

LIVRE BLANC
LA GESTION DES CONNAISSANCES
Groupe de travail sous l'égide du GFII



ALAIN BENSOUSSAN



AVOCATS

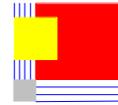


TABLE DES MATIERES

AVANT-PROPOS

1.	INTRODUCTION	6
2.	LE CONCEPT DE "GESTION DES CONNAISSANCES"	7
2.1	APPROCHE	7
2.2	LA PRATIQUE NORMATIVE	8
2.3	LES DEFINITIONS SCIENTIFIQUES	10
3.	LES CONCEPTS	11
3.1	CULTURE DE PARTAGE	11
3.1.1	Appréhender la dimension humaine	11
3.1.2	Gérer les phases de refus	13
3.1.3	Charte comportementale	13
3.2	LA VALEUR DE L'INDEXATION	14
3.2.1	La "trouvabilité" : un défi	14
3.2.2	Intérêt de l'indexation pour les utilisateurs d'information	15
3.2.3	Intérêt de l'indexation pour l'architecture des sites intranet ou des portails	16
3.2.4	Nouvelles applications de l'information indexée	17
3.2.5	Rappel historique sur l'indexation	19
3.2.6	Vocabulaires normalisés ou taxonomies	21
3.2.7	Factiva Intelligent Indexing	25
3.3	MEILLEURES PRATIQUES	28
3.4	TECHNOLOGIE "OBJET"	28
3.4.1	Définition du polymorphisme	29
3.4.2	Définition de l'héritage	29
3.4.3	Encapsulation et abstraction	30
3.4.4	Application des concepts orientés objet à la gestion des connaissances	31

3.5	PROPRIETE DES CONTENUS	32
3.6	EXTRACTION DES CONNAISSANCES	35
3.7	PRODUITS ORIENTES METIERS	37
3.8	GESTION SEMANTIQUE DES CONNAISSANCES	46
3.9	GESTION DU CONTENU VISUEL	48
4.	LES OUTILS	52
4.1	PANORAMA DES TECHNOLOGIES	52
4.1.1	Les outils de capitalisation des connaissances	52
4.1.2	Les outils de travail	53
4.1.3	Les outils de diffusion des connaissances	54
4.2	ILLUSTRATIONS	54
4.2.1	Sensoria Technologie Limited	54
4.2.2	Plutarque	66
4.2.3	iClass	70
4.2.4	CCINET	72
5.	LES MODELES ECONOMIQUES	80
5.1	LES OBJECTIFS	80
5.2	LES CONCEPTS DE VALEUR	80
5.3	LA FACTURATION	84
5.3.1	Les critères de facturation	84
5.3.2	Les modes de facturation	85
5.4	ILLUSTRATION : MODELES DE TARIFICATION LEXISNEXIS	86
6.	LES ASPECTS JURIDIQUES DE LA GESTION DES CONNAISSANCES	92
6.1	DISPOSER DE CONTRATS PROTECTEURS	92
6.1.1	Le contrat de conseil	92
6.1.2	Le contrat de licence d'utilisation de l'outil	93
6.1.3	Le contrat de formation	94
6.1.4	Le contrat de réalisation de développements spécifiques	94
6.1.5	Le contrat de maintenance	95
6.2	GERER LES DROITS SUR L'OUTIL DE GESTION DES CONNAISSANCES ET SON CONTENU	95
6.2.1	Les droits sur le contenu de l'outil : la protection par le droit d'auteur	95
6.2.2	Les droits sur le contenu de l'outil : la protection des bases de données	102
6.2.3	Les droits sur l'outil : la protection des logiciels	108

6.3	DEPLOYER L'OUTIL DANS LE RESPECT DE LA LOI	110
6.3.1	Le respect de la loi "Informatique et libertés"	110
6.3.2	Le respect du droit du travail	113
6.4	ASSURER LA SECURITE DE L'OUTIL	117
6.4.1	Le contrôle de l'utilisation de l'outil	117
6.4.2	Le vol d'informations	119
6.4.3	La mise en place d'une charte de sécurité	120
7.	RECOMMANDATIONS	122
8.	ANNEXES	123

AVANT-PROPOS

La gestion des connaissances constitue un enjeu majeur dans une économie de services puisqu'elle vise à développer le capital intellectuel de l'entreprise, patrimoine et clé de sa compétitivité.

Réuni sous l'égide du Groupement Français de l'Industrie de l'Information (GFII), le groupe de travail¹ "Gestion des connaissances", présidé par Maître Alain Bensoussan, est composé de professionnels aussi bien experts des questions techniques que juridiques.

Ce groupe de travail s'est fixé pour objectif la réalisation du présent livre blanc sur la gestion des connaissances dans l'entreprise².

En effet, l'accélération technologique, la concurrence et la dissémination des compétences conduisent les entreprises qui considèrent leur patrimoine de connaissances comme stratégique, à expérimenter des processus de capitalisation d'expérience et d'aide à la production.

C'est dans ce contexte que le groupe de travail a proposé d'axer sa réflexion sur le knowledge management dans l'entreprise sous l'angle économique, technique et juridique et a émis des recommandations qui pourront guider les entreprises dans leur démarche de gestion du capital intellectuel.

¹ La composition du groupe de travail est présentée en annexe du présent livre blanc.

² Les membres du groupe de travail se sont accordés pour reprendre les travaux au 1^{er} janvier 2005 afin de réaliser une mise à jour du livre blanc.

1. INTRODUCTION

Pour renforcer leur réactivité, leur compétitivité et leur productivité et consacrer l'essentiel de leur temps à la recherche de valeur ajoutée, les entreprises cherchent aujourd'hui à capitaliser, valoriser, mutualiser et partager les connaissances, à stocker, à l'aide d'outils intelligents, le savoir-faire et le savoir-agir et à établir une mémoire collective de l'entreprise.

La nécessité de raisonner en termes de flux et d'échanges, c'est-à-dire de fonder son architecture informationnelle sur les réseaux et l'administration des données, constitue donc pour les entreprises, une préoccupation majeure.

Pour accompagner les producteurs et les utilisateurs d'outils de gestion des connaissances dans leur réflexion, le groupe de travail a mené ses travaux au regard :

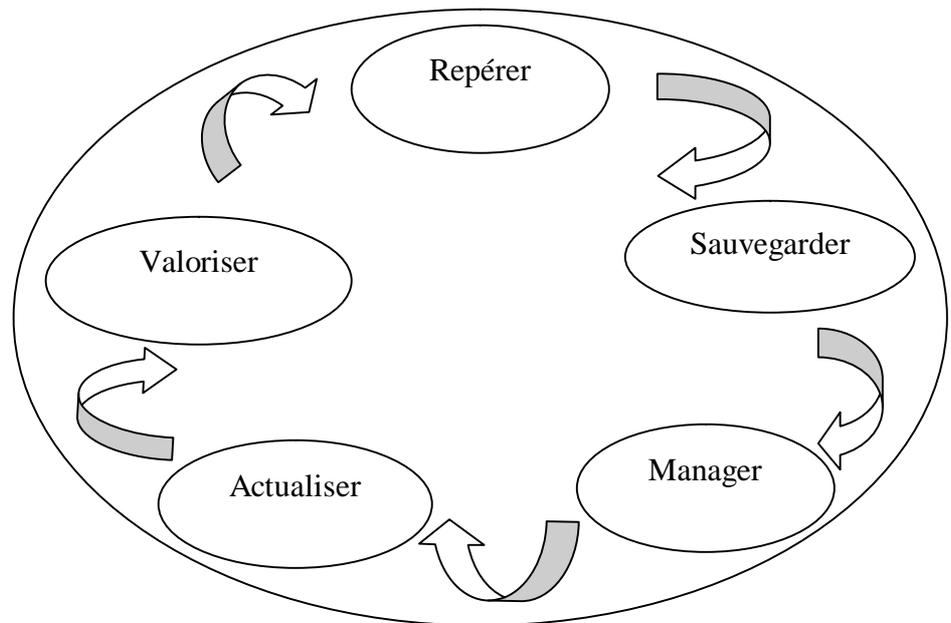
- des fonctions associées à la gestion des connaissances ;
- des attentes des responsables "managers" de la connaissance sur la base des outils actuellement disponibles sur le marché ;
- des aspects juridiques liés à cette problématique ;
- des modèles économiques à développer.

Les recommandations formulées par le groupe de travail ont pour vocation de contribuer à optimiser l'outil de gestion des connaissances au sein de l'entreprise.

2. LE CONCEPT DE "GESTION DES CONNAISSANCES"

2.1 APPROCHE

La gestion des connaissances peut se présenter sous forme d'un cycle de capitalisation d'expérience caractérisé par cinq aspects :



Ce cycle de capitalisation met en évidence trois sources documentaires distinctes, chacune d'elles pouvant être interne ou externe à un agent économique, formalisée ou non et résultant de l'expérience et de l'historique de l'entreprise ou de travaux et réflexions communs à plusieurs entreprises :

La pratique normative (institutionnelle/propriétaire)
Les définitions scientifiques
Les pratiques professionnelles

2.2 LA PRATIQUE NORMATIVE

L'Afnor a publié, en septembre 2000, la norme FD X 50-190 "Outils de management. Capitalisation d'expérience"¹ qui fournit des recommandations pour faciliter la mise en œuvre de dispositifs de capitalisation d'expériences dans les entreprises.

Son cadre général est constitué par la norme ISO 9000:2000 "Systèmes de management de la qualité. Principes essentiels et vocabulaire". La norme ISO 9000:2000, fondée sur huit principes de management de la qualité, envisage les phases de conception, de réalisation et de déploiement dans lesquelles le management de la qualité intervient.

Définir la gestion des connaissances est un art difficile car elle fait appel à deux notions distinctes dont les axes et contenus apparaissent différents, voire éloignés, à savoir, la "gestion" et les "connaissances".

La norme FD X 50-190 définit le terme de "connaissance"² de la manière suivante :

- "Ensemble des représentations, idées ou perceptions acquises par l'étude ou l'expérience".

La norme précise³ que la connaissance comporte deux volets distincts :

- d'une part, une démarche volontaire d'acquisition ;
- d'autre part, une acquisition induite par l'action même.

Cette définition est conceptuelle. La connaissance est en effet synthétisée par la combinaison de différents concepts et de multiples sources d'informations en un tout cohérent, par inférence ou raisonnement déductif.

L'élément moteur de cette définition est l'expérience dont les contours doivent être appréhendés, et la notion d'informations et/ou de données résultant d'une production ou d'une expérience.

Le terme "expérience" est défini par la norme⁴ comme :

- " Savoir acquis par la pratique et/ou l'observation".

Ainsi, alors que l'expérience provient d'une confrontation au réel et à la pratique, la connaissance apparaît comme le résultat d'un effort intellectuel qui inclut l'expérience.

¹ Afnor, "Outils de management. Capitalisation d'expérience", FD X 50-190, septembre 2000.

² Article 3.3.

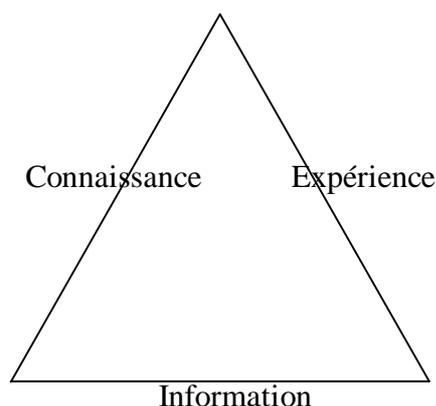
³ Article 3.3.

⁴ Article 3.5.

Le terme "information" est défini par la norme FD X50-190¹ de la manière suivante :

- "Association significative de données portées sur un support selon un code prédéfini".

L'information constitue donc le socle, sur lequel tout repose, selon une triangulation qui peut être la suivante :



L'information se compose de données qui ont été interprétées ou traduites pour en relever la signification ou le motif sous-jacent, étant précisé que pour être "utile", cette information doit être organisée selon une schématique précise, un "code prédéfini" et figurer sur un support.

C'est ce support qui permettra sa capitalisation et ce retraitement de l'information par le prisme de l'expérience et de sa gestion au sein des entreprises.

Ces notions appellent donc un autre concept : la capitalisation.

Le terme "capitalisation" est défini par la norme FD X 50-190² comme :

- "Action d'accumuler volontairement et de manière organisée en vue d'un profit ultérieur".

Ce terme renvoie à un processus spécifique qui, selon la norme³, se décompose en cinq phases, à savoir :

- identification de l'expérience à capitaliser ;
- mise en place des dispositifs nécessaires à la capitalisation ;
- capture de l'expérience à réutiliser ;
- valorisation par transformation en informations utilisables ;
- utilisation du capital ainsi rassemblé.

¹ Article 3.4.

² Article 3.1.

³ Article 4.5.2.

Le management des connaissances se situe donc sous un modèle de capitalisation, un transfert de savoir et des méthodes de capture.

Enfin, la norme met en exergue les liens entre la capitalisation d'expérience et le management de la qualité¹.

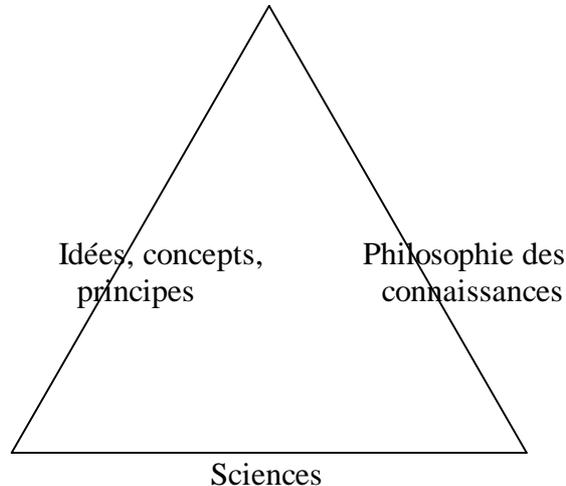
La norme rappelle en effet que "comme d'autres outils de la qualité, la capitalisation d'expérience est une composante essentielle du système de management de la qualité pour l'amélioration des performances de l'organisme en faisant appel à une accumulation d'informations organisées et traduites dans le but d'en assurer la gestion optimale au regard des objectifs de l'organisme"².

La capitalisation d'expérience constitue donc un des dispositifs d'amélioration du système de management de la qualité contribuant à tendre vers le "zéro défaut".

2.3 LES DEFINITIONS SCIENTIFIQUES

Si l'approche normative est un axe de l'approche de la gestion des connaissances, son approche scientifique révèle d'autres aspects permettant d'en tracer plus précisément les contours.

Les définitions scientifiques recouvrent trois domaines :



Cette triangulation "scientifique" permet de confronter trois sciences distinctes :

- la science par la réflexion (philosophie des connaissances) ;
- la démarche scientifique en tant que processus de création ;
- le résultat de cette démarche scientifique : les idées, concepts et principes.

¹ Article 4.6.

² Article 4.6.

Ces résultats de processus, sur le plan juridique, conduisent à réaliser des distinctions entre :

- les processus utilisés et leur partage éventuel ;
- l'appropriation collective et privative des éléments produits.

La problématique est de déterminer si des idées, concepts ou principes peuvent être "privatisés" alors qu'ils sont le résultat d'une production commune ou si l'idée "privative" à l'origine du processus peut être "collective" à terme, grâce au prisme de la gestion des connaissances.

Une réflexion sur l'appropriation privative des connaissances à travers le concept d'idées se développe aujourd'hui.

3. LES CONCEPTS

Une approche conceptuelle de la gestion des connaissances peut être utile à la réussite du programme. Il s'agira notamment d'appréhender les connaissances à travers leur gestion en définissant, en amont, les méthodes permettant de créer une culture de partage au sein de l'entreprise, de gérer les phases de refus et de résistance au programme et d'inciter les membres de l'entreprise à produire du contenu et à utiliser l'outil.

3.1 CULTURE DE PARTAGE

3.1.1 Appréhender la dimension humaine

La gestion des connaissances révèle trois dimensions qui sont en synergie :

- une dimension humaine ;
- une dimension organisationnelle ;
- une dimension technologique.

La dimension humaine ou la "psychologie" des collaborateurs de l'entreprise doit être dirigée vers le partage et la mise en commun des connaissances. C'est le principe essentiel qui, à l'origine, permettra d'appréhender les connaissances à travers leur gestion.

Sous l'angle de la dimension humaine, l'objectif consiste à :

- former les collaborateurs à l'utilisation de l'outil de gestion des connaissances ;
- leur faire accepter le programme de gestion des connaissances ;
- les inciter à transmettre les connaissances dont ils disposent.

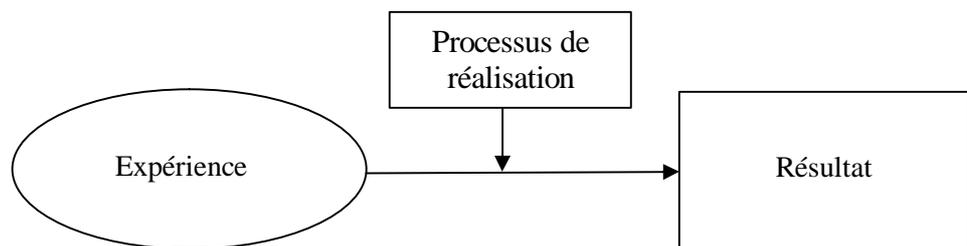
L'enjeu est de transformer la connaissance individuelle en une connaissance collective, accessible à tous les membres de l'entreprise, ce qui implique que la "psychologie" des collaborateurs soit orientée vers une mise en commun et vers une réutilisation par tous, des connaissances.

L'objectif est donc de créer au sein de l'entreprise cette culture de partage, de convaincre les membres de l'entreprise de l'importance du partage des informations, des connaissances et des savoir-faire.

A ce titre, l'incitation (en donnant aux collaborateurs du pouvoir dans le système) et la reconnaissance (en valorisant l'individu à l'origine de la production par l'indication de son nom sur le document fourni, par exemple) font partie des moyens pouvant favoriser la création d'une culture de partage.

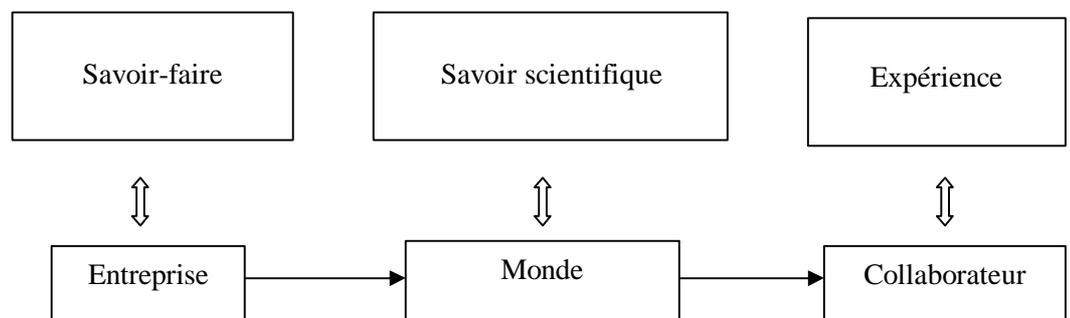
Or, dans la relation "employeur-employé", l'employeur capture le résultat de l'expérience de ses collaborateurs et non le processus ayant conduit à la réalisation de ce résultat.

La difficulté majeure est celle de convaincre le collaborateur que le processus qu'il a suivi pour obtenir le résultat ou la connaissance doit être mis en commun.



Il n'existe pas, à l'heure actuelle, d'outils permettant d' "extraire" les connaissances des collaborateurs.

Leur expérience leur appartient, contrairement au savoir-faire qui appartient à l'entreprise et au savoir scientifique qui appartient à tous.



L'utilisation de systèmes "d'extracteurs" de connaissances contribuerait, sans doute, à accélérer le processus. Mais toutes les expériences ne sont pas forcément extractibles.

3.1.2 Gérer les phases de refus

La réussite d'un programme de gestion des connaissances nécessite de maîtriser et faire franchir très rapidement les phases de refus et d'opposition de chacun.

Il est donc impératif de créer des situations de "confiance". Ceci implique que le processus de capture et d'extraction des connaissances se déroule sans violence. La violence peut en effet être :

- économique (le collaborateur est tenu de faire un travail supplémentaire pour lequel il n'est pas rémunéré) ;
- physique (le collaborateur va devoir se mettre en état de réaliser cette tâche chaque jour ; le fait de passer par un extracteur n'est pas forcément adapté et peut rebuter) ;
- intellectuelle (le collaborateur peut se sentir appauvri, dépossédé).

La nécessité est de lui démontrer que par ce processus il "récupère" plus de connaissances que celles qu'il a acceptées de transmettre et que le système lui est très largement favorable. C'est le "plus plus". Il n'y a pas appauvrissement mais bien enrichissement, contrairement à la première impression du collaborateur.

3.1.3 Charte comportementale

L'élaboration d'une charte comportementale incitant les collaborateurs à utiliser l'outil de gestion des connaissances et à produire du contenu, sans pour autant s'en faire un devoir, peut s'avérer utile.

Il s'agira notamment de mettre en exergue, dans la charte, les avantages qu'un collaborateur peut retirer de la production de contenus destinés à être diffusés au sein de l'outil de gestion des connaissances (reconnaissance, positionnement stratégique au sein de l'entreprise, implication, prime...).

Cette charte aura également pour objet de sensibiliser les collaborateurs-producteurs de contenus aux questions liées :

- à la responsabilité du fait des informations produites.

En effet, il convient d'attirer l'attention du collaborateur sur la nécessité de diffuser des informations fiables dont il assume la responsabilité, en termes de vérification, de valeur et d'exactitude, notamment au regard de certaines catégories ou types d'informations obéissant à des réglementations spécifiques.

- au respect des droits d'autrui. Il s'agira notamment de souligner la nécessité de :

. ne pas diffuser d'informations contraires aux bonnes mœurs ;

. ne pas diffuser d'informations constitutives d'une injure, d'une calomnie ou d'une diffamation à l'égard de toute personne physique ou morale ou de toute institution étatique, gouvernementale ou non ;

. respecter l'ensemble des réalisations et créations d'autrui, et notamment tout élément susceptible d'appropriation privative au regard du Code de la propriété intellectuelle, comme les droits d'auteur ou de toute propriété industrielle, telles que les marques, dessins et modèles ou brevets ;

. rappeler à toute personne mise en cause ou heurtée par la nature des informations contenues dans l'outil, les droits reconnus par la législation en matière de presse et notamment le droit de réponse ;

. respecter l'ensemble de la législation relative à la protection des informations nominatives et notamment la loi Informatique et libertés.

3.2 LA VALEUR DE L'INDEXATION

3.2.1 La "trouvabilité" : un défi

Depuis la naissance de l'imprimerie, parvenir à trouver des documents après qu'ils ont été révisés et classés constitue un véritable défi.

Cette notion de "trouvabilité" revêt une importance accrue dans un environnement où l'information est pléthorique et où, dans un village planétaire, la nécessité de trouver rapidement les informations pertinentes devient indispensable pour la prise de décisions économiques.

Dans le monde économique actuel, le temps est probablement la denrée la plus précieuse. C'est pourquoi les utilisateurs d'information privilégient les outils qui les aident à :

- trouver rapidement les informations nécessaires ;

- gérer la quantité et la qualité des informations qu'ils exploitent régulièrement.

Si le terme d'indexation manque de cachet par rapport aux autres termes employés pour décrire des concepts liés à l'organisation et à la gestion des connaissances, il s'agit néanmoins d'un outil fondamental pour garantir la précision de l'organisation et de la récupération d'informations, en particulier lorsque le nombre de documents traités est important.

Le recours à un langage connu d'indexation granulaire, permettant de faire ressortir les informations les plus pertinentes et d'écarter celles présentant un intérêt secondaire ou nul, est un atout considérable pour les consommateurs d'informations.

De même, les concepteurs de sites et d'interfaces internet peuvent s'appuyer sur des étiquettes d'indexation afin de présenter uniquement aux utilisateurs les informations qui répondent à certains critères, et leur éviter ainsi d'avoir à apprendre l'architecture d'indexation ou la taxonomie.

Les utilisateurs obtiennent donc ce qu'ils cherchent, tandis que le langage d'indexation fonctionne en arrière-plan et demeure invisible pour l'utilisateur.

3.2.2 Intérêt de l'indexation pour les utilisateurs d'information

En quoi ce travail d'indexation, qui se déroule en arrière-plan, est-il si important pour les utilisateurs d'outils de gestion des connaissances ?

Dans un contexte où le contenu disponible augmente, et où cette croissance devrait être exponentielle, ces utilisateurs ne peuvent plus se permettre de diminuer leur productivité en consacrant un temps considérable à rechercher les informations nécessaires pour accomplir leur tâche.

Les entreprises doivent faire face non seulement à une information toujours plus abondante, mais également à la nécessité de fournir l'accès à cette information à un nombre croissant d'utilisateurs, réclamant des applications toujours plus nombreuses.

Pour l'entreprise, ces développements peuvent avoir des répercussions extrêmement favorables, à condition que tous les utilisateurs puissent accéder rapidement et sans frustration aux informations nécessaires à leur travail et à leurs prises de décision.

Dans une large mesure, le développement de l'e-business dépend aujourd'hui des sites Internet et extranet des entreprises.

Or, celles-ci sont confrontées à une même difficulté : veiller à ce que leurs clients et leurs partenaires trouvent facilement sur ces sites les informations utiles, afin de réaliser la vente et de fidéliser la relation.

