Thomas Moerman

Master 2 ICD Spécialités Sciences de l'Information et du Document, parcours PRISME

Rapport de stage

Mission effectuée du 6 avril au 25 septembre 2008
Au sein de la CRCI à l'Agence Régional d'Information Stratégique et Technologique (ARIST)
à Lille

Veille et représentation des connaissances : Le cas du dictionnaire métier et de la veille à l'ARIST

Sous la Direction de :

Monsieur Ismaïl Timimi, responsable universitaire
Monsieur Jean Dufour, responsable professionnel

Soutenu le 17 septembre 2009
Université Charles de Gaulle Lille-3
UFR IDIST

Année universitaire 2008-2009

Remerciements

Tout d'abord je voudrais remercier M. Philippe Domino, directeur de la DRDE (Direction Régional du Développement des Entreprises) et de l'ARIST, et M. Jean Dufour, responsable du Projet Veille et Intelligence Economique, pour m'avoir accueillis dans leur service et de m'avoir fait confiance en me prenant en tant que stagiaire.

Je souhaite remercier également le personnel qui travaille à l'ARIST pour leur accueil, mais également pour leur bonne humeur qui a fait de ce stage une expérience agréable.

Evidemment, je remercie mon tuteur de stage, M. Ismaïl Timimi, pour ces conseils tout au long du stage et pour la rédaction du rapport de stage.

Enfin je voudrais remercier également Florent pour son soutien et ces conseils tout au long du stage et pendant la rédaction du rapport.

Tables des Matières

Introduction **1**

Partie 1 : Présentation du cadre de mon stage **3**

I Présentation de l'ARIST et du PICOM 3

A) Présentation de l'ARIST 3

A-1 La CRCI Nord Pas de Calais 3

A-2 L'ARIST 4

B) Présentation du PICOM 5

B-1 Les pôles de compétitivités 5

B.2 Le Pôle de compétitivité de l'Industrie et du Commerce 7

II La pratique de la veille à l'ARIST et au PICOM 8

A) La veille à l'ARIST 8

B) La veille au PICOM 10

Partie 2 : Développement du dictionnaire métier pour et à l'aide de la veille **13**

I Le premier objectif : la veille technologique 13

A) Qu'est ce que la veille technologique 13

A.1 Définitions 13

A-2 Les objectifs de la veille technologique 14

B) Les outils pour la réalisation de la veille : le cas de Tropes Zoom 15

B-1 Une diversité d'outils logiciels dans les systèmes de veille 15

B-2 Un outil d'aide à l'analyse : *Tropes Zoom* 17

C) La veille technologique sur la RFID et la 3D 20

II La veille comme aide à l'enrichissement du dictionnaire métier 21

A) Le dictionnaire métier : un thésaurus simplifié 21

A.1 Les caractéristiques d'un thésaurus 21

A.2 Le dictionnaire métier 23

B) Méthode d'élaboration du dictionnaire métier 25

B-1 Méthode d'élaboration d'un thésaurus 25

B-2 Méthodes partiellement applicables pour le dictionnaire métier... 26
C) La veille, une méthode d'enrichissement du dictionnaire métier 29
III La veille comme un moyen d'évaluation du dictionnaire métier 31
A) Méthode d'évaluation d'un thésaurus 31
B) Méthodes partiellement applicables pour le dictionnaire métier..... 32
C) La veille, comme moyen de valider le dictionnaire métier 33

Partie 3 : La veille et la représentation des connaissances 35
I Qu'est ce que la représentation des connaissances ? 35
A) Définitions..... 35
A-1 Définitions de « connaissance »..... 36
A-2 Définitions de la représentation des connaissances 37
B) Les types de représentation des connaissances 38
B-1 Les représentations des connaissances en logique 38
B-2 Les représentations des connaissances sous la forme de réseaux . 40
B-3 Les représentations visuelles de connaissances 43
II La place de la représentation des connaissances dans la mise en place et la réalisation d'une veille..... 44
A) La place de la représentation des connaissances dans un contexte de veille 44
B) Les avantages des représentations des connaissances..... 46
C) Les inconvénients/limites des représentations des connaissances 49

Conclusion..... 51
Bibliographie..... 53
Sitographie 55
ANNEXES..... 57

Introduction

L'étude des représentations des connaissances n'est pas nouvelle. A l'époque antique, Aristote a été l'un des premiers à procéder à des classifications hiérarchiques systématiques des connaissances et des concepts. Il divisa la science en trois parties : la science théorétique (la physique, les mathématiques et la théologie), la science pratique (l'économie, l'éthique, la politique et la rhétorique) et la science poïétique qui comprend toutes les activités qui produisent une œuvre. Par la suite, les représentations des connaissances se sont développées essentiellement en biblio-économie et donc par les bibliothécaires. Avec l'apparition de la science de l'information, les représentations de connaissances ont pris une nouvelle dimension en sortant des bibliothèques

De nombreux ouvrages sur les représentations et/ou l'organisation des connaissances sont éditées autant en langue française et en langue anglaise. Ces ouvrages sont essentiellement des ouvrages de linguistiques voir d'informatiques. Un grand nombre de ces ouvrages traitent les représentations des connaissances dans le cadre des bibliothèques ou centres d'informations. Durant nos recherches nous avons constaté que peu d'ouvrages traitent les représentations des connaissances dans un contexte de veille ; Certes il n'est pas obligatoire d'utiliser une représentation des connaissances lors de la réalisation d'une veille mais elle a ces avantages.

Ma mission de stage consistait à la mise en place de la veille pour un pôle de compétitivité qui va du sourcing à l'analyse en passant par l'élaboration du dictionnaire métier que nous considérons comme une représentation de connaissances. Durant notre stage nous avons constaté trois liens entre la veille et le dictionnaire métier : le dictionnaire est une aide à la réalisation de la veille mais inversement la veille est une aide pour l'enrichissement et un moyen de validation du dictionnaire. Ce constat m'a amené à réfléchir sur les liens entre les représentations des connaissances et la veille. Quels sont les liens entre la veille et la représentation des connaissances ? Quelle est la place de la représentation des connaissances dans la mise en place et la réalisation d'une veille ?

Ce mémoire s'organisera de la manière suivante : dans une première partie nous présenterons le cadre de notre stage, ensuite dans une seconde partie intitulé développement

du dictionnaire métier pour et à l'aide de la veille ou nous aborderons notre travail durant le stage en fonction des trois liens entre la veille et le dictionnaire métier que nous avons constaté durant le stage et enfin nous traiterons dans une troisième partie la place des représentations des connaissances dan le cadre de la veille.

Partie 1 : Présentation du cadre de mon stage

I Présentation de l'ARIST et du PICOM

A) Présentation de l'ARIST

L'ARIST, l'Agence Régionale d'Information Stratégique et Technologique, fut mon lieu de stage. Ce service fait partie de la CRCI, la Chambre Régionale de Commerce et d'Industrie du Nord Pas de Calais.

A-1 La CRCI Nord Pas de Calais

Les Chambres de Commerce et d'Industrie (CCI) sont, en France comme dans d'autres pays, des organismes chargés de représenter les intérêts des entreprises commerciales, industrielles et de service d'une zone géographique auprès des pouvoirs publics et de leur apporter certains services. Ce sont des établissements publics. Elles sont administrées par des chefs d'entreprises élus. Elles ont pour mission de contribuer au développement des entreprises, des hommes et des territoires.

La CRCI anime et coordonne le réseau régional des CCI, elle les fédère et les représente auprès des pouvoirs publics, au niveau régional, national et international. Elle développe, avec elles, une stratégie d'ambitions en faveur du rayonnement économique régional. Comme toutes les CRCI, elle est placée sous la tutelle du préfet de région représentant le ministère chargé de l'industrie et le ministère chargé des PME, du Commerce et de l'Artisanat. La CRCI Nord Pas de Calais a été créée en 1964. L'actuel président est Léonce Michel DEPREZ qui est aussi président de la chambre de commerce et d'industrie de Béthune.

La CRCI Nord Pas de Calais se compose de cinq grands services : la direction administrative et financière/ressources humaines/services généraux, la direction de la communication et des relations extérieures, la direction régionale du développement économique et territoriale qui est dotée d'un observatoire économique et territorial qui a pour objectifs d'analyser les besoins des entreprises et des territoires, produire des informations et des avis économiques pour éclairer les décisions des CCI et partenaires, services de l'Etat ou collectivités. Elle atteint cet objectif en contribuant directement à l'intelligence économique territoriale, en proposant la démarche « Stratégie Economique Pour les Territoires » (SEPT) et en travaillant sur la formation. Le quatrième grand service se nomme CCI International. Il a pour mission d'accompagner les entreprises sur les marchés étrangers et de promouvoir l'attractivité de la région et de ses territoires. Enfin, le cinquième et dernier grand service se nomme la DRDE, la Direction Régionale du Développement des Entreprises.

La DRDE se compose de cinq sous-services : le dispositif Action Régionale pour le Développement d'Activités Nouvelles (ARDAN), le pôle environnement et développement durable, le pôle régional création-transmission-reprise, le pôle salons-tourisme et enfin le cinquième sous-service, l'ARIST. Les missions de la DRDE sont de soutenir les entreprises du Nord Pas de Calais dans leur croissance.

A-2 L'ARIST

Les ARIST ont été créées à l'initiative du ministère de l'industrie dans les années 70. A l'origine, chaque chambre régionale de commerce et d'industrie possède son Agence Régionale d'Information Stratégique et Scientifique même si elles ne sont pas toutes de la même dimension. L'agence du Nord Pas de Calais est une des plus importantes de France. L'ARIST Nord Pas de Calais fut créée en 1978. Aujourd'hui, 11 personnes travaillent à l'agence dont le directeur Mr. Philippe Domino qui est également le directeur de la DRDE.

L'ARIST fait partie de la DRDE. L'ARIST a pour mission prioritaire de favoriser le développement des PME-PMI régionales, grâce à l'information technologique et stratégique. Cependant elle effectue d'autres missions : analyser les projets de l'entreprise en tenant compte de son environnement technologique et concurrentiel. Elle aide les PME-PMI à mieux maîtriser l'information (technologie, processus, réglementation, normes,

concurrents, fournisseurs..), les conseille dans le domaine de la propriété intellectuelle, les appuie dans le développement de nouveaux produits et les alerte sur les risques de contrefaçon. Les collaborateurs animent également des groupes de travail au sein des entreprises (résolution de problèmes innovants, méthodes de veille....).

L'ARIST Nord Pas de Calais est divisée en trois pôles : un pôle innovation, un pôle développement partenarial PME, et un pôle veille et intelligence économique. Ces pôles travaillent ensemble, ils ne sont pas cloisonnés. L'ARIST, dans le cadre de sa mission, mène diverses actions : *l'action regroupement* qui apporte une aide et un accompagnement pour des projets de regroupement (achats communs, partage de services,...), l'action *Casciopée* qui propose de projeter les entreprises, d'adapter leur stratégie à l'évolution des marchés sur 3 ans, un programme de soutien de projet TIC d'entreprises. L'ARIST est également membre du réseau EEN (Enterprise Europe Network) qui aide les entreprises du Nord Pas de Calais, à s'ouvrir sur l'Europe ou des entreprises européennes qui veulent s'implanter, à investir dans la région, à travers des partenariats commerciaux et technologiques par exemple.

L'ARIST propose également une veille sur les Brevets et marques. La veille intervient également dans plusieurs actions. L'ARIST travaille avec les PME-PMI mais aussi avec les pôles de compétitivité et des réseaux d'entreprises.

B) Présentation du PICOM

Le travail réalisé durant le stage a été fait en collaboration et pour le PICOM : le Pôle de Compétitivité de l'Industrie et du COMmerce, situé à Entreprises et Cités à Marcq en Baroeul

B-1 Les pôles de compétitivités

En France, les pôles de compétitivité sont des clusters reconnus individuellement par l'Etat, le Comité Interministériel Aménagement et Compétitivité des Territoires définit un pôle de compétitivité :

« *Un pôle de compétitivité est défini comme la combinaison sur un espace géographique donné d'entreprises, de centres de recherche publics ou privés et de centres de formation engagés dans une*

démarche partenariale destinée à dégager des synergies autour de projets communs au caractère innovant. Ce partenariat s'organisera autour d'un marché et d'un domaine technologique et scientifique qui lui est attaché et devra rechercher la masse critique pour atteindre une compétitivité mais aussi une visibilité internationale »

Autrement dit, un pôle de compétitivité est sur un territoire donné, l'association d'entreprises, de centres de recherche et d'organismes de formation, engagés dans une démarche partenariale (stratégie commune de développement), destinée à dégager des synergies autour de projets innovants conduits en commun en direction d'un (ou des) marché(s) donné(s).

La notion de technopole apparait en France dans les années 70, elle a mené à la création de technopoles proprement dites comme Inovallée près de Grenoble et des développements de zones déjà denses comme à Toulouse dans l'aérospatiale. Ces entités peuvent être considérées comme les ancêtres ou les prémices des pôles de compétitivité ou encore des précurseurs.

La politique des pôles de compétitivité a été décidée lors du comité interministériel d'aménagement et de compétitivité des territoires (CIACT) du 13 décembre 2002. Les appels à projets ont été lancés le 2 décembre 2004 et clos le 28 février 2005. De ces appels, 67 projets sont retenus dont six ont une portée mondiale et neuf autres qui ont vocation à l'acquérir. Depuis le CIACT du 5 juillet 2007, le nombre de pôles de compétitivité est passé à 71 pôles dont 7 ayant une portée mondiale et 10 qui ont vocation à le devenir. Ces pôles regroupent 9 000 chercheurs qui travaillent sur 1 000 projets labellisés. Ces pôles sont financés par l'état.

Quatre éléments principaux sont la clé de la réussite du pôle : la première clé c'est la mise en œuvre d'une stratégie commune de développement économique cohérente avec la stratégie globale du territoire. La deuxième clé, ce sont les partenariats approfondis entre acteurs autour de projets. La concentration sur des technologies destinées à des marchés à haut potentiel de croissance est la troisième clé. Enfin la quatrième clé, c'est la masse critique suffisante pour acquérir une visibilité internationale.

Le Nord Pas de Calais possède six pôles de compétitivité : le pôle MAUD, Matériaux et Applications pour une Utilisation Durable, dont les domaines sectoriels sont la chimie et les matériaux pour trois marchés clés : les arts de la table, les industries graphiques

et l'emballage ; le pôle Uptex, dont les domaines sectoriels sont le textile et la VPC, et qui travaille sur les matériaux textiles avancés, les technologies de la perception d'un produit par les sens, la customisation de masse ; le pôle NSL, Nutrition Santé Longévité, qui couvre le secteur de l'agroalimentaire, les biotechnologies et la santé ; le pôle Aquimer, basé à Boulogne sur Mer, dont les domaines sectoriels sont les produits de la mer et l'aquaculture ; le cinquième pôle qui est le seul à porté mondiale du Nord Pas de Calais, basé à Valenciennes, le pôle I-Trans qui couvre le domaine sectoriel des transports terrestres ; enfin le PICOM.

B.2 Le Pôle de compétitivité de l'Industrie et du Commerce

Labellisé par le Comité Interministériel Aménagement et Compétitivité des Territoires (CIACT) le 12 juillet 2005, le PICOM est créé le 2 février 2006. Il regroupe 73 entreprises membres du pôle dont 47 PME, 69 entreprises impliquées dans le pôle dont 47 PME et enfin 37 groupes impliqués dans le pôle dont des groupes français, européens et étrangers (hors Europe). Ces entreprises et groupes couvrent les domaines de la grande distribution, de la vente à distance, de l'industrie technologie. Le PICOM regroupe également 15 laboratoires de recherche, ce qui représente 800 chercheurs, principalement dans les Technologies de l'Information et de la communication (TIC). Nous trouvons comme membres Auchan France, dont le président du conseil de surveillance, Arnaud MULLIEZ, est le président du pôle, France Télécom ou encore CIE IBM France. La difficulté pour le pôle est de faire collaborer l'industrie avec la recherche, et les grandes entreprises avec les PME.

Le PICOM se structure autour de plusieurs comités : le conseil de gouvernance, rassemblant des membres du monde de l'entreprise, de l'enseignement supérieur, de la recherche, de l'état et des collectivités territoriales. Il gère et administre l'association et prend toutes initiatives qui concourent à la réalisation de l'objet de l'association. En son sein un comité stratégique est mis en place, il élabore le projet de politique générale de l'association et met en œuvre le projet arrêté par le Conseil de Gouvernance. Le comité opérationnel prépare et met en œuvre la politique du pôle sur laquel les instances de gouvernances prennent appui. Enfin le comité scientifique et technique est composé de membres des entreprises, d'organisations professionnelles et consulaires et des membres d'établissements d'enseignement supérieur et de la recherche. Il

exerce une mission générale d'animation scientifique et est en lien avec le comité stratégique, il formule des avis et des propositions.

Les trois grandes missions du PICOM sont l'accompagnement des projets de coopérations avec à la fois un appui juridique, une aide à la formulation du projet, un accompagnement financier, l'animation de l'innovation et la veille technologique. Il effectue également des missions dites opérationnelles : veille et prospective, valorisation de la recherche et transfert de technologie, incubation de projets R&D et appui aux start-up, ingénierie de formation et échanges internationaux, communication et création d'événements. L'objectif du PICOM est de faire converger le monde de la technologie et le monde des services. Les organisations des activités de recherches et développement sont orientées sur 3 grands axes métiers : l'optimisation de process (physiques et informationnels), la relation client et la mobilité (une personne peut être un client et un patron) et les nouveaux univers marchands (e-commerce, commerce virtuel, le web 2.0 voir à l'avenir le web 3.0).

La personne, avec qui je travaillais au PICOM, est Mme Emilie OBERT, la responsable veille.

II La pratique de la veille à l'ARIST et au PICOM

A) La veille à l'ARIST

La veille à l'ARIST intervient dans plusieurs actions (*Casciopée, Offensi'veille,...*). Elle est réalisée pour les pôles de compétitivités, des réseaux d'entreprises (des PME d'un même secteur s'associent pour mener des actions qu'elles ne pourraient faire individuellement). La veille peut se faire également directement avec une PME. Ces services sont payant, cela incite les entreprises a s'investir dans le projet ou dans la veille réalisée. A l'ARIST, quatre personnes s'occupent de la veille.

L'ARIST utilise trois sources d'information : le web, les bases de données en ligne comme les bases de données de brevets et la troisième source sont les rapports, études, et autres documents dont certains sont accessibles sur internet. Les documents issus de ces

trois sources forment un corpus bibliographique. Durant mon stage, ma principale source fut le web.

La méthode[1] pour la réalisation de la veille se base sur un ensemble d'outils organisé sous la forme d'une chaîne. L'ARIST n'utilise pas une plateforme d'un même éditeur comme *Digimind* mais une suite d'outils spécialisés dans une tâche qui ne sont pas développés par le même éditeur. Cette chaîne d'outils suit les différentes étapes du système de veille.

Pour le sourcing, l'ARIST utilisait au début tout simplement les moteurs de recherches comme *Google* mais l'agence est arrivée aux limites de ce type d'outils. Il a donc acquis le moteur d'exploration *Squido* (dernière acquisition de l'ARIST), moteur plus performant qui permet de trouver des sources plus rapidement qu'avec *Google* voire que l'on n'aurait pas trouvé avec *Google*. Pour la collecte sur le web, l'agence utilise des robots de capture de l'information avec les logiciels The *Easy Bee* et *WebExt*. The *Easy Bee* est un logiciel qui permet de programmer des agents de navigation et de surveillance, il permet de créer le squelette du robot et *Webext* est l'outil qui permet d'utiliser les robots notamment pour aspirer les données web et les convertir en *XML*. A noter que l'ARIST utilise également *KB Crawl*.

Pour le traitement, l'ARIST utilise deux méthodes d'analyse : une méthode sémantique et une méthode statistique. Pour la première, l'agence utilise le logiciel *Tropes Zoom* pour indexer les documents à l'aide d'un scénario que nous appelons dictionnaire métier fait en prétraitement et pour réaliser une analyse sémantique mais nous verrons plus en détail le fonctionnement de cet outil dans la partie 2. Pour la méthode statistique, l'agence utilise le logiciel Wordmapper, un outil cartographique, qui permet aussi d'indexer les documents. Pour cet outil, on utilise également un dictionnaire composé de mots signifiants, de mots composés, de synonymes et de mots obligatoires que l'on a préalable réalisé.

Enfin la dernière étape, la diffusion se fait différemment si on utilise la méthode sémantique ou la méthode statistique. Pour la première, l'agence peut proposer l'alerte mail avec l'outil PDI qui est associé à *Tropes Zoom*, une analyse qualitative sur un thème qui peut aboutir à un rapport ou une étude, puis l'agence utilise un module web de

[1] Voir annexe 1 p56

catégorisation du corpus en fonction du scénario métier qui permet ensuite d'envoyer par mail une suite d'articles concernant un sujet au client. Ce module est également compatible avec *Tropes Zoom*. Pour la méthode statistique, la diffusion se fait via des cartes produites par *Wordmapper*. Par exemple, on peut obtenir des cartes des thèmes du corpus pour une analyse globale ou une carte des tendances ou encore des réseaux de collaborations. A noter que *Tropes Zoom* édite également des cartes sémantiques, moins utilisées par l'ARIST qui se sert plus souvent de la fonctionnalité de *Tropes Zoom* à pouvoir éditer des graphiques via Excel.

