5	7350				15/19			6/11
4,5	7250					5/9,		
4	6650					19/23		
2,5	4100	21/23					4/6	
2,5	3990				3/5			
2,5	3990				16/18			
2,5	4990					7/9		5/8
3,5	5590							
4	6590	22/24				22/26		
2,5	4990							7/8

# CALENDRIER ZUUU

	Durée (en jours)	Page	Prix NET (en francs)	JAN	FEV	MARS	AVRIL	MAI	JUIN	SEPT	OCT	NOV	DEC
<b>MAÎTRISE D'ŒUVRE ET TRAVAUX</b>													
LES MARCHÉS DE MAÎTRISE D'ŒUVRE DANS LE CADRE DE LA LOI M.O.P.	4	31	8 300			13-16			19-22	12-15			11-14
LA RÉCEPTION DES TRAVAUX ET LES SÉCURITÉS POST-CONTRACTUELLES	2	32	4 980				26-27				12-13		
PRÉPARATION, PASSATION ET GESTION DES MARCHÉS DE TRAVAUX	4	33	8 250			7-10			5-8		3-6		
<b>INFORMATIQUE, TÉLÉCOMMUNICATIONS ET PRESTATIONS INTELLECTUELLES</b>													
LES MARCHÉS DE SERVICES DE TÉLÉCOMMUNICATIONS	1	35	3 120		22					22			
LES MARCHÉS PUBLICS D'INFORMATIQUE	4	38	9 400					16-19				14-17	
PRATIQUE DES MARCHÉS DE PRESTATIONS INTELLECTUELLES	3	39	6 780				3-5						6-8
<b>FOURNITURES, SERVICES ET NETTOYAGE</b>													
LES MARCHÉS PUBLICS D'ASSURANCE	1	34	2 850		23					21			
LES MARCHÉS DE SERVICES DE TÉLÉCOMMUNICATIONS	1	35	3 120		22					22			
PRATIQUE DES MARCHÉS DE NETTOYAGE DE LOCAUX	3	36	6 750		14-16				28-30		2-4		
PRATIQUE DES MARCHÉS DE FOURNITURES COURANTES	3	37	6 350	26-28				15-17		6-8			
<b>SECTEUR HOSPITALIER</b>													
LES ACHATS DES HÔPITAUX PUBLICS	3	22	6 430			20-22						27-29	
<b>DÉLÉGATIONS DE SERVICE PUBLIC</b>													
COMMENT PASSER ET GÉRER LES DÉLÉGATIONS DE SERVICE PUBLIC	2	43	5 100			23-24						25-26	
<b>ACHATS : ORGANISATION, NÉGOCIATION</b>													
ORGANISATION DE LA FONCTION ACHAT	2	40	5 400			2-3				25-26			
LA NÉGOCIATION D'ACHAT	3	41	6 450				17-19					6-8	
<b>RÈGLEMENT FINANCIER / BUDGET ET ADMINISTRATION</b>													
LE RÈGLEMENT FINANCIER DES MARCHÉS	2	15	4 920	24-25			20-21		26-27		18-19		
BUDGET DE L'ÉTAT : DE SA PRÉPARATION À SON EXÉCUTION	2	42	4 740				10-11			18-19			
ORGANISATION ET ÉVOLUTION DE L'ADMINISTRATION FRANÇAISE	2	46	4 420		7-8					4-5			
<b>GESTION DU PATRIMOINE ET APPROVISIONNEMENTS</b>													
PRATIQUE DE LA GESTION DES STOCKS ET DES APPROVISIONNEMENTS	3	44	5 900			20-22							20-22
ORGANISER ET GÉRER LA MAINTENANCE DU PARC IMMOBILIER	2	45	5 300					23-24					29-30

## *ANNEXE 2*

**[Base « Environnement » >>> Offres de formation >>> Organismes ]**

### **Questionnaire utilisateurs**

NOM :..... PRENOM :.....

CELLULE :.....