Les utilisateurs d'information pourront retrouver des documents qui auront été indexés de manière précise et cohérente, à l'aide d'un vocabulaire normalisé.

Ceci est vrai qu'il s'agisse de documents imprimés ou de documents électroniques, et s'applique particulièrement bien au contexte actuel de communications électroniques.

L'accès à un langage normalisé accroît systématiquement l'efficacité du rapatriement d'informations.

Par ailleurs, lorsque le contenu est correctement indexé, les utilisateurs sont assurés de l'exhaustivité de la couverture d'un sujet, et par conséquent des résultats de la recherche.

Ils sont donc moins inquiets de passer à côté d'informations importantes.

Dans une étude présentée lors de la réunion annuelle de l'ASIS en 1996, le docteur Bella Hass Weinberg (Division of Library and Information Science, St. John's University, Jamaica, NY) notait que 10% environ des termes d'indexation attribués manuellement n'apparaissent pas dans le texte intégral.

Elle rapportait également que des études ultérieures avaient démontré que le recours à une indexation s'appuyant sur un vocabulaire normalisé permettait d'augmenter de 10% le texte intégral récupéré.

Les services professionnels de recherche d'informations, tels que Factiva, offrent à l'utilisateur la faculté supplémentaire de cibler sa recherche sur des documents publiés dans une certaine langue, apparaissant dans certaines publications, publiés depuis un certain temps, ou comprenant également d'autres concepts ou noms de sociétés.

Grâce à ces capacités d'indexation supplémentaires et aux fonctionnalités de l'outil, l'utilisateur dispose pour ses décisions professionnelles d'informations particulièrement ciblées et pertinentes.

3.2.3 Intérêt de l'indexation pour l'architecture des sites intranet ou des portails

Les utilisateurs de sites intranet développés en interne se plaignent souvent du temps passé à rechercher des documents, éventuellement déjà connus, ou à naviguer sur le site à la recherche de contenu sur un sujet spécifique.

Afin de répondre à cette exigence de "trouvabilité", les architectes de sites s'efforcent d'affiner leurs systèmes d'étiquetage et d'indexation, afin d'aider l'utilisateur à localiser l'information qu'il cherche.

Les acheteurs de solutions logicielles pour portails sont également très intéressés par les capacités de recherche et de récupération de ces produits, leur objectif étant d'accroître la productivité des collaborateurs de l'entreprise, en permettant à ces derniers d'accéder aux informations nécessaires à leurs tâches professionnelles.

Les fonctionnalités de recherche et de navigation sont considérablement optimisées si l'indexation et la hiérarchisation des sujets (ou taxonomie) apparaissent clairement à l'utilisateur et si elles sont adaptées au type de contenu fourni par les documents sources.

La hiérarchisation des sujets doit être suffisamment explicite pour permettre de décrire le contenu des documents émanant non seulement des divers départements de l'entreprise, mais également de sources externes.

Dans son compte-rendu de la conférence EBIC 2000, TFPL, acteur mondial dans le domaine de la gestion de l'information, cite la phrase de Tom Koulopoulos, président et fondateur de Delphi Group :

- "taxonomies are chic".

Ce rapport affirme également :

- "Face à la nécessité de rationaliser la récupération d'informations dans un contexte de contenu toujours plus abondant, les taxonomies occupent désormais le devant de la scène.

Celles-ci représentent le moyen de fournir le contexte et la structure nécessaires à la recherche et à l'exploitation d'informations et de données, vecteur de développement de l'e-business"¹.

Selon une étude de Forrester², la construction de la hiérarchie des sujets constitue le principal poste de dépenses pour la création d'un premier portail d'entreprise.

Selon les estimations de Forrester, le coût de la constitution d'une hiérarchie de sujets couvrant 500 000 pages internet serait proche d'un demi-million de dollars.

Les entreprises qui ne disposent pas encore d'une taxonomie propre, ou celles qui estiment que leur taxonomie doit se muscler pour traiter un contenu toujours plus volumineux, peuvent gagner du temps et de l'argent en achetant une licence d'utilisation pour tout ou partie d'une taxonomie existante, reconnue et complète.

3.2.4 Nouvelles applications de l'information indexée

Certains clients de Factiva, dotés de systèmes d'information et de gestion des connaissances perfectionnés, créent ou mettent à jour des taxonomies propres à leur entreprise, permettant d'étiqueter les documents et les rapports dans leurs lieux de stockage électroniques.

¹ <http://www.tfpl.com/ebic2000/EBIC2000FR1.htm>.

² Building an Intranet Portal, Forrester Research, janvier 1999.

D'autres évaluent l'intérêt économique de la création d'une telle architecture d'indexation / taxonomie par rapport à un investissement dans un outil automatisé déjà existant.

Factiva a mis au point Factiva Intelligent Indexing, une taxonomie exclusive, organisant hiérarchiquement plus de 1 300 termes sectoriels, géographiques, ou d'actualité et 300 000 codes société, appliquée sans exception à la vaste bibliothèque de contenu proposée par Factiva, riche de plus de 8 000 sources d'information.

Face à la difficulté croissante de trouver des informations dans un contenu interne et externe toujours plus abondant, certains clients de Factiva ont noué un partenariat avec la société afin de concevoir des solutions pour la recherche et la récupération d'informations à partir de sources diverses (archives, catalogues, bases de données, et liens vers des fournisseurs de contenu).

Parallèlement au déploiement de Factiva Intelligent Indexing, les clients se heurtent aux difficultés inhérentes à l'indexation et à l'utilisation d'un vocabulaire normalisé. Factiva propose ainsi différentes solutions de gestion de l'information pour répondre à différents cas de figure :

- le client reçoit de Factiva tous les documents publiés par un ensemble de publications prédéfini.

Les termes d'indexation de Factiva apparaissent sur les documents. Un formulaire figurant sur l'intranet du client permet aux utilisateurs de définir leurs centres d'intérêt dans cet ensemble de publications, afin que les mises à jour soient personnalisées.

Le client sélectionne des articles spécifiques parmi ceux qui lui sont proposés automatiquement, tirant ainsi parti de l'indexation Factiva.

- le client utilise sa propre taxonomie pour classer les documents provenant de sources diverses, dans le cadre d'un produit de business intelligence.

Le client souhaite cesser cette activité et travaille avec Factiva pour adapter sa taxonomie à celle de Factiva. Factiva hébergera le contenu (y compris le contenu non disponible habituellement via Factiva), le filtrera et livrera l'information sur l'intranet de l'entreprise.

Le cas échéant, Factiva peut également élaborer des requêtes complexes supplémentaires basées sur sa taxonomie, afin d'assurer une correspondance totale avec les codes demandés par le client.

- le client souhaite mêler contenu interne et externe et voir son moteur de recherche explorer l'ensemble des sources d'information, qu'elles soient internes ou externes à l'entreprise.

Le produit sera conçu de façon à ce que le moteur de recherche de l'intranet du client envoie deux équations de recherche, l'une vers Factiva et l'autre pour explorer les sites internes.

Les résultats peuvent être présentés globalement ou séparément. Factiva peut mettre à la disposition du client son savoir-faire en matière d'édition, afin de faciliter l'établissement de tables de correspondance entre les équations de recherche de Factiva et celles du client.

Fort de succès de l'indexation des sociétés de Dow Jones Interactive, Factiva devrait proposer prochainement à ses clients son vaste index de sociétés, et en assurer la mise à jour.

Les clients pourront ainsi avoir accès à un index de sociétés standardisé. Pour leurs recherches sur les sociétés, les clients auront la possibilité de définir des subdivisions, qui correspondent à l'indexation utilisée pour les sujets d'actualité.

Les mises à jour parviennent directement sur l'intranet du client. Il est également possible d'acheter une licence d'utilisation de Factiva Intelligent Indexing pour développer des applications propres.

L'architecture d'indexation est suffisamment souple et solide pour servir de taxonomie de référence pour des produits traitant l'actualité et l'information économique.

Très peu d'intervenants dans le domaine de l'information sont à même de rivaliser avec l'expérience et le savoir-faire de Factiva en matière de traitement et d'indexation de contenu multilingue, multimédia et multi-fournisseurs.

Le modèle retenu par Factiva, à savoir l'intégration dans une bibliothèque unique d'une grande quantité et d'une grande diversité de sources de contenu, pouvant faire l'objet de recherches à l'aide d'un langage d'indexation puissant et unique, préfigure les attentes des entreprises en matière d'intranet et de portails.

Factiva propose également des licences d'utilisation sur certains outils destinés à aider les utilisateurs dans leurs recherches d'informations internes.

3.2.5 Rappel historique sur l'indexation

Afin de mieux évaluer l'importance cruciale de l'indexation dans la recherche d'informations à partir des intranet et portails d'entreprises, il est utile de faire un bref rappel historique sur l'indexation.

Depuis longtemps, les utilisateurs se servent d'index pour trouver rapidement les informations dont ils ont besoin. Nous connaissons tous au moins un modèle d'index, celui qui figure généralement à la fin d'un livre.

Cet index permet au lecteur de trouver les références des sujets couverts par l'ouvrage, et comprend généralement une liste alphabétique de sujets ou de noms, suivis du numéro de la page où ils apparaissent. L'accroissement de la quantité d'informations publiées a conduit à la création d'index spécialisés dans certaines disciplines, telles que la chimie, l'ingénierie, l'art et l'architecture ou l'économie.

L'objectif de tels index est d'aider l'utilisateur à trouver des informations à partir des multiples sources de documents imprimés, couvrant ces disciplines.

Le processus d'indexation par sujet consiste en une analyse conceptuelle du document, puis en une traduction de cette analyse en un vocabulaire particulier.

Chaque document se voit attribué certains termes appropriés, issus d'un vocabulaire normalisé, ces termes étant ensuite regroupés dans un index principal, permettant d'accélérer la recherche documentaire et d'aider l'utilisateur à obtenir les informations pertinentes.

L'utilité d'un index par sujet dépendra de plusieurs facteurs :

- la capacité du documentaliste à analyser le sujet traité par le document ;
- sa connaissance de la discipline ;
- sa connaissance des besoins de l'utilisateur pour ce type d'information.

Les premiers index imprimés ont été essentiellement créés pour accompagner des listes de citations ou de résumés de publications.

La quantité de texte pouvant faire l'objet de recherches étant limitée, l'attribution de termes d'indexation appropriés s'avérait cruciale pour diriger l'utilisateur vers les documents pertinents.

Or, force est de constater que l'attribution de termes d'indexation appropriés demeure encore aujourd'hui une préoccupation centrale, et une étape indispensable pour aider l'utilisateur à trouver précisément ce qu'il cherche, dans un volume croissant de documents et de texte intégral.

Les méthodes d'indexation se sont multipliées parallèlement à l'augmentation de la quantité d'informations disponibles électroniquement, et à la popularité croissante des bases de données bibliographiques en ligne.

Le recours à l'intelligence humaine dans le processus d'analyse et d'organisation de l'information a été complété par des programmes d'indexation automatique.

Aujourd'hui, face à la prolifération des ressources internet, l'indexation informatisée à l'aide de balises méta, de vocabulaires normalisés et de sujets-clefs est devenue plus perfectionnée et plus évolutive.

Toutefois, les recherches documentaires les plus efficaces sont généralement obtenues à partir de bibliothèques dont le contenu a bénéficié d'une intervention humaine.

Celle-ci se situe au niveau de l'analyse des concepts et de l'élaboration des règles, mais comprend également le suivi de la précision et de l'exhaustivité du vocabulaire normalisé.

3.2.6 Vocabulaires normalisés ou taxonomies

Comme évoqué précédemment, le président et fondateur de Delphi Group a été cité pour sa déclaration : "Taxonomies are chic".

Mais de quoi s'agit-il précisément ?

Un vocabulaire normalisé est un langage d'indexation, c'est-à-dire un ensemble de termes et de locutions standardisés qu'il est permis d'utiliser dans un système d'indexation pour décrire une famille de sujets ou un domaine d'information.

Les termes vocabulaire normalisé, thésaurus, et schéma de classification sont interchangeables.

Le schéma d'indexation peut prendre le nom de taxonomie s'il s'appuie sur une structure hiérarchisée.

Un vocabulaire normalisé peut être aussi simple qu'une liste alphabétique des termes propres à la description d'une famille de sujet.

De façon générale, les thésaurus sont des ensembles de termes, élaborés et construits de façon très méticuleuse, les termes étant connectés entre eux par des liens du type "élargir le sujet", "affiner le sujet", "sujet associé" ou par des références croisées.

Ces liens illustrent la relation existant entre les termes associés et établissent une structure hiérarchique, ou taxonomie, qui permet d'effectuer des recherches à un niveau plus ou moins général. Les exemples suivants constituent de véritables références en matière de thésaurus et de structure taxonomique :

- INSPEC Thesaurus (IEE Publishing & Information Services) ;

- Medical Subject Headings (MeSH) (National Library of Medicine) ;
- Proquest Controlled Vocabulary¹.

Dans le passé, les documentalistes attribuaient manuellement aux documents les termes du vocabulaire normalisé.

L'introduction de programmes informatiques d'indexation s'est accompagnée de l'élaboration d'ensembles de règles, permettant d'automatiser du moins en partie ce travail d'indexation.

Le nombre de termes d'indexation utilisés et leur degré de spécificité dépendent des directives éditoriales fixées par l'éditeur.

La personne effectuant une recherche peut se référer au même thésaurus que celui utilisé pour l'attribution des termes d'indexation, et choisir les termes susceptibles de produire les résultats les plus pertinents à partir de la base de données ou du système d'information retenu.

Même si le chercheur n'utilise pas sciemment le thésaurus, la précision des résultats de sa recherche se trouvera améliorée si dans les termes utilisés pour sa recherche figurent certains termes d'indexation ou du vocabulaire normalisé.

Dans ce cas de figure, les résultats de sa recherche incluront les concepts associés aux mots, que ces derniers apparaissent ou non dans le texte.

Le chercheur aura également la possibilité d'examiner les termes indexés aux documents récupérés, et de noter les plus pertinents d'entre eux pour de futures recherches répétitives.

L'un des objectifs importants d'un langage d'indexation est de contrôler les synonymes.

Dans un système de recherche de type booléen, les utilisateurs saisissent des mots ou des mots clés liés par l'opérateur "ou", afin d'inclure les différentes possibilités qu'a un auteur de décrire un concept.

Or, une équation de recherche formulée à l'aide d'un terme d'indexation appartenant à un vocabulaire normalisé suffisamment développé permet de couvrir les synonymes.

En effet, à chaque terme du vocabulaire normalisé correspond un concept particulier, quelle que soit la forme sous laquelle il est exprimé dans l'article original.

Face au volume et à la croissance ininterrompue de texte intégral accessible par électronique, l'utilisateur a besoin de toute l'aide possible pour s'y retrouver.

¹ <http://www.proquest.com/hp/Support>.

Ainsi, un contenu indexé à l'aide d'un vocabulaire normalisé perfectionné améliorera de façon spectaculaire l'efficacité de sa recherche.

Un vocabulaire normalisé, créé et mis à jour par des documentalistes connaissant le sujet et les types de document couverts dans le domaine d'information en question, sera dynamique et évolutif.

3.2.6.1 Champs et informations indexées

Dans la plupart des outils de recherche documentaire, chaque document est divisé en sections (ou champs) spécifiques, qui permettront à l'utilisateur d'obtenir des résultats plus précis.

Différents champs seront identifiés en fonction des types de document inclus dans la bibliothèque.

Ainsi, pour l'actualité et les articles économiques, les champs suivants sont généralement identifiés : titre, auteur (ou signataire), nom de la publication, date de la publication, nom de la société, ticker, etc.

S'il s'agit de profils d'entreprise, les champs incluront généralement le nom de la société, le ticker, les codes SIC (ou NAICS) de classification sectorielle, la ville, l'état, le code postal, le nom du principal dirigeant, le chiffre d'affaires et l'effectif total.

Les informations contenues dans ces champs viendront également enrichir certains index spécifiques, tels que celui des tickers ou des codes postaux, qui constituent autant de points d'accès supplémentaires pour les utilisateurs.

3.2.6.2 La référence en matière d'indexation hiérarchique

Si l'on cherche des références dans le domaine de l'indexation, le Medical Subject Headings (MeSH) vient immédiatement à l'esprit.

Il s'agit d'un schéma de classification hiérarchique comprenant 19 000 sujets-clefs et codes principaux, utilisés pour l'indexation des bases de données créées par la bibliothèque nationale de médecine (National Library of Medicine).

La base de données Medline sert de référence dans le domaine biomédical, et représente l'équivalent électronique des Medicus Index, International Nursing Index et International Dental Literature.

L'une des caractéristiques essentielles de cette base de données est la possibilité d'accéder à l'indexation MeSH. Chaque article se voit attribué entre 6 et 15 sujets-clefs, avec un maximum de 3 pour les thèmes clés de l'article. Les articles sont indexés suivant le terme le plus spécifique, afin de garantir un très haut degré de précision lors d'une recherche par sujet.

Les mots-matières, c'est-à-dire les termes couvrant les aspects fréquemment mentionnés d'un sujet (par exemple, effets secondaires ou thérapie) sont ajoutés aux termes MeSH pour identifier clairement le thème principal de l'article.

L'une des fonctionnalités particulièrement puissantes de Medline est la possibilité pour les utilisateurs "d'éclater" une catégorie de termes selon un classement hiérarchique, allant du plus général au plus spécifique.

Cette fonction permet de récupérer l'ensemble des articles traitant du terme général, mais également des termes spécifiques apparaissant sous celui-ci.

Cette fonction est distincte de la fonction "troncature", puisque les termes recherchés ne commencent pas obligatoirement par la même racine.

En "éclatant" un terme, l'utilisateur peut effectuer une recherche sur le terme en question ainsi que sur des termes associés plus spécifiques. Les Medical Subject Headings sont révisés et actualisés en permanence par des spécialistes, maîtrisant chacun un domaine médical spécifique.

Leur travail consiste à rassembler les termes nouveaux au fur et à mesure de leur parution dans des ouvrages scientifiques ou de leur utilisation dans de nouveaux domaines de la recherche, à définir ces termes dans le cadre du vocabulaire normalisé existant, et à recommander leur inclusion dans MeSH.

Ils reçoivent également des suggestions de la part de documentalistes ou d'autres professionnels.

Le temps a prouvé la qualité de l'architecture d'indexation MeSH, laquelle est aujourd'hui reconnue mondialement pour la justesse et la précision de ses résultats.

MeSH apparaît indiscutablement comme le modèle en matière d'indexation, et devrait servir de référence pour évaluer les structures d'indexation utilisées dans d'autres disciplines.

3.2.6.3 Recherche : précision et rappel

La littérature sur les sciences de l'information, et en particulier sur la recherche documentaire via des services électroniques, mentionne fréquemment les notions de précision et de rappel.

La précision se rapporte au nombre de documents pertinents récupérés, tandis que le rappel se réfère au nombre total de documents récupérés.

Ces deux notions sont généralement inversement proportionnelles : plus le nombre total de documents récupérés est élevé, moins la précision (ou le degré de pertinence) sera élevée.

La précision et le rappel sont également influencés par la qualité de la méthode d'indexation et des directives éditoriales définissant le niveau de spécificité qui présidera à l'indexation d'un article.

Si la plupart des index intègre un très haut degré de spécificité, certains sont néanmoins construits autour de concepts très généraux.

Les attentes en matière de rappel varient également selon les utilisateurs.

Ainsi, un responsable des relations publiques voudra absolument récupérer la totalité des documents où figure le nom de son entreprise.

Mais de façon générale, l'utilisateur privilégie la précision, c'est-à-dire un nombre de documents plus restreint mais plus pertinent.

Idéalement, un système d'information devrait être conçu de manière à permettre à l'utilisateur d'affiner facilement sa recherche.

Les schémas d'indexation hiérarchique offrent précisément cette souplesse. L'utilisateur veut être certain que les termes de sa recherche permettront de récupérer l'ensemble des articles traitant d'un sujet donné, et que parallèlement, l'outil de recherche sera suffisamment puissant pour cibler uniquement les articles contenant tous les concepts demandés.

3.2.7 Factiva Intelligent Indexing

La pratique de l'indexation n'est pas nouvelle, comme en témoigne l'histoire de l'indexation des documents imprimés et plus récemment, des documents électroniques.

Mais Factiva a décidé de s'intéresser à nouveau au dossier des techniques d'indexation, dans l'objectif d'améliorer la précision des résultats et de faire face à la surabondance d'information.

Factiva Intelligent Indexing rassemble le meilleur de la tradition et du savoir-faire de Dow Jones Interactive et de Reuters Business Briefing en matière d'indexation.

Factiva Intelligent Indexing combine également les procédés d'indexation automatique les plus en pointe avec l'intervention humaine pour l'analyse et la classification des articles par sujets traités.

Les techniques d'indexation de Dow Jones Interactive ont évolué dans le temps, de l'indexation manuelle de contenu exclusif aux techniques perfectionnées d'indexation automatique de sociétés et de sujets à l'aide de règles, pour répondre à la demande de sa clientèle.

De son côté, Reuters est connu depuis longtemps pour sa méthode d'indexation de contenu à l'aide d'un vocabulaire normalisé, complétée d'une indexation alpha-numérique ou codes hiérarchiques.

Le schéma d'indexation utilisé pour le contenu de Reuters Business Briefing s'appuie sur la méthode d'indexation hiérarchique de la base de données Finsbury Data Services' Textline (Reuters a acheté Finsbury en 1986).

Début 2000, Factiva a élargi et modifié la hiérarchie d'indexation de Reuters Business Briefing pour élaborer la nouvelle hiérarchie d'indexation de Factiva Intelligent Indexing, qui comprenait initialement plus de 650 sujets industries, 280 sujets d'actualité et 370 références géographiques.

La classification sectorielle s'appuie dans les grandes lignes sur le système de classification sectorielle nord-américain (NAICS – North American Industry Classification System), tandis que les codes industries de Factiva correspondent aux codes NAICS ainsi qu'à d'autres normes reconnues.

Il existe cinq niveaux hiérarchiques dans les codes industries, permettant aux utilisateurs de lancer des recherches à un niveau très général ou très spécifique.

Exemple : 13 Ingénierie et fabrication de métaux 135 Véhicules motorisés et pièces détachées 1351 Constructeurs automobiles 135102 Véhicules de tourisme 13510201 Constructeurs de véhicules utilitaires légers

Les documentalistes de Factiva reconnaissent que certains schémas traditionnels de codes hiérarchiques ne sont pas parfaitement adaptés aux secteurs d'activité ou technologies émergents.

C'est pourquoi, l'architecture d'indexation de Factiva Intelligent Indexing pour les secteurs d'activité peut s'enrichir facilement et intégrer ces évolutions.

Les codes pays ou régions suivent la classification ISO.

Les codes sujets d'actualité sont organisés suivant une hiérarchie à quatre niveaux, et sont classés dans les groupes suivants : sociétés, économie, marchés, général et politique, organismes internationaux politiques et économiques, et types de contenu.

Les codes sociétés de Factiva permettent d'identifier environ 300 000 sociétés, constituant ainsi un vaste index des sociétés mondiales, cotées et non cotées.

Les filiales non cotées se voient attribué un code unique, et sont incluses dans l'index des sociétés si la demande des utilisateurs le justifie.

Plutôt que de saisir le nom de la société en texte libre, par exemple Computer Peripherals, Inc., l'utilisateur peut formuler sa recherche à l'aide du code société.

Il s'agit là d'une illustration de l'intérêt frappant et immédiat de l'utilisation de l'index des sociétés pour la précision des résultats.

Cette nouvelle organisation hiérarchique des codes, plus en pointe, servira à indexer tous les produits Factiva.

Il est également prévu que ces informations hiérarchiques figurent à l'écran, de façon à permettre aux utilisateurs d'affiner leurs recherches. Les chercheurs expérimentés pourront ainsi intégrer à leurs équations de recherche les codes Factiva et les termes d'indexation.

En résumé, Factiva Intelligent Indexing est un langage de recherche commun, qui s'applique :

- à un vaste répertoire de sociétés mondiales ;
- à un contenu publié en plus de 20 langues ;
- à un contenu provenant de fournisseurs très divers ;
- à des sujets décrits de façon cohérente (société, secteur d'activité, géographie, actualité).

Ainsi qu'évoqué au paragraphe "Nouvelles applications de l'information indexée", les clients de Factiva ont d'ores et déjà la faculté de s'appuyer sur l'architecture d'indexation perfectionnée de Factiva pour concevoir et personnaliser les solutions d'information mises à la disposition de leurs utilisateurs.

Ils bénéficient en outre de services de Dow Factiva dont la capacité de récupération d'information et la précision sont améliorées.

3.3 MEILLEURES PRATIQUES

L'élaboration et la gestion de fiches contenant les meilleures pratiques, les check-lists, les bases de données de références, les manuels qualité, les actions à entreprendre sont des exemples de gestion des connaissances qui contribuent à réduire les erreurs et la redondance des tâches.

La création d'une base des meilleures pratiques nécessite :

- d'une part, de repérer les meilleures pratiques ;
- et, d'autre part, de les diffuser au sein de l'entreprise.

Il s'agit à ce niveau, de recenser et de "manager" par la valeur.

L'échec étant également source de connaissances, il paraît utile de diffuser en outre des informations, facilement accessibles, par exemple sous forme de fiches, concernant les pièges à éviter (les pratiques à bannir et à proscrire).