La méthode sémantique va être dotée prochainement (probablement début 2010) d'un outil cartographie, *PIKKO*, qui complétera la chaîne d'outils. Un des avantages de cet outil est qu'il sera compatible avec *Tropes Zoom* après une étape de développement pour qu'ils soient à 100% compatibles.

Donc, la veille à l'ARIST est essentiellement technologique et stratégique. Elle a pour objectif d'apporter une aide à l'innovation et à la décision aux PME, directement ou à travers les pôles de compétitivités et les réseaux d'entreprises.

B) La veille au PICOM

L'objectif de l'activité de veille stratégique et technologique, et de la prospective est de favoriser l'émergence de projets collaboratifs par la production et le partage d'information scientifiques et techniques. Cette veille est faite pour les entreprises et non pour une entreprise. En effet, les entreprises membres ou collaborant avec le pôle sont, pour un grand nombre, concurrentes, donc le travail de veille qui leur est destiné doit prendre en compte ce contexte et nous devons rester vigilants sur le contenu des livrables, il faut rester neutre. La veille se fait sur les thématiques choisies par le pôle, des thématiques transversales, communes à l'ensemble ou un groupe d'entreprises. Les principales thématiques de travail portent sur la biométrie, les interfaces homme/machine, les technologies sans contact ou encore la géo-localisation pour la partie technologie et la fidélisation, le comportement client, le sensoriel ou encore les réseaux communautaires pour la seconde partie science humaine et social (SHS). Par exemple, pour la fidélisation, un travail est effectué pour trouver un juste milieu entre la non reconnaissance et le harcèlement du client.

Le pôle a défini quatre grands axes de travail à partir desquels la veille est orientée :

- Le premier axe est intitulé « Prospective de la distribution ». Cet axe est composé de plusieurs dynamiques : tout d'abord une dynamique sociale et démographique c'est-à-dire la question du vieillissement de la population et les consommateurs de demain (génération Y, la génération millénium, etc.), la deuxième dynamique concerne l'évolution des politiques urbaines qui touche la logistique, la troisième dynamique porte sur les évolutions technologiques et la dernière concerne le développement durable avec notamment les contraintes réglementaires.

- Le deuxième axe, « Comportements et modes de consommation », est un axe orienté SHS, comportement client. Nous y trouvons le comportement multi-canal du client, le comportement hédoniste et bien-être c'est-à-dire la personnalisation et l'anticipation des besoins du client. Nous y trouvons également le thème de la consommation durable, du comportement d'achat avec le rapport qualité/prix par exemple. Enfin nous pouvons également y trouver l'économie de la fonctionnalité c'est-à-dire l'idée que l'on ne vend plus un bien mais un service.

- Le troisième axe s'intitule « Univers ubiquitaires ». Cet axe concerne les questions qui vont de la recherche d'information à la destruction ou la réévaluation du bien. L'objectif est d'avoir une continuité dans la relation client. Nous retrouvons, par exemple, les questions sur le multicanal /pluricanal. « Ubiquitaire » désigne ici un environnement où l'utilisateur a accès à un ensemble de services au travers d'interfaces distribuées se voulant intelligentes et dont il est entouré.

- Le quatrième et dernier axe porte sur l'« Organisation des process » c'est-à-dire tout ce qui concerne la logistique, le backoffice, le système d'information, d'un point de vue général tout ce qui n'est pas en interface avec le client. Nous trouvons également le thème de la logistique urbaine qui traite de la question des livraisons dans les centres villes et plus précisément les zones piétonnes, le thème de la customisation ou personnalisation des services ou encore celui du développement durable

notamment pour les transports mais aussi dans le cadre de l'éco-conception des produits avec l'analyse du cycle de vie.

L'information au PICOM est issue de trois sources : les sites grand public comme *le journal du net*, les revues spécialisées pour une information plus qualifiée et plus fiable et enfin les informations issues du relationnel c'est-à-dire lors de salon ou autres manifestation, au sein d'un réseau.

La veille au PICOM utilise quatre principaux outils : le gestionnaire de connaissances pour la collecte l'organisation de l'information à l'aide de robots et de bases à entrées multiples et personnalisées. Un tableau de bord de veille qui permet de rassembler des indicateurs pour un premier niveau d'analyse et de communication. Le troisième outil est un détecteur de tendances émergentes qui permet de donner des vues globales et de recouper l'information pour identifier la donnée latente, le signal faible, l'apparition de nouveau concept grâce à la cartographie notamment. Enfin le quatrième et dernier outil, les outils de synthèse qui permettent de produire des documents clairs et agréables pour faire le point en quelques lignes.

Le PICOM effectue une veille internationale qui passe par la réalisation d'études, d'actions lobbying, de mission à l'internationale, etc. Il a également mis en place une plate-forme de veille collaborative avec l'ARIST ce qui permet de réaliser une veille mutuelle. Cette plate-forme prend la forme d'un blog et a pour objectif d'accompagner le processus de R&D, on y trouve par exemple un panorama de l'actualité. La veille réalisée apparait également dans la lettre d'information du pôle et dans la revue *Tendances et Signaux faibles* éditée par le PICOM. Tous ces outils sont à disposition des entreprises membres. Pour la veille, le PICOM s'est associé au réseau VIGILANCES, aux technopoles DIGIPORT et CIEL ou encore à l'ARIST.

Après avoir présenté le cadre de mon stage, nous allons aborder notre travail sur le dictionnaire métier sur le commerce et l'apport de la veille à ce travail.

Partie 2 : Développement du dictionnaire métier pour et à l'aide de la veille

Dans cette partie, nous aborderons mon travail durant le stage, autour de l'élaboration du dictionnaire métier sur le commerce mais aussi de la réalisation d'une veille technologique sur deux sujets : les puces RFID et la 3D dans le cadre du commerce. Nous le traiterons en fonction des trois objectifs de cette veille constatés durant le stage : la veille proprement dite, l'aide à l'enrichissement du dictionnaire métier et un moyen de validation des termes du dictionnaire métier.

I Le premier objectif : la veille technologique

A) Qu'est ce que la veille technologique

A.1 Définitions

Il n'existe pas de définition consensuelle de la notion de veille mais il est néanmoins possible de la définir comme un processus qui vise à être à l'écoute des informations relatives à un domaine donné dans lequel une entreprise ou une organisation exerce son activité ou dans lequel elle pourrait l'exercer. Selon la norme XP X 50-053 l'AFNOR définit la veille comme une activité continue et en grande partie itérative visant à une surveillance active de l'environnement technologique, scientifique, etc. pour en anticiper les évolutions. Avant tout la démarche de veille doit prendre en compte l'ensemble des évolutions liées à l'environnement de l'entreprise. D'un point de vue fonctionnel, la veille se décompose en cinq étapes : la définition/redéfinition du sujet, l'identification et sélection des sources d'information autrement dit le sourcing, la collecte et la sélection des informations, l'analyse et l'organisation des informations, la synthétisation et la mise en perspective et enfin la diffusion/communication de l'information

Il n'existe pas non plus de définition consensuelle pour la veille technologique. D'un point de vue de l'entreprise, la veille technologique est une activité mise

en œuvre par l'entreprise pour suivre les évolutions susceptibles d'influer sur le devenir de son métier. Autrement dit une veille technologique est constituée par l'ensemble des techniques visant à organiser de façon systématique la collecte, l'analyse, la diffusion de l'exploitation des informations techniques utiles à la sauvegarde et à la croissance des entreprises. La revue *La recherche* donne une définition d'un point de vue opérationnel de la veille technologique. Selon cette revue, elle « est le moyen pour l'entreprise de faire émerger les éléments stratégiques de la masse d'information disponible aujourd'hui. Ni espionnage industriel ni réalisation d'un état de l'art purement spéculatif dans un domaine technique restreint, la veille est avant tout destinée à éclairer les responsables de l'entreprise dans la résolution des problèmes industriels auxquels ils sont confrontés ».

Si l'on veut qu'une veille technologique fonctionne, qu'elle donne des résultats, il est crucial que la procédure soit bien définie à l'intérieur de l'entreprise, qu'elle soit soutenue par la direction de l'entreprise, que la culture d'entreprise soit participative à tous les niveaux hiérarchiques et que l'information soit toujours diffusée aux bonnes personnes et départements.

A-2 Les objectifs de la veille technologique

Les objectifs de la veille technologique sont la détection de signaux faibles, la recherche d'informations utiles aux décisions et aux actions, autrement dit l'apport d'une aide aux décisions et aux actions à caractère stratégique pour l'entreprise. Pour Daniel Rouach, « les signaux isolés, sporadiques... doivent être détectés par les veilleurs technologiques, c'est l'objectif essentiel de la veille technologique ».

Donc la veille technologique a une véritable importance au sein de l'entreprise, notamment en période de crise comme celle que les entreprises subissent actuellement. Elle leur permet de trouver de nouvelles pistes ou d'anticiper sur de nouvelles technologies, de trouver de nouvelles applications à leur produit ou alors de savoir s'il est possible d'utiliser de nouvelles technologies dans leur domaine, par exemple pour le commerce. En effet il faut savoir que, pour qu'une nouvelle technologie soit intéressante, il faut connaître son potentiel d'application. La veille concurrentielle peut s'apparenter à la veille technologique mais seulement si l'on s'intéresse à l'outil industriel du concurrent et non à ses clients sinon on ferait de la veille commerciale voire marketing. Cette partie

concurrentielle de la veille technologique se traduit essentiellement par la surveillance des brevets et des publications scientifiques de recherches fondamentales et appliquées. Cependant cette partie peut aussi conduire à des dérives notamment l'espionnage industriel c'est-à-dire la consultation d'informations noires.

Après que nous ayons défini la veille technologique, nous abordons les outils qui permettent cette veille, en s'intéressant en particulier à un logiciel : *Tropes Zoom.*

B) Les outils pour la réalisation de la veille : le cas de Tropes Zoom

B-1 Une diversité d'outils logiciels dans les systèmes de veille

Le premier des outils utilisés pour réaliser une veille est avant tout l'humain. En effet, il intervient à chaque étape et c'est l'humain qui réalise la veille, les outils logiciels ne sont qu'une aide. Il n'existe pas encore d'outil capable de remplacer l'humain dans un système de veille. Cependant si le seul outil du système de veille était l'humain, les veilles seraient longues et deviendraient inutiles car une des caractéristiques de la veille c'est d'avoir rapidement une vision de l'avenir (les nouvelles technologiques, les nouvelles tendances, etc.), d'apporter une aide à l'anticipation, à l'élaboration de stratégies (commerciale, marketing, etc.). Donc la vitesse de réalisation joue un rôle important dans la réalisation d'une veille. Les sources d'informations, qui sont de plus en plus nombreuses et de plus en plus riches, sont l'autre raison de l'indispensabilité des outils d'aide à la veille. Le nombre de publications scientifiques est en perpétuel augmentation et surtout de plus en plus d'entre elles sont accessibles sur internet, sur des bases de données payantes ou gratuitement sur des archives ouvertes, archives qui se développent de plus en plus notamment dans les pays anglo-saxons. Les outils logiciels aident à la collecte, à l'indexation et à l'analyse de l'information. Cette croissance perpétuelle des ressources a conduit également au développement d'outils de plus en plus automatisés, ce qui permet aux veilleurs, à temps de travail égal, d'accroître le nombre de sources d'information à surveiller et d'améliorer l'analyse des variations de contenu. A noter que dans certain cas, la réalisation d'études notamment, l'utilisation d'outils peut être secondaire, surtout si on a une bonne connaissance des sources d'informations.

Il existe une large gamme d'outils logiciels d'aide à la veille. Nous pouvons différencier deux types d'outils : ceux qui proposent un ensemble d'outils qui englobe tout le

système de veille autrement dit des plates-formes de veille comme les suites *Digimind* ou *AMI Software* qui permettent à plusieurs utilisateurs de créer différentes bases de surveillance, de capitaliser le contenu collecté, de l'enrichir par leurs commentaires et d'amorcer des analyses de premier niveau grâce à des modules de cartographie de l'information. Elles se composent donc de plusieurs modules allant de la recherche de sources d'informations à la diffusion, le plus souvent par mail, en passant par l'extraction ou la collecte, l'indexation, l'aide à l'analyse et des outils de présentations des résultats de la veille comme les outils cartographiques tel que *Word Mapper* qui sont aussi utilisés pour l'analyse de l'information.

La deuxième catégorie d'outils concerne ceux spécialisés dans une des étapes du système de veille. On trouve par exemple les métamoteurs de recherche qui permettent d'interroger simultanément les bases de données de différents moteurs de recherche et d'en compiler les résultats, les requêtes d'interrogation pouvant être régulièrement mises à jour par le logiciel. *SQUIDO,* qui est utilisé à l'ARIST, est un métamoteur de recherche, appelé également moteur d'exploration. Cet outil explore le web en allant de lien en lien jusqu'à un certain niveau défini lors du paramétrage de la recherche. Il propose de renouveler la recherche à intervalle régulier et envoie une alerte par mail quand l'exploration est finie. On trouve également des outils de surveillance de pages web qui permettent aux utilisateurs d'êtres alertés des changements intervenus sur certain sites web en fonction de critères comme la présence de mots clés. *KB Crawl* est l'un des outils de surveillance de page web le plus connu. Il permet de surveiller les pages web en programmant des crawls, en aspirant les pages web, en utilisant des mots clés et en délimitant les zones à surveiller pour écarter les bannières publicitaires ou les menus qui peuvent contenir des mots clés.

Enfin il existe une multitude d'outils qui ne sont pas forcément développés pour des applications de veille. On peut trouver des logiciels de traduction pour traduire soit les requêtes soit les résultats, des aspirateurs de sites, des aggrégateurs de flux RSS, des éditeurs de flux RSS, des outils de mapping comme *MindManager* pour présenter des résultats de veille de manière plus visuelle, etc. A noter que les outils de cartographies sont de plus en plus utilisés comme aide à l'analyse de l'information. Ils permettent une vision plus claire des idées ou informations importantes présentes dans un corpus de documents, corpus qui sont composés de plus en plus de documents ; c'est pourquoi les outils de cartographie connaissent une augmentation du taux d'utilisation dans les systèmes de veille.

Donc nous pouvons soit utiliser une plate-forme de veille avec toute les fonctionnalités nécessaires pour chaque étape du système de veille ; soit plusieurs outils, d'éditeurs différents, spécialisés dans une des étapes et que l'on associe pour former une chaîne de logiciels comme cela se fait à l'ARIST. Cependant il ne faut pas oublier que ces outils ne sont qu'une aide et que l'humain reste encore indispensable durant la réalisation d'une veille.

Après avoir vu de façon général les outils d'aide à la réalisation d'une veille, nous allons nous intéresser à un outil en particulier : *Tropes Zoom.*

B-2 Un outil d'aide à l'analyse : *Tropes Zoom*

-Présentation et fonctionnement de *Tropes Zoom*

Pourquoi le logiciel *Tropes Zoom* ? Premièrement, ce logiciel se situe au cœur de ma mission de stage car le dictionnaire métier que je dois réaliser doit être intégré dans cet outil. Deuxièmement, *Tropes Zoom* est l'outil que j'ai le plus utilisé durant mon stage autant pour réaliser les veilles sur la RFID, que comme aide à l'élaboration du dictionnaire métier. Le logiciel *Tropes Zoom* est un logiciel d'analyse sémantique créé en 1994 par Pierre Molette en association avec l'université de Paris VIII et depuis édité par *Acétic.* Cet outil est composé de deux modules : *Tropes* pour l'analyse et *Zoom* pour la recherche et récemment un troisième outil appelé *Robot* pour la collecte sur le web (pas utilisé à l'ARIST).

Le module *Zoom*[2] est un moteur de recherche sémantique. Il indexe automatiquement un corpus de documents soit autour des informations qui structurent les documents soit autour d'un scénario que nous appelons dictionnaire métier : ma mission consiste à son élaboration. Un mot suffit pour retrouver tous les documents contenant ce mot, lui et ses équivalents. Cette fonction d'indexation identifie avec précision les informations contenues dans le texte. *Zoom* est également doté d'une fonction recherche qui propose, après avoir posé une requête en langage naturel, une liste de critères pertinents qui ont émergé des documents indexés ou du dictionnaire métier. *Zoom* propose de visualiser le document soit dans un navigateur internet soit sous format XML.

[2] Voir annexe 2 p57

Le module *Tropes*[3] est un outil d'analyse sémantique. Le moteur d'analyse de *Tropes* utilise une architecture en série décomposée en six étapes principales. Le texte ou le corpus de textes sont chargés depuis *Zoom* puis ils subissent une analyse morphologique c'est-à-dire repérer les mots qui composent le texte, c'est la première étape.

Ensuite les phrases subissent une analyse syntaxique et les formes une analyse lexicale, c'est la deuxième et troisième étape. Cette étape permet d'affecter des catégories grammaticales (les verbes, les connecteurs, les modalisations, les adjectifs qualitatifs, les pronoms personnels, les substantifs et les noms propres) à chaque mot et de trouver une unité cognitive de base, la proposition. Ce traitement n'utilise pas de règles de grammaire, il se fonde sur une analyse des probabilités d'apparition d'une syntaxe, pour chaque mot ambigu, à l'intérieur de chaque proposition. Ce traitement d'analyse lexicale est nécessaire car les ambiguïtés sont si fréquentes qu'il est souvent impossible de comprendre un mot sans l'avoir préalablement affecté dans la bonne catégorie grammaticale.

La quatrième étape, l'analyse sémantique, *Tropes* l'effectue en deux traitements : l'attribution de métacatégories et la classification des références. Pour le premier traitement, *Tropes* utilise un dictionnaire général dont les mots sont classés dans des catégories sémantiques. La classification des références est l'étape la plus complexe du traitement d'un texte par le logiciel. Elle s'appuie sur un réseau sémantique et une logique probabiliste d'attribution de classes d'équivalents sémantiques aux mots ambigus. La levée d'ambiguïté sémantique est effectuée en calculant, pour chaque mot, la meilleure probabilité d'apparition d'un contexte.

L'analyse cognitivo-discursive est la cinquième étape. C'est une des caractéristiques de *Tropes*. Par cette analyse, *Tropes* extrait les propositions qui caractérisent le texte, sans se répéter. On appelle ces propositions des « propositions remarquables » L'aboutissement de cette analyse est la structure fondamentale de la signification (SFS) qui est composée de propositions remarquables qui assurent le lien entre le début et la fin de l'action. L'analyse se fait par le calcul d'un score d'appartenance à la SFS. Cela permet à l'analyste de jouer sur le taux de contraction du texte et le logiciel permet de régler la quantité de propositions affichées. Toutefois les propositions remarquables, tout comme le SFS, ne constituent qu'un résumé du texte ou d'un corpus de textes.

[3] Voir annexe 3 p59

Enfin la sixième et dernière étape, les statistiques, *Tropes* effectue des statistiques sur la répartition en sous-catégories des métacatégories sémantiques. Pour les verbes, par exemple, le calcul portera sur la répartition des catégories factif, statif et réflexif. Ces statistiques vont permettre de diagnostiquer les catégories de mots considérées comme les plus fréquentes et pour déterminer le style général du texte

-Analyse critique de *Tropes Zoom*

Tropes Zoom est un outil professionnel. Si on veut le maîtriser à 100%, une formation peut être nécessaire cependant la documentation fournie avec le logiciel peut suffir pour apprendre à utiliser le logiciel. Durant mon stage, nous n'avons utilisé qu'une partie des fonctionnalités proposées par *Tropes Zoom*. Du point de vue d'un nouvel utilisateur qui découvre le logiciel, *Tropes Zoom* apparait comme un outil très complet, avec de nombreuses fonctionnalités. L'interface est claire, compréhensible. Les boutons ne sont pas obscures, nous comprenons rapidement quelles sont les fonctionnalités qui y sont reliées : soit l'icone est explicite soit le nom de la fonctionnalité est indiqué. Certaines fonctions sont plus évidentes à saisir que d'autres. Les fonctions concernant la recherche, l'indexation pour la sélection des documents sont très simples à appréhender alors que les fonctions cartographiques, par exemple, ne sont pas évidentes à comprendre lors d'une première approche (signification des couleurs, des formes, etc.). Toutefois l'accompagnement de l'utilisateur est bien intégré. Des onglets « expliquer » sont présents à chaque action de l'utilisateur (lors de l'indexation, l'édition de carte, etc.) et l'aide est clairement visible.