POSTE ET DESCRIPTIF DU POSTE :.....

.....

.....

#### **1- Quelle utilisation faites vous de cette base ?**

.....

.....

.....

.....

.....

.....

#### **2- Actuellement, la base vous paraît elle : (vous pouvez cocher plusieurs cases)**

Bonne

Insuffisante

Potentiellement source d'erreur

Opaque (les informations sont difficiles à relier entre elles)

#### **3- A l'état actuel, quels sont les points forts de la base ?**

.....

.....

.....

.....

.....

.....

**4- Actuellement, quels sont les points faibles et insuffisances de la base ?**

.....  
.....  
.....  
.....  
.....  
.....  
.....

**5- En l'état actuel, les descripteurs (ou « intitulés de stage ») vous paraissent ils ?**

- Révélateurs du contenu des stages
- Partiellement révélateurs des stages
- Non révélateurs du contenu des stages

**6- Quels sont les changements envisageables qui vous viennent d'emblée à L'esprit ?**

.....  
.....  
.....  
.....  
.....

**7- En somme, quels sont les représentations (les souhaits) que vous pouvez vous faire du nouveau « système documentaire » ?**

.....  
.....  
.....  
.....  
.....  
.....  
.....

# *ANNEXE 3*

Retranscription des « codes secteurs » présents dans PARADOX 3.1

**01/ Techniques de management et organisation des services**

- 01-01 Encadrement d'une équipe
- 01-02 Gestion de conflit
- 01-03 Elaboration et conduite de projet de service
- 01-04 Marketing
- 01-05 Gestion du temps (délégation)
- 01-06 Conduite de réunion
- 01-07 Encadrement technique
- 01-08 Méthodologie de projet
- 01-09 Management et organisation

**02/ Outils de gestion des ressources humaines**

- 02-01 Statut FPT
- 02-02 Paie
- 02-03 Entretien évaluation
- 02-04 Fiches postes et métiers
- 02-05 organisations paritaires
- 02-06 Recrutement (externe, interne)
- 02-07 Formation

**03/ Outils de gestion financière**

- 03-01 Comptabilité et gestion
- 03-02 Finances publiques
- 03-03 Economie

**04/ Communication relationnelle**

- 04-01 Communication individuelle
- 04-02 Communication collective
- 04-03 Communication écrite
- 04-04 Relation à l'utilisateur
- 04-05 accueil

**05/ Techniques appliquées et réglementation**

- 05-01 Voiries et infrastructures
- 05-02 Entretien, petites réparations
- 05-03 Bâtiment
- 05-04 Electricité (habitation)
- 05-05 Sécurité incendie
- 05-06 Automobile
- 05-07 Patrimoine
- 05-08 Environnement, espaces verts et eau
- 05-09 Aménagement du territoire
- 05-10 Action économique et emploi
- 05-11 Génie urbain
- 05-12 Culture
- 05-13 Sport
- 05-14 Techniques de secrétariat
- 05-15 Permis de conduire
- 05-16 Techniques audiovisuelles
- 05-17 Transport
- 05-18 Assurances
- 05-19 Marchés publics

- 05-20 Bibliothèques, archives, document
- 05-21 Construction
- 05-22 Reprographie
- 05-23 Cuisine, couture
- 05-24 Lecture rapide
- 05-25 Gestion des stocks
- 05-26 Laboratoire
- 05-27 Tourisme, hôtellerie

**06/ Droit**

- 06-01 Droit civil et droit social
- 06-03 Droit immobilier
- 06-04 Europe
- 06-05 Droit administratif
- 06-06 Droit privé

**07/ Connaissance de l'environnement médico-social**

- 07-01 Toxicomanie
- 07-02 Alcoologie
- 07-03 Tabagisme
- 07-04 Sida
- 07-05 Violences institutionnelles dans le travail
- 07-06 Ecoute
- 07-07 Environnement social et médico-social