De plus, s'agissant des meilleures pratiques, la difficulté se situe au niveau de la restitution et non de l'extraction des connaissances.

De même, l'accumulation de connaissances nécessite une sélection. Cet aspect est un élément déterminant.

Enfin, le principe de "l'effet paillason" qui consiste à confondre la propriété d'un objet avec celle qui en est proche¹ est à prendre en compte : en matière de gestion des connaissances, plus on est compétent, plus on perçoit les erreurs introduites dans l'espace de connaissance. A l'inverse, plus on s'écarte de son domaine de compétence, plus "l'effet paillason" est à craindre. Le système devrait permettre d'éviter cet effet.

3.4 TECHNOLOGIE "OBJET"

Les langages "objet" permettent d'appréhender le même objet sous plusieurs angles.

Ces technologies sont parfois utilisées dans les programmes de gestion des connaissances.

Dans les technologies "objet", il y a deux principes forts :

- le principe de polymorphisme ;
- le principe de l'héritage.

¹ Comme le fait tout paillason demandant d'"essuyer vos pieds" alors qu'il ne vient à l'idée de personne d'enlever chaussures et chaussettes pour obtempérer, chacun comprenant bien que ce sont les semelles des chaussures qui doivent être essuyées.

3.4.1 Définition du polymorphisme

Le polymorphisme peut être défini comme l'objet qui peut prendre différentes formes au cours de sa vie. Ainsi, un objet A peut se transformer en objet B.

Dans les langages de programmation orientée objet (POO), le polymorphisme permet d'appliquer à un objet une méthode qui restera applicable quelle que soit la transformation de l'objet.

L'avantage de ce concept est d'accroître la pérennité du code d'écriture du système d'information en réduisant à l'extrême les opérations de maintenance évolutives et ce, même si un objet A se transforme en un objet B.

Ce concept est donc au cœur de la gestion des connaissances dont le système d'information est un des éléments, avec la dimension organisationnelle qui infère la gestion et la dimension humaine qui permet la capture des connaissances.

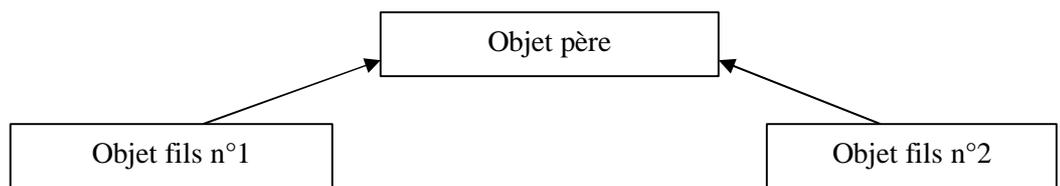
Le polymorphisme est une extension de la notion d'héritage : on utilise l'héritage afin d'appliquer le polymorphisme.

3.4.2 Définition de l'héritage

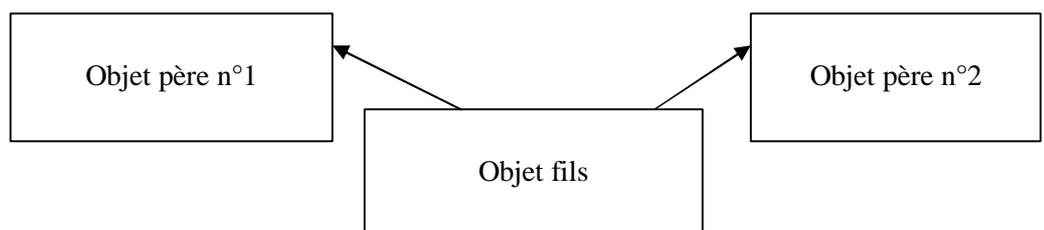
L'héritage consiste à permettre aux objets d'un niveau inférieur d'hériter des comportements et/ou propriétés des objets d'un niveau supérieur.

L'héritage s'organise de façon hiérarchique, où un objet "père" donne à ses objets "fils" des propriétés propres à son type d'objet.

L'héritage est dit "simple" lorsque l'objet "fils" hérite d'un seul objet "père" au même niveau de filiation.



L'héritage est "multiple" lorsque l'objet "fils" hérite pour un même niveau de filiation de plusieurs objets "pères".



L'avantage de ce concept est de réduire considérablement la réécriture de code en héritant de l'implémentation des comportements ou propriétés des objets supérieurs, en partant du plus général au plus particulier (grand ensemble vers sous-ensembles).

3.4.3 Encapsulation et abstraction

Dans la conception objet, l'utilisateur a une vue restreinte des objets qu'on lui présente.

Il n'accède qu'à des portes qui lui donnent la possibilité de modifier un comportement dans l'objet, sans en connaître les modes d'implémentations (le "comment ?").

Ces portes se nomment des "interfaces" ; elles dissimulent le contenu de l'objet. Dans le même ordre d'idées, peu importe de savoir comment fonctionne une voiture, en termes de moteur. L'essentiel est d'avoir le moyen de rouler (clé de contact, volant...).

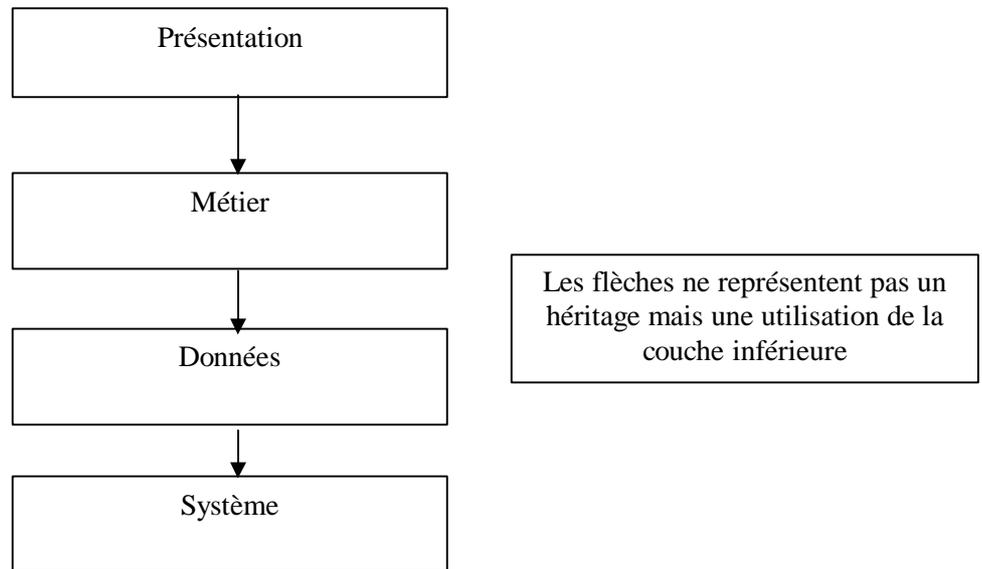
L'utilisateur n'a pas accès à l'implémentation ; on lui offre les moyens de manipuler l'objet et on lui garantit que les méthodes seront toujours les mêmes et ce, même si l'objet évolue, ou si le cœur change. L'utilisateur aura toujours un volant pour sa voiture, même s'il change le moteur.

Au niveau de l'abstraction, la conception objet va bien plus loin dans le principe.

En plus d'une abstraction des propriétés et de la façon dont sont conçus les comportements, on peut affiner et encapsuler par sous-ensembles complets : un sous-ensemble contient un agrégat d'objets. Le tout rassemblé offre un service complet.

Aussi, dans la conception objet, on répartit les grandes fonctionnalités par couches, qui sont hiérarchiques.

Chaque niveau ne peut interroger que la couche située en-dessous : par exemple, un système pourrait contenir un ensemble de sous-systèmes élaborés en couche :



Cette conception isole chaque grande fonction du système qui a pour objectif une meilleure stabilité et une possibilité d'évolution : une évolution d'une couche n'impacte pas forcément une autre couche.

La communication entre les couches s'effectue par des interfaces, objets qui offrent les méthodes d'accès à la couche inférieure et supérieure. De plus, le cœur de la couche pourrait changer, mais son accès rester identique.

Par exemple, pour un moteur de recherche, l'interface aura des méthodes de type : `rechercher_doc()`, `indexer_doc()`, `supprimer_doc()`.

Il incombe aux méthodes d'appeler les bons objets de leur couche, tout cela reste invisible à l'utilisateur qui utilise le moteur par l'intermédiaire de l'interface. Il peut arriver dans le cycle de vie de l'application, qu'on veuille changer de moteur de recherche, jugé peu pertinent. L'utilisateur accèdera toujours à la même interface, bien que la couche ait changé de moteur.

3.4.4 Application des concepts orientés objet à la gestion des connaissances

Dans la gestion des connaissances, l'un des points clés est de permettre aux néophytes comme aux spécialistes, d'accéder et/ou de restituer le niveau de connaissance requis par rapport à un contexte particulier.

Comment les différents concepts décrits précédemment peuvent-ils permettre d'atteindre ces objectifs d'accès et/ou de restitution ?

La notion de polymorphisme permet de gérer les homonymies sur des concepts identiques en termes de signifiants et de résultats, même s'ils ne sont que voisins en termes de moyens et s'ils sont différents au niveau du détail de l'action.

Ainsi, pour les néophytes, on s'attachera d'abord au plus haut niveau d'abstraction en se concentrant sur le signifiant et sur le résultat, alors que pour les spécialistes, on descendra jusqu'au niveau de détail le plus fin.

Par exemple :

a) Le signifiant "Avancer" est identique au plus haut niveau d'abstraction : on avance aussi bien sur terre que dans l'eau (ou dans l'air).

b) Le résultat "Avancer" est identique au plus haut niveau d'abstraction : on se déplace d'un point A à un point B en ayant la tête en avant (sinon on reculerait).

c) Le moyen "Avancer" est voisin : on bouge ses membres.

Or, le détail de l'action "Avancer" est différent : il se traduit par des mots distincts qui sont sauter, marcher (ou voler).

Pour sa part, la notion d'héritage simple ou multiple favorise les raisonnements inductifs ou déductifs selon que l'on parle du niveau de concept le plus ou le moins élevé.

La notion d'encapsulation permet de masquer la complexité d'un domaine, lorsque le fonctionnement détaillé du domaine n'intéresse que le spécialiste.

La notion de niveau d'abstraction offre la possibilité de constituer des systèmes d'information à la fois autonomes dans leur fonctionnement et interopérables entre eux.

3.5 PROPRIETE DES CONTENUS

Le processus de gestion des connaissances implique la réalisation de quatre actions, à savoir :

- capter les connaissances ;
- structurer les connaissances ;
- diffuser les connaissances ;
- mettre à jour les connaissances.

La première action consiste à apporter à l'entreprise un maximum d'informations pour les transformer ultérieurement en connaissances.

Ces informations peuvent provenir de diverses sources, tant internes qu'externes à l'entreprise.

Ces sources d'informations sont, par exemple, l'accumulation des pratiques internes des collaborateurs, l'analyse de documents publiés par des concurrents, l'exploitation de la presse notamment professionnelle.

La source permet l'identification de l'émission, la datation des informations et le recueil des données propres à l'information (origine, information primaire ou transformée, évolution de l'information...) indépendamment de son contenu.

Ainsi, plus la source de l'information est renseignée, plus l'information peut être captée avec sécurité et plus la base de connaissances sera pérenne et fiable.

La seconde action consiste à formaliser aussi bien des idées, intuitions, pratiques, expériences, savoir-faire (connaissances tacites) que des documents, normes, règles, procédures (connaissances explicites) pour permettre leur diffusion par le biais, notamment, de médias électroniques tels que web, intranet, internet, CD-Rom et e-learning.

Les contenus ainsi diffusés pourront faire l'objet d'une protection par le droit d'auteur notamment.

Le droit d'auteur réserve en effet sa protection à toutes les créations intellectuelles originales, c'est-à-dire celles qui portent l'empreinte de la personnalité de leur auteur et ce, sans considération de genre, de forme d'expression, de mérite ou de destination.

L'article L.112-2 du Code de la propriété intellectuelle dresse une liste des œuvres de l'esprit éligibles à la protection par le droit d'auteur (œuvres graphiques, cinématographiques, plastiques...) sans que cette énumération soit limitative.

Ainsi, tout auteur d'une œuvre de l'esprit satisfaisant à la condition d'originalité bénéficie de la protection accordée par le droit d'auteur, laquelle lui permet d'interdire toute reproduction, totale ou partielle, de son œuvre sans son accord.

La reproduction totale ou partielle d'une œuvre originale, sans l'accord de l'auteur, est constitutive d'une contrefaçon, laquelle s'apprécie en fonction des points de ressemblance.

Toutefois, si le droit d'auteur protège toute création originale dans sa forme, sont expressément exclues du champ de la protection par le droit d'auteur, les idées qui demeurent libres et ne peuvent faire l'objet d'une protection privative¹.

Ainsi, la Cour de cassation a jugé que le Code de la propriété intellectuelle ne protège pas les idées exprimées mais seulement la forme originale dans laquelle elles le sont.

L'application concrète de ce principe est que toute méthode ou concept abstrait sera considéré comme relevant du domaine des idées et, par conséquent, échappera à toute protection par le droit d'auteur dès lors que cette méthode n'aura pas atteint un niveau suffisant de formalisation.

Ainsi le contenu diffusé peut être considéré comme une création intellectuelle.

Toutefois en matière de logiciel, si cette création est réalisée par un salarié dans le cadre de l'exécution de son contrat de travail ou dans le cadre de l'exercice de ses fonctions, le droit accorde une dévolution automatique des droits de propriété à l'employeur.

Pour les autres créations intellectuelles, l'employeur qui entend exercer ses droits patrimoniaux sur les créations salariées, devra conclure un contrat de cession de droits avec les auteurs salariés, conformément aux dispositions du Code de la propriété intellectuelle.

A défaut, les salariés sont considérés comme des auteurs et, en tant que tels, leurs œuvres sont protégées par le Code de la propriété littéraire et artistique.

De même, si le contenu diffusé a une source externe et constitue une œuvre de l'esprit, son auteur bénéficiera de la protection accordée au titre du droit d'auteur.

La diffusion de ce contenu dans l'entreprise nécessitera une cession des droits de propriété intellectuelle constatée par écrit et définissant précisément l'étendue, la destination, le lieu, la durée et la nature des droits cédés².

En outre, des créations non protégées par le Code de la propriété intellectuelle ne sont pas pour autant "réutilisables" librement.

¹ Une réflexion sur l'appropriation privative des connaissances à travers le concept d'idées se développe aujourd'hui.

² Article L.131-3 du Code de la propriété intellectuelle : "La transmission des droits de l'auteur est subordonnée à la condition que chacun des droits cédés fasse l'objet d'une mention distincte dans l'acte de cession et que le domaine d'exploitation des droits cédés soit délimité quant à son étendue et à sa destination, quant au lieu et quant à la durée".

Ainsi, constitue un agissement de concurrence déloyale ou parasitaire le fait d'utiliser le travail ou la production d'un tiers sans réaliser ses propres investissements et en bénéficiant ainsi des investissements de ce tiers.

3.6 EXTRACTION DES CONNAISSANCES

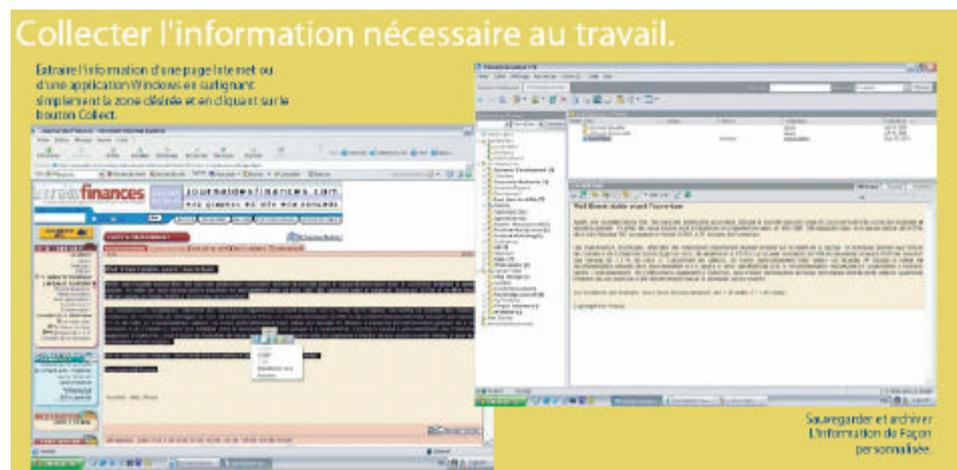
Les entreprises ont besoin de méthodes et d'outils capables d'extraire les connaissances pertinentes, de les stocker et de les diffuser.

A cet effet de nombreux projets de recherche se développent. Des extracteurs de connaissances commencent même à apparaître sur le marché. On citera pour exemple le produit Quantum de la société Entopia.

Entopia commercialise depuis maintenant deux ans Quantum, une application client-serveur capable de capturer, d'indexer à la volée et de constituer dynamiquement une base de connaissances partagée et ce, sans perturber l'utilisateur dans son travail quotidien.

En effet, l'interaction requise est minimale puisqu'il suffit à l'utilisateur d'activer d'un simple clic le module "collecter" de Quantum à partir de la barre de tâches de sa station de travail.

Ensuite, comme pour un copier-coller, l'utilisateur "collecte" tout ce qu'il considère utile et nécessaire à son travail qu'il s'agisse d'une page web ou d'une partie d'une page web, d'un fichier ou d'une partie d'un fichier et ce, quel que soit le format du fichier : PDF, XML, Microsoft Word, PowerPoint, Excel, Access, etc.



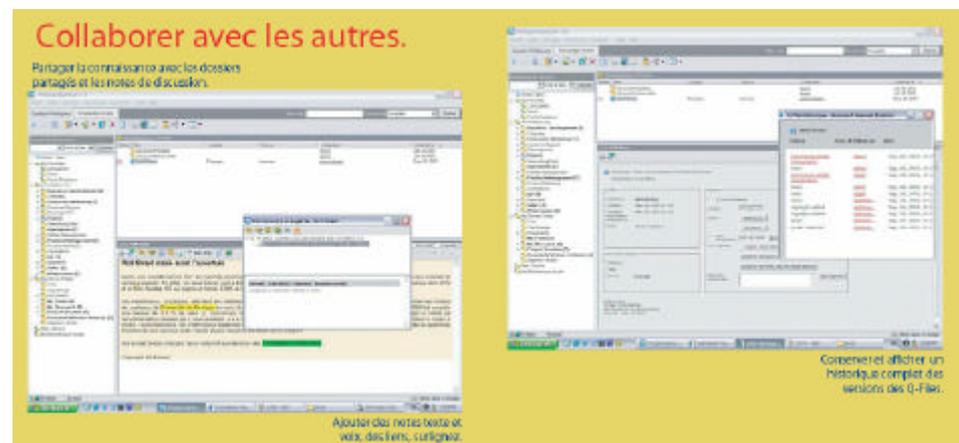
Les données ainsi collectées sont analysées, résumées et indexées par le moteur sémantique de Quantum. Les données disposent alors d'une signature sémantique qui est la matérialisation de la connaissance extraite.

Cette signature alimente ensuite une base de connaissances en temps réel. A tout moment, les données collectées peuvent être enrichies de commentaires, surlignages ou méta-données détaillées.

L'extracteur de connaissance balaye à nouveau les données pour mettre à jour à la base de connaissances.

A partir du module "collaborer" de Quantum, l'utilisateur peut partager ses données avec d'autres utilisateurs dans l'entreprise.

Au fur et à mesure que le partage se fait, l'enrichissement des données par les commentaires, les surlignages ou les méta-données détaillées est systématiquement analysé et géré par Quantum pour alimenter la base de connaissances.



Il est important de noter que des fonctionnalités de contrôle d'accès permettent aux propriétaires des données de gérer l'accès à leurs informations.

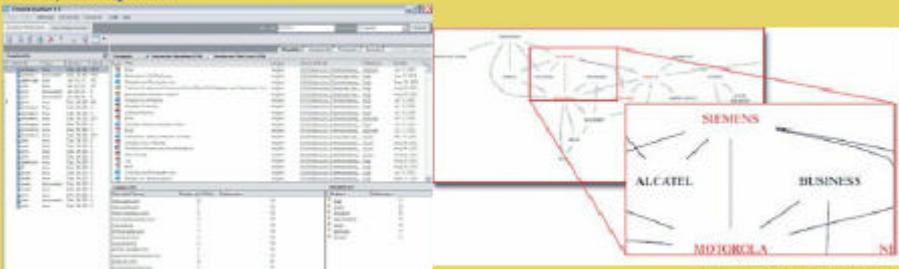
Cependant, et même si le fichier initialement collecté est désigné comme "privé", la signature sémantique du fichier est disponible et elle est associée aux experts appropriés.

Cette disponibilité de la signature et l'association aux experts appropriés est essentielle au module "capitaliser" de Quantum.

En effet, le module "capitaliser" de Quantum transforme les connaissances en capital intellectuel pour l'entreprise. Ainsi, à partir d'une simple requête en langage courant, Quantum est en mesure d'identifier et de classer par pertinence les ressources, les experts et tout contributeur qui pourraient posséder des informations sur un sujet donné.

Capitaliser sur la connaissance cachée dans l'entreprise.

Localiser l'information critique, experts, sources, utilisation des documents avec Ento (Patent Knowledge Location)



Utiliser le module E-Map pour créer une représentation graphique automatique des liens entre les différents concepts clés contenus dans les documents sélectionnés.

Query Management permet aux utilisateurs de faire une recherche dans une requête sauvegardée, de la modifier et de la mettre à jour.

3.7 PRODUITS ORIENTES METIERS

Les entreprises recherchent, parmi les outils de gestion de connaissances commercialisés sur le marché, des produits adaptés à leur cœur de métier et correspondant à leurs besoins.

Le produit Lingway Patent Suite est proposé par la société française Lingway spécialisée dans la conception, le développement et l'implémentation de logiciels basés sur la linguistique.

Lingway Patent Suite est une offre orientée Information Brevets qui comprend :

- des outils de sélection (le filtrage et la catégorisation), notamment pour réaliser automatiquement des "Patent Reviews" ;
- des outils de gestion d'une base de données de résumés et de textes de brevets propres à l'entreprise, comprenant des fonctions de recherche multilingue avancées ;
- des outils d'analyse et de structuration de cette base de données ;
- des outils d'aide à la lecture, visant à faciliter et accélérer la lecture d'un brevet ;
- des outils de synthèse et de suivi de la production de brevets.

La mise en œuvre de ces différentes fonctions est basée sur des moteurs linguistiques allant bien au-delà des techniques documentaires classiques. Plusieurs fonctionnalités de Lingway Patent Suite ont été réalisées en collaboration avec des offices nationaux de brevets, au premier rang desquels l'INPI avec lequel Lingway coopère depuis plusieurs années.

Le produit bénéficie également des travaux réalisés dans le cadre du projet ePatent, qui réunit les offices anglais (UKPO), espagnol (OEPM) et autrichien (OEPA) ainsi que la société Jouve et l'INPI dans un projet mené dans le cadre du programme eContent financé par la Commission européenne.

- les outils de sélection

Dans une entreprise, les départements chargés de la Propriété industrielle sont généralement abonnés à de nombreuses bases de données externes qu'ils interrogent régulièrement avec des profils documentaires sous forme de requêtes booléennes classiques.

Ce processus permet de télécharger localement les nouvelles notices et résumés sur les sujets intéressant la société.

Cette sélection initiale est généralement volontairement trop large, afin d'être certain de ne pas laisser échapper des brevets importants pour les diverses activités de l'entreprise.

Se pose donc un double problème, qui consiste, d'une part, à filtrer les notices pour ne retenir que celles intéressant effectivement la société et, d'autre part, à les catégoriser c'est-à-dire à les répartir correctement dans des catégories propres à l'entreprise, correspondant à ses activités, ses thèmes de recherche, voire son organisation.

Pour ce faire, il est nécessaire de s'adapter aux spécificités propres à chaque société, ce qui représente un coût non négligeable quand il est traité par des experts.

Lingway Patent Suite propose un outil permettant de filtrer, catégoriser et structurer ce flux de documents brevets entrant dans l'entreprise afin, d'une part, de stocker ces documents dans une base de données locale et, d'autre part, de produire un document de synthèse (une "Patent Review") pouvant être directement communiqué aux départements de Recherche de l'entreprise. L'outil permet :

- la prise en compte d'un plan de classement (ou plan de catégorisation, ou taxonomie) propre à la société ;
- l'apprentissage du module à partir d'un historique de notices déjà catégorisées ;
- le filtrage proprement dit, consistant à écarter les notices et résumés qui ne sont pas pertinents ;
- la catégorisation, c'est-à-dire la répartition des notices et résumés dans les différentes catégories propres à la société, qui sont généralement de l'ordre de quelques dizaines ;
- la détection, consistant à identifier les termes clés dans les documents, à partir d'une simple liste fournie par la société ;
- la mise en forme du résultat de ce traitement - qui peut être réalisé périodiquement, par exemple toutes les semaines - produit un document XML (ou HTML) prêt à être diffusé dans la société.

- les outils d'archivage et de recherche

Les documents téléchargés localement depuis les serveurs externes, résumés ou brevets complets, peuvent être archivés dans une base documentaire locale dotée de fonctionnalités de recherche avancées.