D'un point de vue technique, *Tropes Zoom* est un outil performant proposant une analyse fine. On considère qu'un mot sur quatre est ambigu, c'est l'une des principales difficultés que l'on rencontre lorsque l'on fait une analyse de contenu. *Tropes* est capable de détecter et résoudre les mots composés, nombres, césures, abréviations, etc. En effet durant mon stage, nous n'avons pas constaté d'erreur de reconnaissance de mots composés ou d'abréviations. Toutefois *Tropes* ne détecte pas les fautes d'orthographe ni les fautes de style. Il considère que le texte ou le corpus de textes traités ont été corrigés, comportent une ponctuation et respectent les conventions typographiques de la langue française ou anglaise. Ne pas détecter les fautes d'orthographe peut poser problème car si les textes traités proviennent de blogs, les risques de fautes sont plus importants que dans les publications scientifiques. Enfin, l'un des points fort de *Tropes Zoom*, c'est sa complémentarité avec d'autres outils comme *Excel* pour éditer des graphiques, l'outil *PDI* pour éditer des rapports

et les envoyer aux clients ou encore dernièrement, l'outil cartographique PIKKO (complémentarité possible mais pas encore faisable).

C) La veille technologique sur la RFID et la 3D

Durant mon stage, nous avons réalisé une veille sur la RFID (Radio Frequency IDentification) et la 3D dans le cadre du commerce pour le PICOM. Pour le premier, l'objectif était de fournir de l'information sur des technologies RFID et des applications qui en découlent, concernant l'usage de la RFID par les clients après leur passage en caisse. Cette requête m'a amené à faire des recherches sur la protection des données se situant dans la puce RFID, sur la désactivation/mise en sommeil et la réactivation de la puce ou encore sur l'internet des objets. Cette veille a abouti à l'envoi de trois synthèses correspondant à 26 articles et d'un tableau récapitulatif des acteurs et de leur produit[4]. Concernant la veille sur la 3D, elle avait pour objectif la rédaction d'articles[5] dont le premier portait sur l'évolution de la 3D dans le cadre du commerce en 2009 : les nouveautés technologiques, les nouveaux concepts et les perspectives d'avenir. Ce travail m'a conduit à faire des recherches sur la 3D, la réalité augmentée ou encore les interfaces en 3D. Ces deux veilles ont été réalisées avec *Tropes Zoom*.

Durant mon stage, pendant la réalisation des veilles, nous avons constaté que ces veilles avaient finalement trois objectifs. La veille proprement dite, technologique dans le cadre de mon stage, est le 1e objectif. Le second c'est l'aide à l'enrichissement du dictionnaire métier et le troisième, la veille comme un moyen de valider les termes du dictionnaire métier. Ces trois objectifs forment une boucle. En effet on utilise le dictionnaire métier pour réaliser la veille qui, elle, aide à enrichir et valider le dictionnaire.

Ayant vu le premier objectif : la veille technologique, nous allons aborder le 2e objectif : l'aide à l'enrichissement du dictionnaire métier et à travers cela, l'élaboration du dictionnaire.

---

[4] Voir annexe 4 et 5 p61 et 71

[5] Voir annexe 6 et 7 p74 et 75

II La veille comme aide à l'enrichissement du dictionnaire métier

A) Le dictionnaire métier : un thésaurus simplifié

Tout d'abord nous allons rappeler ce qu'est un thésaurus et les grandes caractéristiques d'un thésaurus.

A.1 Les caractéristiques d'un thésaurus

Historiquement, l'emploi du terme thésaurus remonte au XVIe siècle mais le thésaurus que nous utilisons aujourd'hui comme un outil n'apparait que dans les années cinquante. L'ISO (Organisation International de normalisation) donne une définition d'un thésaurus selon sa fonction et/ou sa structure.

- Du point de vue de sa fonction, un thésaurus est un instrument de contrôle de la terminologie utilisé pour transposer, traduire en un langage artificiel plus strict c'est-à-dire un langage documentaire, le langage naturel employé dans les documents par les indexeurs ou les usagers. Le thésaurus permet aussi de faire l'action inverse : revenir au langage naturel à partir du langage documentaire.

- Du point de vue de sa structure, le thésaurus est un vocabulaire contrôlé, organisé et dynamique, constitué de termes ayant entre eux des relations sémantiques et génériques, et qui, de façon exhaustive, s'applique à un domaine particulier de la connaissance.

Généralement la définition ISO fait autorité mais il existe une multitude de définitions de thésaurus comme celle-ci : un thésaurus est une liste d'autorité de descripteurs et non descripteurs obéissants à des règles terminologiques propres et reliés entre eux par des relations sémantiques. Cette liste permet de traduire en un langage artificiel dépourvu d'ambiguïté des notions exprimées en langage naturel. On peut aussi le définir ainsi : Le thésaurus est une liste structurée de concepts destinés à représenter de manière univoque le contenu des documents et des questions dans un système documentaire déterminé. Toutes ces définitions sont une reformulation plus ou moins proche de la définition de l'ISO.

La structuration et la hiérarchisation sont des caractéristiques fondamentales des thésaurus. Il existe quatre principaux types de relations sémantiques qui structurent un thésaurus. Tout d'abord la relation hiérarchique qui se fonde sur de la « supériorité » et de la « subordination ». Cette relation sert à localiser des concepts généraux ou spécifiques à l'intérieur d'une liste ordonnée. Cette hiérarchisation permet d'avoir un important facteur d'amélioration du rappel (allant du spécifique au générique) ainsi que du degré de précision (allant du générique au spécifique). La relation dit d'équivalence est la deuxième relation de base que l'on peut trouver au sein d'un thésaurus. La norme ISO donne comme définition : « la relation entre descripteurs et non descripteurs lorsque l'on considère que, à des fins d'indexation, deux termes ou plus renvoient au même concept ». Dans une classe de termes équivalents, un seul terme est retenu pour représenter le concept, il est toujours choisi en fonction des besoins des utilisateurs du thésaurus. Le troisième type de relation qui apparait dans le thésaurus est : la relation associative. Les normes ISO considèrent qu'une relation associative est possible entre deux descripteurs « s'ils sont mentalement associés au point qu'il soit nécessaire de mettre en évidence dans le thésaurus le lien qui les relie, la raison étant que celui-ci révélera d'autres termes susceptibles d'être employés pour l'indexation et la recherche ». En langage clair, la relation associative existe entre des termes étroitement liés par le sens mais non hiérarchiquement et qui ne sont pas en relation d'équivalence.

Les descripteurs et les non descripteurs sont une autre caractéristique du thésaurus. C'est l'unité de représentation d'un contenu qui est de nature conceptuelle. On considère que le descripteur est la clé d'accès aux documents dont il synthétise le contenu. Les descripteurs sont en général considérés comme des mots clés. Les non descripteurs désignent des concepts identiques ou voisins de ceux représentés par les descripteurs et sont susceptible d'être utilisés par les futurs utilisateurs.

Enfin nous rappelons que l'objectif premier du thésaurus est de permettre l'indexation de documents afin de les retrouver plus facilement car on indexe un document dans le but de le retrouver. Cependant le thésaurus a aussi un rôle de structuration d'un domaine de connaissance, autrement dit une représentation des connaissances d'un domaine.

Après avoir vu les caractéristiques d'un thésaurus, nous allons maintenant situer le dictionnaire métier par rapport au thésaurus.

A.2 Le dictionnaire métier

Comme nous le disons dans le grand *I Veille technologique,* le dictionnaire métier a pour but d'être intégré dans le logiciel *Tropes Zoom* pour indexer des documents et les analyser dans le cadre d'une veille.

Le dictionnaire métier est une liste de termes qui peuvent prendre la forme de mots ou d'expressions. Donc on peut considérer ces termes comme des descripteurs. De plus, le dictionnaire métier est une liste de termes structurés. Cependant une seule relation sémantique est présente dans le dictionnaire : la relation hiérarchique. En effet le dictionnaire métier ne comporte pas de relation associée mais une forme de relation d'équivalence différente de celle d'un thésaurus. Nous trouvons dans le dictionnaire métier des synonymes de mots clés mais aussi les termes sous leur forme au singulier et au pluriel. Pour les termes ayant des accents, nous trouvons dans le dictionnaire une version avec et sans accent. Ces termes synonymes ne se différencient pas des termes que nous pouvons appeler descripteurs (voir extrait). Donc le dictionnaire métier possède une des caractéristiques fondamentales du thésaurus, il est structuré, hiérarchisé. Toutefois il est dépourvu de non-descripteurs.

Le dictionnaire métier possède des caractéristiques de base du thésaurus, il est composé de descripteurs et surtout il est structuré. Cependant nous ne pouvons considérer le dictionnaire métier comme un thésaurus car il ne prend pas en compte certaines relations sémantiques comme la relation associative ou la relation par équivalence dont le dictionnaire métier possède une version qui ressemble un peu à celle du thésaurus mais, comme nous l'avons vu, n'est pas la même chose. L'autre raison qui fait que nous ne pouvons ne pas considérer le dictionnaire métier comme un thésaurus est qu'il est dépourvu de non-descripteurs étant donné qu'il ne possède pas de relation par équivalence comme dans un thésaurus. Donc, en possédant les bases d'un thésaurus mais ne possédant pas toutes ces caractéristiques et en ayant le même objectif de représenter un domaine de connaissance voir de l'organiser, nous pouvons considérer le dictionnaire métier comme un thésaurus simplifié.

Exemple : Extrait du dictionnaire métier sur le commerce

Les termes en italiques sont les équivalents.

Type de commerce

Commerce de détail
Commerce
Commerce de detail
Commerce de détail
Detaillant
Détaillant
Discount
Dépôt vente
Depot vente
Dépôt vente
Depots ventes
Dépôts ventes
Depot-vente
Dépôt-vente
Hard discount
Hard discount
Hard-discount
Maxi discompte
Maxi discount
Maxidiscompte
Maxidiscount
Vente au detail
Vente au détail

Est-ce que cette simplicité fait du dictionnaire métier un outil plus pratique que le thésaurus ? Chacun de ces outils ont leurs avantages et leurs défauts. Toutefois, il faut faire attention au contexte de chaque outil car nombre de défauts sont souvent dus aux contraintes imposées par les logiciels, dans lesquels on intègre un dictionnaire métier ou un thésaurus. Par exemple, le dictionnaire métier ne possède pas de relation associative car le logiciel *Tropes Zoom* ne prend pas en compte ce type de relation sémantique. Donc comparer un thésaurus au dictionnaire métier n'est pas une chose facile. Toutefois, nous pouvons dire que le dictionnaire a un contenu plus concret car il est utilisé pour indexer des documents (le plus souvent des articles scientifiques) en fonction des termes de l'article se trouvant également dans le dictionnaire métier. Le thésaurus, lui, indexe les documents en fonction du contenu, avec des mots clés représentatifs du contenu, mots clés pas forcément présents dans le contenu. C'est la grande différence entre le thésaurus et le dictionnaire et la grande difficulté pour l'élaboration du dictionnaire. En effet les termes qui le composent doivent être utilisés dans les articles indexés, ce qui n'est pas le cas pour le thésaurus. Le dictionnaire métier est vraiment fait pour la réalisation d'une veille. Cependant, le thésaurus est plus

pertinent pour l'indexation, pour retrouver les documents et la représentation des connaissances. Donc nous pourrions songer qu'à l'avenir, un outil comme *Tropes Zoom* puisse incorporer un outil de représentation des connaissances plus évolués de type thésaurus, ce qui améliora l'analyse de contenu et donc la veille réalisée avec ce logiciel.

Après avoir situé le dictionnaire, nous allons maintenant aborder l'élaboration du dictionnaire métier.

B) Méthode d'élaboration du dictionnaire métier

B-1 Méthode d'élaboration d'un thésaurus

Pour développer un thésaurus, deux méthodes sont possible. Une première méthode que je qualifie de « classique » où l'on part littéralement de rien et la méthode de mise à jour où l'on adapte un thésaurus déjà existant en fonction du système d'information dans lequel il sera intégré. Mais avant de commencer le travail d'élaboration, il faut mener un travail de conception et de réflexion pour savoir quel type de thésaurus on a besoin de manière à choisir la méthode d'élaboration du thésaurus et afin de garantir la pérennité du thésaurus.

La méthode dite « classique » se compose de quatre étapes : collecte de termes, comparaison des termes, hiérarchisation des termes et la validation du thésaurus. Pour la collecte des termes, il faut déterminer le domaine de connaissance qui sera couvert par le thésaurus et en établir les limites. Il faut établir les sous-domaines et distinguer ceux qui vont être traités en profondeur et ceux qui resteront marginaux. La collecte de termes peut se faire à partir de trois sources : les mots issus du langage naturel c'est-à-dire du fond documentaire existant ou de la littérature du domaine de connaissance que l'on a préalablement défini, ceux issus d'un langage normalisé c'est à dire des terminologies ou thésaurus validés déjà existants et enfin la source orale de l'expérience de spécialistes du domaine dont traite le thésaurus mais cela peut être aussi des futurs utilisateurs. La deuxième étape de comparaison et de mise en relation des termes permet de faire un premier tri des termes collectés. Lors de cette étape, on classe les termes et on les regroupe dans les sous-domaines, cela permet de repérer les groupes de termes qui expriment le même concept. La comparaison permet de voir quels sont les termes les plus pertinents pour pourvoir différencier les descripteurs et les non-descripteurs. La hiérarchisation, la troisième étape, est

un travail de réflexion pour couvrir au mieux le domaine de connaissance que doit traiter le thésaurus tout en restant cohérent avec le public à qui on destine le thésaurus. Cette étape est une des plus importantes étant donné que la structure est une caractéristique fondamentale du thésaurus. La dernière étape est la validation du thésaurus par des spécialistes du domaine de connaissance traité et l'application de différents tests pour vérifier la pertinence des descripteurs.

La méthode dite de « mise à jour » reprend les étapes de la première méthode mais là on doit mettre à jour un thésaurus qui permet son adaptation à un fonds documentaire déjà existant ou à un domaine de connaissance. . Il faut savoir qu'il est rare voire très rare de trouver un thésaurus qui correspondrait parfaitement à notre fond documentaire ou à un domaine de connaissances. La plupart du temps il faut adapter, mettre à jour le thésaurus en ajoutant ou en supprimant des descripteurs voire en ajoutant ou en supprimant des branches entières du thésaurus. Cette méthode est plus rapide que la méthode dite classique car, en mettant à jour un thésaurus, les étapes de hiérarchisation et de collecte de terme sont moins longues étant donné que ce travail est en partie fait par les concepteurs du thésaurus que l'on adapte. Quand on utilise cette méthode pour mettre en place un thésaurus, il ne faut pas oublier d'avertir les concepteurs du thésaurus de la démarche même s'ils mettent gratuitement et librement leur thésaurus sur Internet par exemple.

Après avoir décrit les méthodes d'élaborations d'un thésaurus, nous allons maintenant voir si elles sont applicables entièrement, partiellement ou pas du tout pour le dictionnaire métier.

B-2 Méthodes partiellement applicables pour le dictionnaire métier

Tout comme pour le thésaurus, avant l'élaboration proprement dite du dictionnaire métier, nous avons défini le domaine de connaissance du commerce. Avec mon tuteur de stage M. Jean Dufour et Mme Emilie Obert du PICOM, nous avons établi les orientations que devaient prendre le dictionnaire métier qui tourne autour des axes de la veille faite au PICOM présenté dans la partie 1. A partir de là nous pouvions commencer l'élaboration du dictionnaire métier.

La méthode de mise à jour peut être utilisée pour le dictionnaire métier cependant je n'ai pas appliqué cette méthode durant mon stage. Une des principales raisons de ne pas avoir choisi cette méthode est que je n'ai pas trouvé de thésaurus sur le commerce ou sur une partie du domaine de connaissance du commerce. Par contre au début du stage, j'ai réalisé un état de l'art, commencé dans le cadre d'un projet tutoré, des terminologies, des thésaurus et autres dictionnaires, essentiellement sur internet. Le résultat des recherches fut que nous n'avons pas trouvé de thésaurus sur le commerce d'un point de vue général. Nous avons surtout trouvé des terminologies et des dictionnaires sur des sujets faisant parties du domaine de connaissance du commerce comme le marketing, le e-commerce, le commerce équitable, etc. Finalement nous avons trouvé un thésaurus proche de ce que l'on cherchait c'est-à-dire un thésaurus général sur le commerce : le thésaurus de Delphes[6] édité par l'ARIST de la Chambre de Commerce et de l'Industrie de Paris, dont la dernière mise à jour date de 2004 et qui traite de l'entreprise, de ses marchés, ses finances et sa gestion dans l'environnement français et international. Ce thésaurus fut très intéressant même si la structure de base ne correspondait pas à ce que j'avais besoin. Donc la méthode de mise à jour est théoriquement possible mais je ne l'ai pas utilisée durant mon stage.

Concernant la méthode dite « classique », elle n'est que partiellement applicable pour le dictionnaire métier. La première étape que j'ai réalisée fut de mener des recherches d'études, d'articles et autres documents, majoritairement sur internet, pour me familiariser avec le monde du commerce. Ces recherches ont été faites simplement avec Google. Puis j'ai approfondi ces recherches sur certains thèmes : le e-commerce, la logistique, le marketing, l'innovation et les technologies dans le commerce, sur l'environnement et le développement durable. En parallèle de ces recherches, j'ai continué le sourcing sur le commerce commencé par M. Dufour. Ce travail de recherche est le même pour le thésaurus que pour le dictionnaire quand on ne connait pas le domaine qu'il couvrira.

La phase collecte est très semblable entre la méthode d'élaboration du thésaurus et celle du dictionnaire métier. Tout d'abord les sources des termes sont les mêmes. Ils proviennent de terminologies et de thésaurus, d'articles collectés et visionnés via *Tropes Zoom* et d'entretiens avec des spécialistes même si je n'ai pas utilisé cette dernière source cité durant mon stage. Ces trois sources sont théoriquement applicables pour le dictionnaire. Cependant, cette étape ne se déroule pas forcément au même moment, un gros

[6] http://www.arist.ccip.fr/Thesaurus-Delphes-2004-Extraits.pdf
http://www.arist.ccip.fr/partenaires.htm#delphes

travail de collecte de termes est réalisé au début pour les deux outils. Durant mon stage, j'ai constaté que, pour le dictionnaire métier, ce travail de collecte se prolonge car d'une part la base documentaire se renouvelle rapidement, à chaque nouvelle collecte faite par les robots, et d'autre part, l'élaboration du dictionnaire impose un travail important de recherche de synonymes, pour que l'indexation soit la plus pertinente possible et pour éviter de passer à côté de documents importants pour la veille.

La phase de comparaison et de mise en relation est peu importante voire inexistante lors de la réalisation du dictionnaire métier étant donnée qu'il est dépourvu de non-descripteurs. La réflexion sur le choix des descripteurs et des non-descripteurs ne se fait pas pour le dictionnaire métier. Par contre, comme je l'ai déjà dit, un gros travail sur les synonymes est réalisé car, je le rappelle, les termes ou expressions qui composent le dictionnaire doivent pourvoir être retrouvés dans les articles et *Tropes Zoom* ne fait aucune supposition, de déclinaison par lui-même à partir des termes du dictionnaire. Donc il faut intégrer dans le dictionnaire toutes les déclinaisons des termes ou expressions (singulier, pluriel, avec ou sans accent, etc.). Ce travail, qui est assez long à faire, est obligatoire pour que la veille réalisée à l'aide du dictionnaire soit pertinente et qu'on ne passe à côté d'aucune informations, de documents et autres signaux faibles qui répondraient aux objectifs de la veille. Après la collecte des termes issus des terminologies et des dictionnaires, on en a fait une liste dans un tableur Excel que l'on a décortiqué en faisant le tri entre les termes qui correspondaient aux orientations du dictionnaire. Par exemple nous avons supprimé la plupart des termes boursiers, financiers, de comptabilité. En même temps, on commençait à faire un travail de mise en relation de ces termes : les termes liés au marketing, à la logistique, au type de commerce, aux technologies, etc. Donc, dans l'élaboration du dictionnaire, cette étape de comparaison et de mise en relation est quasi inexistante si on considère que la mise en relation peut être faite durant l'étape de hiérarchisation.

Concernant, l'étape de la hiérarchisation, elle est applicable pour le dictionnaire métier. En effet après la collecte des termes et les premières mises en relation décrite dans le paragraphe précédent, j'ai effectué un premier travail de hiérarchisation, le même travail que l'on aurait réalisé pour un thésaurus. Cependant à l'inverse d'un thésaurus où après avoir déterminé les éléments racines on garde une certaine stabilité, la hiérarchisation d'un dictionnaire est plus évolutive, à cause de *Tropes Zoom* qui impose une entrée pour chaque terme ou expression. Cette contrainte a pour conséquence qu'il faille rajouter des éléments racines durant le développement du dictionnaire pour éviter la double

entrée. Le terme « technologie » illustre bien cette contrainte car dans un thésaurus il pourrait être présent dans plusieurs éléments racines mais étant donné que cela n'est pas possible dans le dictionnaire métier, le terme « technologie » devient un élément racine à part entière. Cette contrainte a aussi un avantage car elle élimine les ambiguïtés, en effet on ne risque pas de rencontrer deux fois le même mot dans des contextes différents.