**08/ Connaissance du fonctionnement familial**

- 08-01 Suivi de l'enfant
- 08-02 Développement de l'enfant
- 08-03 Diététique
- 08-04 L'adoption
- 08-05 L'abandon
- 08-06 L'adolescence
- 08-07 La maltraitance
- 08-08 La famille
- 08-09 La famille et l'enfant
- 08-10 Le comportement de l'enfant

**09/ Pratiques d'intervention sociale**

- 09-01 Accompagnement social des familles
- 09-02 Expulsion, surendettement
- 09-03 Immigration
- 09-04 agrément des assistantes sociales
- 09-05 Développement social local

**10/ Bureautique et informatique**

- 10-01 Bureautique
- 10-02 Word
- 10-03 Excel
- 10-04 SGBD
- 10-05 Environnement
- 10-06 PAO et images
- 10-07 Spécifiques et applicatifs
- 10-08 Informatique
- 10-09 Transmissions
- 10-10 Nouvelles technologies

# *ANNEXE 4*

TS4 APPROCHE PSYCHIQUE RELATIONNELLE

TS5 ANALYSE SYSTEMIQUE — 7-158-200-300-451

7-125-158-196-200-323-353-300 — TS5 ANALYSE TRANSACTIONNELLE

TS5 APPROCHE GORDON — 158-200

TS5 GESTALT — 63-158-200-274-323

TS5 PNL — 7-63-158-196-200-287-323-451-507  
*EP PROGRAMMATION NEURO-  
LINGUISTIQUE*

TS4 APPROCHE PSYCHOCORPORELLE — 477  
RELATIONNELLE

TS5 LANGAGE CORPOREL — 200-477

TS5 SOPHROLOGIE — 55-63-158-200-451

TS4 APPROCHE PSYCHOSOCIOLOGIQUE  
RELATIONNELLE

TS5 ASSERTIVITE — 84-179-196-200-353  
*EP AFFIRMATION DE SOI*

58-158-171 — TS5 GESTION DES CONFLITS — 200-223-249-451-478

TS5 GESTION DE LA VIOLENCE — 58-158-246-315-358

TS2 MANAGEMENT DES HOMMES — 179-247

TS3 ACCOMPAGNEMENT DE CHANGEMENT — 77-158

59-74-77-107-110 — TS3 ANIMATION D'EQUIPE — 124-158-176-179-215-247-248-323-353-374

391

TS4 RESPONSABILISATION D'EQUIPE — 77-315

TS4 TRAVAIL EN EQUIPE — 7-77-125-439

Langage Programmation avancée en	C++	501
	C++	268-501
Données	CADASTRALES	158-454
	CADRE réglementaire	6-49-293*
Développement des	CAPACITES	
	CARENCE intra-familiale	158-321
Gestion des	CARRIERES	407
	CERTIFICAT d'aptitude aux soins à domicile	437
	CERTIFICAT management des équipes et des hommes	376
	CESS psychologie du vieillissement	180
Accompagnement de Conduite de	CHANGEMENT	77-158
	CHANGEMENT	58-120-171
	CHANT	58
Prise en Prise en	CHARGE	
	CHARGE de la douleur	476
Prise en	CHARGE psychomotrice	476
	CHARGES sociales	77
Physique	CHIMIE	
	CHOIX de solution informatique	59
	CHSCT	247
	CLAVIER alphanumérique	158
Compte	CLIENT	223
	CODE	6-158
Négociation	COLLECTIVE	475*
	COLLECTIVITE	
Fonctionnement du Trésorerie du Réseau	COMITE d'entreprise	475*
	COMITE	247
	COMITE	247
Premier Méthode de	COMMUTE	
	CONFIGURATION	85-158-356
Premier Méthode de	CONTACT avec un PC	38-70-133-158-175-223-289-319-412
	COMMANDE	501
	COMMERCE électronique	268
	COMMUNICATION écrite	158-456-459-478-485
	COMMUNICATION	158