En effet, la recherche par la taxonomie devient rapidement insuffisante dès que le nombre de brevets archivés devient significatif et que l'on ressent le besoin de pouvoir faire des recherches plus ciblées, éventuellement en plusieurs langues.

Lingway Patent Suite propose un moteur documentaire conçu pour pouvoir être utilisé directement par les ingénieurs de recherche, sans avoir à utiliser un langage documentaire spécialisé (familier des documentalistes, mais généralement peu connu et mal utilisé par les ingénieurs de recherche). Cet outil permet :

- l'archivage et la gestion des documents entrant, ou la connexion au logiciel de gestion documentaire existant ;

- la recherche en langage naturel : le logiciel interprète la question en utilisant un dictionnaire, la complète automatiquement avec des termes synonymes ou voisins et la traduit dans le langage documentaire de recherche dans la base de données ;

- la mise au point interactive d'une question : l'utilisateur a la possibilité de réviser la liste de mots voisins et synonymes déterminés par le système ;

- la correction orthographique : les fautes de frappe ou d'orthographe sont corrigées ;

- la recherche multilingue : à partir d'une question posée en une langue, par exemple en français ou anglais, le système retrouve les brevets dans diverses langues (français, anglais, allemand, espagnol) ;

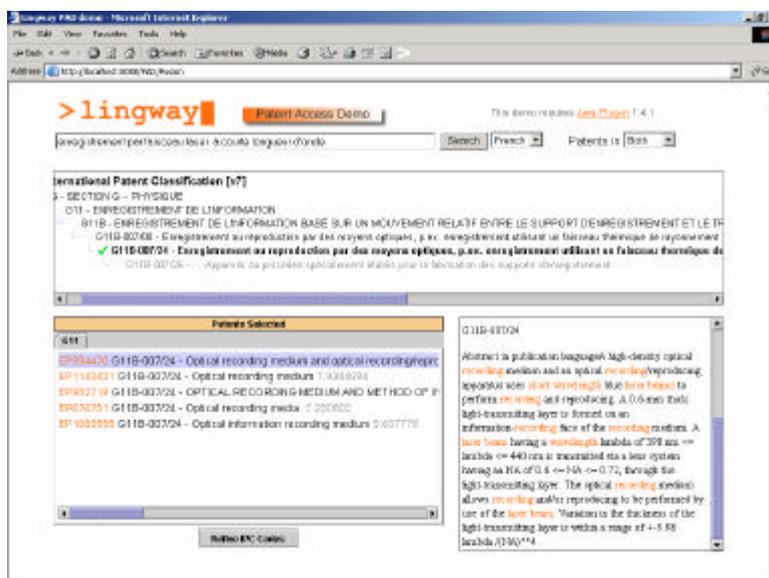
- la détection automatique de langue : le système repère automatiquement dans quelle langue est posée la question.

Pour les grandes bases de données, couvrant de nombreux domaines, il peut être utile d'utiliser la Classification internationale de brevets (CIB) lors de la recherche.

Dans ce cas, la question initiale en langage naturel est traduite en parallèle vers la CIB et vers les textes des résumés, et les résultats obtenus par ces deux recherches sont fusionnés en utilisant des heuristiques de classement particulières.

L'exemple ci-dessous illustre ces possibilités. On utilise ici l'interface combinant l'accès au texte des résumés et à la Classification internationale de brevets.

Une question est posée en français "enregistrement par faisceau laser à courte longueur d'onde" : on voit que le système s'est positionné sur une entrée correcte dans la CIB et a affiché un premier résumé, en anglais, qui répond à la question. On remarque notamment la reconnaissance des mots composés (longueur d'onde = wavelength).



- les outils d'analyse et de structuration

Les brevets sont des textes particulièrement difficiles à lire, notamment par ceux qui n'en n'ont pas l'habitude.

En effet, le langage est souvent un mélange de styles technique et juridique qui complique d'autant plus singulièrement la lecture que le rédacteur initial du brevet a souvent recours à de longs développements ne facilitant pas la compréhension rapide du brevet.

D'autre part, les brevets sont de plus en plus une source incontournable d'informations pour la veille scientifique et stratégique, et les informations que l'on peut tirer de l'analyse systématique des fonds de brevets sont un atout précieux pour le management de la Recherche dans les entreprises industrielles.

Ces besoins conduisent donc à vouloir structurer de plus en plus le texte des brevets pour pouvoir ensuite réaliser divers traitements automatiques sur une information structurée.

Même si la tendance, depuis plusieurs années, va dans le sens de structurer de plus en plus fortement les brevets, avec des DTD SGML puis XML de plus en plus complètes, cet effort de structuration ne touche pas, ou très peu, le contenu textuel lui-même.

Lingway Patent Suite propose donc des outils de structuration XML des textes de brevets permettant :

- s'il n'est pas déjà réalisé, le balisage XML de la structure générale du brevet :

- abstract ;
- references ;
- claims ;
- description ;
- background.

- le balisage XML d'un certain nombre d'entités :

- les personnes citées ;
- les organisations et sociétés citées ;
- les références à d'autres brevets ;
- les technologies clés pour l'entreprise.

- le balisage XML de certaines phrases clés :

- les phrases décrivant l'objet du brevet ;
- les phrases décrivant les insuffisances de l'état de l'art ;
- les phrases décrivant les avantages de l'invention.

Le résultat de ce traitement est un fichier XML balisé, qui comporte également un indice de confiance sur les balises posées.

Ce balisage se fait à partir de lexiques (par exemple, liste de sociétés, liste de technologies clés) et à partir de règles linguistiques qui repèrent automatiquement des entités qui ne sont pas encore connues dans le lexique.

- les outils d'aide à la lecture

Faciliter la lecture des brevets est un besoin souvent exprimé par les responsables de Propriété industrielle, qui constatent que les ingénieurs de recherche sont souvent peu motivés pour les lire.

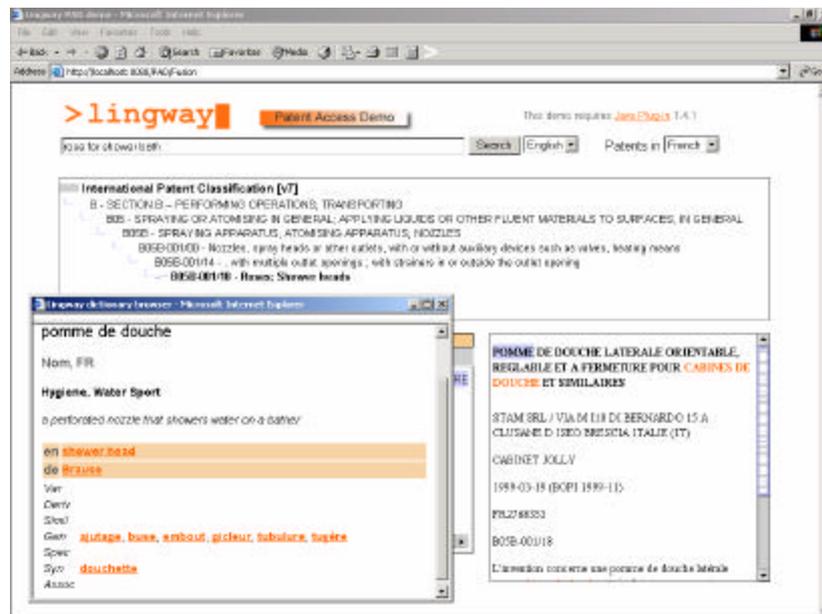
Lingway Patent Suite propose donc divers outils pour rendre cette lecture plus facile.

- la consultation interactive du dictionnaire multilingue. Cette fonction permet, en cliquant sur un mot, d'ouvrir une fenêtre sur le dictionnaire, afin d'obtenir :

- une définition du terme ;
- sa traduction dans différentes langues ;
- un ensemble de mots voisins, qui peuvent donner des idées pour de nouvelles recherches documentaires.

La copie d'écran ci-dessous illustre cette possibilité : après avoir posé la question "rose for shower bath", on trouve un résumé qui parle de "pomme de douche".

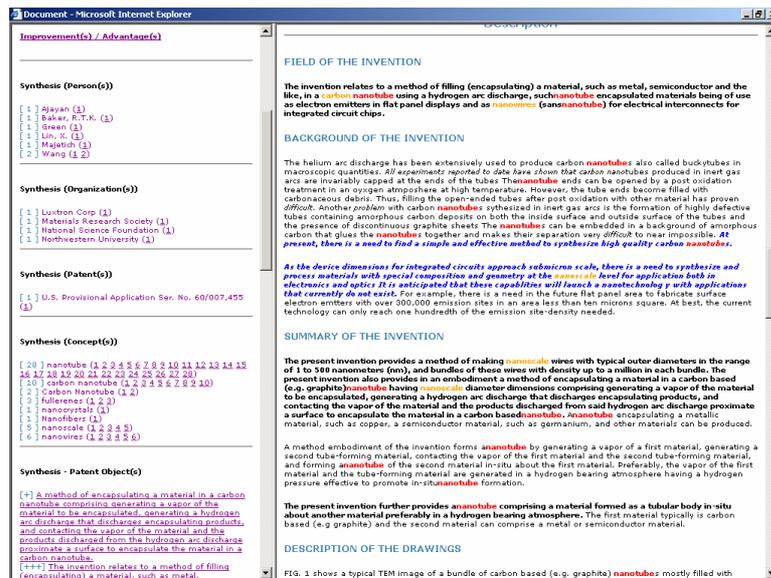
Un lecteur anglophone qui ne comprendrait pas ce sens du mot "pomme", clique dessus, ce qui ouvre la fenêtre correspondant au bon mot, où l'on voit la traduction "shower head".



- la lecture assistée. Cette fonction permet de naviguer facilement dans le texte du brevet, grâce à une présentation en deux colonnes indiquant à gauche les éléments identifiés (personnes, sociétés, technologies, etc.) avec les liens vers leurs occurrences dans le texte du brevet présenté dans la colonne de droite. Cette fonction, qui suppose d'avoir préalablement réalisé la structuration XML décrite plus haut, permet notamment :

- . la colorisation du texte, qui fait apparaître dans des couleurs différentes les diverses zones repérées lors du marquage XML. Cette fonctionnalité permet d'appréhender plus rapidement le contenu du texte, par exemple en différenciant rapidement les phrases décrivant l'objet du brevet, celles qui décrivent les insuffisances de l'état de l'art, et celles qui mettent l'accent sur les avantages de l'invention.

- . des fonctions de tri sur les entités nommées, permettent également de repérer rapidement certains points significatifs, par exemple en soulignant quels sont les technologies ou les auteurs les plus cités dans le brevet.



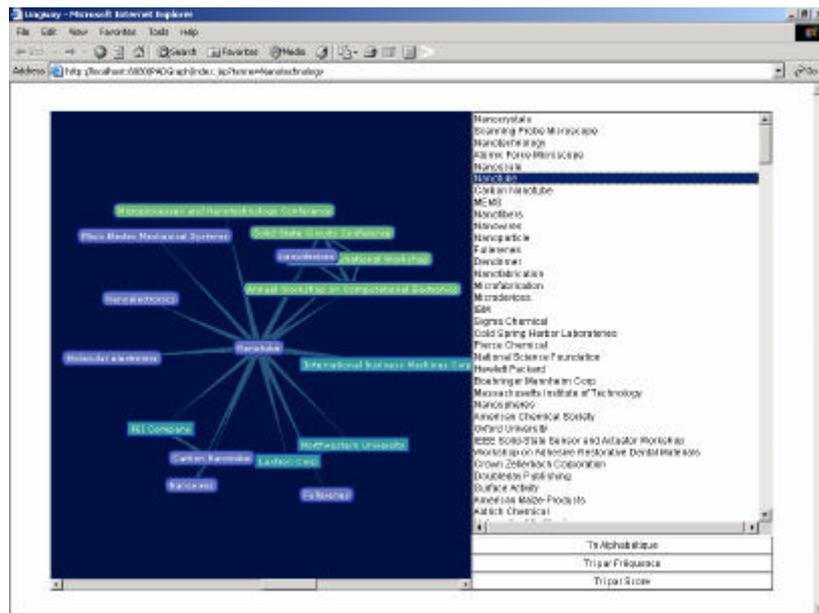
- . enfin, une fonction de "résumé", qui permet de construire des résumés ou des fiches de synthèse à partir de la structure XML, en sélectionnant, de manière paramétrable, les "n" phrases les plus importantes dans chaque catégorie, les "n" technologies les plus citées, etc. Depuis 2004, cette fonctionnalité est incluse dans l'offre de Questel.Orbit.

- les outils de synthèse et de visualisation

Les outils d'aide à la lecture présentés ci-dessus sont essentiellement destinés à faciliter la compréhension d'un ensemble limité de brevets, lus individuellement. Un autre besoin vise à obtenir une représentation globale d'un grand nombre de brevets, que l'on ne veut pas lire dans le détail, mais que l'on souhaite analyser pour repérer des tendances, des corrélations entre sociétés, personnes, technologies, etc.

Pour cela, Lingway Patent Suite propose un outil de visualisation permettant :

- le calcul des relations de proximité entre les entités identifiées lors de la structuration en XML des brevets ;
- l'affichage du graphe des relations, la nature des objets étant identifiée par une couleur (dans l'exemple ci-dessous, les technologies apparaissent en violet, les sociétés en bleu et les conférences en vert) ;
- une navigation aisée, soit à partir du graphe, soit à partir de la liste des entités extraites.



A titre également d'illustration des produits orientés métiers, Ciceron, filiale des Editions Lamy est éditeur de logiciels de gestion et prestataire de services pour les cabinets d'avocat et les services juridiques d'entreprises (plus de 10 000 utilisateurs en Europe du monoposte aux réseaux de plusieurs centaines de postes multisites).

Ciceron intervient dans la conception de logiciels et de solutions techniques, dans le conseil à la formation informatique auprès des professionnels du droit, indépendants ou au sein des entreprises, dans la mise en place des réseaux informatiques et l'assistance après-vente.

Cicéron commercialise une gamme de logiciels de gestion avec notamment :

- Cicéron Avocat (en mode Client/Serveur) : ce logiciel permet de rationaliser la gestion et le suivi des dossiers, de partager l'information en temps réel et d'optimiser le traitement workflow du cabinet ;
- Cicéron OnLine (en mode A.S.P.) : à destination des cabinets d'avocat de petite taille, Cicéron OnLine est un ASP (Application Service Provider) permettant de gérer son cabinet à distance et en toute sécurité : gestion des dossiers clients, consultation des agendas, gestion des temps passés, saisie de diligences, facturation, comptabilité générale, formulaires Lamy, recherche plein texte ...
- Cicéron Juris : outil de workflow, ce logiciel s'adresse aux services juridiques des entreprises. Il permet de gérer les dossiers en ce qui concerne la base de données des intervenants, les différents domaines et spécialités juridiques, leur suivi en termes d'informations historique, financière, statistique... Cicéron Juris peut être couplé à Cicéron GED, service de gestion électronique de documents et au GPJ, Guide des Professions Juridiques.

Cicéron propose également une gamme complète de services et des modules complémentaires tels que : assistance, maintenance, audit, formation, intranet/extranet, couplage téléphonique informatique, gestion électronique de documents, service de sauvegarde en ligne, intégration de formulaires.

Stetho est, quant à lui, un logiciel métier dédié aux acteurs des services de santé au travail en charge de la prévention des risques et au suivi médical des salariés de toutes les entreprises quel que soit leur forme ou statut juridique. Stetho assiste aujourd'hui quelques 3 000 intervenants dans leurs tâches quotidiennes de prévention que sont les études de poste de travail, les visites médicales, la mise en place des protocoles de surveillance médicale et toutes celles qui relèvent de la gestion de base (convocation, certificats divers, statistiques).

Stetho répond aux attentes réglementaires et associe dans son utilisation tous les acteurs d'un service de santé au travail, secrétaires médicales, intervenants pluridisciplinaires - technicien HSE et ergonomes - et médecins du travail.

Il intègre un fonds documentaire, le Lamyval, constitué d'extraits choisis des ouvrages des Éditions Lamy. Cette source d'informations permet aux acteurs concernés de disposer en temps réel de toutes les données réglementaires les plus fréquentes.

Stetho s'inscrit dans les architectures informatiques actuelles et répond aux contraintes de gestion centralisée ou distribuée des divers services de santé au travail.

3.8 GESTION SEMANTIQUE DES CONNAISSANCES

Des "moteurs" d'analyse sémantique sont proposés sur le marché.

Il s'agit notamment de moteurs s'appuyant sur un référentiel de concepts, sorte de thésaurus liant termes et expressions pour recontextualiser le vocabulaire dans le texte.

Ces outils favorisent l'accès aux connaissances à des non-professionnels du domaine¹.

Lingway KM est une suite d'outils dédiés à la gestion des connaissances. Cette suite est proposée par la société française Lingway spécialisée dans la conception, le développement et l'implémentation de logiciels basés sur la linguistique.

Lingway KM offre les principales fonctionnalités suivantes :

- crawler web ;
- filtres acceptant la plupart des formats courants ;
- moteur de recherche sémantique en langage naturel :
 - . ontologie standard de 100 000 concepts ;
 - . module de dialogue ;
 - . mise en évidence des hits avant la lecture des réponses ;
- correction orthographique des questions ;
- recherche cross-langage :
 - . français, anglais, allemand, espagnol, néerlandais et portugais en cours.
- recherche combinée full-text / taxonomie ;
- extraction d'entités nommées ;
- extraction de terminologie ;
- extraction de phrases importantes ;
- résumé automatique et aide à la lecture rapide ;
- clustering et cartographie ;
- consultation et gestion des dictionnaires.

¹ Exemples d'éditeurs : Albert, Arisem, Convera, Lexiquet, Sinequa, Temis, etc.

Lingway KM répond à une philosophie ouverte et modulaire. Ses principaux modules sont les suivants :

- moteur de recherche sémantique full text (Fulty) :

moteur de recherche sémantique full-text en langage naturel, incluant des possibilités de recherche multilingue. Fulty permet d'obtenir plus rapidement un nombre plus grand de documents pertinents pour la recherche.

- moteur de recherche via une taxonomie (Tacsy) :

module de recherche et catégorisation de textes (phrases, questions...) par rapport à une taxonomie. Ce module est particulièrement utile pour interroger de grandes taxonomies comme la Classification internationale des brevets, la nomenclature INSEE, etc., et pour coder des verbatims .

- outils d'aide à la lecture et la compréhension (Xtirp) :

module d'analyse, structuration XML et visualisation de textes, permettant le marquage de la structure générale d'un texte, l'extraction d'entités nommées (personnes, organisations, produits, lieux, etc.), et des phrases importantes sous un point de vue donné, pour en faciliter la lecture, créer des bases de données extraites, faire des synthèses automatiques et réaliser des cartographies dynamiques.

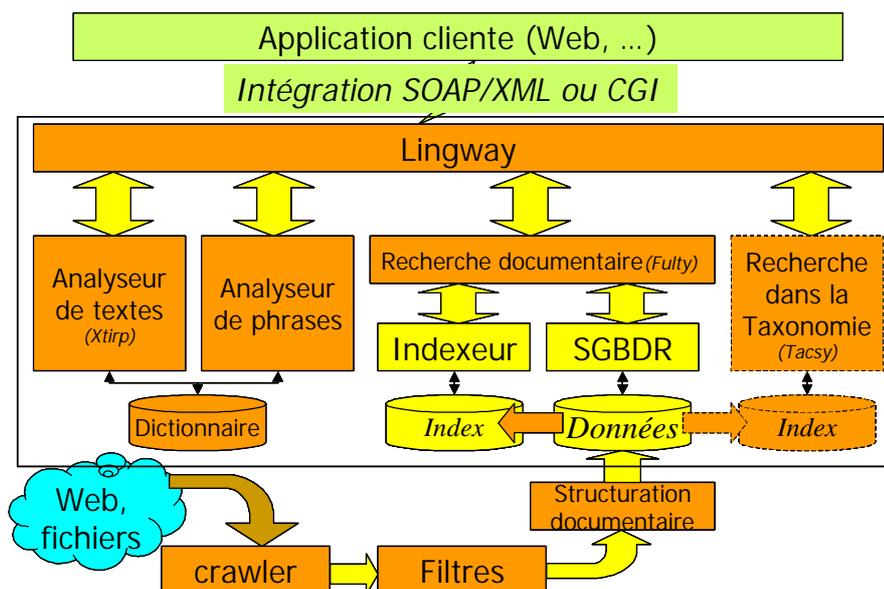
- moteur de génération (Geny) :

logiciel permettant de générer un texte, avec diverses variantes en plusieurs langues et plusieurs styles, à partir de données structurées.

- module de gestion des dictionnaires (Alexy) :

module de gestion des ressources linguistiques (terminologies, lexiques, etc.) gérant les données utilisées et/ou produites par les différents modules. Il existe en mode intégré pour une consultation aisée de mots et termes en contexte.

Le schéma suivant montre l'architecture générale de Lingway KM.



Les bases de données (base relationnelles et base d'index) sont gérées en standard par des logiciels open-source (Lucène et MySQL). Des connecteurs sont possibles vers la plupart des logiciels du marché (Verity, Hummingbird, Oracle, etc.)

3.9 GESTION DU CONTENU VISUEL

Le contenu visuel occupe aujourd'hui une place prépondérante et représente un vecteur central de transmission de l'information.

La description de ce contenu par des techniques d'analyse d'images est moins subjective que la seule description habituelle par des mots clés (quand elle existe).

D'autre part, étant indépendante de la langue de recherche, elle devient primordiale pour l'exploration efficace d'un flot multimédia.

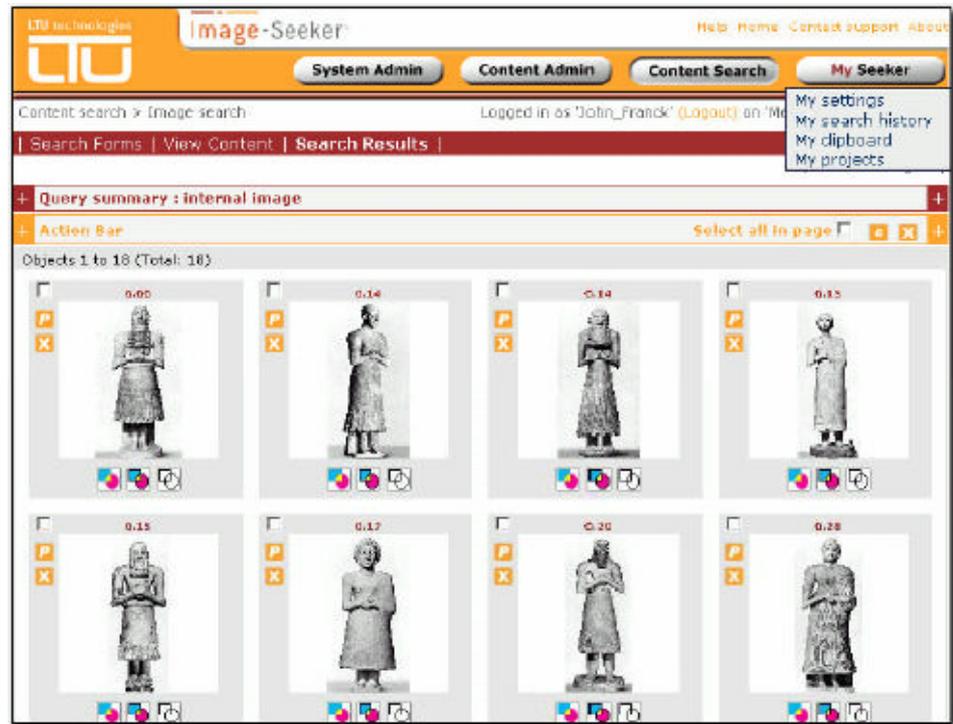
L'analyse des images satellites, la biométrie et la problématique de la détection et de l'analyse de similarité de visages sont à l'origine de la gestion du contenu visuel.

Par la suite, l'explosion du multimédia et la prolifération sur le net des données visuelles, licites ou illicites, ont largement contribué à renforcer la demande en matière de gestion du contenu visuel.

Dans ce domaine, le dynamisme des programmes de recherche et tout particulièrement de ceux de l'Inria qui conduisent à la création d'entreprises sont à saluer.

Ainsi, l'exemple de la jeune société française LTU technologies et son logiciel vedette "Image-Seeker" confirme la maturité et la viabilité de la technologie.

En effet, "Image-Seeker" est en passe de devenir la référence pour les grandes entreprises privées et les organisations du secteur public comme l'INPI.



"Image-Seeker" est un système complet de gestion des images avec indexation et recherche par le contenu. Il simplifie les procédures de recherche et facilite la navigation dans les grandes bases de données d'images.

Le cœur de la technologie repose sur un analyseur à forte sensibilité capable d'indexer, de reconnaître et de comparer des images à partir de leurs composantes visuelles.

Qu'il s'agisse d'une photographie, d'un dessin, d'une illustration ou de tout autre document visuel numérisé, le système produit en temps réel la description automatique de son contenu.

En fonction des besoins exprimés par l'utilisateur, la description peut être :

- absolue : réponse "il s'agit de personnes" à la question "qu'est ce que c'est ?» ;

- relative : réponse "ceci ressemble à cela" à la question "qu'est ce qui ressemble à cela ?".

La segmentation de l'image constitue la première étape du processus d'analyse.