Donc, l'élaboration du dictionnaire métier suit partiellement la méthode dite « classique » pour créer un thésaurus. On retrouve au cours de la conception du dictionnaire métier l'étape de collecte, de hiérarchisation et de la mise à jour, que je n'ai pas abordée dans cette partie car je le ferai dans la sous partie suivante, mais pas l'étape de la comparaison à cause de l'absence des non-descripteurs. Toutefois dans cette comparaison, il ne faut pas oublié de prendre en compte le contexte de ces outils c'est-à-dire que le dictionnaire est formaté pour *Tropes Zoom*.

Après avoir vu comment élaborer le dictionnaire métier, nous allons aborder une méthode d'enrichissement du dictionnaire grâce à la veille, méthode que j'ai découverte durant mon stage.

C) La veille, une méthode d'enrichissement du dictionnaire métier

Nous avons vu que, pour élaborer un dictionnaire métier, on reprend une bonne partie de la méthode de conception du thésaurus. Nous avons également exposé le constat que la veille durant mon stage avait 3 objectifs. Nous allons aborder ici le deuxième objectif. En effet durant mon stage, pendant que nous réalisions la veille sur les puces RFID ou celle sur la 3D dans le cadre du commerce, le premier objectif de la veille que nous avons abordé précédemment, nous avons constaté un usage secondaire possible de la veille : l'enrichissement du dictionnaire métier. Il faut d'abord préciser qu'au moment de commencer la veille, le dictionnaire n'était pas terminé. L'exemple qui illustre le mieux cet usage et sur lequel j'ai le plus travaillé, étant donné qu'on réalisait une veille technologique, est celui de l'élément racine « Technologie ».

En consultant et sélectionnant des documents sur la RFID ou sur la 3D, par l'intermédiaire de *Tropes Zoom*, nous avons trouvé d'autres technologies semblables ou

proches de la RFID ou de la 3D comme le *Visual Computing*[7] mais auquel nous n'aurions pas pensé ou tout simplement que nous ne connaissions pas. Ces technologies sont ajoutées au dictionnaire métier. La veille permet aussi de peaufiner le dictionnaire métier car elle permet de toujours mieux connaître le domaine de connaissance qui est dans le cadre de mon stage : le commerce. Le travail sur les synonymes illustre également bien l'enrichissement du dictionnaire métier grâce à la veille. Pour rappel, pour chaque concept, il faut intégrer dans le dictionnaire tous ces synonymes et les expressions utilisées pour l'exprimer. Pour pourvoir réaliser ce travail, la veille est d'une grande aide car la recherche et la consultation des documents, dans mon cas, des articles via *Tropes Zoom*, permet de repérer des synonymes et autres expressions que l'on peut ajouter dans le dictionnaire. Cette méthode est très pratique lorsque nous ne maîtrisons pas le domaine de connaissance. A noter que cet exemple est très spécifique au dictionnaire car comme nous l'avons vu, *Tropes Zoom* ne fait aucune supposition à partir des termes ou concepts du dictionnaire qui indexe les documents en fonction du contenu, donc il faut intégrer un maximum de synonymes et d'expressions utilisés pour exprimer les concepts.

Cependant cette méthode a ces limites car elle dépend de la veille effectuée. En effet, en reprenant les veilles sur la RFID et la 3D que nous avons réalisées, nous avons essentiellement enrichi l'élément racine « technologie » et un peu celui des « acteurs » mais les autres éléments n'ont pas été enrichis ou très peu. Elle comporte aussi un risque car la veille peut être réalisée alors que le dictionnaire métier n'est pas fini. La veille peut ne pas être optimale et être plus longue à réaliser, le dictionnaire étant peu affiné, un plus grand nombre de document est à traiter.

Donc la veille peut apporter une aide non négligeable pour l'enrichissement du dictionnaire métier mais il faut prendre le risque de faire une veille sans que le dictionnaire métier, soit fini et qui n'aidera qu'à compléter les éléments racines en rapport avec le sujet de la veille. Nous avons découvert cet usage secondaire de la veille dans le contexte de l'élaboration du dictionnaire et l'utilisation de *Tropes Zoom.* Mais est-ce que cette méthode est applicable dans un autre contexte ? Il serait intéressant de se demander si dans le cadre de l'élaboration d'un thésaurus faisant parti d'un système de veille, la méthode que nous avons décrite est applicable.

---

[7] Le visual computing : analyse, traitement et présentation visuelle d'information dans le but de modeler en temps réel l'interaction des utilisateurs avec les ordinateurs et autres terminaux de façon intuitive et aussi réelle que possible

Après que avoir défini et décrit l'élaboration du dictionnaire en faisant un parallèle avec le thésaurus, nous allons aborder le troisième objectif de la veille et la validation du dictionnaire métier tout en continuant de faire un parallèle avec le thésaurus.

III La veille comme un moyen d'évaluation du dictionnaire métier

A) Méthode d'évaluation d'un thésaurus

Un thésaurus, avant son utilisation, doit être validé par des spécialistes, validation qui peut se dérouler tout au long du processus d'élaboration du thésaurus. Il n'existe pas de méthode unique pour valider le thésaurus.

Cette évaluation ou validation peut passer par des entretiens, sous la forme de discussions avec les spécialistes tout au long de l'élaboration du thésaurus pour connaître leur avis, ou à la fin de l'élaboration lors de l'étape validation. Ces entretiens peuvent donner des indications sur le choix des descripteurs et la validation des descripteurs. L'analyse des discours des spécialistes peut aider à peaufiner le thésaurus pour l'adapter au mieux aux futurs utilisateurs. En effet, un thésaurus est réalisé en fonction du fonds ou du domaine de connaissance traité par le thésaurus mais aussi en fonction des futurs utilisateurs. Pour le thésaurus, c'est une garantie de pérennité.

Pour l'évaluation et la validation du thésaurus nous pouvons également réaliser des tests. Par exemple, si le thésaurus est réalisé dans le but d'indexer un fond documentaire, on indexe un échantillon, une partie du fond. Ensuite, on formule des requêtes sous la forme de descripteurs et non-descripteurs, puis on évalue la pertinence des réponses en calculant le taux de silence et le taux de bruit. Ces tests doivent être effectués par la personne qui réalise le thésaurus mais aussi par les futurs utilisateurs. En effet, pendant qu'ils le testeront, nous pourrons observer comment ils utilisent le thésaurus et s'ils rencontrent des difficultés car si les futurs utilisateurs ne savent pas manipuler le thésaurus, celui-ci ne peut être validé.

Donc la meilleure méthode pour valider un thésaurus c'est les tests. Cette méthode des tests est la plus simple à mettre en place et la plus pertinente pour mettre en évidence les lacunes ou les erreurs dans le thésaurus.

B) Méthodes partiellement applicables pour le dictionnaire métier

En théorie, les méthodes décrites dans le paragraphe précédent pour le thésaurus sont applicables pour le dictionnaire métier. Comme pour le thésaurus, nous pouvons faire appel à un spécialiste pour faire valider le dictionnaire à la fin du travail d'élaboration mais aussi tout au long de l'élaboration. Dans le cadre de mon stage, le spécialiste a été Mme Emilie Obert, responsable veille au PICOM. A noter qu'il est même conseillé que l'élaboration du dictionnaire métier (s'applique aussi pour le thésaurus) se fasse en collaboration avec un spécialiste.

Toutefois le dictionnaire métier, à l'instar du thésaurus, n'est pas seulement réalisé en fonction du public mais il est d'abord réalisé en fonction du contenu des documents. En effet, comme je l'ai déjà expliqué, l'indexation se fait en fonction des termes se trouvant dans les documents et le dictionnaire métier. Donc, le choix des termes dépend plus de leur présence dans les documents que de leur utilisation par les futurs utilisateurs. Cependant, il faut essayer de garder une certaine logique dans le dictionnaire métier, de manière qu'un utilisateur puisse trouver ce qui correspond à sa requête, sa problématique dans le dictionnaire métier.

Comme pour le thésaurus, nous pouvons également réaliser des tests pour valider le dictionnaire métier, sachant que ces tests ne se font qu'avec un fond documentaire connu. Un des premiers tests que l'on peut faire est de voir si chaque terme est utilisé pour indexer un document, si un terme du dictionnaire n'est relié à aucun document, ce terme doit être mal formulé. Ensuite nous pouvons évaluer la pertinence des documents indexés pour un terme. Ce test permet de valider les équivalents et de voir aussi si le terme est trop générique, car si un terme indexe un grand nombre de documents c'est qu'il est possible de peaufiner le dictionnaire. Ces tests peuvent être utilisés tout au long de l'élaboration du dictionnaire pour le valider au fur et à mesure. Donc, comme pour le thésaurus, les tests sont une des méthodes d'évaluation la plus simple à mettre en place et la plus pertinente pour mettre en lumière les lacunes et les erreurs.

Après avoir décrit les méthodes d'évaluation d'un thésaurus et de leur faisabilité pour le dictionnaire métier, nous allons aborder le 3e objectif de notre veille : un moyen de valider les termes du dictionnaire métier.

C) La veille, comme moyen de valider le dictionnaire métier

Durant mon stage nous avons fait le constat que la veille peut être un moyen pour valider les termes du dictionnaire métier. En effet les tests décrits dans la partie précédente peuvent prendre la forme d'une veille ou se faire durant une veille. Ce travail de validation peut se réaliser durant l'élaboration du dictionnaire. En faisant une veille, nous testons le choix des termes, leur formulation et/ou leur situation (équivalents ou termes).

Par exemple, dans le cadre de la veille sur la RFID, Mme Obert avait proposé des mots correspondants à la requête qu'elle nous avait exposée : « activation », « désactivation », « mettre en sommeil », etc. Nous avons intégré ces mots dans le dictionnaire métier en tant qu'équivalent du terme « RFID ». Ensuite nous avons poursuivi notre veille mais nous avons constaté que la place de ces mots n'était pas pertinente, qu'ils n'apportaient rien. Nous avons donc changé ces mots en termes auxquels nous avons ajouté des équivalents prenant la forme d'une déclinaison de formulation : « activer la puce », « activer les puces », « désactiver la puce », « désactiver la puce RFID », « activer la puce RFID », « activation de la RFID », etc. Ces termes sont placés comme sous terme de « RFID ». Ce changement permet de retrouver plus facilement les documents ayant pour objet le sujet qui correspond à ces termes, sinon les documents seraient noyés dans des dizaines voire des centaines de documents et donc plus longs à retrouver.

Cependant cette méthode dépend des documents collectés. En effet le commerce est un domaine de connaissance large, donc il est possible que les documents collectés ne couvrent pas tout ce domaine et donc ne permettent pas de tester toutes les parties du dictionnaire métier. Cette méthode dépend aussi de la veille effectuée. En effet les veilles que nous avons réalisées ne couvrent pas tout le dictionnaire donc la validation ne touchera également qu'une partie du dictionnaire.

Partie 3 : La veille et la représentation des connaissances

Nous avons vu que le dictionnaire métier est un outil permettant de réaliser une veille notamment lors de l'indexation et la recherche de documents autrement dit la recherche d'information. Nous avons également constaté, durant et dans le cadre du stage, que le dictionnaire métier, intégré à *Tropes zoom*, et la veille avaient trois types de relations : le dictionnaire comme outil d'aide à la réalisation d'une veille qui est elle-même, une aide à l'enrichissement du dictionnaire et enfin la veille comme un moyen pour la validation des termes du dictionnaire métier. Nous pouvons affirmer que le dictionnaire métier est une représentation et également une forme d'organisation des connaissances. Donc quels sont les liens entre la veille et la représentation des connaissances ? Quelle est la place de la représentation des connaissances dans la mise en place et la réalisation d'une veille ?

Nous allons voir tout d'abord ce qu'est la représentation des connaissances et nous ferons aussi un point sur l'organisation des connaissances. Ensuite nous expliquerons quelle est la place de la représentation des connaissances dans la mise en forme et la réalisation d'une veille et nous aborderons également les avantages et les inconvénients.

I Qu'est ce que la représentation des connaissances ?

A) Définitions

La représentation des connaissances désigne un ensemble d'outils et de procédés destinés d'une part à représenter et d'autre part à organiser le savoir humain pour l'utiliser et le partager. Les connaissances ne sont pas forcément représentées par des mots et des phrases. Mais qu'est ce que la connaissance ?

A-1 Définitions de « connaissance »

Il existe une multitude de définition de la notion, du concept de « connaissance ». Le dictionnaire Larousse définit la « connaissance » comme une activité intellectuelle visant à avoir la compétence de quelque chose. Dans le domaine de la gestion des connaissances, elle est définie comme correspondante à l'appropriation et l'interprétation des informations par les hommes. Pour D. CRIE[8] « la connaissance est donc le résultat d'une information traitée, compréhensible et assimilable par un être humain. C'est une manière de comprendre, de percevoir, elle régit les rapports entre les afférences cognitives de l'individu et le monde extérieur ». « La connaissance est possédée par les être humains ou les agents inanimés en tant que vérités et croyances, perspectives et concepts, jugements et attentes, méthodologies et savoir faire. La connaissance est utilisée pour recevoir l'information, la reconnaître et l'identifier ; l'analyser, l'interpréter et l'évaluer ; synthétiser, faire des hypothèses et décider ; s'adapter ou adapter, planifier, et piloter » pour K Wiig[9]. Ou encore « la connaissance est l'habilité individuelle à tirer des distinctions à l'intérieur d'un domaine collectif d'action, basée sur une appréciation du contexte ou d'une théorie, ou des deux » pour H.TOUSKAS[10]. Par exemple en philosophie, on distingue traditionnellement trois types de connaissance : la connaissance propositionnelle est le fait de savoir qu'une certaine proposition est vraie, par exemple, "savoir que la Terre est ronde" ; la connaissance objectuelle, aussi appelée acquaintance, est le fait de connaître une chose particulière, par exemple, "connaître Paris", le savoir faire est le fait d'être capable de réussir une action, par exemple, "savoir faire des crêpes". Donc il existe une multitude de définitions de la notion de « connaissance » en fonction des matières (philosophie, économie, etc.). Stéphane GORIA, dans sa thèse[11] a recensé plus de 35 définitions différentes de la notion.

En entreprise la connaissance correspond à un capital de compétences, d'expertises que détiennent les hommes et les femmes dans différents domaines professionnels (exemples : marketing, R&D, ingénierie, production, logistique,

---

[8] Crie D, 2001. *NTIC et extraction des connaissances*, paru dans : Les cahiers de la Recherche, avril, 2001

[9] Wiig K., 1999. *Knowledge Management: An emerging Discipline Rooted in a Long History*, paru dans : Knowledge Management, sous la direction de D. Chavel & C. Despres, Editions Theseus, Paris

[10] Tsoukas H., 2005. *Complex Knowledge: Studies in Organizational Epistemology*, Editions Oxford, New York

[11] Voir bibliographie

approvisionnements, commercial, juridique, ...) constituant le cœur de métier de l'entreprise. Ces compétences doivent être gérées et capitalisées pour améliorer l'efficacité globale de l'entreprise, la sécurité et la fiabilité des opérateurs et des traitements de connaissances, ainsi que l'accessibilité des connaissances par les usagers, notamment avec l'aide des technologies (dites technologie de la connaissance ou technologies cognitiques).

A-2 Définitions de la représentation des connaissances

Une représentation des connaissances est généralement comprise comme un ensemble de signes élaborés de telle manière qu'il puisse être utilisé comme une codification de certaines perceptions du monde (intellectuelles ou physiques). Pour R.J. Brachman et H.J. Lesvesque[12] la représentation des connaissances a pour objet la « description du monde de telle manière qu'un agent intelligent puisse déduire de nouvelles conclusions à propos de l'environnement par une manipulation formelle des descriptions ». De manière analogue, pour O. Mariño Drews[13], une représentation des connaissances est une « modélisation des différents éléments du monde réel et la détermination des procédures d'interprétation faisant le lien entre le monde et le modèle, tant au moment de l'acquisition de connaissances et l'élaboration du modèle, que pendant la manipulation de la représentation (pour donner des explications) et, finalement, lors de l'application des résultats du modèle du monde ».

La représentation des connaissances est, selon notre approche, un instrument d'aide à la communication et d'aide au raisonnement. Cet instrument s'appuie sur certains postulats et fondements culturels qui forment des conventions qui peuvent être remises en cause, mais qui sont à l'origine de tout développement de nouvelles représentations. J. C. Binwal & Lalhmachhuana[14] définissent la représentation des connaissances comme « un ensemble de conventions syntactiques et sémantiques qui rendent possibles la description de choses ». Nous noterons, à titre indicatif, qu'une représentation de connaissances, d'un point de vue psychologique, peut être considérée aussi comme un processus ; c'est-à-dire que,

[12] Brachman R.J. & Levesque H.J., 1985. *Readings in knowledge representation*, Editions Morgan Kaufman, Los Altos

[13] Mariño Drews O., 1993. *Raisonnement classificatoire dans une représentation à objets multipoints de vue*, thèse, Informatique, Université Joseph Fourier – Grenoble 1

[14] Binwal J.C. & Lalhmachhuana, 2001. Knowledge Representation : Concept, Techniques and The Analytico-Synthetic Paradigm, paru dans : Knowledge Organization, n°281. pp 5-16

pour accéder à leur compréhension, il nous faut développer un processus de réflexion et d'interprétation complexe. Ce que l'on attend d'une représentation de connaissances c'est qu'elle permette de guider ou de développer un raisonnement par inférence c'est-à-dire une opération mentale, ou jugement, qui consiste à tirer une conclusion d'une série de propositions reconnues pour vraies, conclusions qui sont tirées à partir de règles de bases

B) Les types de représentation des connaissances

Nous pouvons dégager trois grands types de représentations des connaissances : les représentations des connaissances en logique, les représentations sous forme de réseaux et les représentations visuelles des connaissances.

B-1 Les représentations des connaissances en logique

Pour Y. Delmas-Rigoutsos et R. Lalement[15], « la logique est une passerelle entre deux « mondes » : le monde des énoncés, univers formel, structuré par la forme même des énoncés, et celui des objets de la pensée (ou mathématiques), structuré par des relations qu'ils entretiennent.». La logique est donc tout d'abord un système ou un art, qui permet d'assurer un échange d'opinions et de représentations du monde, tout en permettant, s'il le faut, leur discussion et justification. Elle s'appuie toujours sur un langage (logique) constitué de signes, quelle que soit la logique utilisée. Donc une logique est bien une forme particulière de représentation de connaissances, car sa définition s'accorde parfaitement avec les définitions proposées pour la représentation des connaissances, définitions que nous avons exposé dans le paragraphe précédent. Les signes ou symboles du langage sur lequel s'appuie une logique, forment un alphabet. Ils représentent des variables, des constantes ou encore des fonctions logiques.

La première représentation des connaissances dédiées aux raisonnements logiques auquel on peut être confronté c'est l'assignation de valeur de vérité à des énoncés. « De manière très simple nous pouvons résumer cette idée au fait d'associer à un énoncé

---

[15] Delmas-Rigoutsos Y. & Lalement R., 2000. *La Logique ou l'Art de raisonner*, Editions Le Pommier-Fayard

particulier appelé proposition, une valeur de vérité : vrai ou faux, en fonction d'un certain jugement, fondé sur des règles logiques, qu'on lui porte »[16]. Donc sous une forme logique assez simple, les domaines de connaissances peuvent être appréhendés en termes de propositions sur lesquelles on peut statuer sur leur vérité ou leur fausseté. Toutefois pour pouvoir qualifier un énoncé de vrai ou faux, il faut tout de même l'exprimer à travers un langage. Les langages logiques ont pour ambition de rendre compte d'une partie des connaissances et des expressions employées pour définir et utiliser ces connaissances par les humains. Parmi celles-ci, il y en a beaucoup qui n'ont pas pour objet des propositions seulement mais caractérisent des individus. Nous n'allons pas ici détailler tous les langages car d'une part c'est complexe et d'autre part cela ne serait pas pertinent par rapport à notre sujet.

Donc une logique peut se définir comme un système formel qui définit les expressions bien formées d'un langage associé à la logique, c'est le composant syntaxique de la logique. Elle se définit aussi à partir d'un univers d'interprétation considéré comme l'espace sémantique pour le système formel. Enfin la logique de type classique se définit par une fonction interprétative qui définit des principes de correspondance entre les expressions bien formées du système formel et les entités de l'univers d'interprétations. Cette représentation permet d'effectuer des raisonnements censés être objectifs au regard des prémisses et des règles qui le définissent. La formulation la plus simple, permettant une représentation du monde avec le type de systèmes abordé ci-dessus, prend le nom de logique des propositions.

Il existe plusieurs types de logiques. Tout d'abord nous avons la logique de prédicats qui permet d'augmenter le potentiel d'expression sémantiques des systèmes de représentations logiques. Les langages ou logiques de prédicats ont pour but de parler des objets des énoncés en utilisant certaines de leurs propriétés sous la forme de prédicats. Sous l'étiquette prédicat sont regroupés des faits simples, des propriétés et des relations existant entre des objets dont on souhaite se servir pour développer un raisonnement objectif et éventuellement démontrer une propriété particulière ou justifier une argumentation. Nous avons ensuite les logiques modales qui permettent d'expliciter l'aspect contextuel et pragmatique des connaissances exprimées, notamment les notions de point de vue, de temporalité et le concept de modalité. La logique de traits est le troisième type de logique.

[16] Voir bibliographie

Elle se fonde sur deux langages : l'un « syntaxique » et l'autre « sémantique » et sur la structure de traits qui se définit à travers un petit nombre de caractéristiques distinctives nommées traits et qui peuvent être grammaticaux, sémantiques,... Cette structure se présente sous la forme d'un graphe qui comporte une racine à partir de laquelle partent des liens ou relations binaires orientées. Enfin la logique terminologie, dérivée de la logique des prédicats, elle se caractérise en termes de rôles, de concepts et d'individus, qui correspondent respectivement à des relations binaires qui caractérisent des individus ou des concepts, à des ensembles d'individus et à des entités élémentaires. Leur principale fonction est de rendre compte de connaissances qui sont présentables en terme de classification hiérarchique, de caractériser des relations de hiérarchies. Simplement, elle permet de représenter la connaissance terminologique d'un domaine d'application d'une manière formelle et structurée.

B-2 Les représentations des connaissances sous la forme de réseaux

Comme nous l'avons vu et comme nous le verrons il existe de nombreuses façons de représenter les connaissances. L'un des types les plus connus et les plus répandus concerne la représentation sous forme de réseaux. Il peut être un moyen simple de travailler par analogie, c'est-à-dire par ressemblance, avec le treillis complexe qui est censé correspondre à la représentation d'un domaine de connaissance et sur un plan plus général la représentation du monde de chaque individu. Il existe plusieurs formes de représentations des connaissances en réseaux.

Les cartes conceptuelles ou cartes de concepts (plus rarement carte de connaissances) sont, avant tout, des représentations graphiques et visuelles de compréhension de connaissances sous forme de réseaux. De manière générale une carte conceptuelle est une représentation de concepts reliés sémantiquement entre eux. Nous traitons les cartes conceptuelles dans cette partie sur les réseaux et pas dans la partie suivante sur les représentations visuelles des connaissances car l'idée de structure en réseau est fondamentale dans les cartes conceptuelles. Cette forme de représentation a été proposée

pour la première fois par J-Stewart & al[17] pour rendre compte de la perception d'une discipline par des personnes. Une carte conceptuelle est faite sur un plan, soit en deux dimensions soit en trois dimension, sur lesquelles s'agrègent des concepts, comme des bulles ou des étiquettes contenant un nom, à travers l'élaboration d'un réseau plus ou moins complexe de liens entre ces concepts. Ces cartes ont pour principaux avantages de permettre d'estimer la compréhension d'une personne pour un sujet donné, de pourvoir se faire assez facilement une idée de la complexité d'un sujet voir d'un domaine de connaissance, de se rendre assez facilement compte des parties d'un domaine (ou sujet) mal maîtrisé ou ignoré. Il permet à un personne néophyte et ne connaissant rien d'un domaine de connaissance de se faire une idée, de commencer à connaître ce domaine. Par exemple, la prise de note d'un étudiant peu prendre la forme d'une carte conceptuelle en partant du sujet du cours puis en créant un réseau autour avec des mots clés ou des points important du cour. Des logiciels comme *Wordmapper* ont été développés pour réaliser des cartes conceptuelles à partir d'un texte ou d'un corpus de textes. D'autres outils comme *Mindmanager* permettent de réaliser des cartes manuellement et pas à partir de textes.

La deuxième forme de représentation des connaissances par réseau c'est l'ontologie. Le terme « ontologie » recouvre trois approches distinctes qui ont chacune, soit en partie, soit totalement, pour objet la catégorisation et donc la représentation de connaissances. La première est philosophique et définit l'ontologie comme une partie de la réflexion métaphysique sur l'être en tant qu'être. Elle étudie à la fois l'existence des choses, les relations qui lient les choses à l'esprit, les relations qui lient les choses, dont l'homme, a l'univers et la manière dont on peut catégoriser les choses pour mieux les appréhender. La deuxième approche est orientée vers la formalisation des connaissances et considère l'ontologie comme un ensemble de théories et de règles dédiées à la classification des choses de l'univers. La troisième approche est plutôt technologique et considère l'ontologie comme un moyen, constitué d'un réseau de concepts, qui a pour objet la couverture des connaissances nécessaires à l'élaboration d'une tâche donnée ou la couverture d'un domaine particulier de connaissances. Une quatrième approche est possible, dans laquelle un vocabulaire commun destiné à des chercheurs qui ont besoin de partager des informations dans un domaine particulier. Elle inclut des définitions des concepts de base et des relations

[17] Stewart J. & al, 1979. Concept maps: A tool for use in biology teaching, paru dans : The American Biology Teacher, n°413. pp 171-175

entre eux, qui soient interprétables par un automate. L'intérêt des ontologies est d'avoir une capacité de recherche d'information en minimisant les ambiguïtés et en apportant un degré de précision suffisant aux recherches, d'avoir également une capacité de filtrage qui permet d'examiner une grande quantité d'informations et de diriger l'information intéressante en fonction des filtres.

Les classifications sont la troisième forme de représentation des connaissances en réseau. Une classification est un système organisé et hiérarchisé de classement des connaissances dans un domaine particulier, voire de la totalité des connaissances pour un certain intervalle de temps. Les principaux utilisateurs des classifications sont les spécialistes des sciences de la nature, les bibliothécaires et les philosophes. Il existe de nombreux types de classifications comme la classification aristotélicienne des connaissances qui représente la connaissance à partir de catégories distinctes : les catégories essentielles liées à la substance et les catégories accidentelles qui sont transitoires ou alors la classification kantienne des connaissances. En bibliothéconomie, la classification décimale de *Dewey* est la classification la plus répandue. Elle est basée sur une approche épistémologique visant à classer l'ensemble du savoir humain selon une division hiérarchique. La classification permet de situer les connaissances les unes par rapport aux autres. Elle facilite le rangement des exemplaires physiques, qui est une de ses caractéristiques. On peut également citer la classification décimale universelle (CDU) qui est issu de la classification de Dewey, ou encore la classification de la Bibliothèque du congrès (LCC).

Le thésaurus est la quatrième forme de représentation des connaissances en réseau. Nous décrirons très brièvement ce qu'est un thésaurus l'ayant déjà abordé dans la partie 2. Rappelons juste que selon la norme ISO 2788, un thésaurus est un outil de représentation des connaissances, standardisé et comportant des descripteurs et des liens particuliers et que cette forme de représentation est essentiellement utilisée dans les centres de documentation et les bibliothèques.

Enfin nous allons aborder deux autres types de représentations des connaissances en réseaux, moins connus. Les structures Qualia, tout d'abord, proposent un système de représentation des connaissances exprimées surtout à l'aide de noms. Ces structures utilisent pour représenter un concept donné, quatre rôles : le rôle constitutif pour la relation entre un objet et ses constituants, le rôle formel pour identifier les distinctions de l'objet considéré en fonction d'un point de vue ou d'un domaine, le rôle Télique pour les

intentions et les fonctions de l'objet et enfin le quatrième rôle, l'agentif qui permet de mieux identifier l'origine des objets. Les représentations en réseau sous la forme de plusieurs points de vue, le deuxième type, est une représentation qui part du constat qu'un même ensemble de connaissances peut, selon la personne qui l'utilise ou l'utilisation qu'elle compte en faire, être interprété de manières complètement différentes. Cela peut être par exemple une représentation en réseaux à l'aide de facettes qui se présente sous la forme de réseaux où chaque concept est considéré selon plusieurs facettes ou angles de vue différents.

B-3 Les représentations visuelles de connaissances

La représentation visuelle de connaissances repose sur six variables distinctes qui sont la taille, la valeur, le grain, la couleur, l'orientation et la forme. A partir de ces six variables nous pouvons interpréter ces représentations graphiques. Les principaux objectifs des représentations visuelles sont d'aider, par le biais d'un support visuel, à la découverte d'information ou de connaissances, à raisonner ou prendre des décisions, à communiquer des informations, à enregistrer des informations et enfin à expliquer des problèmes. La représentation visuelle des connaissances peut prendre la forme de diagrammes logique (une des plus simples), des graphes complexes c'est-à-dire une forme plus complexe du diagramme voir un assemblage de diagrammes, de réseaux, de cartes ou encore de symboles. Ces représentations graphiques peuvent être en deux dimensions cependant la technique en trois dimension est de plus en plus utilisée, elles peuvent employer des métaphores comme des structure arborescentes ou des paysages. Les représentations visuelles de connaissances sont les représentations qui sont les plus orientées et en interactions vers les utilisateurs.

Nous venons de voir qu'il existe une multitude de représentations de connaissances : en logique, en forme de réseau et visuelle. Cependant il existe un nombre, non négligeable, de représentations dérivées de tous les types de représentations des connaissances citées ci-dessus. Le dictionnaire fait partie de cette catégorie car comme nous l'avons montré dans la partie deux le dictionnaire est un thésaurus simplifié. Cette large gamme de représentations permet un grand nombre d'applications dans le domaine de la

documentation dont la veille. Mais qu'elle est la place des représentations des représentations des connaissances dans un système de veille ?

II La place de la représentation des connaissances dans la mise en place et la réalisation d'une veille

Dans cette partie nous allons tout d'abord situer la place de la représentation des connaissances dans la mise en place et la réalisation d'une veille. Ensuite nous verrons les avantages puis les inconvénients/limites des représentations des connaissances dans un contexte de veille

A) La place de la représentation des connaissances dans un contexte de veille

Les représentations de connaissances peuvent avoir une place importante dans un système de veille. Elles sont présentes quasiment tout au long du système de veille. Elles apportent une aide aux veilleurs et donc à la réalisation de la veille. Inversement la veille peut avoir une influence sur les représentations de connaissances.

Les représentations des connaissances peuvent être impliquées dans plusieurs étapes d'un système de veille. La plus évidente est celle de la classification des informations ou documents après leur collecte. Ce classement permet de retrouver facilement et rapidement les documents dans la base de documents collectés. Donc pour le veilleur qui recherche de l'information sur un sujet, le commerce électronique par exemple, utiliser une représentation des connaissances, un thésaurus ou une ontologie ou un de leur dérivé, lui permet de trouver les documents qui ont pour objet le commerce électronique, plus rapidement. A noter que plus la représentation des connaissances est affinée plus les sujets des requêtes peuvent être précis. En reprenant l'exemple du commerce électronique, si un veilleur fait une recherche sur les technologies 3D dans ce secteur d'activité, si la représentation des connaissances utilisées s'arrête au niveau de « commerce électronique », le veilleur risque de se retrouver face à un grand nombre de documents non pertinents, d'avoir un taux de bruit élevé. Toutefois si la représentation des connaissances utilisée est plus affinée, plus précise, le veilleur se retrouve devant les documents dont l'objet porte sur

le sujet recherché. A noter qu'en générale cette étape de classification est fait par un logiciel dans lequel est intégrée une représentation des connaissances.

La représentation des connaissances permet la classification, l'indexation des documents dans une base de données comme nous l'avons vu dans le paragraphe précédent. Ensuite la représentation des connaissances permet de rechercher de l'information dans ces bases de données. En science de la recherche de l'information (la RI), cette tâche est défini comme « l'ensemble des méthodes, procédures et techniques permettant, en fonction de critères de recherche propres à l'usager, de sélectionner l'information dans un ou plusieurs fonds de documents plus ou moins structurés »[18]. Les représentations des connaissances sont une de ces techniques, méthodes. Inconsciemment, lorsque que nous faisons une recherche sur un moteur ou métamoteur de recherche, nous utilisons une représentation des connaissances. En effet, après la collecte des pages web par leurs robots, ils classent ces pages de façon à permettre de les sélectionner en fonction des requêtes des utilisateurs. Toutefois chaque moteur de recherche utilise ses propres représentations de connaissances qui ne sont pas uniquement guidées par une volonté de classer les pages web pour que les réponses aux requêtes soient les plus pertinentes possibles (un taux de bruit et de silence bas) mais elles sont aussi influencées par d'autres facteurs, commerciales entre autres (liens sponsorisés,...).

Les représentations des connaissances peuvent aussi apporter une aide à l'analyse de l'information notamment en donnant une meilleure visualisation des informations d'un texte ou d'un corpus de textes. Ce cas concerne essentiellement les cartes conceptuelles qui permettent au veilleur d'avoir une vision des points importants, des idées principales d'un corpus de textes. En fonction des sujets traités, le veilleur peut réaliser une carte conceptuelle lui montrant les tendances par exemple. C. Ware[19] a montré que les combinaisons entre les images et les mots apportent une aide à la compréhension et au raisonnement. Ces combinaisons apportent une aide pour une meilleure compréhension des idées que l'on veut générer et que l'on veut faire passer, véhiculer.

---

[18] Institut national des techniques de la documentation coordonnée par Boulogne Arlette, *Vocabulaire de la documentation*, ADBS éditons, Paris : 2004.

[19] Ware C., 2003. *Information Visualization : Perception for Design*, Editions Morgan Kaufmann Publishers, San Francisco

Le deuxième grand apport des représentations des connaissances concerne l'étape de la diffusion/communication du système de veille. Les représentations visuelles des connaissances sont les principales représentations des connaissances utilisées dans ce cas. Elles apportent une véritable aide aux veilleurs pour présenter les résultats de sa veille à son client. En général les chefs d'entreprises n'ont pas toujours le temps de lire les rapports. C'est pourquoi en présentant les résultats de la veille au chef d'entreprise sous une forme visuelle à l'aide d'une représentation visuelle des connaissances, cette représentation lui permet de prendre connaissance et de comprendre les résultats plus rapidement et plus simplement, même si cela ne remplace pas les explications du veilleur. Ces représentations s'intègrent parfaitement dans le système de veille car elles sont généralement éditées par des logiciels qui sont considérés comme des logiciels de veille. Donc les représentations des connaissances apportent une aide pour améliorer la compréhension des informations, résultats de la veille communiqués par le veilleur.

Les représentations de connaissances apportent une aide au veilleur à différentes étapes de la veille : indexation et recherche des informations, l'analyse et la communication. Toutefois cette relation n'est pas en sens unique. En effet la veille peut influencer sur les représentations des connaissances. Dans le cadre de mon stage, la veille apparaissait comme une aide à l'enrichissement du dictionnaire métier et un moyen de valider ce dictionnaire. Nous pensons que ce constat peut être applicable pour d'autres représentations des connaissances, mais pas à toutes celles que nous avons présentées, car elles ne sont pas toutes utilisables dans un système de veille comme la plupart des classifications qui sont surtout utilisées dans les bibliothèques.

B) Les avantages des représentations des connaissances

Le principal avantage des représentations des connaissances est de permettre d'appréhender, de découvrir un domaine de connaissance que nous ignorons. Ce genre de situation est très courant. Un veilleur peut être confronté à une multitude de domaines de connaissances au cours des différentes études et veilles qu'il doit mener. Par exemple, l'ARIST travaille avec les pôles de compétitivité et est donc confrontée au domaine de connaissances qui s'y rattache c'est-à-dire le textile, la santé, le commerce, l'industrie de la mer,...Les représentations des connaissances sont des repères nécessaires dans des domaines

de connaissances que l'on ne maîtrise pas. Par exemple le thésaurus MESH[20]permet de découvrir le domaine de la santé même si l'utilisation d'un dictionnaire reste nécessaire car en règle général un thésaurus ne comprend pas de définition mais juste des notes d'applications pour supprimer les ambiguïtés.

L'autre grand principal avantage des représentations des connaissances est qu'elles permettent de faire face à la perpétuelle augmentation du nombre de documents auxquels nous avons accès notamment sur internet ou sur des bases de données, des archives ouvertes accessibles sur internet. En effet la production de publications scientifiques est en perpétuelle augmentation ce qui est principalement dû à la nécessité de produire des publications pour pourvoir continuer à faire de la recherche d'où l'expression « produire ou périr ». Ces sources sont très utilisées pour la veille technologique. Ces publications scientifiques ne sont pas les seuls documents en augmentation. En effet, *Google* estime à un billion le nombre de pages web et 40 milliards le nombre de pages qu'il indexe[21]. Parmi la multitude de contenu que l'on peut trouver sur internet le blog prend une place de plus en plus importante. En effet le nombre de blogs est également en perpétuelle augmentation et ces blogs touchent de nombreux sujets et donc peuvent être utilisés par plusieurs types de veille (technologique, marketing, d'opinion ou encore d'image). Internet n'est pas le seul accès à l'information, de nombreuses revues grands publics et spécialisées, au format papier existent toujours, le numérique et internet ne les a pas fait disparaître même si la plupart de ces revues ont des sites internet. Donc face à toutes ces sources et au nombre grandissant de documents et d'informations, les représentations de connaissances jouent un rôle de filtrage notamment en indexant les documents et en éliminant nombre de documents non pertinents par rapport au sujet de la veille. On retrouve cette action de filtrage notamment avec les ontologies comme nous l'avons vu dans précédemment, mais aussi avec les thésaurus. Cette action de filtrage permet de résoudre partiellement le problème du veilleur de fournir toujours plus rapidement les informations les plus pertinentes qui sont en général noyées dans d'autres informations notamment sur internet. En effet l'objectif du veilleur est de repérer à l'avance les fameux signaux faibles, les futurs évolutions (du marché, technologiques,...) mais cet objectif est rendu difficile à atteindre à cause du nombre toujours

---

[20] Arborescence du thésaurus MESH (Medical Subject headings), thésaurus réalisé par la National Library of medicine : http://www.chu-rouen.fr/ssf/arborescences.html (consulté le 08/09/09).

[21] http://www.infos-du-net.com/actualite/14118-nombre-pages-internet.html (consulté le 09/09/09).

plus important d'informations accessibles sur le web principalement. Donc les représentations de connaissances permettent de traiter rapidement un maximum d'informations ou de documents avec leurs fonctions d'indexation et de filtrage de façon à permettre au veilleur d'avoir accès plus rapidement aux informations qui l'intéressent.

La possibilité de faire évoluer/de mettre à jour, applicable pour la plupart des représentations de connaissances que nous avons présentées, peut être considérée comme un avantage. En effet, les domaines de connaissances ne sont pas figés, ils évoluent grâce aux avancés de la recherche, au développement de nouvelles technologies,... Les mises à jour des représentations permettent de suivre les évolutions. Cependant pour certain domaine de connaissance, il est difficile de suivre les évolutions. Le domaine de connaissances de médecine, par exemple, évolue très vite et donc les mises à jour sont fréquentes mais ils n'arrivent pas toujours à suivre l'évolution du domaine. Les représentations sont également modulables voir personnalisables en fonction du sujet et/ou du destinataire de la veille. Par exemple, le destinataire ne désire que des informations en français alors que d'autres veulent exclusivement des informations en anglais ou le sujet de la veille peut nécessiter autant des informations en français qu'en anglais donc il faut des représentations des connaissances bilingues voir multilingues. L'avantage est que nous pouvons récupérer une représentation des connaissances pour l'adapter à un sujet ou un destinataire sans devoir en reproduire une à chaque fois.

Donc les avantages des représentations de connaissances sont d'appréhender, de découvrir, de comprendre un domaine de connaissances qui nous est inconnu, de permettre de mieux le délimiter, de contrôler son sujet. Elles apportent également une aide pour gérer les documents et les informations toujours plus nombreux. Enfin, la modularité, la possibilité de mise à jour est également un avantage des représentations des connaissances.

C) Les inconvénients/limites des représentations des connaissances

Tout d'abord, la principale limite est que les représentations de connaissances ne sont pas des références absolues. D'une part ces représentations sont généralement réalisées avec un langage normalisé c'est-à-dire l'ensemble de termes qui compose les terminologiques, ces termes correspondant à des concepts ont été validés, se sont des termes simplifiés : verbe à l'infinitif, expression la plus simple possible,... Elles n'utilisent pas le langage naturel c'est-à-dire le langage « ordinaire », à quelques exceptions près. Par exemple, le dictionnaire métier, sur lequel nous avons travaillé durant mon stage, est réalisé en langage naturel car l'indexation se faisait en fonction de la présence des termes dans les documents autrement dit les termes utilisés par l'auteur et pas une indexation à base de mots clés représentant les idées et les concepts développés dans les documents. L'utilisation du langage normalisé limite la représentation des concepts. D'autre part, nous avons vu dans la sous partie précédente que les représentations des connaissances peuvent se mettre à jour et nous avons considéré cela comme un avantage. Cependant cette fonctionnalité montre que le veilleur ne peut s'appuyer uniquement sur la représentation d'un domaine de connaissances car ils peuvent être ou devenir caducs donc les mises à jour sont obligatoires pour que la représentation des connaissances et les résultats de la veille restent pertinents. Il doit continuer à être attentif à l'actualité du domaine pour prendre en compte les évolutions du domaine ce qui nécessite un travail continu.