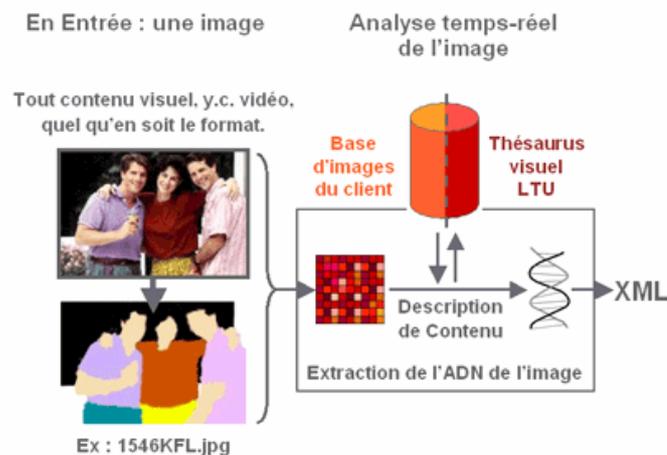
La technologie LTU permet de découper l'image en fonction des différents objets qui la composent.

La deuxième étape est l'indexation : pour chaque image préalablement segmentée, le système extrait un identifiant unique, également appelé signature numérique ou ADN de l'image.

Cet ADN est une représentation optimisée des attributs visuels de l'image, tels que la forme, la couleur, la texture, la configuration spatiale, etc.

Elle possède des propriétés d'invariance à la qualité, la taille de l'image, la luminosité, les contrastes, la distorsion, la rotation et la translation d'objets ainsi qu'au changement d'échelle.

L'image est donc codée sous la forme d'un vecteur de valeurs compacte qui code efficacement toutes les informations relatives au contenu de l'image.



La technologie LTU permet également de passer de l'ADN à la description sémantique.

Pour cela et au cours du processus de description, le contenu de l'ADN est analysé par les différents modules experts afin d'être reconnu relativement à la base de connaissances.

Cette base de connaissances peut être :

- une base interne permettant une description absolue du contenu ;
- une base externe permettant une description relative du contenu.

Les modules experts utilisent des algorithmes sophistiqués de reconnaissance de formes, tels que des réseaux de neurones, des fonctions à base radiale, des estimations Bayésiennes, et des «Supports Vector Machine».

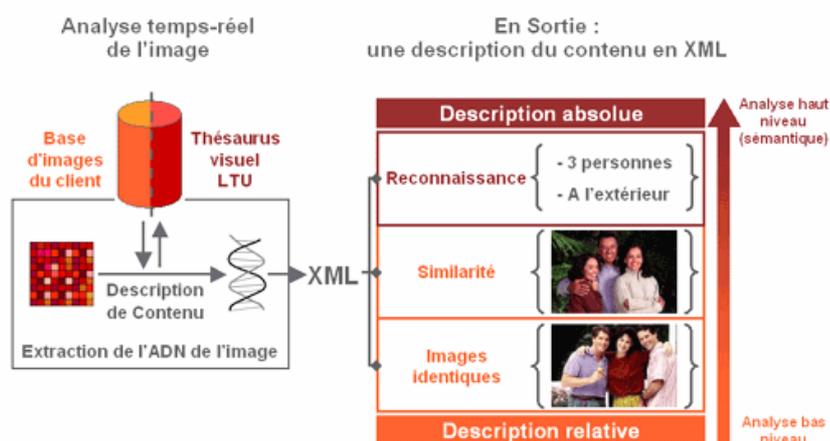
Le processus de reconnaissance a été développé par LTU technologies pour reproduire de manière statistique le comportement humain, sur la base de nombreux tests psychosensoriels menés par LTU technologies.

Il est également important de souligner que le moteur d'analyse a des facultés d'apprentissage.

Il peut apprendre des catégories d'objets, intégrer des variations d'apparence et enrichir ainsi la base de connaissances.

Il est aussi économiquement intéressant de "propager" les méta-données d'une base d'images bien renseignées à l'ensemble des résultats obtenus par des recherches par similarité visuelle.

Appelé "automatic keyword suggestion", le système se présente alors comme un système semi-automatique : proposition par la machine puis validation par l'humain.



Il est important de préciser que la société LTU technologies est capable de développer des signatures "ADN" spécifiques pour des domaines d'applications donnés.

Nous citons, par exemple, la signature "propriété industrielle" développée pour la problématique du portail Plutarque.

Le moteur d'analyse est accessible par une interface web standard ou pleinement intégrable dans un système existant.

Tous les formats d'images sont supportés par "Image-Seeker" (jpg, gif, bmp, tif, png, psd, jpeg2000 etc.).

Les séquences vidéo sont également prises en compte et indexées par Image-Seeker et peuvent être recherchées par images clefs.

Enfin, comme le texte et l'image sont complémentaires, Image-Seeker est capable de combiner une recherche sémantique et une image.

Cette combinaison des deux étant meilleure que le meilleur des deux, on dispose alors d'un système calculant la similarité sémantico-visuelle et capable de lever les ambiguïtés qu'un moteur sémantique seul ou qu'un moteur image seul ne pourrait traiter.

Ce type de complémentarité texte/image est d'ores et déjà en production, puisque la technologie LTU est aussi disponible en OEM.

4. LES OUTILS

Les outils vont permettre d'organiser, d'une part, les connaissances elles-mêmes (les noter, les archiver, les classer...) et, d'autre part, d'en gérer la diffusion dans l'entreprise. Les différents outils de gestion des connaissances peuvent être regroupés en trois grandes catégories :

- les outils de capitalisation des connaissances ;
- les outils de travail ;
- les outils de diffusion des connaissances.

4.1 PANORAMA DES TECHNOLOGIES

4.1.1 Les outils de capitalisation des connaissances

Parmi ces outils, on peut citer :

- les serveurs de connaissances ;
- les systèmes de gestion électronique de documents (GED) ;
- les systèmes de datawarehouse ;
- les annuaires de messagerie.

Les serveurs de connaissances permettent de déterminer :

- le mode de structuration des informations et des connaissances intégrées au sein de l'outil ;
- le mode d'accès à ces connaissances, lequel met en œuvre les hyperliens mais aussi d'autres techniques comme les thésaurus.

Les systèmes de gestion électronique de documents (GED) permettent de disposer, sous forme numérique et indexée, de connaissances jusqu'alors seulement présentes sous forme analogique (papier, voix, images, etc...) et souvent non indexées, ce qui rend la capitalisation difficile et accroît les réticences des utilisateurs.

Les systèmes de datawarehouse centralisent les connaissances extraites de différentes applications et bases de données rattachées à l'environnement de production notamment pour faciliter la prise de décisions.

Enfin, l'annuaire de messagerie rassemble l'ensemble des informations concernant les membres de l'entreprise et permet de créer un document réunissant les collaborateurs par compétence.

A titre d'illustration, Factiva Fusion est un outil de gestion et d'enrichissement du contenu qui répond à l'un des enjeux majeurs des entreprises, à savoir l'intégration de contenus structurés et non structurés issus de sources disparates dans un seul format "Trouvable".

Factiva Fusion "normalise" les données internes et externes à l'entreprise dans un seul format XML rendant possible leur recherche et leur intégration dans les logiciels d'entreprise existants.

Le contenu est ensuite "catégorisé" au moyen d'une taxonomie personnalisée selon les souhaits du client.

Les entreprises peuvent au choix utiliser la taxonomie propriétaire de Factiva ou utiliser leur propre taxonomie.

De même, Verity Inc. dont le siège se trouve à Sunnyvale en Californie, est présente en France depuis plus de dix ans.

Verity propose des solutions logicielles qui permettent aux entreprises d'optimiser le retour sur investissement de leur capital intellectuel grâce à ses technologies de recherche, de classification, de catégorisation et de personnalisation de l'information.

Les solutions Verity sont utilisées pour mieux partager l'information dans les entreprises, pour faciliter les ventes sur les sites d'e-commerce et pour des activités B2B sur des places de marché en ligne.

En outre, la technologie de Verity est le composant clé de nombreuses applications e-business.

4.1.2 Les outils de travail

Les outils de travail regroupent notamment :

- les systèmes de groupware ;
- les systèmes experts.

Les systèmes de groupware permettent aux collaborateurs de travailler en commun sur un document, et ce de manière organisée.

Les systèmes experts sont composés de logiciels exploitant dans un domaine particulier des connaissances explicites et organisées pouvant se substituer à un expert humain.

4.1.3 Les outils de diffusion des connaissances

Ils regroupent notamment :

- les serveurs web ;
- les forums et les newsgroup ;
- le système de workflow ;
- l'intranet de l'entreprise ;
- l'e-learning.

Le mode "push" des serveurs web est particulièrement utile en ce qu'il permet d'envoyer des informations et des alertes vers les utilisateurs suivant des critères prédéfinis.

Les forums et les newsgroup permettent de dialoguer, ce qui peut conduire à enrichir la base de connaissances de l'entreprise.

Les systèmes de workflow coordonnent, sur la base de règles, l'interaction entre des personnes, des informations et des applications informatiques.

L'intranet permet de regrouper l'ensemble des informations et connaissances de l'entreprise.

L'e-learning permet une mutualisation des connaissances à travers une formation à distance (formation didactique, intuitive et on line).

4.2 ILLUSTRATIONS

Certains des membres du groupe de travail ont présenté l'outil de gestion des connaissances qu'ils utilisent en interne ou qu'ils ont développés.

4.2.1 Sensoria Technologie Limited

Sensoria Technology a développé une technologie radicalement nouvelle de reconnaissance de formes et d'apprentissage non supervisé, dynamique et adaptatif.

Cette technologie s'appuie principalement sur des algorithmes d'apprentissage modélisant certains principes neurobiologiques.

Elle est appliquée en priorité, sous la marque Influo, au domaine des moteurs de recherche intelligents à apprentissage autonome et indépendant du signal.

Selon la définition qu'en donne la société, Influo iSearch est "un Moteur de Recherche de toute nouvelle génération livré en rack logiciel pour des applications totalement autonomes."

Il doit son efficacité, entre autres, à sa Tête de lecture neuromimétique, une des principales pièces qui le composent.

Et à ce titre, il constitue une brique fondamentale d'une bonne solution de gestion des connaissances.

En effet, l'analyse par Sensoria des questions que pose la gestion des connaissances identifie trois types de problèmes :

- un problème technique, lié à l'extraction de connaissances ;
- un problème humain, lié à la difficulté de motiver les utilisateurs ;
- un problème économique, lié entre autres aux coûts récurrents des solutions.

Parce que, selon Sensoria, ces trois points représentent en réalité trois facettes distinctes d'un seul et même problème, ils peuvent être étudiés à travers le prisme de sa technologie neuromimétique, sur laquelle s'appuient ses moteurs Influo.

Les réponses apportées par ces derniers ont été présentées dans le cadre du groupe de travail et ont contribué à enrichir la réflexion.

- un problème technique, lié à l'extraction de connaissances
- connaissances explicites / connaissances tacites

Les études sur la gestion des connaissances font souvent référence aux notions de "connaissances explicites" versus "connaissances tacites".

On reconnaît par là le fait très simple que certaines connaissances sont facilement appréhendées par des interlocuteurs différents, ne seraient-ce que parce qu'elles ont été exprimées. Ce sont les connaissances explicites.

D'autres, au contraire, demeurent tacites, c'est-à-dire confinées à la conscience d'une poignée d'interlocuteurs, voire de leurs seuls créateurs.

Selon une étude déjà ancienne du cabinet Ernst & Young, 95% des connaissances de l'entreprise demeureraient tacites, tout simplement parce que, ayant été exprimées oralement, elles n'existent finalement que de manière informelle "dans la tête" de leur énonciateur.

Cependant, la question mérite d'être élargie. Comment en effet être certain qu'une notion manipulée par l'un, même dans un document écrit et correctement stocké, sera appréhendée de manière adéquate par l'autre interlocuteur ? C'est bien pour cette raison que les listes contrôlées, vocabulaires normalisés et autres thésaurus ou taxinomies ont été mises en place dans nombre de solutions.

Sans même parler des aspects contraignants de ces procédures, qui ont un impact direct sur la motivation des collaborateurs, force est de constater qu'elles ont obligatoirement un temps de retard sur la réalité.

Il s'agit là d'une des faiblesses à porter au débit des outils d'aide à l'extraction manuelle de connaissance, lesquels, comme on l'a vu, se montrent par ailleurs très utiles.

D'une manière générale, les outils à disposition gèrent mal les aspects dynamiques et adaptatifs nécessaires.

- le problème : le traitement du langage, plus grand dénominateur commun

Un producteur d'information insère de la connaissance dans le système dans une langue, avec des termes, un style, une clarté qui lui sont propres.

L'utilisateur du système ou le consommateur est plus ou moins proche de la pensée du producteur, ce qui donnera lieu à une interprétation d'autant plus sujette à variation que la distance culturelle est grande.

Or, à cette distance culturelle (il s'agit bien de perpétrer la culture de l'entreprise voire même, avant tout, de l'insuffler) s'ajoute la distance dans le temps puisque aussi bien il s'agit de documents écrits parfois à plusieurs années d'intervalle.

Une notion pour l'un ayant souvent le défaut de résonner différemment, même légèrement, pour un autre, le challenge réside dans l'établissement de ponts entre tous les concepts véhiculés à travers les documents.

C'est-à-dire, ni plus ni moins, dans la constitution d'une sorte de thésaurus évolutif pour chacun des concepts de connaissance utilisés.

On notera qu'il convient de s'intéresser aux notions utiles parce que réellement maniées, et non pas à des concepts issus d'une quelconque liste extérieure à l'entreprise.

Or, un concept est avant toute chose un mot (ou un groupe de mots) écrit, doté d'un ensemble de liens à d'autres mots, ces liens étant eux-mêmes pondérés (ayant une importance relative).

Dans l'idéal, il faudrait pouvoir, dans un texte donné, repérer automatiquement les groupes de mots lorsqu'ils apparaissent.

Or, ceci pose un problème lié directement à la variabilité du langage. Pour s'en convaincre, on pensera aux difficultés que posent les mots "outils" (articles, conjonctions, etc.) en français.

Les exclure totalement rend des documents inaccessibles par les groupes de mots tels que *or noir*.

Les conserver produit a contrario un "bruit" (*présidence + république* n'est pas *présidence de la république*) gênant pour l'efficacité du système, comme le confirment de simples requêtes avec les moteurs de recherche les plus connus.

Finalement, à y regarder de plus près, la manière dont sont construits les groupes de mots ne semble pas aussi évidente. Il n'existe pas réellement d'ensemble de règles cohérent et infallible.

Illustration

L'extraction de connaissance commence par celle des groupes de mots

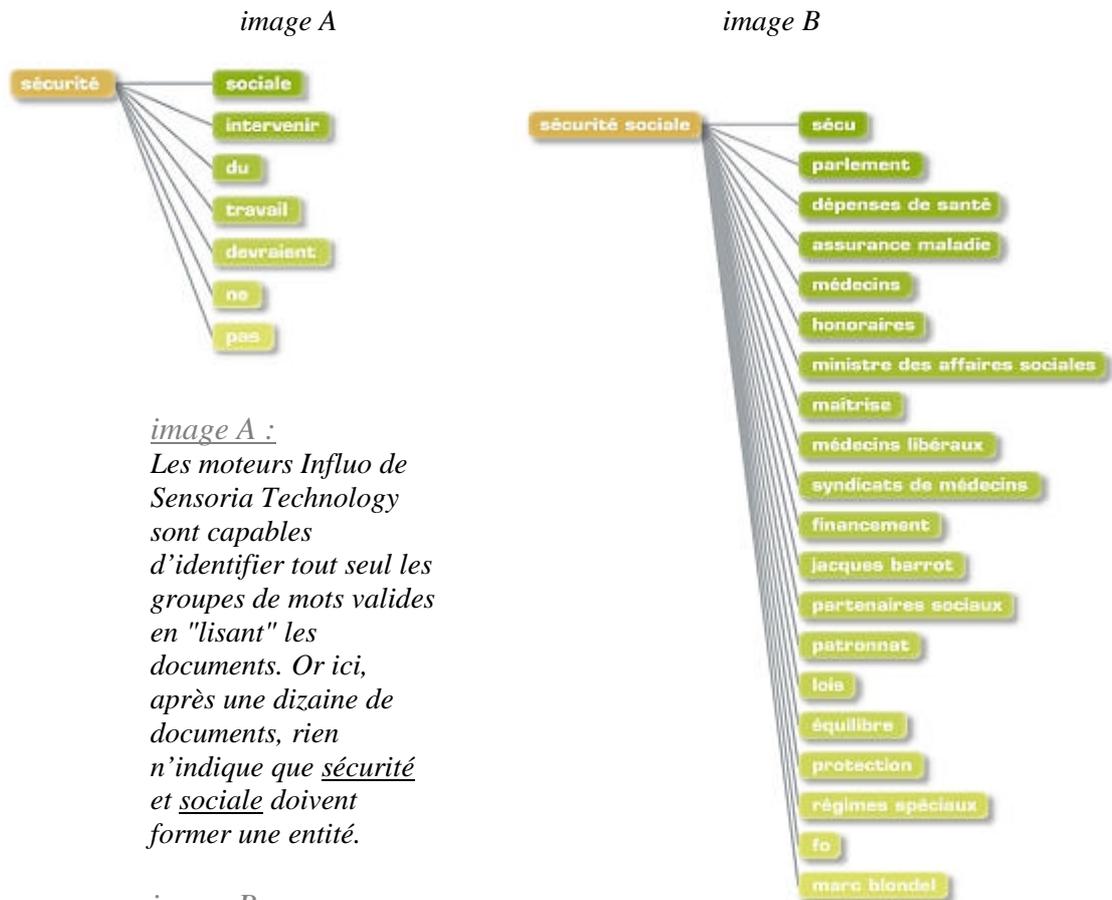


image A :

Les moteurs Influo de Sensoria Technology sont capables d'identifier tout seul les groupes de mots valides en "lisant" les documents. Or ici, après une dizaine de documents, rien n'indique que sécurité et sociale doivent former une entité.

image B :

Après un temps, le moteur Influo a décidé seul que sécurité sociale a un sens, que l'on peut observer ici. On note qu'aucune règle ne peut prédire que sécurité sociale et sécu sont liés et quasi synonymes. Or, cela touche directement à l'aspect vivant de la langue et de la connaissance.

- rappel et précision mis à mal

Sensoria Technology, en fondant ses moteurs Influo sur sa technologie neuromimétique, reconnaît que le sens et la forme (l'orthographe) des mots et des groupes de mots sont si intimement liés, que traiter l'un indépendamment de l'autre conduit à l'impasse. Un traitement neuromimétique du langage, au contraire, permet une interprétation correcte des textes soumis. Ainsi, les moteurs Influo sont capables d'analyser dans un ensemble de textes une forme telle que $tf + \text{espace} + l$ ($tf\ l$) comme une variante de $tf\ l$, de $TF\ l$ et, finalement, de $TF\ l^1$. Sans une telle aptitude, on comprend que les notions de rappel et de précision, présentées plus haut dans ce document, puissent être mises à mal².

¹ Cas réel de plusieurs dizaines de milliers d'articles de quotidien.

² Rappel ou précision, rappel et précision : question de pertinence. Sensoria Technology ; 2004.

En effet, les documents dans lesquels n'apparaît qu'une forme "bruitée" telle que *tf I* ne pourront pas être retrouvés avec l'orthographe *TFI*. Les performances de l'outil en la matière risquent bel et bien d'introduire une zone "d'incompréhension" supplémentaire dans le système de gestion des connaissances. Et ce, parce qu'il est impossible de prévoir comment un texte sera écrit *réellement*.

- vie et mort de la connaissance

Il est possible d'envisager la connaissance sous un angle statique. Mais, à l'évidence, la connaissance est une chose vivante.

Une notion, par exemple, peut prendre de l'importance au fil du temps, ou en perdre. Un outil pratique doit donc apporter une aide sur ce plan.

Une indexation manuelle, aussi fine soit-elle, s'appuiera avec bénéfice sur un outil capable d'identifier automatiquement des notions puis un ensemble de liens entre ces notions, puis enfin de mettre à jour automatiquement la force, voire l'existence même de ces liens.

Même et surtout lorsqu'il s'agit de liens naissant, ce que l'on appelle alors des "signaux faibles", dont la détection de l'émergence au sein des documents de l'entreprise semble un point important de la gestion des connaissances.

Or, un signal faible est un signal peu amplifié, et donc peu répété. Il s'agit donc d'identifier un thème d'une importance certaine alors même qu'il se retrouve dans un nombre de documents très restreint. Par définition, les classiques méthodes statistiques s'adaptent difficilement à la faible répétition.

Les méthodes d'analyse linguistique ou s'appuyant sur des thesaurus conviennent peu aux situations d'imprévu, caractéristiques des signaux faibles où, par définition, quelque chose de nouveau est en train de se constituer.

Sur le plan technique, la détection automatique des contextes sémantiques est une des grandes difficultés qui se présentent aux outils de gestion des connaissances. Sensoria Technology et ses moteurs Inluo démontrent que cela est possible dorénavant.

- un problème humain, lié à la difficulté de motiver les utilisateurs

Créateurs, producteurs, consommateurs ou utilisateurs se trouvent tous confrontés au même problème potentiel : lorsque les contenus sont disponibles, comment les retrouvent-t-on, de quelle manière y accède-t-on ? Or, si la solution n'est pas efficace, si elle s'avère lente ou rébarbative (voire, pire encore, les trois), elle a peu de chance de séduire durablement les utilisateurs et d'aider l'entreprise à surmonter la difficulté à instaurer une culture du partage.

A contrario, les utilisateurs s'estiment en droit de s'interroger :

- que puis-je espérer trouver ?
- ai-je le droit à l'hésitation ?
- où se trouvent les documents utiles ?

La gestion des connaissances vise entre autres à rendre accessible les notions et les concepts manipulés par l'entreprise à l'ensemble de ses collaborateurs et, par définition, à ceux qui en sont les moins familiers.

Comme, par exemple, aux nouveaux arrivants. Il va de soi que, lorsqu'on ne connaît pas a priori ce que renferme une base d'information ou de connaissance, la question de savoir à l'avance si on a une chance de trouver des informations est importante.

C'est ce qui rend agréable à tout un chacun, dans la vie courante, la réponse rapide et avisée d'un bibliothécaire :

- Question : "Avez-vous quelque chose sur *ce sujet* (le sujet qui nous intéresse) ?" ;

- Réponse : "Malheureusement, notre fonds n'est pas bien riche sur *ce sujet* ; etc."

La frustration guette l'utilisateur qui s'investirait dans une recherche infructueuse sans avoir été prévenu de cet échec prévisible.

De même que sa motivation risquerait de s'user au fil de recherches où il aurait lui-même à faire le tri dans une information pertinente, mais trop riche.

Autrement dit, l'assistance à la recherche est une fonctionnalité nécessaire.

Cette assistance conduit à un paradoxe, qui rend à première vue les objectifs des utilisateurs en partie contradictoires avec ceux des promoteurs de l'application de gestion des connaissances.

Autant les premiers désirent avant tout trouver vite ce qu'ils cherchent, autant les seconds souhaitent que les utilisateurs non seulement consomment, mais aussi produisent information et connaissance.

Lorsque le promoteur devient éditeur et l'utilisateur consommateur, c'est-à-dire lorsque le premier tente d'établir une relation économique avec le second, le paradoxe s'accroît puisque la rentabilité¹ repose sur l'utilisation durable et répétée de la plate-forme (outil + base de connaissance-information).

¹ voir Chapitre V, Les modèles économiques.

Ce paradoxe apparent se résout naturellement dans le cercle vertueux proposé par les outils tels que les moteurs Influo.



Illustration

Les moteurs Influo et le cercle vertueux de la gestion des connaissances

Un outil efficace de gestion des connaissances doit permettre tout autant de trouver rapidement ce qu'on cherche que de s'enrichir au passage.

Les promoteurs des applications se doivent également d'envisager la question des technologies de gestion des connaissances sous un angle beaucoup plus direct et pragmatique.

- un problème économique, lié entre autres aux coûts récurrents des solutions.

Toutes les technologies au service de la gestion et, plus particulièrement, de l'extraction des connaissances, ne se valent pas sur le plan économique.

Plusieurs paramètres entrent en ligne de compte, et notamment la rapidité et la simplicité de mise en œuvre, ainsi que le coût total d'exploitation.

Ce dernier montre un visage bien différent selon qu'il sera nécessaire ou non de procéder à des mises à jour fréquentes de l'outil.

La vitesse de réaction aux nouveaux termes, acronymes, prénoms et autres, a également une influence directe sur le coût réel d'une solution.

Sur tous ces points, Sensoria Technology et les moteurs Influo apportent la réponse de l'automatisation et de l'adaptation permanente. Leur philosophie se voit d'ailleurs résumée par la définition même de la marque Influo : *plug & find*.

La technologie neuromimétique de Sensoria confère donc aux moteurs Influo leur fonctionnement entièrement automatique et autonome.

On retiendra leur capacité d'apprentissage sans fin (*never-ending learning*).