Le deuxième grand inconvénient des représentations des connaissances est que leurs élaborations sont lourdes et longues à réaliser notamment pour celles utilisées pour l'indexation et l'aide à l'analyse alors que pour les représentations visuelles des connaissances, le travail est moindre car essentiellement réalisé par des outils. Le travail est lourd et long car une grande partie de l'élaboration des représentations se fait manuellement comme pour la réalisation d'un thésaurus dont nous avons décrit la méthode d'élaboration dans la partie 2 ou pour une ontologie. Cependant des outils ont été développés pour aider à l'élaboration, ils allègent un peu le travail manuel qui reste encore majoritaire dans la réalisation des représentations. L'utilisation de terminologie et des représentations de connaissances déjà élaborées permet également d'alléger le travail manuel dans l'élaboration d'une représentation des connaissances. Généralement, l'élaboration d'une représentation est considérée comme contraignante notamment par les veilleurs mais une fois élaborée elle reste un outil très utile au sein d'un système de veille. Ce travail est également lourd car il nécessite un travail continu de mise à jour du fait qu'une représentation des connaissances

suit les évolutions du domaine de connaissances. Nous pourrons oser aller jusqu'à dire que l'élaboration d'une représentation des connaissances n'est jamais finie à cause de ce travail en continu.

La plupart des représentations des connaissances sont orientées, elles sont élaborées en fonction d'un point de vue. En effet elles sont faites en fonction du sujet ou du destinataire de la veille, on peut dire que les représentations sont personnalisables. Un domaine de connaissances peut être traité différemment par plusieurs thésaurus. Par exemple le choix des termes peut varier selon le destinataire ou en fonction du sujet, l'approfondissement de certain point du domaine de connaissance par rapport aux autres. Le domaine de connaissance du commerce sera représenté différemment si la représentation est destine au secteur marketing ou si elle est destinée au secteur logistique, idem pour les sujets de la veille. Le veilleur doit faire attention lors de l'élaboration d'une représentation des connaissances (à l'exception des graphiques), il doit prendre en compte soit le sujet de la veille soit aux personnes à qui est destinée la représentation ou les livrables produis à l'aide de la représentation. Si le veilleur utilise une représentation des connaissances déjà élaborée, il doit être attentif aux mêmes points cités ci-dessus. Donc un des inconvénients des représentations des connaissances, c'est qu'ils peuvent représenter un domaine de connaissance de différentes façons, de points de vue, avec différentes interprétations.

Les contraintes proviennent aussi des logiciels dans lesquels peuvent être intégrés les représentations des connaissances. En général ils imposent un ensemble de contraintes comme pour le dictionnaire métier qui est intégré dans *Tropes Zoom* qui impose des limites au niveau des relations ou qui interdit la double entrée. Cependant certaines de ces contraintes peuvent également être un avantage. Par exemple l'impossibilité de la double entrée pour le dictionnaire métier dans *Tropes Zoom* permet d'éliminer un grand nombre d'ambiguïtés. Enfin les inconvénients peuvent aussi provenir des différents types de représentations des connaissances comme le thésaurus, l'ontologie, les représentations visuelles,... Par exemple l'ontologie est une représentation des connaissances très rigide, rigoureuse des significations entre les différents concepts, ce qui a pour conséquence que les mises à jour sont difficiles à réaliser. Nous retrouvons des contraintes pour chaque type de représentations des connaissances que nous avons cité.

Conclusion

Une représentation des connaissances est une représentation et une organisation du savoir humain dans une optique d'utilisation et de partage. Il existe une multitude de types de représentations des organisations : des représentations logiques, des représentations en réseaux comme les thésaurus ou les ontologies et enfin les représentations visuelles des connaissances. Dans le cadre de notre stage, nous utilisons le dictionnaire métier, une représentation des connaissances en réseaux et pour être plus précis un thésaurus simplifié.

Les représentations des connaissances peuvent tenir une place important dans un système de veille car elles sont présentes dans plusieurs étapes de la veille : collecte, recherche, aide à l'analyse et communication. Les représentations apportent une aide conséquent au veilleur notamment pour connaître, appréhender un domaine de connaissance qui lui est inconnu et de trouver rapidement des informations pertinentes parmi la quantité toujours plus importante d'informations accessibles sur internet, principalement, grâce à leur capacité de filtrage et d'indexation. Elles sont également d'une précieuse aide pour présenter les résultats d'une veille grâce notamment aux cartes conceptuelles. Inversement la veille influence les représentations des connaissances. Elle peut être une aide à l'enrichissement de la représentation et un moyen de valider les termes qui composent la représentation des connaissances. Toutefois les représentations des connaissances ne sont pas indispensables lors de la réalisation d'une veille. En résumé les liens entre la veille et les représentations des connaissances sont multiples et chaque partie peut apporter quelque chose à l'autre.

Concernant mon stage, dans son ensemble, je considère qu'il est réussi car il m'a permis de réaliser une veille dans un cadre professionnel, de participer à chaque étape de la veille : le sourcing, la collecte, l'analyse et la communication.il m'a permis de découvrir un nouveau domaine de connaissance, le commerce, que je ne connaissais pas, notamment la partie technologie qui m'a beaucoup intéressée. Toutefois je considère le stage également comme une demi-réussite car concernant l'élaboration du dictionnaire je ne pourrais aller aussi loin dans l'élaboration que je l'aurais voulu, notamment j'aurais à peine abordé la version anglaise du dictionnaire que je considère aussi importante que la version française voir plus car les ressources en langue anglaise sont plus importante que en langue française.

Bibliographie

-AITCHISON Jean, GILCHRIST Alan, *Construire un Thesaurus : manuel pratique*, ADBS éditions, Paris : 1987 (2e éditions), 217p.

-DIENG Rose, *Méthodes et outils pour la gestion des connaissances*, DUNOD, Paris : 2000, 302p.

-GENSEL Jérôme, 1995, *Contraintes et représentation de connaissances par objets : Application au modèle Tropes*, Thèse de doctorat, université Grenoble 1, http://tel.archives-ouvertes.fr/tel-00005046/fr/

- GHIGLIONE Rodolphe, LANDRE Agnès, BROMBERG Marcel, MOLETTE Pierre, *L'analyse automatique des contenus*, DUNOD, Paris : 1998, 156p.

-GORIA Stéphane, 2006, *L'expression du problème dans la recherche d'informations : Application à un contexte d'intermédiation territoriale*, Thèse de doctorat, université de Nancy 2, http://tel.archives-ouvertes.fr/tel-00011918/fr/

-KAYSER Daniel, *La représentation des connaissances*, ed Hermès, Paris : 1997, 308p.

-LAUREILHE Marie Thérèse, *Le thesaurus, son rôle, sa structure, son élaboration*, Presses de L'E.N.S.B, Villeurbanne, 1981 (2e édition), 88p.

- ROUACH Daniel, *La veille Technologique et l'intelligence économique*, PUF, Paris : 2008 (4e édition), coll Que sais-je ?, 128p.

- CONDAMINES Anne, *Terminologie et représentation des connaissances*, INRP, Lyon : 1994, 17p, http://documents.irevues.inist.fr/bitstream/2042/23235/1/DIDASKALIA_1994_5_35.pdf

Sitographie

Site web du PICOM : http://www.picom.fr/

Site web CRCI : http://www.nordpasdecalais.cci.fr/fr/accueil.aspx

Site web *Tropes Zoom*: http://www.acetic.fr/tropesfr.htm

Extrait du thesaurus Delphes : http://www.arist.ccip.fr/Thesaurus-Delphes-2004-Extraits.pdf

Liens pour se procurer le thésaurus Delphes : http://www.arist.ccip.fr/partenaires.htm#delphes

ANNEXES

Sommaire des annexes

ANNEXE 1 : Schéma de la méthodologie de la veille de l'ARIST 59

ANNEXE 2 : Zoom 60

ANNEXE 3 : Tropes 62

ANNEXE 4 : Synthèses issues de la veille sur la RFID 64

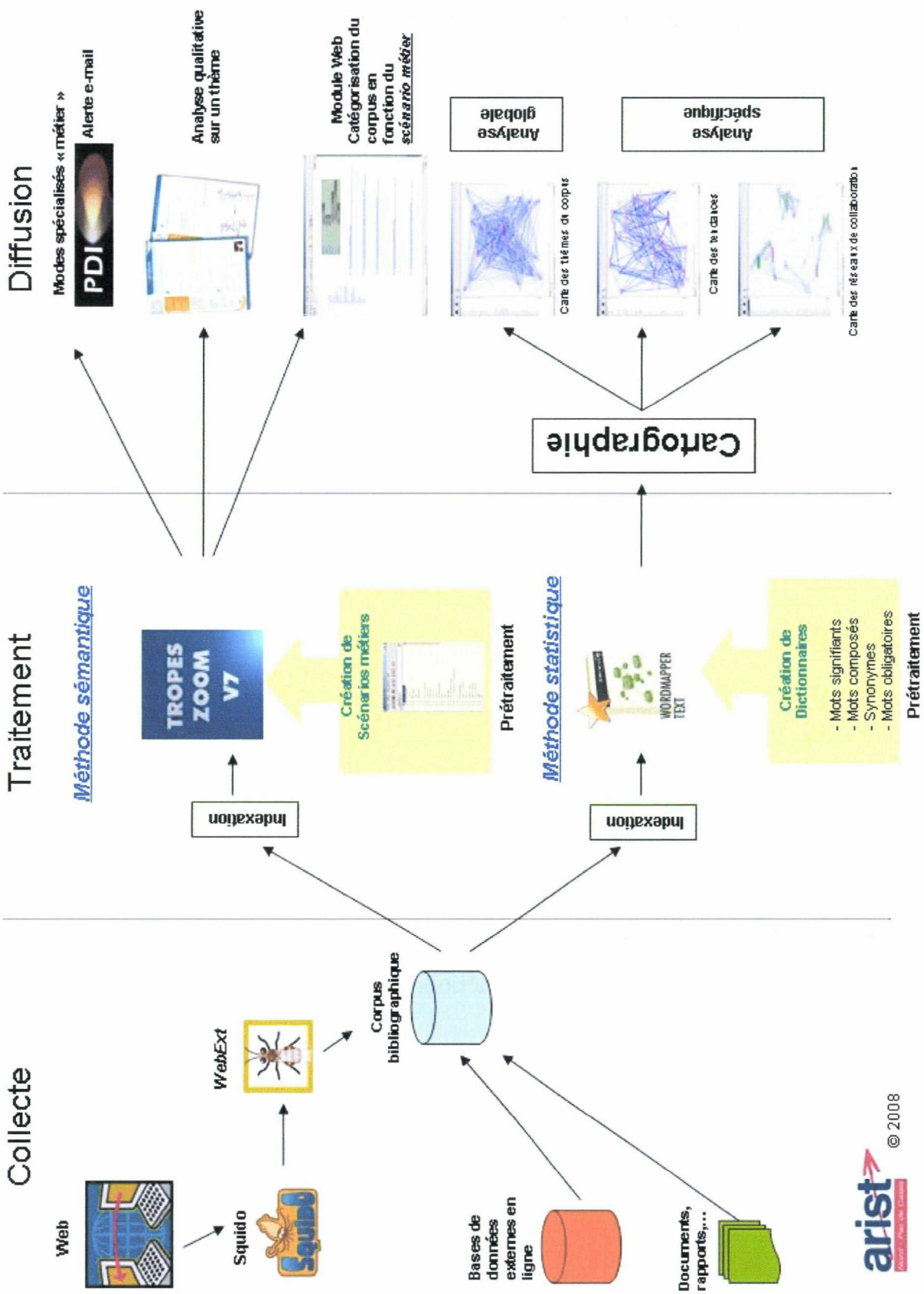
ANNEXE 5 : Tableau d'acteurs RFID et de leurs produits 74

ANNEXE 6 : Article sur l'évolution de la 3D dans le cadre du commerce en 2009 77

ANNEXE 7 : Article sur la solution *Pachube* 78

ANNEXE 8 : Les éléments racines du dictionnaire métier 79

ANNEXE 1 : Schéma de la méthodologie de la veille de l'ARIST



ANNEXE 2 : Le logiciel Zoom

Quelques captures d'écran pour montrer un aperçu du logiciel Zoom :

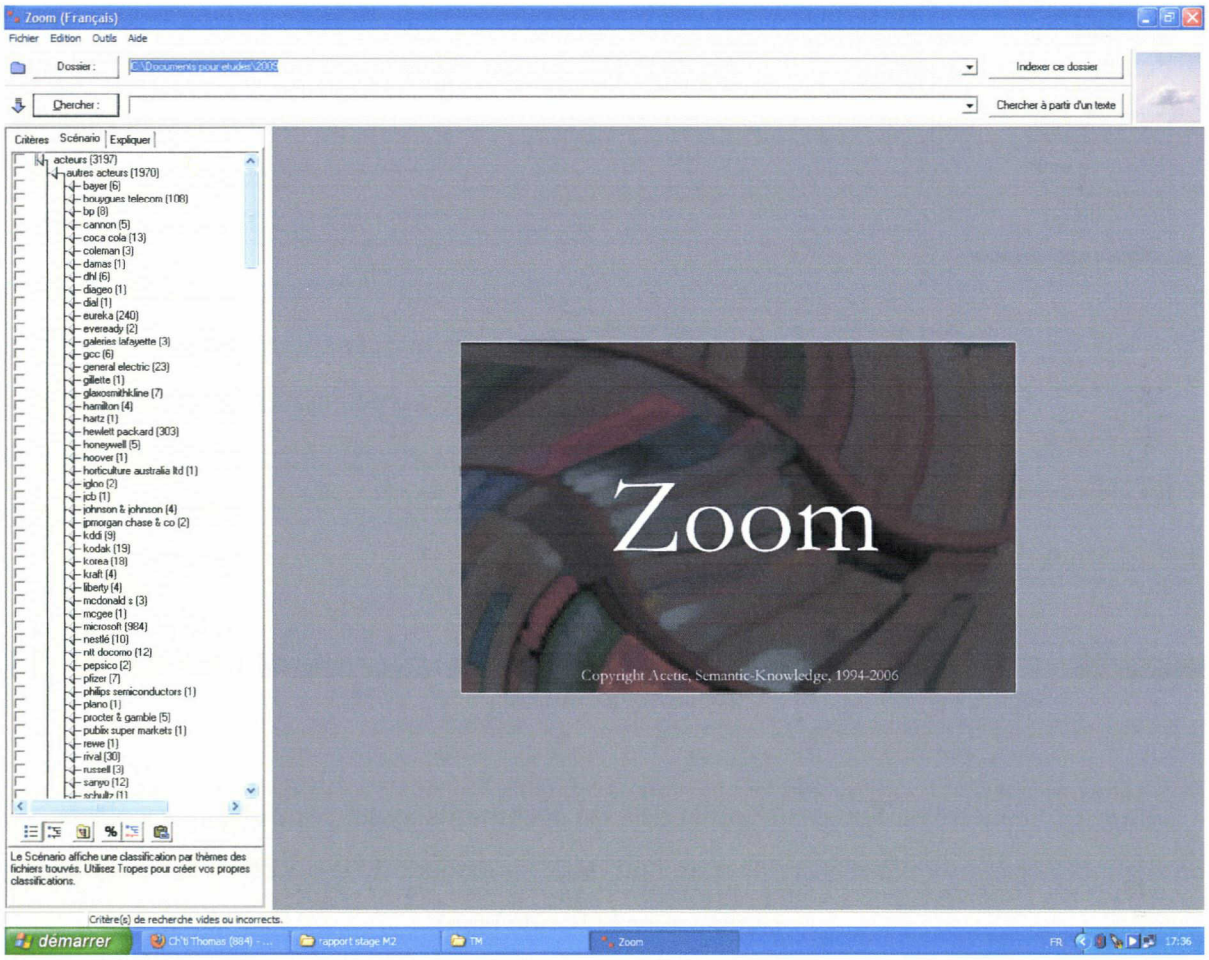


Image 1 : Fenêtre d'accueil de Zoom, le dictionnaire métier se trouve dans la colonne à gauche.

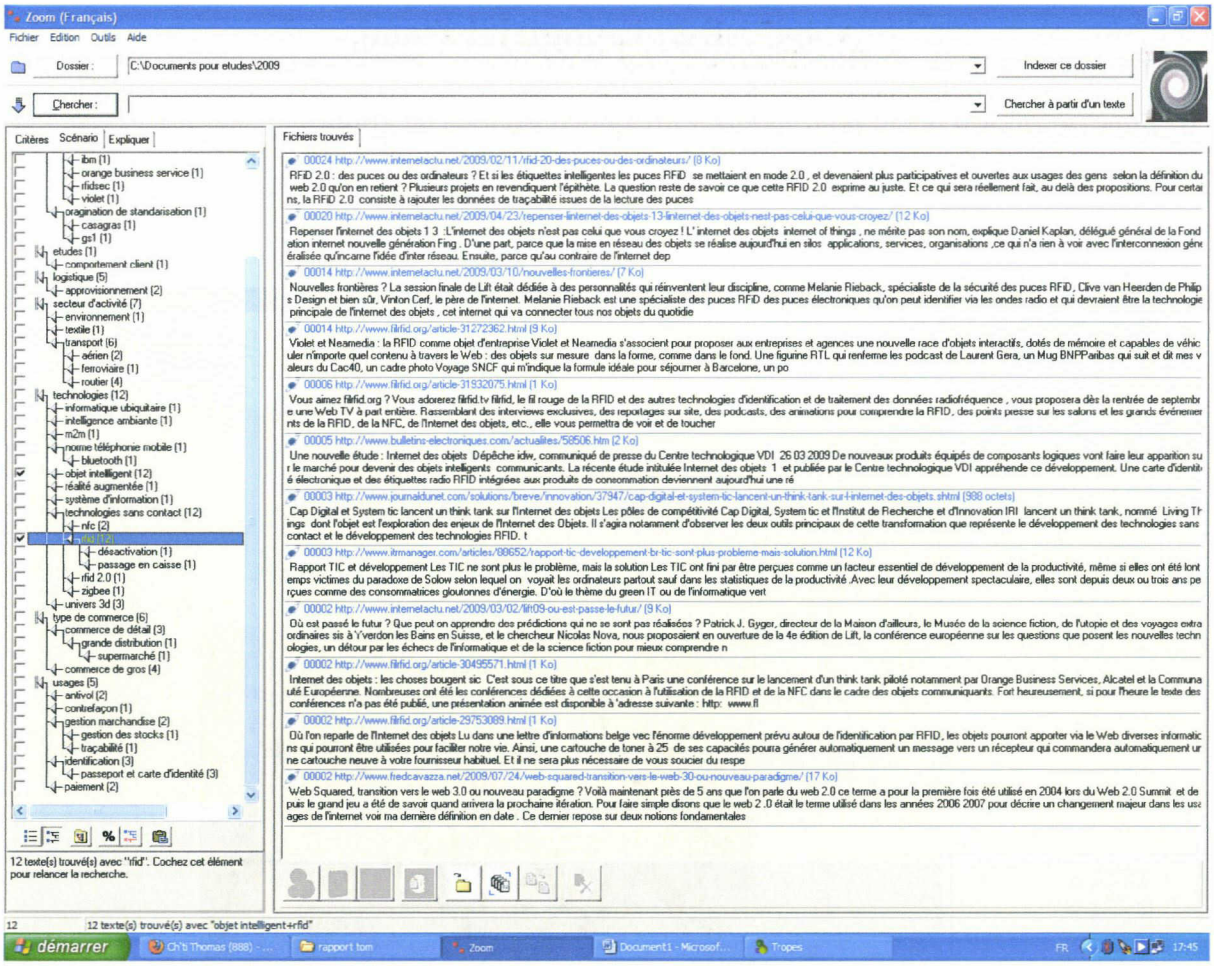


Image 2 : Fenêtre ou l'on peut voir la liste de documents ayant pour objet les « objets intelligents » (première case cochée) et ou l'on parle également de la « RFID » (deuxième case cochée avec la bande bleu)

ANNEXE 3 : Le logiciel Tropes

Quelques captures d'écran pour montrer un aperçu du logiciel Zoom :

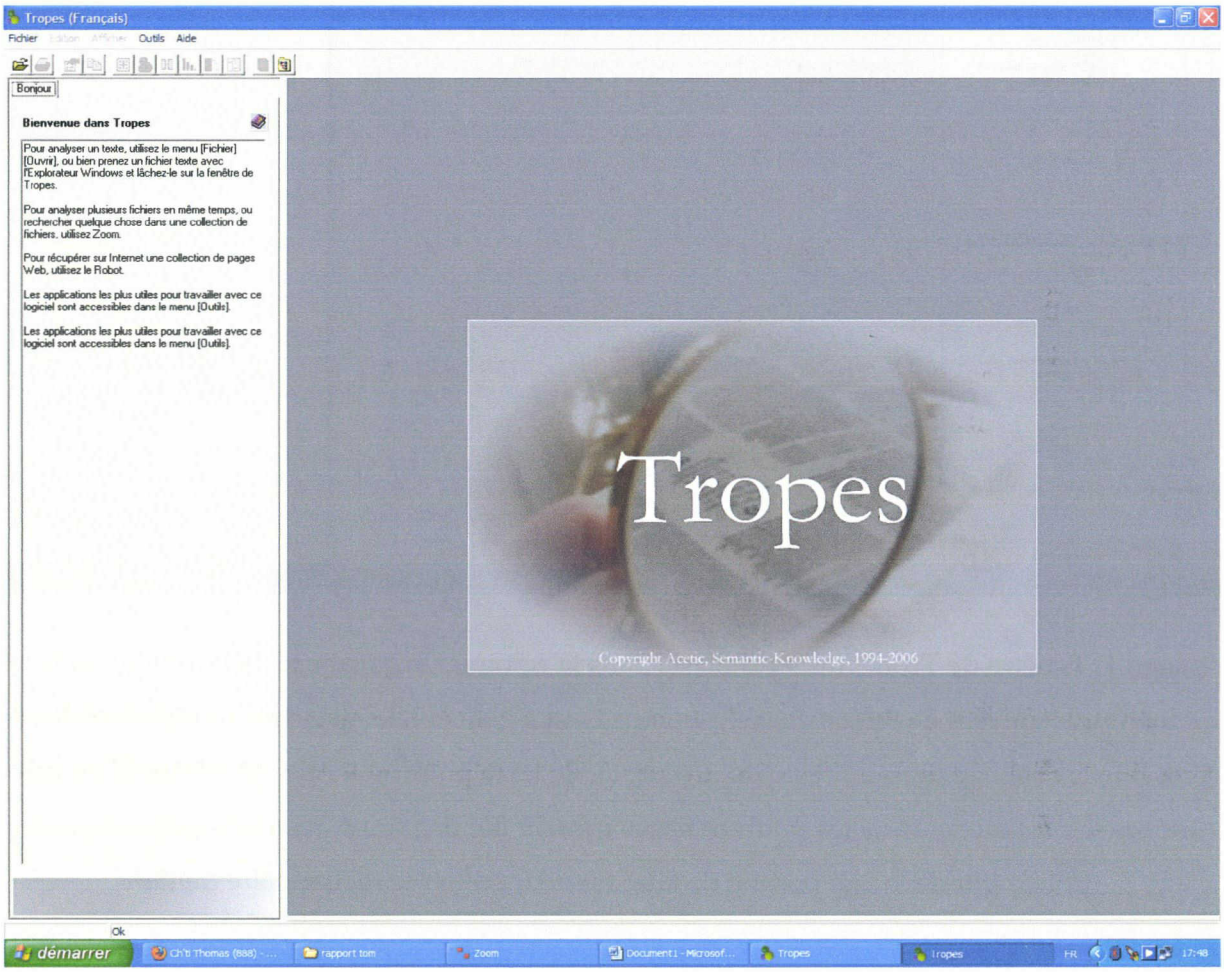


Image 3 : Fenêtre d'accueil du logiciel Tropes

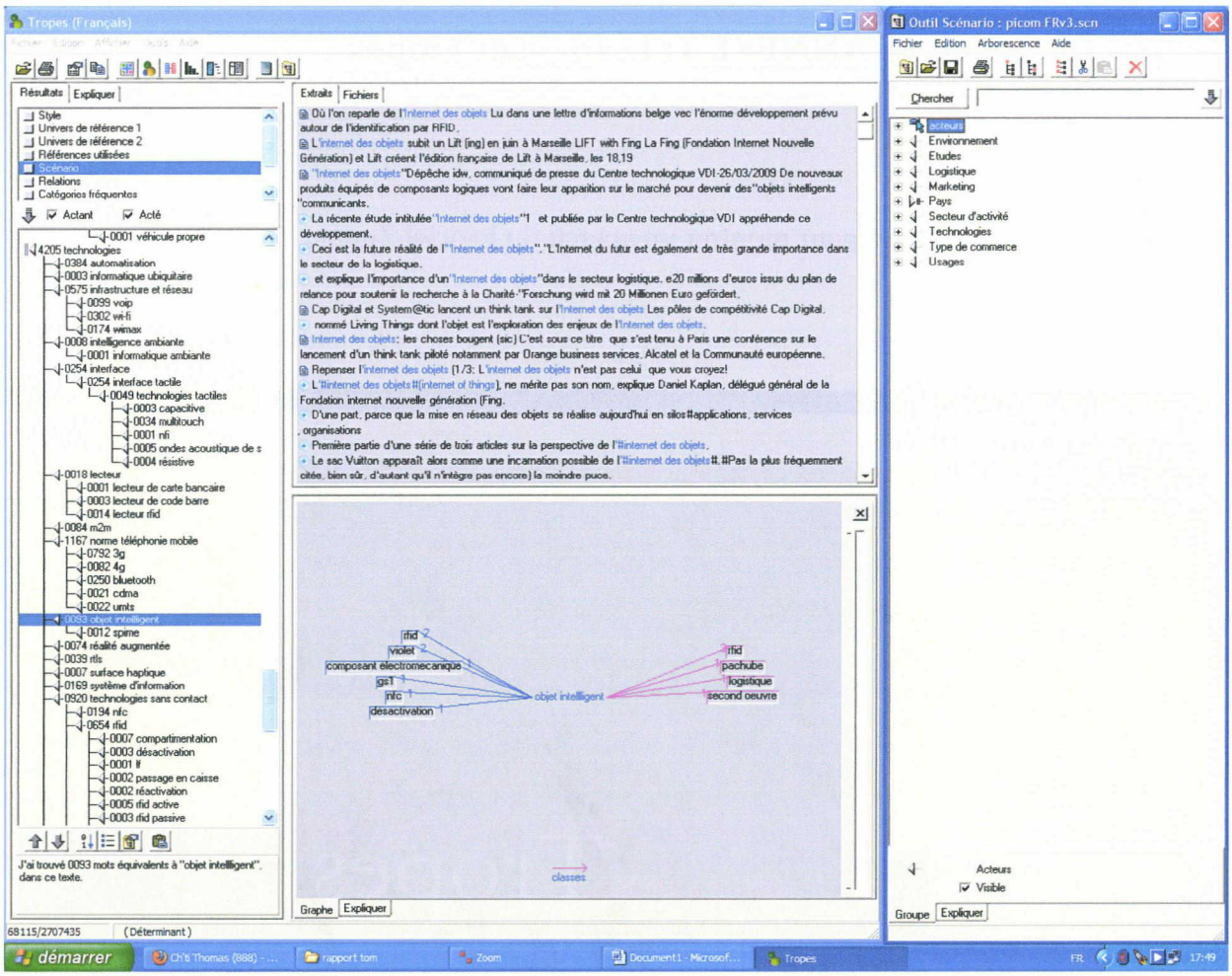


Image 4 : Fenêtre de Tropes après analyse, dans la colonne de gauche le dictionnaire métier ou les autres univers de référence (sélection en haut à gauche), le rectangle au milieu en haut présente le contexte dans un texte ou un corpus de texte pour un terme ; le rectangle du bas montre les actants et les actés pour un terme (possibilité de faire d'autres graphiques) et à droite l'interface qui permet de faire ou de modifier le dictionnaire métier.

ANNEXE 4 : Synthèses issues de la veille sur la RFID

- Nous avons réalisé trois synthèses durant la veille sur la RFID :

Synthèse RFID : 03/06/09

- Les sociétés Violet et Neamedia ont développé des objets interactifs. Ces objets sont équipé de puces RFID ce qui permet de les relier à la plateforme violet via le réseau internet autrement dit à du contenu numérique. Ils sont développés une technologie permettant l'association de contenu numérique à un objet physique. Cette technologie rentre dans le cadre de l'internet des objets, thème que je trouve intéressant à approfondir par rapport à votre demande tout comme celui d'objet intelligent. Ces deux thèmes sont, en plus, très liés à la RFID. Application : exemple d'objet, le *Mir:ror,* lecteur de tag RFID qui lance l'application associé à ce tag. Les applications sont nombreuses car cette technologie est applicable sur de nombreux objets : marketing, communication, etc. On pourrait aussi envisager d'appliquer cette technologie à des objets du quotidien comme un micro-onde et que grâce à la puce RFID on est accès à de la documentation numérique en ligne sur l'objet comme le mode d'emploi voir des services comme le SAV et à terme permettre un dialogue entre le client et le commerçant. Je trouve que cette technologie est une possible réponse à votre demande sur la RFID. (http://www.filrfid.org/article-31272362.html)

- Alcatel-Lucent lance le kit *Tikitag.* Ce kit, permettant de relier un objet à une application, se base sur la technologie NFC. Il se compose d'autocollants équipés de puces NFC, d'un lecteur NFC et d'un accès au site web de Tikitag pour pourvoir créer une application. Usages : cette technologie permet d'avoir accès à du contenu supplémentaire sur un tableau dans un musée par exemple, en passant son Smartphone équipé d'un lecteur NFC devant une œuvre pour voir la page Wikipédia de l'artiste sur son appareil. Ici, il serait intéressant de voir la relation entre internet mobile et la RFID. On pourrait appliquer cette solution sur un objet du quotidien comme pour le paragraphe précédent, c'est-à-dire avoir du contenu supplémentaire sur l'objet acheté.

(http://pro.01net.com/editorial/391914/un-kit-de-developpement-pour-connecter-les-objets-aux-applications-web/ et http://www.atelier.fr/usages/10/30092008/les-objets-communicants-deviennent-une-realite-pour-le-grand-public-37225-.html)

L'acteur Airtag a développé un outil du même type. Ce basant sur la technologie RFID et NFC. Le kit est composé d'un lecteur, de tag et d'outil de développement logiciel qui permet de programmer un objet. Comme pour les paragraphes précédents cette solution permet d'associer du contenu en ligne à un objet. Airtag a aussi développé une carte de fidélité équipée de la technologie sans contact NFC, elle simplifie le système de fidélisation et peut amener d'autres clients dans ce système. Les applications de cet outil sont d'ordre marketing.
(http://www.businessmobile.fr/actualites/technologies/0,39044306,39383746,00.htm et http://www.atelier.fr/mobilite/10/02012008/sans-contact-fidelise-client-airfid-nfc-35769-.html)

Les solutions présentées ci-dessus sont les solutions qui répondent aux mieux (pour l'instant) à ce que vous nous avez décri lors de notre dernière réunion.

- La société Twinlinx a développé une puce intégrant la technologie RFID et NFC. Cela permet la lecture d'un tag à courte (dix centimètre) et à longue distance (un mètre). Elle est envisagée dans des applications autour du mobile. Cela peut être intéressant car un large public pourra lire ces puces, autant ceux ayant un lecteur NFC que ceux ayant un lecteur RFID. Twinlinx a aussi développé une solution pour équiper les mobiles (surtout les Smartphones) d'un lecteur NFC grâce à une étiquette électronique. Cette transformation en lecteur NFC d'un mobile permet le payement, l'accès a du contenu via l'internet mobile, etc. Cette solution complète celle décrite dans les premiers paragraphes.
(http://www.atelier.fr/applications/10/04022008/prix-de-l-innovation-tic-paca-2008-rfid-web-20-mobilite-electronique-35935-.html et http://www.atelier.fr/usages/10/09092008/blu-link--twinlinx--nfc--37120-.html)

Nokia développe aussi un Smartphone équipé d'un lecteur RFID, le mobile devient une interface entre l'objet, le service et le consommateur.
(http://www.atelier.fr/applications/10/23042008/rfid-mobile-nokia-internet-des-objets-36445-.html).

- Des chercheurs américains ont découvert un moyen de sécuriser les puces à moindres coût ; Ils utilisent la mémoire de type SRAM présent dans les puces RFID. Cette technique se fie seulement aux principes fondamentaux de cette mémoire. A noter que ce type de mémoire n'est pas encore présent dans toutes les puces RFID actuel, les chercheurs parlent d'une anticipation d'une à plusieurs année. Cette technique permet de sécuriser plus simplement les puces RFID sans ajouter un système en plus dans la puce. Comme je n'ai pas trouvé de solution qui permet à un particulier de réveiller ou réactiver une puce RFID, la technique décrite ci-dessus peut être une bonne alternative et peut rassurer les clients sur la protection de leur vie privée. (http://www.bulletins-electroniques.com/actualites/58415.htm)

- ThimgMagic a développé un lecteur RFID permettant de rechercher des objets équipés de tag RFID par rapport à une liste pour des inventaires par exemple. Pour l'instant utilisé dans le monde professionnel, on peut envisager des applications sous la forme de services pour les clients dans une grande surface (liste de course, ...).
(http://www.internetactu.net/2008/12/23/rfid-le-moteur-de-recherche-du-monde-reel-suite/)

- Pachube est un outil qui permet de connecter et partager en temps réel les données d'un capteur, quel qu'il soit. Cet outil, qui prend la forme d'un service web, se base sur la technologie RFID, NFC et la technologie mobile. Cet outil se situe dans le cadre de l'internet des objets mais aborde aussi le machines to machines ou M2M (Est-ce-que c'est un sujet qui vous intéresse ?). De nombreuses applications sont possibles car on a la possibilité de créer une application en fonction de nos besoins, donc des applications sont possibles dans le commerce, dans le marketing par exemple.

(http://www.internetactu.net/2009/02/05/pachube-des-applications-pour-linternet-des-objets/)

La société SmartBlade a lancé une étagère intelligente. Elle est équipée d'une antenne Led flash, elle détecte les produits équipés de tag RFID. Elle permet une meilleur gestion du rayonnage et des stocks, et évite les rayons vides, mais on peut

aussi envisager des applications marketing, de communication ou encore de présentation de produit. (http://www.filrfid.org/article-24372582.html).

Enfin, l'institut pour la protection et la sécurité des citoyens a développé un réseau de transpondeurs RFID dans une ville d'Italie pour permettre à un mal voyant ou non voyant, équipé d'un Smartphone/lecteur RFID et d'une canne spéciale, de mieux circuler et de fournir des informations de navigation à travers une ville comme dans l'article. Cependant, on pourrait appliquer cette technologie dans les centres commerciaux pour les rendre encore plus accessibles. (http://www.atelier.fr/mobilite/10/03012008/ipsc-projet-pilote-ville-italienne-permet-malvoyants-reperer-technologie-rfid-35782-.html)

Synthèse : 31/07/09

- *Artfest international,* spécialiste de la commercialisation, de la vente et de l'impression de peintures, de photos de célébrités en édition limitée dédicacées et autres souvenirs, a complété le développement et l'acquisition de nouveaux logiciels et puces RFID. C'est un système de traçage et d'authentification de produit. Les puces RFID sont attachées derrières les tableaux ou photos, elles contiennent des informations sur l'origine, l'authenticité et sur les propriétés intellectuelles du produit. Ici la RFID est utilisé comme moyen de protection pour les clients (authenticité) que pour les artistes. J'ai sélectionné cet article car c'est l'un des premiers qui montre les puces RFID comme outil de stockage de données (comme une clé USB) et pas juste une passerelle entre l'objet et du contenu en ligne. Cet article se rapproche de votre demande de se concentrer sur la puce en elle-même. Cependant il n'évoque pas de notion ou concept de compartimentation et il n'est peut être pas assez précis au niveau technique, pas assez de détails sur l'outil. Il serait peut être intéressant d'approfondir les recherches sur cet acteur.
http://www.morerfid.com/details.php?subdetail=Report&action=details&report_id=5814&display=RFID

- *Nokia* a lancé son nouveau téléphone portable le *Nokia 6216 classic* au WIMA Conference à Monaco en avril 2009. Ce nouveau mobile est le troisième équipé de la technologie NFC mais c'est le premier à l'avoir incorporé non pas dans le téléphone lui-même mais sur la carte SIM. Ce déplacement apporte un élément de sécurité notamment pour les transactions sans contact. Par exemple le particulier pourra, quand il changera de téléphone, déplacer son portefeuille électronique (« mobile wallet »), se trouvant dans la carte SIM, en toute sécurité et facilement lorsqu'il placera sa carte SIM dans son nouveau téléphone. Il permet de protéger les données sur les cartes de crédits qui sont stockées sur la carte SIM. Ce téléphone ne sera disponible qu'au quatrième trimestre 2009. J'ai sélectionné cet article car je pense que les portables équipés de la technologie NFC et/ou RFID est pour l'instant et (je pense pour les années avenir) les seul moyens que les particuliers ont à leur disposition pour utiliser les applications utilisant la RFID et NFC et notamment les applications que vous avez évoqués.
http://www.computerworld.com/s/article/9132037/Nokia_puts_mobile_wallet_on_SIM_card?intsrc=news_ts_head et http://www.slashphone.com/nokia-introduces-nokia-6216-classic-for-nfc-contactless-payment-235545
A noté que *Sony Ericsson* a annoncé que la technologie RFID arrivera dans les téléphones en 2010. *Apple* a déposé des brevets qui indiquent qu'ils envisagent d'implanter un lecteur RFID dans l'*Iphone*. Cette technologie permettra à l'*Iphone* d'identifier des objets ou la nature des bâtiments autour de l'utilisateur. Par exemple : « Dans un musée par exemple, ce logiciel récupèrerait une base de données depuis une connexion Dock à l'entrée. Puis, face à une œuvre, l'*IPhone* pourrait renseigner le visiteur à son sujet ». Un autre brevet indique l'utilisation d'un système de reconnaissance de visage afin de le coupler à une application de contrôle d'accès à des fonctionnalités. On retrouve le même principe avec la reconnaissance des empruntes digitales qui permet de lancer une application par doigt (l'index lance l'application A, le majeur la B, etc.). Enfin *Apple* a déposé un brevet concernant la technologie de retour sensoriel sur l'écran multitouch, l'usager obtient une réaction lorsqu'il touche l'écran grâce à un système de capteurs piézoélectroniques.
http://www.businessmobile.fr/actualites/technologies/0,39044306,39701172,00.htm

http://www.igeneration.fr/iphone/de-nouveaux-brevets-pour-les-iphone-et-touch-8320

- *UPM Raflatac,* leader mondial dans la production de puces RFID HF (haute fréquence) et UHF (ultra haute fréquence), a lancé une nouvelle puce appelé *RaceTrack NFC* RFID dont la capacité/contenance de la mémoire et la sécurité des données connaissent des améliorations. Cette puce utilise une mémoire EEPROM c'est-à-dire un type de mémoire morte. Cette mémoire permet de garder des informations enregistrées même si elle n'est plus alimentée en électricité. Et ce même courant électrique suffit pour effacer le contenu de la mémoire. La sécurité/protection des données a été amélioré grâce à l'utilisation de l'algorithme 3DES pour l'authentification et une protection anti-clonage/copie des informations qui se trouvent sur la puce. Dans l'article, il précise que cette puce peut permettre de nouvelles applications mais sans préciser lesquelles.
http://www.morerfid.com/details.php?subdetail=Report&action=details&report_id=5468&display=RFID

- *UPM Raflatac* a également lancé des nouvelles versions de leurs gammes de puces RFID (*DogBones, ShortDipole, Frog, GEM*) qui sont encore plus performantes. Elles sont équipées d'une mémoire EPC 96-bit ce qui permet d'améliorer la sensibilité et la fiabilité de la lecture des puces. L'amélioration de la sensibilité, donc de la précision, permet d'aborder de nouvelles applications. De plus, elles fonctionnent même à proximité de matériaux métallique ou liquide, alors que certaine puce ne fonctionne pas dans les même conditions, cela est du aux ondes lors des connexions entre la puce et le lecteur. La nouvelle version de la puce *GEM,* développé pour la gestion des inventaires, peut maintenant faire l'inventaire de petits bien, objets comme les bijoux ou les montres mais l'on pourrait élargir aux produits, alimentaire ou non, vendus dans les supermarchés.
Dogbone = applications pour le supply chain management pour produit fini ;
ShortDipole = applications pour le supply chain management pour produit standard ;
Frog = pour des applications qu'il appelle « orientation-sensitive » ;

GEM = puce utilisé pour la gestion des inventaires notamment pour les petits objets.
http://www.morerfid.com/details.php?subdetail=Report&action=details&report_id=5354&display=RFID

- *RFIDSpan Technology* a annoncé le lancement d'un système multiplexeur (circuit permettant de concentrer sur une même voie de transmission différents types de liaisons) UHF RFID Intelligent. Cet outil concerne plutôt la logistique et le back office, sans lien direct avec le client mais indirectement car améliorera l'approvisionnement dans les rayons. Il permet de mettre en réseau plus de 259 antennes RFID ce qui permet de quadriller un espace comme un supermarché (back et front office). Il apporte une solution aux deux problèmes majeurs des systèmes de localisation et de pistage de biens/unités/objets en temps réel et des étagères intelligentes. Cet outil permet de réduire le nombre de lecteur RFID et simplifie le système de communication par RF. Il permet d'optimiser les étagères intelligentes comme les *SmartBlades* développées par *Intellident* en les reliant sous forme de réseau dont les informations transmises ou reçues par les étagères sont centralisées vers un nombre restreint de lecteurs/émetteurs RFID voir un seul lecteur/émetteur (produit actuellement disponible).
http://www.morerfid.com/details.php?subdetail=Report&action=details&report_id=5458&display=RFID