Mais leur présentation et leur démonstration devant le groupe de travail ont été également l'occasion de présenter et d'illustrer quelques définitions souvent rencontrées en gestion des connaissances et touchant à l'aspect pratique des choses :

- la catégorisation, parfois appelée indexation, est l'utilisation d'un plan de classement défini par des experts pour y ranger les documents et les restituer par l'interrogation de ce même plan de classement ;

- par opposition, la classification automatique est l'utilisation de contextes hiérarchisés formant quasi-partition de la base (c'est-à-dire ayant le plus faible recouvrement possible entre eux), et entièrement issus de l'analyse automatique du fonds. De ce fait, aucune de ces "classes" ne peut être vide. L'identification de ces contextes de manière dynamique (constamment remise en cause) est fondamentale pour une gestion des connaissances efficace. Aucune classe ne pouvant plus devenir obsolète, la question de leur "nettoyage" ne se pose plus ;

- ces mêmes contextes peuvent être utilisés pour grouper les documents retournés pour une requête donnée ; on parle alors également de clusterisation (le mot anglais pour regroupement) ;

- les contextes peuvent s'appliquer à la requête uniquement, dans un premier temps, dans le but d'éliminer toute ambiguïté ; il s'agit alors, pour le moteur, de faire des suggestions afin d'affiner la requête ;

Illustration :

Contextes de "Principe" : l'identification de contextes sémantiques, une fonction fondamentale



Pour être utiles, les contextes doivent avoir une signification dans le fonds concerné, c'est-à-dire être extraits automatiquement et dynamiquement. Ici, ils aident à préciser une requête trop générale.

- la navigation sémantique (ou cartographie) s'appuie sur la proposition, par l'outil, d'un ensemble de termes formant contexte de la requête ou étant de simples termes liés sémantiquement aux mots de cette dernière. L'utilisateur a ici la possibilité de naviguer dans la connaissance du système de manière, entre autres, à se faire une idée à l'avance (sans avoir à lire les documents eux-mêmes) de la richesse de la base sur le sujet choisi. De manière également à découvrir d'éventuels liens inconnus de lui. On peut alors réellement parler de gestion des connaissances ;

- au cœur de la gestion des connaissances, on retrouve une nouvelle fois la capacité du système à suggérer des termes liés à ceux de la requête, mais avec pour objectif, cette fois, de retrouver des documents ne contenant pas les mots exacts. Ce sont les inférences : une requête *vache folle* retourne des documents ne contenant que *ESB* (encéphalite spongiforme bovine). Un tel comportement s'obtient avec plus ou moins de bonheur par l'intermédiaire d'un thésaurus dans les systèmes classiques, dès lors que le cas a été prévu à l'avance par les experts. Comme on l'a vu, l'intérêt d'un système dynamique est, entre autres, de n'être pas bloqué par l'apparition de nouvelles notions comme, ce qui est très commun, les noms propres ou, dans notre exemple, les acronymes ;

- les contextes, encore eux, peuvent s'appliquer aux documents ; il s'agit alors d'identifier les différents thèmes d'un document. Ce qui permet notamment, lorsque l'utilisateur présente un document au système, de proposer des documents similaires, correspondant à un thème choisi parmi la liste. Cette fonction est couramment dénommée *more like this*. Toute la difficulté réside évidemment dans cette capacité à identifier les thèmes d'un document ;

- enfin, la dernière définition, celle du profil de requête, est propre aux moteurs *Influo*¹. Il s'agit de fournir immédiatement de l'information sur l'information de manière à gagner du temps. Le profil de requête participe à l'assistance à l'utilisateur et tente, avec d'autres indicateurs, de lui éviter les pertes de temps.

Propre également aux moteurs *Influo* et ultime degré de l'assistance à l'utilisateur, la correction orthographique en contexte poursuit l'objectif de lui conserver son droit à l'hésitation, tout en assurant qu'il est bien compris.

Car, lorsqu'il s'agit de mettre en commun des contributions individuelles (ou de groupes de travail), il serait paradoxal de présupposer que tous les collaborateurs possèdent déjà la connaissance de tous les concepts maniés dans l'entreprise.

Au contraire, pour un réel ralliement de tous les collaborateurs, il est peut-être bon d'admettre qu'une gestion efficace de la connaissance aille parfois jusqu'à la découverte de l'orthographe de ses concepts.

Dans ce cas, la correction orthographique de la requête, si tant est qu'elle est efficace parce que, comme toujours, en contexte, devient plus qu'un simple complément du couplage des variantes dans les documents (cf. : *tf 1 / tf1 / TF 1 / TF1*).

Elle fait du système de gestion de connaissance un intermédiaire fiable et compétent entre tous les collaborateurs.

¹ Moteurs de recherche : où est la technologie ? ; Sensoria Technology ; 2004

Illustration

Correction orthographique en contexte : très loin d'un gadget

principe de précaution

SAUF CONTEXTUEL

AUTEUR

DATE avant après mars 2000

iSearch

RAPPORT DE REQUETE 320 résultats

vosre requête : principe de précaution

ambiguité corrigée : principe de précaution

termes associés : OGM - effet de serre - Kyoto

autres suggestions : principe de prudence

32 PAGES

résultats affichés : 1 à 32

Trier le grain de l'ivraie

...Il est vrai que les avis liés aux OGM ne sont pas le **principe de précaution**

Les américains en face

...face à Bruxelles, dont le **principe de précaution** agace fortement des représentants américains

La terre souffre

...entre autres les gaz à effet de serre, les désaccords entre spécialistes de la cause **principe de précaution**

La libération des échanges rogne l'espace individuel

...se dessine aujourd'hui, au grand dam des anti-globalisation. Le marchandage autour de la libéralisation des échanges, où le **principe de précaution**, voire-même le respect de la diversité culturelle, devraient...

AIDE : PROFIL DE REQUETE

pertinence

76-100 %	35
50-74 %	45
25-49 %	150
1-24 %	90

nbre documents

Répartition des documents en fonction de leur pertinence. Si la barre verte est minoritaire, le sujet de la requête est peu traité dans ce fonds.

dynamique	spécialisation
9 - 98 %	2 %
Indique les notes de pertinence extrêmes.	indique la proportion du fonds mobilisé par la requête.

TRI

- pertinence
- date
- auteur

09/2000 Roland Levasseur 98 %

04/2000 Joan Riomsky 98 %

05/2002 Jacques de Brière 97 %

11/1999 Catherine Bergé 96 %

Une correction orthographique puissante s'avère extrêmement utile lorsqu'elle est pratiquée "en contexte". Mais elle requiert une technologie très particulière, comme celle des moteurs Influo de Sensoria Technology.

4.2.2 Plutarque

4.2.2.1 Portail de recherche d'information en propriété industrielle

Plutarque est le portail de recherche d'information en propriété industrielle proposé par l'Institut National de la Propriété Industrielle (INPI).

On dit de Plutarque qu'il est intelligent car il comprend et interprète les questions et permet un accès simple et direct aux titres de propriété industrielle pertinents (brevets, marques, dessins et modèles et leurs jurisprudences) des cinquante pays qu'il fédère.

L'accès à ces contenus est rendu intuitif et personnalisable par l'agrégation de technologies sophistiquées. En effet, Plutarque dispose d'outils linguistiques et sémantiques associés à un module de recherche et de navigation par l'image.

L'intelligence de Plutarque est la simple matérialisation des relations entre les connaissances extraites des typologies utilisateurs et de l'analyse formelle de leurs questions avec la connaissance du contenu propriété industrielle extraite par les experts des offices.

Deux espaces Plutarque sont à la disposition des utilisateurs :

- un espace public, gratuit, limité en terme de contenu et de services ;
- un espace privatif, sur abonnement, qui offre l'intégralité des contenus et des services. Cet espace est complètement personnalisable.

4.2.2.2 Technologies de recherche : représentation formelle de la question

Plutarque propose une large palette de fonctionnalités et différentes technologies de recherche pour répondre aux attentes et aux besoins des utilisateurs.

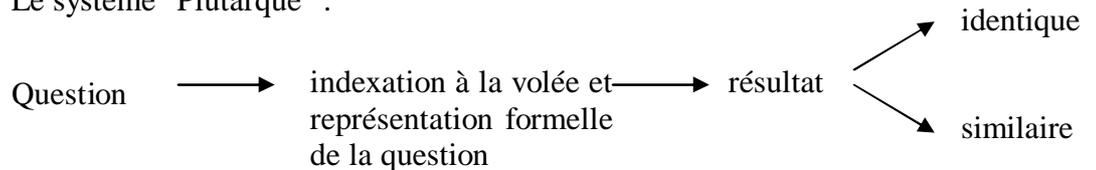
On distingue notamment :

- le module de recherche langage naturel développé par la société Lingway : ce module de recherche est capable de générer une requête booléenne pertinente pour les contenus brevets, marques et dessins et modèles à partir de la représentation formelle d'une question posée en langage courant. L'utilisateur n'a donc pas à maîtriser un logiciel d'interrogation sophistiqué ;

- le module de recherche et de navigation par l'image développé par la société LTU Technologies : ce module de recherche est capable d'identifier les marques figuratives, les dessins ou les modèles identiques ou similaires à la représentation formelle d'une question image ;

- le module "formulaire" ou "expert" qui utilise les algorithmes du moteur documentaire GTI Jouve et qui permet les requêtes par mots clefs à travers un formulaire de saisie ou à partir d'une stratégie de recherche construite par un professionnel de la documentation.

Le système "Plutarque" :



Pour le module langage naturel, la représentation formelle de la question repose sur :

- une analyse morphologique de la question ;
- une analyse syntaxique (identifier des groupes de mots...) de la question ;
- une analyse sémantique de la question,

et ce, quelle que soit la langue de la question (anglais, allemand, français ou espagnol).

La représentation formelle de la question est alors enrichie et pondérée par :

- le réseau des concepts génériques (Lingway) ;
- le réseau des concepts spécialisés en propriété industrielle (INPI).

Le résultat est une requête booléenne qui permet, au choix, d'obtenir une information bibliographique ou texte intégral qui répond parfaitement à la question.

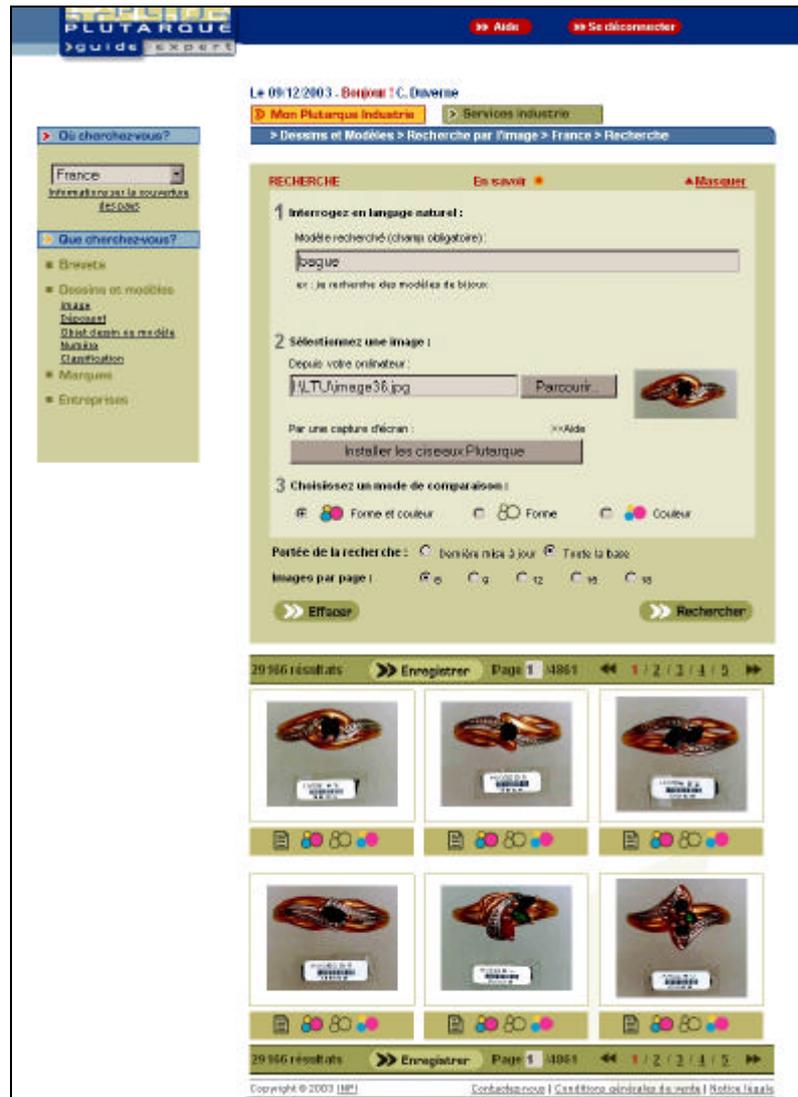
Pour les fonds étrangers, une traduction à la volée est également disponible.

En ce qui concerne le module de recherche et de navigation par l'image, la représentation formelle de la question se fait à partir des éléments graphiques suivants :

- couleur ;
- forme ;
- texture.

La représentation formelle de la question est alors comparée aux représentations formelles des images des bases de données. Les images identiques ou similaires à la question sont alors affichées. A partir des réponses images, l'utilisateur peut, au choix, obtenir une information bibliographique ou un texte intégral.

Enfin, le module langage naturel et le module de recherche et de navigation par l'image ont été croisés. Ainsi, une question en langage naturel en entrée, peut être associée à une image.



4.2.2.3 La nature de l'information de propriété industrielle : connaissance du contenu

L'information de propriété industrielle est une information riche, variée et complexe :

- elle existe sous une grande diversité (juridique, langue...) ;

- elle se caractérise par une diversité de volumes (dépôts, enregistrements, délivrances, oppositions...);
- elle se caractérise par une grande complexité (différence entre un déposant et un inventeur, marque renouvelée, inscription au registre, brevet sous priorité...);
- elle est évolutive (en fonction des technologies, des marchés..).

Par ailleurs, pour faciliter le travail des offices, l'information de propriété industrielle est structurée, standardisée et indexée :

- norme internationale ST80 pour les champs d'un brevet ;
- codes de classification des domaines techniques (CIB) ;
- pour les marques et les dessins et modèles : codes de classification des produits et services (Nice et Locarno) ;
- pour les marques figuratives : codes de classification des éléments figuratifs (Vienne).

Cette indexation est réalisée par des experts en propriété industrielle au sein même des offices. Elle est utilisée sur le plan international et correspond à des codes alphanumériques.

L'ensemble de l'indexation représente la connaissance extraite du contenu.

C'est cette connaissance qui est rapprochée de la représentation formelle de la question.

4.2.2.4 Les utilisateurs : connaissance de la typologie client

Aujourd'hui, les clients des offices sont tous des utilisateurs professionnels des produits et services d'information qui alimentent un consommateur final d'information de propriété industrielle (étudiant, chercheur, cadre dirigeant d'une PME, personnel de service juridique, personnel de service financier, personnel de service marketing ou commercial...).

Il est intéressant de préciser qu'il n'est pas rare de constater qu'au sein des grandes entreprises, seuls quelques professionnels de l'information de propriété industrielle alimentent plusieurs milliers de consommateurs finals.

Dans le cadre d'une mission de service public, un office est tenu de servir tous les publics et non pas seulement un public particulier. C'est sur ce principe que l'INPI a mis en œuvre son portail Plutarque.

La connaissance du consommateur final a donc été extraite de l'étude et de l'analyse du comportement du professionnel de l'information de propriété industrielle. Ainsi, dès qu'un utilisateur s'identifie sur Plutarque, un premier niveau de personnalisation s'opère au niveau de l'interface pour la typologie donnée de l'utilisateur. L'utilisateur est alors libre d'affiner cette personnalisation.

Le comportement est associé aux attributs quantitatifs et qualitatifs les plus pertinents du professionnel qui sont :

- son entreprise ;
- son adresse ;
- son budget ;
- sa politique d'accès à l'information (intranet...) ;
- sa participation à des salons ;
- les journaux lus ;
- son équipement ;
- sa consommation ;
- le nombre de consommateurs finals servis (...).

Tous ces attributs sont archivés dans une base de connaissance CRM qui alimente un analyseur multidimensionnel et un outil graphique statistique.

Des comportements similaires sont ainsi révélés et des typologies créées.

4.2.3 iClass

La solution globale iClass, développée par le Bureau Van Dijk est destinée à être utilisée par toute personne ayant un rôle dans le cycle de décision, c'est-à-dire par tout utilisateur potentiel de l'information stratégique : direction générale, direction de la stratégie, direction des systèmes d'information, direction de l'organisation, direction marketing, direction juridique.

L'information stratégique est définie et hiérarchisée, elle est analysée selon quatre niveaux (résumé, synthèse, liens, informations brutes) par des experts métier et elle est structurée et présentée selon un mode qui permet également de visualiser l'information manquante.

iClass gère trois niveaux d'information.

Une cellule est la plus petite unité documentaire et représente le croisement des trois axes. Chaque cellule constitue un point d'accès à l'information. La cellule est composée de deux niveaux d'information :

- le cartouche (niveau 1) : premier niveau de description des informations stratégiques ; il mentionne le sujet (titre, élément du résumé...) et les critères de référence associés. Une codification spécifique indique l'existence d'actualités et/ou de liens ;

- la synthèse (niveau 2) : synthèse de l'ensemble des types d'information liés à la cellule. Chaque synthèse est structurée en chapitres, eux-mêmes composés d'analyses, de liens et d'actualités.

Le troisième niveau de la base de connaissance est représenté par l'ensemble des sources d'information liées à une cellule (documents sources, liens, informations associées).

Techniquement, iClass est une solution 100% Microsoft, basée sur SQL Server.

La version 3.0 d'iClass offre un certain nombre de fonctionnalités et d'outils parmi lesquels une base de données multidimensionnelle, des tableaux croisés dynamiques, un outil de comparaison de données, un module de veille semi-automatisée, une gestion des mises à jour, la gestion des actualités, la gestion des profils, un module de statistiques.

Des modules de workflow et de signature électronique ont été introduits. La messagerie est également intégrée.

Par ailleurs, des services associés sont proposés, notamment l'alimentation des bases de données, les mises à jour des bases, la gestion des contenus, la maintenance applicative, l'hébergement sécurisé et la formation des utilisateurs.

La méthodologie iClass se décline en quatre étapes :

- élaboration du schéma directeur et du plan détaillé pour sa mise en œuvre ;
- mise en œuvre fonctionnelle et technique de la solution ;
- veille et collecte de l'information. Suivi de la solution ;
- hébergement et diffusion de l'information.

Pour mettre en œuvre ce système de veille et de gestion des informations stratégiques, des consultants réalisent une première étape d'audit sur place consistant en une étude des besoins et attentes (entretiens utilisateurs, tables rondes...) de façon à définir le plan d'action fonctionnel, organisationnel, technique et budgétaire.

Il est procédé à une analyse des typologies d'information, à une cartographie des sources d'information et de leur format qui donnent lieu à un plan de classement, un plan de collecte et un plan de reprise, une analyse des profils utilisateurs (lecteur, éditeur, administrateur) qui permet de définir la procédure de fonctionnement et de gestion des accès. La mise en œuvre de la méthodologie suivie passe également par :

- une analyse des cycles de mises à jour ;
- une analyse des flux ;
- une analyse des modes de restitution ;
- une analyse des besoins comparatifs ;

- une analyse des besoins en termes de veille ;
- une analyse des supports et modes de collecte des avis utilisateurs ;
- une étude technique ;
- la mise en place d'outils de surveillance ;
- l'élaboration d'un plan de formation et d'un plan de communication ;
- comités de pilotage / groupe projet...

La méthodologie s'appuie sur des fiches projets spécifiques, qui comprennent pour chaque domaine une liste d'action précise constituée des éléments suivants : objectifs, méthodologie, prérequis, planification, calendrier, acteurs, charges humaines, charges matérielles, budget prévisionnel.

A l'issue de la première étape, le client possède 10 fiches projets détaillées (fiche projet n°1 : modalités de pilotage et plan de communication ; fiche projet n°2 : stratégie d'entreprise ; fiche projet n°3 : plan de formation ; fiche projet n°4 : plan d'action technique et interfaces ; fiche projet n°5 : plan d'action organisationnel et fonctionnel ; fiche projet n°6 : élaboration du plan de classement ; fiche projet n°7 : réseau de veille ; fiche projet n°8 : modalités ASP ; fiche projet n°9 : reprise de l'existant ; fiche projet n°10 : gestion des accès aux sources d'informations).

L'objectif est également de participer à une démarche qualité. Ainsi, une réunion avec la direction de la qualité, la prise en compte des règles de qualité et la définition du plan qualité du projet est effectuée. Ceci aboutit à la rédaction d'un SLA (charte qualité), d'une note de cadrage du projet, de tableaux de bord de suivi, de bilans.

La deuxième étape consiste en la mise en œuvre fonctionnelle et technique de la solution : création de l'instance iClass, procédure de gestion des sources, alimentation des bases. Lors de cette phase, il s'agit également de développer les interfaces avec les systèmes d'information existants.

Après l'étape de mise en œuvre de la solution fonctionnelle et technique, il s'agit de faire vivre la solution globale de gestion des informations stratégiques, à l'appui du réseau de veille.

Parallèlement, il s'agit de gérer les modalités d'hébergement et de maintenance de la solution.

4.2.4 CCINET

CCINET, l'extranet des 179 chambres de commerce et d'industrie, est développé par l'Assemblée Française des Chambres de Commerce et d'Industrie (ACFCI) et fédère 32 intranets qui correspondent à la mosaïque des métiers exercés par les chambres.



Figure 1 : Page d'accueil de CCINET

Le modèle économique de CCINET a été établi pour permettre aux chambres, quelle que soit leur taille, de rentabiliser leur contribution à partir d'un gain de temps individuel situé entre 70 et 75 minutes par an et par utilisateur :

- pour la Chambre de Commerce et d'Industrie d'Ambert, 72 minutes par an gagnées par individu ;
- pour la Chambre de Commerce et d'Industrie de Toulouse, 74 minutes par an gagnées par individu.

En corollaire, la contribution financière est plus élevée pour les chambres ayant un poids économique plus important.

Les fondamentaux de CCINET sont les suivants :

- placer l'individu au cœur de l'architecture du système (l'annuaire joue, à cet égard, une importance stratégique : chaque individu a son profil utilisateur) ;
- donner à l'individu du pouvoir dans le système (c'est l'individu qui doit déclarer ses fonctions et ses compétences. Il a également le pouvoir de produire de l'information) ;

- être débarrassé des contraintes techniques (les informations sont produites via des formulaires qui, dès qu'ils sont validés, mettent à jour en temps réel les pages web du site).



Figure 2 "Mon CCINET"



Figure 3: consultation de la fiche individuelle

La contrepartie de ce pouvoir individuel et de cette confiance donnée a priori est que toute information et toute action sont signées.

En ce qui concerne les annuaires, ceux-ci sont adaptés à chaque type de recherches.

Il existe :

- un annuaire transversal ;
- des annuaires régionaux par métiers et thématiques.

Les recherches s'effectuent :

- par coordonnées, recherches inversées (nom, prénom, mél, téléphone, fax...);
- par compétences et fonctions.

L'objectif est de trouver aisément les homologues et experts et d'améliorer la qualité du conseil.

Une recherche permet de constituer instantanément une liste de distribution pour envoyer un message électronique en une seule fois à l'ensemble des personnes présentes dans le résultat.

Annuaire

Recherche par coordonnées Recherche par compétences Recherche par fonctions Recherche avancée

Il y a 3 membre(s) de CCINET correspondant à vos critères de recherche

Elus Collaborateurs Autres ou non précisé

Pour envoyer un e-mail groupé :
 1. Sélectionnez vos destinataires en cochant les cases
 2. Préparez votre message en cliquant sur "Créer un message"

Membres 1 à 3 Tout sélectionner

Nom - Prénom	Téléphone	C(R)CI	E-mail
HAIGH Roger	05 53 05 08 11	Perigueux	<input type="checkbox"/> r.haigh@ecole-de-savignac.com
VEILLE Pierre	0240446315	Pays-de-la-Loire (C.R.C.I)	<input type="checkbox"/> p.veille@paysdelaloire.cci.fr
ZONANA André	0491393484	Marseille	<input type="checkbox"/> andre.zonana@marseille-provence.cci.fr

Tout sélectionner

Les champs en gras sont obligatoires.

De : CCINET Admin (ccinet-admin@cdci.cci.fr)

A : J'ajoute/modifie les emails destinataires (emails séparés par des ";" ou ",") :
 r.haigh@ecole-de-savignac.com; p.veille@paysdelaloire.cci.fr; andre.zonana@marseille-provence.cci.fr;
 Je veux être également destinataire de ce message pour en garder une trace.

Copie cachée : J'ajoute/modifie les emails destinataires cachés (emails séparés par des ";" ou ",") :

Objet :

Corps du message : Le 2/12/2003

Fichiers attachés à l'email
 Lien(s) vers le(s) fichier(s) physique(s)

Attention : la taille maximale de l'ensemble des fichiers est fixée à 300Ko.
 Merci de ne pas attacher plusieurs fichiers de même nom.

Figure 4 :
Recherche dans l'annuaire
Et contact direct par mail

CCINET offre également à chacun le droit de produire.

Le rôle du management est, à cet égard, essentiel. Il s'agit :

- d'aider l'utilisateur à se placer (comme consommateur mais aussi et surtout comme producteur) ;
- jouer sur deux ressorts (la valorisation et l'obligation) ;
- et sur deux axes (individuel et collectif).