- *Netsize* (leader dans le commerce et des communications mobiles) et *AIRTAG* (premier fournisseur de solution marketing sans contact) ont développé une solution de marketing mobile. Basée sur la solution *AIRTAG Pad* et l'outil de gestion de campagnes marketing mobiles de *Netsize*, la solution permet aux clients d'adhérer à la mailing liste et aux services mobiles d'un commerçant ou distributeur. *AIRTAG PAD* est un terminal point de vente multimédia sans contact intégrant l'ensemble des technologies cartes et mobiles (dont NFC). Il permet au distributeur d'enrichir leurs bases de données clients. Le *mMarketing* de *Netsize* complète le terminal *AIRTAG* en permettant au distributeur d'utiliser le mobile pour informer, interagir avec leurs clients. Cela leur permet de gérer des mailings listes et lancer des compagnes marketings sous forme de contenus textes, SMS, MMS ou pages internet mobiles. Un tableau de bord détaillé

permet de mesurer et contrôler l'efficacité des campagnes en temps réel. Cet outil donne une idée de ce que sera le marketing à l'avenir qui passera probablement de plus en plus par les téléphones mobiles et internet
http://www.morerfid.com/details.php?subdetail=Report&action=details&report_id=5489&display=RFID

- *Confidex* a lancé un service de personnalisation de tag RFID dans l'optique de faciliter le déploiement de ces tags. Il inclut l'encodage ou le pré-encodage du tag, le visual marking et des outils de fixation. Ce service utilise également une technologie visual information qui permet de détecter l'ID du tag sans devoir utiliser un lecteur RFID. La personnalisation permet de s'adapter à un large éventail de cas, de situation comme le commerce et peut être le cas vous nous avez décri. *Confidex* propose de fournir des tags pré-encodés qui permettent un gain de temps et de ressources. Le pré-encodage se base sur les codes EPC. Je trouve l'approche de l'encodage des puces intéressante car permet de s'adapter à un grand nombre de situation ou au cas par cas de manière de s'adapter au mieux au client. Je pense aussi que si on peut encoder/programmer un puce, on peut arriver à compartimenter une puces RFID et a y stocker des informations accessibles pour les clients. C'est probablement une piste à approfondir.
http://www.morerfid.com/details.php?subdetail=Report&action=details&report_id=5890&display=RFID

- *Tego Revolutionizes* a développé un tag RFID intelligent appelé *TegoTags*. La particularité de ce tag RFID est d'avoir une mémoire élevée (high memory). Il peut stocker plus de 2500 fois la quantité des données qui est standard dans l'industrie aujourd'hui. *TegoTags* délivre de l'information sans contact/fil en utilisant les équipements de lecture de tag RFID standard ce qui permet une grande accessibilité de ces puces notamment dans les lieux ou les connexions internet, wifi, les réseaux téléphoniques (3g,...) ne sont pas disponible. *TegoTags* est utilisé en aéronautique, cher Boeing qui y stock les historique de maintenance, des informations sur le personnel responsable de la maintenance et la sécurité pour un model d'avion

donnée Selon Boeing, les TegoTags permettent aux techniciens de faire des réparations efficacement soit dans les entrepôts soit à distance sur les pistes d'atterrissages. Pour *Tego Revolutionizes,* sa puce pourrait simplifier et rendre plus écologique les marquages des produits à travers la chaine d'approvisionnement réalisé par les industriel de la vente aux détails. Cette a l'air vraiment intéressant car permet de stocker un nombre important d'informations et peut servir de base à l'application que vous nous avez décrite, avoir une puce RFID avec un espace de stockage suffisant pour le mode d'emploi du produit tagger, SAV, la garantie, etc. Cependant la protection des informations et la compartimentation ne sont pas évoquées.

http://www.morerfid.com/details.php?subdetail=Report&action=details&report_id=5418&display=RFID

Synthèse : 05/08/09

- *Dairy Queens,* chaine de restaurant américain, teste un système de fidélisation de clients basé sure la RFID appelé *Mobiquitous* développé par *Carmel* et *Tethrball.* Les clients participant à ce programme utilisent une étiquette RFID collé à leur téléphone portable. Le client reçoit une offre de remise par SMS qu'il peut utiliser dans les espaces de ventes participant au programme. Le client peut soit présenter son étiquette RFID au lecteur au niveau de la caisse, soit il peut imprimer leur coupon de réduction dans une cabine individuelle. Quand le client effectue une des tâches décrites précédemment, son numéro identifiant est automatiquement transmit au système *Mobiquitous.* Le commerçant peut cibler ces promotions pour un client en se basant sur les achats qui ont été fait par le client et les succès des précédentes promotions. Le commerçant peut voir quel trafic de client est généré en prévision de la nouvelle campagne de coupons de réduction. Ce système apporte également des informations démographiques et savoir combien de client sont venus à la boutique. Ce système ne nécessite aucun achat ou téléchargement d'une application pour téléphone portable.

http://www.rfidupdate.com/articles/index.php?id=1808

- Une chaine de magasin américain, *TOP Food and Drug,* a également développé un programme, appelé *TOP Connection,* de relation client qui utilise la

technologie RFID. La plateforme logicielle est développée par *Accelitec*. Le système se compose de la manière suivante : des lecteurs RFID sont placés au niveau des terminaux de payement et chaque client possède soit une carte RFID soit une étiquette à coller sur leur téléphone portable. Chaque puce RFID contient un identifiant client unique. Ce matériel RFID est fourni par *UPM Raflatac*. Tous les informations sur les clients client sont cryptées et conservées sur un serveur *Accelitec* sécurisé. Le cryptage des données permettent de protéger les informations bancaire et sur l'identité des clients durant la lecture de la puce RFID et les transactions notamment au moment du payement. Lors de l'inscription au programme, les clients peuvent soit donner un nombre d'information de base soit donner des informations plus détailler ce qui permet de hautement customiser les offres et les communications marketing.

http://www.morerfid.com/details.php?subdetail=Report&action=details&report_id=5623&display=RFID

ANNEXE 5 : Tableau d'acteurs RFID et de leur produit

Entreprise/laboratoire	nationalité	Dirigeant	Contact mail	Site web	Produit/solution	Description du produit/solution
Violet	Français	Rafi Haladjian	contact@violet.net	http://www.violet.net	Violet et Neamedia se sont associés. MIR:ROR et des objets communicants personnalisé et répondant aux besoins des clients	MIR:ROR est un lecteur de tag RFID relié à un ordination par port USB qui lance des applications associé à ce tag comme avoir accès à du contenu en ligne. Les objets communicants sont des objets équipés de puces RFID. Violet commercialise aussi le célèbre lapin Nabaztag
Neamedia	Français		info@neamedia.fr	http://www.neamedia.fr/		
Alcaltel-Lucent	Franco-américaine	Philippe camus et Ben Verwaayen	Contact directement sur le site web	http://www.touchatag.com/	Touchatag (anciennement Tikitag)	Kit permettant de relier un objet à une application en se basant sur la technologie NFC. Il se compose d'autocollants équipés de puces NFC, d'un lecteur NFC et d'un accès au site web Touchatag pour pourvoir créer les applications
Airtag	Français	Jérôme Leroyer	Contact directement sur le site web	http://www.airtagkit.com/ ; http://www.airtag.com/-Francais-.html	Airtagkit	Kit permettant de relier un objet à une application (du contenu en ligne) en se basant sur la technologie NFC et RFID; le kit se compose d'un lecteur, de tags et d'outils de développement logiciel qui permet de programmer un objet
Twinlinx	Français	Jacek Koalski	Contact directement sur le site web	http://www.twinlinx.com/	mise au point d'une puce intégrant la technologie NFC et RFID ; MyMax sticker (anciennement Blu link)	MyMax sticker est une solution pour transformer n'importe quel mobile équipé de bluetooth en lecteur de tags NFC grâce à des étiquettes électroniques
Thingmagic	Américain	Tom Grant	salesteam@thingmagic.com	http://www.thingmagic.com/	Tool Link	Lecteur RFID qui permet de détecter des objets équipés de tag RFID par rapport à une liste
Intellident	Anglais	?	contact@identag.net	http://www.intellident.co.uk	SmartBlade	Etagère intelligent équipé d'une antenne led flash qui détecte les objets ayant un tag RFID avec précision, il permet à l'utilisateur de localiser des objets ou inversement de constater leur absence (évite les rayon vides). Cet outil est assez précis car grace à la fonction Smartsearch permet de trouver un objet (livre, etc) qui peut être petite taille mais qui est équipé d'une puce RFID. Cette fonction situe l'objet dans l'étagère à 30 cm près.
IPSC (une institut européenne)	UE	Stephan Lechner	ipsc@jrc.it	http://ipsc.jrc.ec.europa.eu/home.php	réseau de transpondeurs RFID	Ce réseau de transpondeurs RFID permet de s'orienté et de recueillir des informations dans un environnement (ville, supermarché, etc.). Le projet pilote comportait 2500 transpondeurs. Pour l'instant il faut être équipé d'une canne qui active les transpondeurs quand il passe à proximité et d'un téléphone portable préalablement enrichi d'une base de données recensant géographiquemet les référence de chaque transpondeur qui quand ils sont activés communique son code identifiant au mobile. Ce produit est assez précis, il permet de renseigner l'utilisateur sur sa position exacte et son environnemlent direct.

Entreprise/laboratoire	nationalité	Dirigeant	Contact mail	Site web	Produit/solution	Description du produit/solution
University of Massachusetts Amherst	Américain	Holcomb, D. Burleson, W. Fu, K	kevinfu@cs.umass.edu, burleson@ecs.umass.edu, dholcomb@ecs.umass.edu	http://www.eecs.berkeley.edu/~holcomb/rfid_sec-FERNS-camera-ready.pdf	Sécurisation des puces	Découverte d'un nouveau moyen de sécuriser les puces RFID en utilisant la mémoire de type SRAM présent dans les puces, solution simple et peu couteuse.
La ville de Salerno en association avec IBM	amérocano italien	Vincenzo De Luca (Salerno), Samuel J.Palmisano (IBM)		http://www-304.ibm.com/jct03001c/press/us/en/pressrelease/27557.wss	concept de cité intelligente	Idée de rendre une ville intelligente sans passer par de nouvelles infrastructures mais en utilisant des puces RFID installé dans toute la ville pour donner des infos supplémentaire sur les mobiles des touristes ou alors rendre la ville plus accessible aux mals voyants
Artfest international	Américain	Edward Vakser	info@artfestinternational.com	http://ArtfestInternational.com	Traçage et authentification de produit	Système d'autnentification et de traçage se composant d'un logiciel et de puces RFID qui sont attachées aux produits (ici dernières les œuvres d'art), c'est un moyen de protection zet de garantir l'authenticité d'un objet.
Nokia	Finlande	Olli-Pekka Kallasvuo	Contact directement sur le site web	http://www.nokia.com/	téléphonie mobile : le modèle 6216 classic	Le nokia 6216 classic incorpore la technologie NFC dans la carte SIM. Cette innovation garantit une meilleur sécurité des données notammment bancaire et permet, lors d'un changement de portable, de déplacer facilement son porte monaie électronique, par exemple, qui se trouve sur la carte SIM.
UPM Raflatac	Finlande	?	Contact directement sur le site web	http://www.upmraflatac.com/europe/fra/	Puce RaceTrack NFC RFID	Puce dont la mémoire et la sécurité ont subi une amélioration notamment de capacité pour la mémoire. Cette puce a une mémoire de type EEPROM dit mémoire morte. La sécurité par authentification et l'anti clonage des puces (et des données qu'elle contient) a été renforcé à l'aide de l'algorithme 3DES.
					Noyuvelle version des puces Dogbones, ShortDipole, Frog et Gem	Toutes ces puces sont équipées de mémoire EPC 96-bit ce qui améliore leurs sensibilitées (la précision) et leurs fiabilitées. Elles sont capables de fonctionner proche de matériaux métalllique ou liquide). La puce Gem, développée pour les inventaires, peut maintenant le faire sur de petits biens. Les puces DogBones et ShortDipole sont développées pour des applications liées au supply chain management et la puce Frog pour des applications qu'ils appellent "orientation sensitive".
RFIDSpan Technology	Américaine	Albert Hong	info@rfidspan.com	http://mysite.verizon.net/rfidspan/product.html	iSwitch-6000 and iSwitch-3000,	Système multiplexeur UHF et RFIDintelligent permet de mettre 256 antenne en réseau soit autant de puces. Cet outil permet de limiter le nombre de lecteur RFID et symplifie le système de communication par RF; Il permet d'optimiser les étagère intelligente cité précédemment en les organisant sous forme d'un réseau relié à quelque voir un lecteur RFID

Entreprise/laboratoire	nationalité	Dirigeant	Contact mail	Site web	Produit/solution	Description du produit/solution
Netsize/AIRTAG	Français/Français	Stanilas Chesnais / Jérôme Leroyer	Contact directement sur le site web pour les deux	http://www.netsize.com http://www.airtag.com/fr-airtag-pad-interactive-shopper-marketing.html	Solution de marketing mobile	se présente sous la forme de borne et soit de cartes soit téléphone portable. Il permet au commerçant d'enrichir leur base de donnée client pour pourvoir lancer des campagnes marketings sous forme de contenu texte, SMS, MMS ou page internet mobile. Un tableau de bord permet de vérifier l'efficacité des campagnes en temps réel
Confidex	Finlande	Reijo Mäihäniem	contact@confidex.fi	http://www.confidex.fi	Personnalisation de tag RFID	Confidex a lancé un service de personnalisation de tag qui l'encodage ou le pré-encodage du tag, le visual marking et des outils de fixations. Le client peut personnaliser lui-même les tag cependant Confidex propose de pré-encoder les tags ce qui offre un gain de temps et de ressource.
Tego	Américain	Timothy Butler	info@tegoinc.com	http://www.tegoinc.com	Puce RFID TegoTags	L'étiquette TégoTag a une mémoire à capacitée élevée. Il peut stocker plus de 2500 fois la quantité des données qui est standard dans l'industrie aujourd'hui. Compatible avec les lecteurs de tags RFID standarts, les TéoTags transmettre leurs informations sans contact/fil. Pour Tego , sa puce pourrait simplifier et rendre plus écologique les marquages des produits à travers la chaine d'approvisionnement réalisée par les industriel de la vente aux détails.

ANNEXE 6 : Article sur l'évolution de la 3D dans le cadre du commerce en 2009

Article sur l'évolution de la 3D/réalité augmentée :

La technologie 3D est présente dans beaucoup de domaines ou secteurs comme le commerce. De plus elle est associée à d'autres technologies, notamment dans la téléphonie mobile avec la possibilité de visionner des vidéos 3D sans l'utilisation de lunette spéciale[22]. En effet de nombreux interfaces 3D sont actuellement en développement comme le cube gCubik du NICT (Institut national japonais des TIC)[23] ou la table tactile 3D de la DCNS (actuellement à usage militaire), ces technologies présagent l'évolution des interfaces homme/machines en 3D et pourraient amener à des applications dans le marketing ou la relation client. La 3D est également employé par le e-commerce comme un moyen de présenter les produits mais aussi de virtualisation du commerce appelé « v-commerce » par l'intermédiaire d'univers virtuels comme Second life ou par la création de galerie marchande 3D comme celle de Tamkari[24]. La réalité augmentée, l'autre technologie en plein expansion, utilise la 3D comme dans l'application d'IKEA[25] qui permet de visualiser des meubles virtuels dans un environnement réel. Cette technologie s'est aussi développée dans la téléphonie mobile permettant d'avoir des informations complémentaire sur un objet ou un service[26]. Même si pour l'instant une grande partie des applications apparaissent comme des gadgets[27], de nombreux acteurs misent sur la 3D et la réalité augmentée, Google, Mozilla et Microsoft en tête[28], et nous nous orientons vers un internet, des solutions marketing (panneau publicitaire 3D, rayonnage intelligent, publicité papier etc.), des relations clients (carte de fidélité intelligente, nouvelle interface[29], etc.) et un back/front office (logistique, etc.) en 3D.

[22] « La 3d sans fil arrive », *Usine nouvelle*, 28/05/2009, n°3149, p49
[23] http://www.clubic.com/actualite-277736-gcubik-une-nouvelle-interface-3d-japonaise.html
[24] http://www.metaverse3d.com/2009/01/02/tamkari-une-galerie-marchande-3d/
[25] http://www.loria.fr/~gsimon/ra/
[26] http://www.hebiflux.com/blog/2009/05/26/est-ce-enfin-le-debut-des-applications-de-realite-augmentee-sur-mobile/
[27] http://www.techeblog.com/index.php/tech-gadget/topps-introduces-3d-live-baseball-cards
[28] http://www.geekandhype.com/2009/05/16/microsofts-2019-future-vision/
[29] http://www.laval-virtual.org/index.php?option=com_content&task=view&id=156

ANNEXE 7 : Article sur la solution *Pachube*

Pachube : 19/06/09

Pachube prend la forme d'un réseau social, d'une plateforme de partage de données. L'idée de ce réseau est de collecter, d'agréger des informations de différents capteurs (station météo, compteurs de consommation, objet communicants, capteur lié à l'internet des objets, etc.). Pour cela, il faut s'inscrit sur le site web de Pachube et créer son application. On peut trouver sur Pachube une application pour évaluer les conséquences de l'occupation humaine sur une habitation, une pour connecter ses immeubles et visualiser les dépenses énergétiques, relier des utilisateurs d'une marque de vêtements entre eux ou encore contrôler sa maison à distance. Le potentiels d'application semble important voir infini. La particularité de Pachube est de pourvoir collecter des informations de capteurs situés dans des univers virtuel comme *Second Life*. Cette particularité permet de relier des objets virtuels à des objets réels. J'ai conscience que celle solution s'écarte un peu de la thématique RFID, quoique ces puces pourraient devenir ou être des capteurs, et se rapproche plus de l'internet des objets et du M2M (machines to machines) mais je trouve le principe intéressant avec un grand potentiel d'application. Pour le commerce, cette solution pourrait être un outil marketing pour étudier les pratiques des clients grâce à des capteurs sur certain produit en vente. Cette solution touche aussi le e-commerce et le v-commerce grâce à la possibilité d'interagir avec des objets dans des univers virtuel.

- http://www.internetactu.net/2009/02/05/pachube-des-applications-pour-linternet-des-objets/
-http://www.atelier.fr/usages/10/23012009/pachube-interagir-mondes-virtuels-capteurs-immeubles-objets-sols-37737-.html
- http://wangxiang.wordpress.com/2009/03/22/premiers-essais-avec-pachube/ (concerne plus la partie virtuel)
- http://www.pachube.com/

ANNEXE 8 : Les éléments racines du dictionnaire métier

Veuillez trouver ci-dessous les éléments racines sur 2 niveaux.

- Acteurs
  - Autres Acteur
  - Enseignes étrangères
  - Enseignes françaises
  - Fournisseurs de solutions
  - Organisation de standardisation
  - Pôle de compétitivité
- Environnement
  - Déchets
  - Réglementation
- Etudes
  - Comportement client
  - Multicanal
- Logistique
  - Approvisionnement
  - Centrale d'achat
  - Chaîne logistique
  - Entrepôt
  - Logistique urbaine
  - Palette
- Marketing
  - Crowdsourcing
  - Ethnomarketing
  - Géomarketing
  - Marketing communautaire
  - Marketing de permission
  - Marketing direct
  - Marketing global
  - Marketing mobile
  - Marketing one-to-one
  - Marketing orienté
  - Marketing sensoriel
  - Marketing viral
  - Mix marketing
  - Neuromarketing
  - Webmarketing
- Pays
  - Afrique
  - Amérique
  - Asie
  - Europe
  - Moyen-Orient
  - Océan Indien
  - Océanie
- Secteur d'activité
  - Activité bancaire

- Agriculture
- Agroalimentaire
- Arts ménagers
- Assurance
- Batiment
- Bijouterie
- Chimie
- Communication
- Electricité
- Electronique
- Emballage
- Loisir
- Textile
- Transport

Technologies

- Automatisation
- Informatique ubiquitaire
- Infrastructure et réseau
- Intelligence ambiante
- Interface
- Lecteur
- M2M
- Norme téléphonie mobile
- Objet intelligent
- Réalité augmentée
- RTLS
- RTSC
- Surface haptique
- Système d'information
- Technologies sans contact
- Télématique
- Univers 3D
- Visual computing

Type de commerce

- Commerce de détail
- Commerce de gros
- Commerce organisé
- Modèle économique

Usages

- Antivol
- Contrefaçon
- Espace de vente intelligent
- Gestion marchandise
- Identification
- Paiement