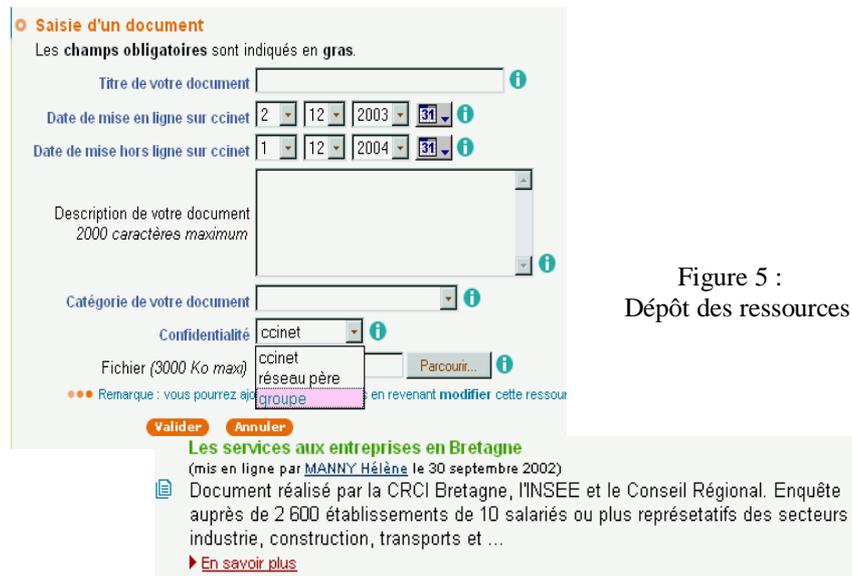


Figure 5 :
Dépôt des ressources

Figure 6 : signature des ressources

En ce qui concerne les obligations :

- sur le plan collectif, en cas de non-contribution au financement du système, il y a un accès restreint pour l'ensemble des collaborateurs du client ;

- sur le plan individuel, chacun doit déclarer dans l'annuaire sa fonction et ses compétences pour pouvoir accéder aux listes de diffusion et à l'intranet métier.

Par ailleurs, pour inciter à la production, CCINET a mis en place un système de valorisation :

- sur le plan collectif, une fiche d'action met en valeur un résultat collectif tout en valorisant les individus qui ont créé la fiche et ont mis en œuvre l'action ;

- sur le plan individuel, pour toutes les ressources, le nom de celui qui les a fournies apparaît en premier.

Identification de l'action	
Titre de la fiche	Plate-forme d'initiative locale
Numéro de l'action	87
Nom de la C(R)CI à l'origine de l'action	Paris
Rôle(s) de la C(R)CI	Chef de file ou coordonnateur Partenaire technique Partenaire financier Partenaire humain
Date de début	03/05/2000
Date de fin	[Aucune]
Objectif de l'action	Accélérer le développement des jeunes entreprises grâce à un p
Description de l'action	
Résultats	Attendus : Soutenir la création de 25 entreprises en 2002. Obtenus : Soutenir la création de 15 entreprises en 2000, 20 entr 2002
Outils mis en oeuvre	
Cible(s)	PME/PMI TPE (très petites entreprises) Porteurs de projets
Conditions de	
Cadre juridique et financier	
Budget global	110 Euros TTC
Contact opérationnel de l'action	
C(R)CI	Paris
Créateur de la fiche	
C(R)CI	Paris
Domaines, Spécialités, Réseaux	
Domaine(s) d'intervention	Création/Transmission d'entreprise

Figure 7 : consultation d'une fiche d'action – aperçu -

Le besoin de relais intermédiaires entre le management et l'utilisateur, le besoin de délégation et d'organisation, l'importance des chaînons manquants sont déterminants.

Pour cela, l'imbrication entre les éléments techniques et les éléments organisationnels est rendue volontairement très forte dans CCINET et au travers de deux types de relais humains :

- des correspondants des chambres : une chambre ne peut pas être déployée s'il n'y a pas au moins un correspondant formé ;
- des animateurs métier : un intranet métier ne peut pas être techniquement ouvert s'il n'y a pas d'animateur.

Ces relais humains mettent en oeuvre deux principes clés d'organisation :

- chaque C(R)CI est responsable des données qui lui sont directement liées. Cette responsabilité répartie sur trois rôles : l'utilisateur lui-même, l'administrateur principal et l'administrateur de messagerie locale ;
- l'ACFCI valide la cohérence d'ensemble des informations et donc la qualité de l'ensemble. La responsabilité est répartie sur trois rôles : l'animateur de réseau, l'animateur délégué et l'administrateur fédéral.

En synthèse, la responsabilité de la mise à jour des informations s'articule autour de six rôles :

- l'administrateur C(R)CI est responsable des coordonnées des élus et collaborateurs de sa chambre ;
- la modification des coordonnées et compétences liées à l'utilisateur sont dévolues à l'utilisateur ;
- l'administrateur de messagerie C(R)CI est responsable des coordonnées liées aux individus de son annuaire de messagerie ;
- le support utilisateur au niveau local de la chambre est assuré par le correspondant CCINET ;
- l'organisation du contenu de CCINET et des échanges d'information est de la responsabilité de l'animateur réseau et des animateurs délégués ;
- la validation de l'ensemble de données liées à CCINET est de la responsabilité de l'administrateur fédéral ACFCI.

En ce qui concerne les listes de diffusion :

- l'animation des forums questions/réponses via la messagerie électronique est contrôlée par un animateur par liste ;
- pour participer, il faut s'abonner ;
- il y a un e-mail par liste, identique à tout abonné ;
- une liste par métier, par groupe, par projet ;
- chaque message est archivé (recherche chronologique et recherche par mots clés).

Ce système permet un partage rapide et efficace des savoir-faire et un gain de temps au quotidien.

D'une façon plus générale, l'objectif est de faire de l'e-mail un partenaire indissociable de CCINET et inciter les utilisateurs à aller naviguer dans CCINET.

Par exemple, la page d'accueil d'un intranet réseau/projet/régional ou groupe peut être directement envoyée par l'animateur aux abonnés concernés.

Elle se transforme ainsi en une newsletter arrivant directement dans leur boîte aux lettres et permettant, par de simples clics à l'intérieur du message, d'aller dans CCINET.



Figure 8 : Page d'accueil et newsletter animateur

Au cas où l'animateur n'aurait pas envoyé cette newsletter, une lettre hebdomadaire est adressée automatiquement et récapitule toutes les nouvelles ressources déposées pour chaque réseau projet et groupe auquel l'utilisateur est inscrit. Un simple clic à l'intérieur du message permet d'accéder immédiatement à l'information.

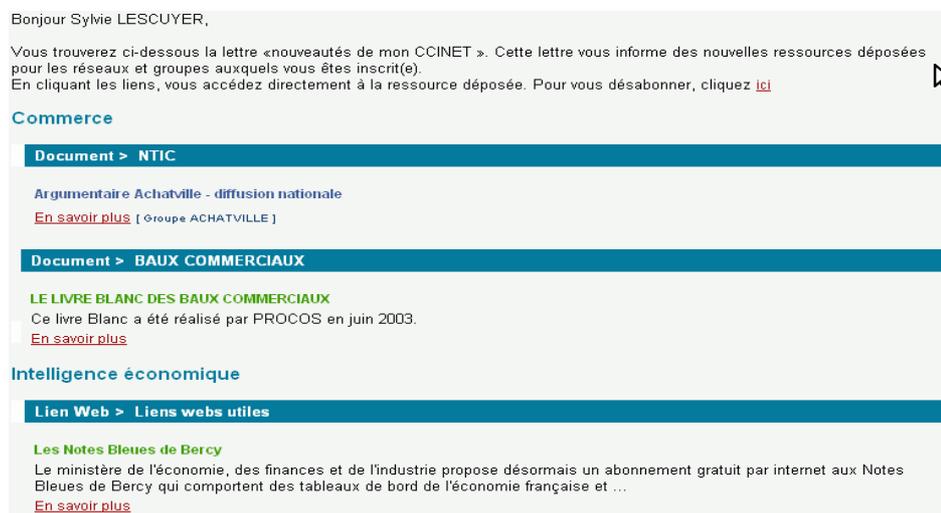


Figure 9 : Newsletter automatique

En résumé, CCINET a donc pris les options suivantes :

- donner du pouvoir à l'individu reconnu par sa hiérarchie ;
- créer des relais entre le management et l'individu ;
- organiser un couplage des fonctions de publication et de diffusion.

En ce qui concerne la collecte des fiches d'action, elle est menée sous forme d'enquête, dont sont informés au préalable les directeurs généraux ainsi que les répondants.

L'enquête applique les méthodes classiques du marketing direct :

- le répondant doit s'assimiler à la cible ;
- le message doit être simple et percutant ;
- une date limite de réponse est imposée ;
- des relances sont mises en œuvre ;
- un engagement de restitution de la part de l'enquêteur est requis ;
- la collecte doit être perçue comme étant utile sur le plan individuel et collectif.

5. LES MODELES ECONOMIQUES

5.1 LES OBJECTIFS

En matière de gestion des connaissances, le groupe de travail a identifié plusieurs modèles économiques parmi lesquels :

- le modèle économique de l'acquisition des connaissances ;
- le modèle de la gestion de la base de connaissances ;
- le modèle de la diffusion de la base de connaissances.

Il est apparu nécessaire d'évaluer :

- le coût de l'acquisition et de la capture des connaissances ;
- le coût de la valeur des connaissances ;
- le coût de la diffusion des connaissances ;
- le coût de la valorisation des connaissances ;
- le retour sur investissement (temps gagné par an et par individu grâce à l'outil de gestion des connaissances)...

et de réaliser le modèle économique par rapport aux :

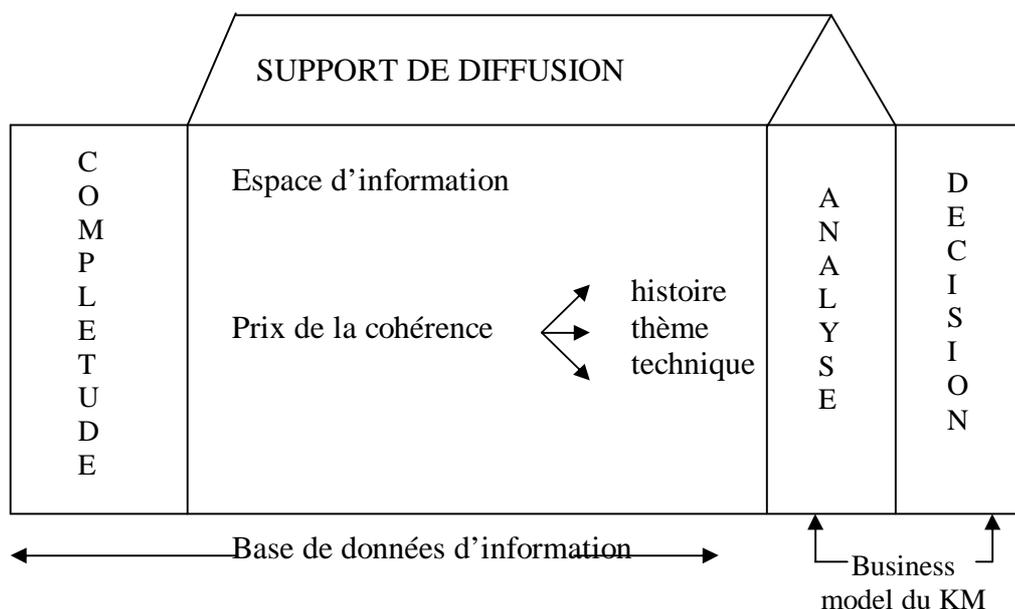
- moyens internes (l'entreprise) ;
- moyens externes (le marché).

5.2 LES CONCEPTS DE VALEUR

Traditionnellement, on distingue quatre types de valeur :

- la valeur d'usage (théorie générale Adam Smith) ;
- la valeur de création (valeur coût) ;
- la valeur de marché (valeur liée à la concurrence sur le marché) ;
- la valeur d'option (liée au risque).

En matière de gestion des connaissances, la mesure de la valeur nécessite en plus de prendre en compte des concepts de valeur spécifiques qu'il est possible de présenter schématiquement :



Quatre valeurs ressortent de ce schéma :

- la valeur intrinsèque ;
- la valeur objective ;
- les gains de productivité ;
- le retour sur investissement.

- La valeur intrinsèque

La valeur intrinsèque comprend la valeur de la complétude, la valeur de l'information et la valeur de l'analyse de l'information.

La valeur de la complétude est très élevée car il est difficile de rassembler toutes les informations pertinentes. Le développement du coefficient de complétude et du niveau de sécurité est donc important.

Dans un cabinet d'avocat par exemple, il est de plus en plus fréquent de vendre de la complétude et non de l'information, la complétude étant déterminante¹.

¹ L'un des moyens permettant d'assurer la complétude consiste à utiliser des micros bases à l'intérieur des grandes bases.

Par ailleurs, comme dans de nombreuses autres professions, le client détient l'information et souhaite donc uniquement que lui soit apportée une solution. Ainsi il y a bien une distinction à faire entre la valeur de l'information et la valeur de l'analyse de cette information.

Par exemple, un abonnement à un journal (source d'information) constitue une valeur d'information qui comprend :

- le coût de production de l'information ;
- le coût de sa mise en ligne ;
- les redevances dues à l'éditeur ;
- la marge du diffuseur.

Si l'internet conduit à une démocratisation de l'accès à l'information et peut donner l'impression que l'information est dévalorisée (gratuité fréquente de l'information), les outils de gestion des connaissances permettent de valoriser cette information grâce à son analyse.

- la valeur objective

La valeur objective de l'information a un coût d'opportunité en ce qu'elle permet de faire un choix.

Il en est ainsi, par exemple, de l'information boursière reçue en exclusivité ou avant l'ouverture des marchés, ou encore de l'information exclusive permettant de réaliser ou d'éviter un investissement.

La valeur peut être déterminée, en outre, en fonction de la nature dominante, ou non, de l'information. Il est ainsi envisageable de pondérer la valeur de l'information par le poids de l'auteur dans sa capacité à émettre une opinion.

Il pourrait être proposé aussi d'imaginer un business model dans lequel la valeur de l'analyse serait déterminée par l'exactitude de cette analyse. La valeur dépendrait donc de la décision prise.

Le business model de la décision peut être défini comme suit :

Coût de l'analyse = temps passé + licence d'utilisation

- les gains de productivité

Concernant les gains de productivité, la matrice d'analyse coût/bénéfice suivante est proposée :

- salaire horaire moyen des utilisateurs du portail intranet ;
- nombre d'utilisateurs ayant accès à l'intranet ;

- temps gagné dans la recherche d'informations hebdomadaires (comparer l'accès à des informations disparates par rapport à une base de données) ;
- nombre de semaines de travail effectif ;
- coûts annualisés portail + achat d'informations ;
- retour sur investissement.

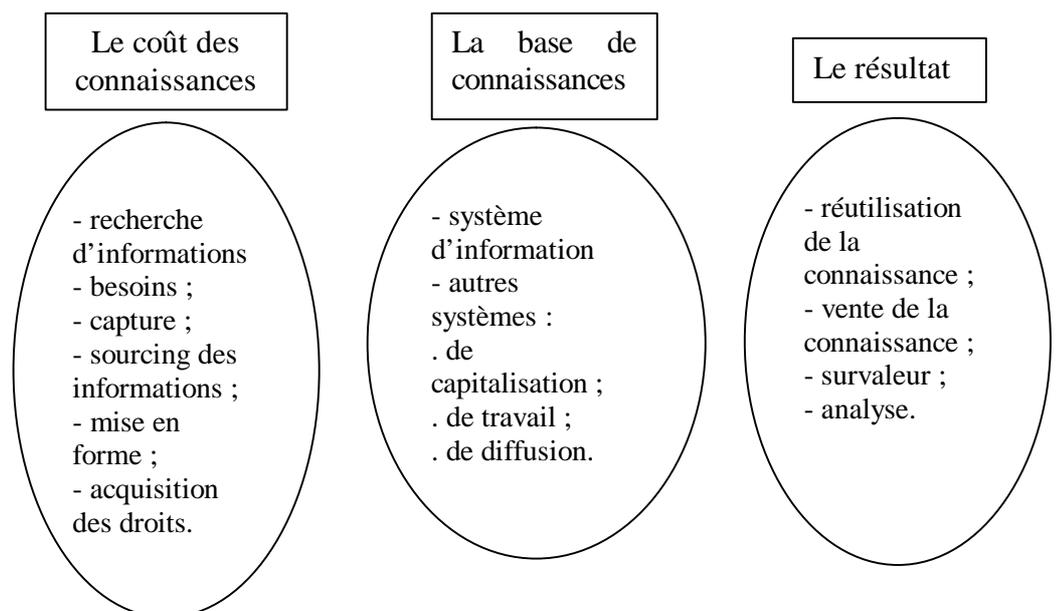
A partir d'un certain niveau le système peut être "nuisible" et ne permet pas de réaliser des gains de productivité. En effet, il apparaît que plus la base de données est intelligente, moins elle est accessible.

En outre, plus l'information disponible est importante, plus le temps de réalisation est court et plus l'information disponible est importante, plus le temps nécessaire à l'analyse se réduit.

On peut également constater que :

- le temps gagné grâce à l'information disponible est perdu en raison de sa non-pertinence ;
- les gains sont inversement proportionnels à la cohérence.
- le retour sur investissement

La somme des coûts des connaissances et des coûts de gestion de la base doit être inférieure au prix du résultat vendu.



Il faut également tenir compte du Maintien en condition opérationnelle (MCO), c'est-à-dire du maintien dans le temps de la pertinence de la connaissance, ce qui suppose :

- la mise à jour ;
- la consolidation ;
- la veille.

5.3 LA FACTURATION

La prestation proposée consiste à offrir des solutions d'accès à des contenus. La question des critères à prendre en considération pour déterminer :

- le prix à payer (unités de mesure et leur agrégation) ;
- le mode de facturation,

se pose donc.

5.3.1 Les critères de facturation

Les unités de mesure varient en fonction du type de connaissances fournies.

- les résultats obtenus

De manière générale, il y a en effet un frein psychologique à payer un résultat qui peut être considéré comme une non réponse.

A titre d'illustration, on peut prendre l'exemple d'une entreprise qui souhaite déposer une marque. Elle devra vérifier si le signe choisi est disponible et pour cela il lui est nécessaire de commander une recherche d'antériorité de marque. Lorsqu'aucune antériorité n'est identifiée l'entreprise peut être réticente à régler les frais de cette recherche. Or, l'absence de réponse constitue en tant que telle une information essentielle sur une marque.

- le critère de l'utilisation

Plusieurs éléments peuvent être pris en compte, tels que :

- l'écran et le téléchargement (en kilooctets) ;
- la question ;
- le couple recherche/réponse ;
- l'usage limité.

- la sécurité

La sécurité notamment vis à vis de la concurrence est essentielle et doit être valorisée.

Il s'agit notamment de garantir qu'aucune autre personne n'aura accès aux données consultées par l'entreprise. Le fait de verrouiller l'information garantit la confidentialité totale.

- le nombre d'utilisateurs et leur qualification ;
- la qualité des liens ;
- la nature des documents (difficulté de disponibilité...) ;
- la rareté (le prix de la valeur dépend de sa rareté. Or, la définition de la rareté dans un espace de connaissance n'est pas aisée) ;
- la proximité de la décision (l'urgence) ;
- les chrono prix (association du prix et d'une période de temps prédéfinie) ;
- l'accessibilité (permanence ou non) ;
- le critère de pertinence ;
- la qualité du document (haute définition...) ;
- la révision de la réponse ;
- les forums (qui consistent à réunir les experts).

5.3.2 Les modes de facturation

Les modes de facturation suivants ont été identifiés :

- possibilité d'offrir un système d'information à une entreprise n'ayant pas les moyens d'investir dans une telle solution, et ce contre un pourcentage sur son chiffre d'affaires ;
- droit d'accès (login, password...) ;
- droit d'usage ;
- forfait ;
- bouquet ;
- tranche ;
- pay per view (paiement à l'article) ;
- porte-monnaie électronique ;
- débit de cartes bancaires.

Il est difficile de mettre en place un système de "coût à la volée", c'est-à-dire lorsque le paiement se fait en même temps que la fourniture de la prestation.

5.4 ILLUSTRATION : MODELES DE TARIFICATION : PRESENTATION D'EXEMPLES MIS EN ŒUVRE SUR LE MARCHÉ ET AU SEIN DU GROUPE LEXISNEXIS

LexisNexis est une entreprise réputée dans le domaine de l'hébergement et la diffusion en ligne de grands contenus juridiques, de presse et d'information sur les entreprises. On désigne aujourd'hui son métier sous le nom d'agrégateur de contenu.

Basée à Dayton (Ohio) et en tant que contractante de l'US Air Force, LexisNexis a créé en 1966 le premier système de recherche électronique permettant des recherches en texte intégral (full-text search engine). En 1973, la société a introduit ce système de recherche auprès des juristes américains et cette commercialisation a bouleversé leurs manières d'effectuer des recherches et d'analyser l'information. Le développement d'un moteur de recherche sur texte intégral rendait possible des recherches sur de gros volumes d'information.

En 2004, plus de 25 000 bases de données juridiques sont disponibles recouvrant plusieurs pays dont la France avec celles du JurisClasseur.

En 1980, LexisNexis a utilisé ce nouveau savoir-faire au domaine de la presse et a progressivement élargi son offre pour compter aujourd'hui plus de 11 000 bases de données au sein desquelles on retrouve bien évidemment l'essentiel de la presse internationale et française indispensable aux décideurs.

Une si longue expérience dans le « en-ligne » a permis à LexisNexis, comme à ses concurrents, d'expérimenter différents modes de tarification.

Il y a probablement trois idées-forces qui sous-tendent toute politique de tarification pour les produits diffusés en ligne :

1. Technologie et équipement du marché
2. Les besoins du marché et son niveau d'équipement
3. Les exemples de tarification au fil du temps
4. Modes de paiement
5. Les critères complémentaires
6. Les micro paiements

Avant de détailler ces points, retenons qu'il n'y a donc aucun mode de tarification pérenne et qu'il est nécessaire d'anticiper.

1. Technologie et équipement du marché

Si l'on se rapporte à cet aspect du point de vue du client, il n'est rien de commun entre la période antérieure et postérieure à Internet. Auparavant, il était nécessaire de disposer de terminaux dédiés (Lexis en proposait un rouge magnifique), ou d'un Minitel pour la France. Ces terminaux étaient conçus pour fonctionner à travers des réseaux spécialisés à vitesse plus ou moins grande (type Transpac).

L'introduction progressive des PC, si elle a bien permis l'apparition progressive d'application à accès distant de type « client-serveur » ne modifiait pas le caractère profondément professionnel et coûteux de ces outils.

Internet, standard unique ouvre la voie à une démocratisation profonde de la transmission d'information ou de logiciels et bouleverse par-là même les modèles de tarification établis.

Si l'on se place du côté du fournisseur de l'information, on comprend que les systèmes de facturation (communément appelés back-offices) qui débouchent sur une facture adressée au client, doivent perpétuellement s'adapter à l'évolution technologique. Toutefois, et pour tous les acteurs du marché, ce poste de dépense étant extrêmement élevé, une forte réflexion précède toujours la décision d'investir pour son évolution.

2. Les besoins du marché et son niveau d'équipement

Avec internet, et c'est totalement nouveau, tout consommateur final (end-user) peut devenir un client potentiel. Il est donc indispensable de bien adapter son modèle de tarification au public auquel on destine le service ou le produit. Pour le grand public, le système de micro-paiement doit impérativement être proposé, mais avec un succès très relatif. Tout client final n'est prêt à régler par carte de crédit que si le service en question lui est personnel.

Dans le cas qui nous préoccupe ici, la gestion de la connaissance dans le milieu professionnel, on s'aperçoit qu'il s'agit d'un mode encore mal adapté qui représente une part infime des revenus des agrégateurs. Toutefois, on ne peut que prédire son développement dans les 5 prochaines années.

Concluons ce chapitre en ajoutant qu'Internet perturbe fortement, depuis son origine, l'économie de l'information dans tous les domaines : films, musique, journaux,.....

Dans ce contexte difficile d'une gratuité considérée comme un droit acquis par l'utilisateur final, on comprendra que l'éducation du marché constituera le « grand œuvre » des éditeurs de tous poils pour rappeler que la chaîne de l'information à un coût dont l'utilisateur doit supporter une part, même infime.

3. Exemples de tarification au fil du temps

Comme la plupart des acteurs, LexisNexis est passé au fil du temps et en fonction de la technologie, d'une facturation à l'heure, puis au volume d'information transféré (mesuré en Ko), puis à la question posée, combinée avec un coût à l'impression des documents..

Aujourd'hui, les acteurs utilisent d'ailleurs majoritairement comme base de tarification l'article ou l'unité documentaire diffusé.

On notera par contre que le mix parfait entre l'information elle-même et la technologie reste encore à trouver.

4. Modes de paiement

La plupart des intervenants du marché proposent, dans le cadre de leurs contrats, les modes suivants :

Le « *pay as you go* » : l'utilisateur paie effectivement ce qu'il a consommé

Le forfait contrôlé : le client dispose, pour une durée limitée (12 mois la plupart du temps), d'un accès au service pour un volume convenu en avance. Si ce volume est dépassé, les prestataires mettent en œuvre des modalités de coupure automatiques ou manuelles, voire de repassage automatique au « *pay as you go* ».

les « unités » : le client prépaye, comme pour un porte monnaie électronique, un certain nombre d'unités. Lorsque son compte ne contient plus assez d'unité, il est obligé de réapprovisionner son compte.

5. Les critères complémentaires

La créativité permanente des différents intervenants et la segmentation progressive du marché, qui dépasse désormais les seuls professionnels de l'information (documentalistes, knowledge managers,...) ont abouti progressivement à l'introduction de tarifications par :

Segment de clientèle ;

Type de produits : recherche d'archives, panoramas de presse ;

Durée du contrat en privilégiant les clients qui s'engagent sur le long terme ;

Par nombre de destinataires directs, voire indirects, de l'information ;

Personnalisation des interfaces et outils de recherche.

6. Les micro-paiements

Voici quelques années, le Minitel permettait, dans un contexte français spécifique, une facturation à l'heure variant en fonction du « palier » retenu par l'éditeur (3613, 3614, 3615,...). Ce système apportait une grande sécurité puisque c'était l'opérateur téléphonique – alors unique – qui utilisait la facturation téléphonique pour ajouter les montants dûs aux titre des services minitel. Il était alors simple, tant pour le professionnel que le particulier d'accéder ponctuellement à des services en ligne.

A l'heure actuelle, les paiements effectués par carte bancaire ou par porte-monnaie électronique, type « Click & Buy » représentent une part infime du chiffre d'affaire réalisé par les éditeurs de contenus. On peut penser que, progressivement, les outils de paiement sur Internet seront parfaitement sécurisés et lèveront les dernières inhibitions des clients soi disant potentiels. La véritable difficulté semble plutôt résider, en ce qui

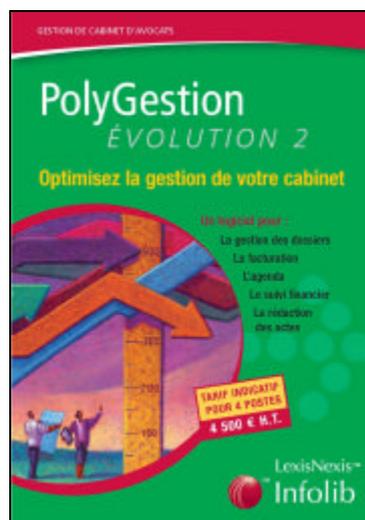
concerne les contenus, dans l'éducation progressive d'un marché qui a oublié que l'information avait un coût.

7. Modèles de tarification au sein de LexisNexis-JurisClasseur

LexisNexis est représenté en France non seulement par son activité de diffusion en ligne d'information juridique et de presse telle que nous l'avons décrite ci-dessus, mais aussi par des activités d'édition (le JurisClasseur) et de SSII (Infolib) qui apportent un éclairage complémentaire sur les modèles tarifaires mis en place dans ces deux activités.

Comme tout éditeur, JurisClasseur doit faire face au délicat équilibre entre son support de diffusion traditionnel, le papier, et Internet. Que ce soit dans le secteur de la presse et de l'édition, il est toujours délicat de trouver des modes de tarification incitatifs, mais permettant une migration des clients sans à-coup pour l'équilibre économique des acteurs. Ce couplage papier/en-ligne est mis en place de manière très différente selon les acteurs du marché : certains font appel, particulièrement dans la presse, à la publicité pour équilibrer la quasi-gratuité de leur titre de presse. D'autres initiatives visent à démarquer au maximum les contenus figurant sur les deux supports ou à réaliser des fonctionnalités (recherche, navigation, liens, fichiers multimédias, recherche d'archives,..) susceptibles d'apporter des avantages différents et permettant au contenu « en ligne » d'avoir une valeur supplémentaire facilement appréhendable par le client. Lorsque cette différence n'existe pas ou n'est pas perçue, la mise en œuvre d'outils de facturation, aussi variés fussent-ils, ne suffit pas à elle toute seule à apporter une solution pérenne et satisfaisante.

Enfin, l'activité de vente de logiciels, activité standard de la SSII Infolib, nous fait revenir aux modèles traditionnels de la vente de logiciels, à savoir, le nombre de licences vendues en fonction du nombre de postes installés.



8. Conclusion

Sur un sujet en permanente évolution (voir plus haut), il est extrêmement délicat, voire prétentieux, de conclure. Les initiatives qui naissent en France ou à l'étranger sont toutes marquées par le souci de coller aux besoins d'un marché qui devient, Internet aidant, plus complexe à appréhender, plus exigeant en termes de service, plus imprévisible et moins fidèle.

Dans ce contexte mouvant, les services en ligne (particulièrement les contenus) accentuent la porosité entre marchés professionnels (B to B) et marchés des particuliers (B to C), contraignant par là même les acteurs du marché à investir dans des solutions de paiement complexes et coûteuses dont le retour sur investissement peut s'avérer parfois risqué, voire hypothétique.



T A R I F S

	FORMAT	ARCHIVES	BIBLIOTHEQUE	FICHIER
Presse nationale en Français				
Les Echos	Texte intégral	Depuis le 04/01/1993	NEWS	LESECH
Le Monde	Texte intégral	Depuis le 01/01/1990	NEWS	MONDE
La Tribune	Texte intégral	Depuis le 01/04/2002	NEWS	TRIBN
Le Figaro	Texte intégral	Depuis le 09/01/1997	NEWS	FIGARO
Libération	Texte intégral	Depuis le 01/01/2000	NEWS	LIBERTN
L'Humanité	Texte intégral	Depuis le 01/01/2002	NEWS	HUMANT
L'Express	Texte intégral	Depuis le 25/07/2002	NEWS	LXPRES
Le Point	Texte intégral	Depuis le 07/01/1995	NEWS	LPOINT
L'Expansion	Texte intégral	Depuis le 01/03/2001	NEWS	XPSION
Le Nouvel Economiste	Texte intégral	Depuis le 01/09/2003	NEWS	NOUVEC
Le Temps (Suisse)	Texte intégral	Depuis le 01/01/1998	NEWS	TEMPFR
Presse régionale en Français				
Sud Ouest	Texte intégral	Depuis le 07/05/1994	NEWS	SOUEST
Le Télégramme de Brest	Texte intégral	Depuis le 01/02/2002	NEWS	TELBST
Bref Rhône Alpes	Texte intégral	Depuis le 01/08/2002	NEWS	BREFRA
Principaux journaux européens				
Financial Times (UK)	Texte intégral	Depuis le 02/01/1982	NEWS	FINTME
The Times (UK)	Texte intégral	Depuis le 01/07/1985	NEWS	TTIMES
The Economist (UK)	Texte intégral	Depuis le 04/01/1975	NEWS	ECON
The Independent (UK)	Texte intégral	Depuis le 19/09/1988	NEWS	INDPNT
The Guardian (UK)	Texte intégral	Depuis le 14/07/1984	NEWS	GUARDN
The Mirror (UK)	Texte intégral	Depuis le 29/05/1995	NEWS	MIRROR
The Daily Telegraph (UK)	Texte intégral	Depuis le 30/10/2000	NEWS	TELEGR
Frankfurter Allg. Zeitung (All)	Texte intégral	Depuis le 02/01/1993	NEWS	FAZ
El Mundo (Esp)	Texte intégral	Depuis le 01/06/2002	NEWS	MUNDO
El País (Esp)	Texte intégral	Depuis le 19/04/1996	NEWS	ELPAIS
Cinco Dias (Esp)	Texte intégral	Depuis le 01/01/1998	NEWS	CINCOD
Expansion (Esp)	Texte intégral	Depuis le 02/01/1998	NEWS	RECOLS
La Stampa (Ita)	Texte intégral	Depuis le 02/01/1992	NEWS	LSTMPA
Il Sole 24 Ore (Ita)	Texte intégral	Depuis le 02/01/1991	NEWS	ILSOLE
De Telegraaf (PB)	Texte intégral	Depuis le 02/01/1999	NEWS	DETELE
Principaux journaux américains				
The Washington Post (US)	Texte intégral	Depuis le 01/01/1977	NEWS	WPOST
The New York Times (US)	Texte intégral	Depuis le 01/01/2002	NEWS	NYT
Los Angeles Times (US)	Texte intégral	Depuis le 01/01/2002	NEWS	LAT
USA Today (US)	Texte intégral	Depuis le 01/01/2002	NEWS	USATDY
Chicago Tribune (US)	Texte intégral	Depuis le 01/01/1985	NEWS	CHTRIB

6. LES ASPECTS JURIDIQUES DE LA GESTION DES CONNAISSANCES

Le groupe de travail a également consacré son étude aux aspects juridiques liés à la réalisation et à l'utilisation d'un outil de gestion des connaissances.

En effet, pour que cet outil permette une capitalisation efficace du savoir-faire et du savoir-agir de l'entreprise, pour qu'il constitue un véritable outil d'aide à la production intellectuelle et à la décision, il faut un encadrement juridique fort.

6.1 DISPOSER DE CONTRATS PROTECTEURS

Les principaux contrats sont :

- le contrat de conseil ;
- le contrat de licence d'utilisation de l'outil ;
- le contrat de formation ;
- le contrat de réalisation de développements spécifiques ;
- le contrat de maintenance de l'outil.

6.1.1 Le contrat de conseil

Il existe sur le marché une gamme très étendue d'outils de gestion des connaissances. Il peut être utile, en amont, de faire appel à un prestataire chargé de conseiller le client sur les produits qui paraissent les plus adaptés à ses besoins.

Le contrat de conseil aura donc pour objet de définir les conditions et les modalités selon lesquelles le prestataire s'engage à exécuter les prestations d'étude et de conseil demandées par le client sur la base d'un cahier des charges.

Dans la majeure partie des cas, les études et les conseils donnent lieu à l'élaboration d'un rapport écrit qui sera remis au client.

Le prestataire doit fournir un document compréhensible par l'utilisateur en langage clair. En effet, l'utilisateur non averti n'est pas capable de valider une étude qui ne serait accessible qu'aux seuls spécialistes.

Il est donc important d'envisager les conditions d'exécution d'un contrôle de pertinence de la prestation réalisée avec les objectifs définis à l'origine. Dans cette hypothèse, il est essentiel de prévoir une procédure qui permettra ce contrôle de conformité. Pour cela, il conviendra de définir:

- une méthode d'évaluation des études et des conseils ;
- un délai pour les vérifier ;
- une procédure de reprise des éléments erronés.

Dans le cadre d'une non-conformité de l'étude aux objectifs assignés, il sera nécessaire de mettre en place une procédure permettant de gérer cette non-conformité, en prévoyant éventuellement le remboursement des sommes versées ou encore l'organisation d'une procédure de conciliation.

En outre, si le rapport final ainsi que l'analyse qui l'a précédée font preuve d'originalité au sens du Code de la propriété intellectuelle, ils pourront faire l'objet d'une protection par le droit d'auteur. Il conviendra donc de gérer au sein du contrat les droits de propriété sur l'étude et de se préoccuper du statut des résultats de cette étude.

Par ailleurs, le client est tenu à une obligation de collaboration. Cette obligation est d'autant plus importante que le succès d'une étude dépend de la communication d'informations précises par le client sur son organisation et son fonctionnement.

Ce devoir de collaboration du client, qui prend naissance dès la phase précontractuelle, se prolonge au cours de l'exécution de la prestation.

Une fois l'étude réalisée, le client va entrer dans une phase de sélection des produits et des prestataires.

Il est alors recommandé d'adresser aux prestataires présélectionnés des prérequis techniques ainsi que des prérequis juridiques.

La formulation de prérequis juridiques permettra en effet au client de connaître les engagements que son futur cocontractant sera prêt à prendre, par exemple en termes de responsabilité.

Cela permettra également au client, au vu des réponses, de mettre en place une stratégie de négociation contractuelle.

Les prérequis juridiques pourront en effet être présentés comme constituant des principes substantiels pour le client qui devront donc figurer dans le contrat qui sera signé avec l'entreprise retenue.

Quel est ce contrat ? Il s'agit le plus souvent du contrat de licence d'utilisation de l'outil de gestion des connaissances.

6.1.2 Le contrat de licence d'utilisation de l'outil

La finalité de ce contrat est la mise à disposition d'un droit d'utilisation.

Ce contrat devra donc comporter notamment une clause relative à l'étendue de la licence qui énumèrera précisément les droits concédés à l'utilisateur et déterminera les contours :

- droit de reproduction cédé ;
- droit d'utilisation cédé ;
- droit d'adaptation cédé ;
- droit de représentation cédé.

En pratique, l'outil fait le plus souvent l'objet d'une simple cession du seul droit d'utilisation, à titre bien entendu non exclusif.

Un article "garantie de jouissance paisible" précisant que l'éditeur garantit le licencié contre toutes les actions en contrefaçon dirigées contre l'outil pourra également être prévu dans le contrat.

6.1.3 Le contrat de formation

Ce contrat de formation définit les conditions dans lesquelles le prestataire fournit au client des prestations de formation.

Ce contrat pourra comporter des clauses relatives :

- au déroulement de la formation (préparation, locaux, moyens logistiques, feuille de présence...);
- aux conditions d'annulation ;
- à la propriété des supports de formation ;
- à l'assurance qualité formation ;
- aux conditions financières.

6.1.4 Le contrat de réalisation de développements spécifiques

Ce contrat définit les relations entre le client et l'éditeur concernant la réalisation de développements adaptés aux besoins spécifiques du client.

En ce qui concerne l'acquisition d'une œuvre soumise au droit d'auteur (les développements spécifiques), pourra figurer au sein du contrat une clause de cession de droits de propriété littéraire et artistique. Les conditions de cette cession devront y être précisées. Il s'agira de déterminer l'étendue, la nature, la destination et la durée des droits cédés.

Une clause "recette" (ou réception) devra également être stipulée. La phase de réception permet de vérifier la conformité des développements aux spécifications approuvées par le client. Elle se décompose en deux étapes :

- une recette provisoire correspondant au contrôle de conformité sur test ;
- une recette définitive correspondant au contrôle de conformité en charge.

En outre, une clause "garanties" pourra, d'une part, déterminer la période durant laquelle le prestataire s'engage à corriger toute anomalie de fonctionnement ou non-conformité aux spécifications et, d'autre part, prévoir une garantie d'évolution et de compatibilité des développements spécifiques.

6.1.5 Le contrat de maintenance

Une fois l'outil de gestion des connaissances intégré dans l'entreprise, les droits sur son contenu gérés et les démarches légales concernant son déploiement réalisées, des dysfonctionnements peuvent toutefois apparaître, d'où la nécessité de mettre en place, en amont, une maintenance en condition opérationnelle qui pourra être assurée en interne si l'entreprise dispose des ressources nécessaires ou, dans le cas contraire, par une entreprise extérieure.

Dans cette dernière hypothèse, un contrat de maintenance devra être conclu et indiquer notamment la nature de la maintenance (préventive et/ou évolutive et/ou corrective) et les modalités d'intervention.

6.2 GERER LES DROITS SUR L'OUTIL DE GESTION DES CONNAISSANCES ET SON CONTENU

6.2.1 Les droits sur le contenu de l'outil : la protection par le droit d'auteur

Le producteur ou l'éditeur, lorsqu'il réalisera l'outil de gestion des connaissances, devra respecter les droits que pourraient détenir des tiers sur les éléments intégrés dans cet outil. Il en est de même des utilisateurs amenés à produire du contenu destiné à être diffusé au sein de l'outil.

6.2.1.1 Les œuvres protégées par le droit d'auteur

Toute œuvre originale, c'est-à-dire reflétant la personnalité de son créateur ou marquée de l'empreinte de ce dernier, à condition qu'elle soit formalisée, est protégée par le droit d'auteur, et ce quels qu'en soient son genre, sa forme d'expression, son mérite ou sa destination¹.

Cette protection est accordée par la loi du seul fait de la création de cette œuvre, et donc indépendamment de toute procédure éventuelle de dépôt.

L'article L. 112-2 du Code de la propriété intellectuelle énumère, de façon non exhaustive, les œuvres protégées par le droit d'auteur.

Parmi cette énumération on retrouve :

- les écrits à caractère littéraire, artistique, scientifique, journalistique ;
- les dessins, photos, illustrations, plans et croquis ;
- les logiciels ;
- les compositions musicales.

La jurisprudence a reconnu qu'une page de présentation et d'accueil d'un site web pouvait, dès lors qu'elle est originale, être protégée par le droit d'auteur².

¹ Article L. 112-1 du Code de la propriété intellectuelle.

² Tribunal de commerce de Paris, 9 février 1998, Expertises, juillet 1998, page 236.

Par analogie, la page de présentation et d'accueil de l'outil de gestion des connaissances pourrait sans doute bénéficier de la protection par le droit d'auteur, sous réserve d'originalité.

6.2.1.2 Le titulaire des droits sur l'œuvre originale

Aux termes de l'article L. 111-1 du Code de la propriété intellectuelle, l'auteur d'une œuvre de l'esprit jouit sur cette œuvre, du seul fait de sa création, d'un droit de propriété incorporelle exclusif.

L'auteur est donc titulaire de l'ensemble des droits sur l'œuvre originale créée par lui.

Mais, en est-il de même lorsque cette œuvre lui a été commandée par un tiers ? En est-il de même lorsque cette œuvre a été réalisée dans le cadre de l'exécution de son travail et d'après les instructions de son employeur ? Par ailleurs, comment appréhender cette question de la propriété incorporelle sur l'œuvre lorsque plusieurs personnes ont concouru à la réalisation de cette œuvre ?

- L'œuvre de commande

Le fait de commander et/ou de financer la réalisation d'une œuvre ne confère pas au donneur d'ordre de droits de propriété intellectuelle sur cette œuvre.

Par conséquent, même dans le cadre d'une œuvre spécifique réalisée sur commande de l'utilisateur et conformément aux exigences de celui-ci, en l'absence de stipulations contractuelles spécifiques conformes aux exigences du Code de la propriété intellectuelle en matière de cession de droits d'auteur, les droits sur ladite œuvre demeurent la propriété exclusive de l'auteur.

- L'œuvre composite

L'œuvre composite¹ est l'œuvre "nouvelle" à laquelle est incorporée une œuvre préexistante sans la collaboration de l'auteur de cette dernière².

L'exploitation de l'œuvre composite ne peut avoir lieu sans l'accord du titulaire des droits sur l'œuvre préexistante³.

¹ Tel pourrait être le cas d'une note rédigée par un membre de l'entreprise à partir de documents intégrés dans l'outil de gestion des connaissances.

² Article L. 113-2 du CPI.

³ L. 113-4 du CPI.

Non seulement le titulaire des droits sur l'œuvre composite doit obtenir l'autorisation du titulaire des droits sur l'œuvre originale pour exploiter l'œuvre composite, mais, de plus, il ne peut l'exploiter au-delà des termes (limites territoriales, autorisations pour certains modes d'exploitations seulement, etc...) dans lesquels cette autorisation lui a été donnée.

- L'œuvre collective

Aux termes de l'article L. 113-2, alinéa 3 du Code de la propriété intellectuelle :

- "Est dite collective l'œuvre créée sur l'initiative d'une personne physique ou morale qui l'édite, la publie et la divulgue sous sa direction et son nom et dans laquelle la contribution personnelle des divers auteurs participant à son élaboration se fond dans l'ensemble en vue duquel elle est conçue, sans qu'il soit possible d'attribuer à chacun d'eux un droit distinct sur l'ensemble réalisé".

Trois conditions doivent être réunies pour qu'une personne, physique ou morale, puisse se voir déclarer titulaire de l'ensemble des droits sur une œuvre collective, étant précisé que la jurisprudence adopte une position plutôt extensive quant à la reconnaissance de l'existence d'une œuvre collective :

- cette personne doit avoir eu l'initiative de l'œuvre ;
- elle doit l'avoir éditée ou publiée et donc divulguée ;
- il doit y avoir eu fusion des contributions des différents auteurs aboutissant à l'impossibilité d'attribuer à chacun d'eux un droit distinct sur l'œuvre¹.

- L'œuvre de collaboration

Aux termes de l'article L. 113-2 alinéa 1^{er} du Code de la propriété intellectuelle, l'œuvre de collaboration est celle "à la création de laquelle ont concouru plusieurs personnes physiques".

Pour que l'œuvre puisse être qualifiée d'œuvre de collaboration, il faut :

- que deux ou plusieurs personnes, physiques en principe, aient participé à la création de l'œuvre ;
- que ces personnes l'aient fait de façon concertée, dans une communauté d'esprit et dans un but commun.

Si ces conditions sont réunies, l'œuvre de collaboration est alors la propriété commune des auteurs qui doivent exercer leur droit d'un commun accord².

¹ Cour de cassation, civil 1^{ère}, 3 juillet 1996, JCP Ed. G 1996, IV, n° 2008.

² L 113-3 du CPI.

- L'œuvre, création de salariés

L'employé jouit sur son œuvre, du seul fait de sa création, d'un droit de propriété intellectuelle exclusif et opposable à tous, même si l'œuvre a été créée dans le cadre de l'exécution de son contrat de travail ou d'après les instructions de son employeur, sous réserve que le contrat de travail ne comporte une clause de cession des droits d'auteur.

Parmi toutes les œuvres de l'esprit, seuls les logiciels échappent à cette règle. En effet, les droits patrimoniaux sur les logiciels et leur documentation créés par un ou plusieurs employés "dans l'exercice de leur fonction ou d'après les instructions de leur employeur sont dévolus à l'employeur qui est seul habilité à les exercer"¹.

- L'œuvre, création d'agents publics

En vertu d'un avis du 21 novembre 1972 du Conseil d'Etat, les créations réalisées par des agents de l'Etat dans le cadre de leur fonction sont dévolues à la personne publique pour les besoins du service public. Cet avis s'applique également aux agents contractuels de l'Etat.

Cette position est confortée par l'article 18 du projet de loi relatif au droit d'auteur et aux droits voisins dans la société de l'information² aux termes duquel "(...) le droit d'exploitation d'une œuvre créée par un agent de l'Etat dans l'exercice de ses fonctions ou d'après les instructions reçues est, dès la création, cédé de plein droit à l'Etat".

Cependant, certaines créations réalisées par les agents de l'Etat peuvent contenir ou intégrer des créations préexistantes. Or, ces créations ou œuvres préexistantes peuvent être protégées par le droit d'auteur, sous réserve de leur originalité, ce qui implique l'obtention de l'autorisation pour leur utilisation par le titulaire.

- Conclusion sur la titularité des droits de propriété intellectuelle

Il conviendra donc, à chaque fois que le producteur ou l'utilisateur souhaitera diffuser au sein de l'outil de gestion des connaissances une œuvre réalisée par un tiers d'acquiescer les droits auprès du titulaire de ces droits.

6.2.1.3 Les droits de l'auteur et de l'utilisateur d'une œuvre originale

L'ensemble des droits sur l'œuvre originale appartient à l'auteur qui détient donc sur celle-ci l'intégralité des droits moraux et des droits patrimoniaux³.

¹ L. 113-9 du CPI.

² Portant transposition de la directive du 22 mai 2001 relative à certains aspects du droit d'auteur et droits voisins dans la société de l'information.

³ L. 121-1 CPI.

A ce titre, l'auteur jouit sur son œuvre :

- dans le cadre de son droit moral :
 - du droit à la paternité de son œuvre : l'auteur peut exiger que son nom figure sur son œuvre et sur toute reproduction ou représentation de celle-ci ;
 - du droit de divulgation ;
 - du droit de repentir ou de retrait de son œuvre, à condition d'indemniser préalablement l'éventuel cessionnaire des droits sur cette œuvre ;
 - du droit au respect de son œuvre qui entraîne notamment pour conséquence que l'œuvre ne peut être altérée ni dans sa forme, ni dans son esprit par un tiers ;
- au titre des droits patrimoniaux :
 - d'un monopole de reproduction : seul l'auteur a le droit de fixer matériellement sur un support son œuvre en vue d'une communication ultérieure indirecte au public ;
 - d'un droit de représentation : seul l'auteur a le droit de communiquer au public son œuvre.

L'auteur jouit sur l'ensemble de ces droits d'un monopole exclusif, pendant une durée de 70 ans à compter de son décès, si l'auteur est une personne physique, et à compter de la date de divulgation si l'auteur est une personne morale¹.

Toutefois, le Code de la propriété intellectuelle reconnaît certains droits à l'utilisateur de l'œuvre².

Il prévoit ainsi que l'utilisateur peut :

- représenter l'œuvre dans le cercle de famille (notion interprétée restrictivement par la jurisprudence) à condition que cette représentation soit privée et gratuite ;
- reproduire l'œuvre pour son usage privé, sauf les logiciels et les bases de données électroniques pour lesquels l'exception de copie privée ne peut jouer ;
- procéder à des analyses et courtes citations de l'œuvre ;
- réaliser des revues de presse.

¹ L 123-1 du CPI.

² L. 122-5 du CPI.

Les droits accordés à l'utilisateur par les textes sont donc extrêmement limités. Très souvent, ces droits légaux accordés à l'utilisateur seront insuffisants pour l'usage que l'utilisateur souhaite faire de l'œuvre. Il conviendra alors que l'utilisateur de l'outil de gestion des connaissances acquiert les droits nécessaires à l'utilisation de l'œuvre qu'il entend faire.

6.2.1.4 Acquisition de droits sur une œuvre originale

Le Code de la propriété intellectuelle exige¹, pour que la cession d'un droit de propriété intellectuelle puisse s'opérer valablement, qu'il s'agisse du droit de reproduction, du droit de représentation ou de tout autre droit, que cette cession obéisse aux conditions suivantes :

- chacun des droits cédés doit faire l'objet d'une mention distincte dans l'acte de cession ;
- le domaine d'exploitation des droits cédés doit être délimité quant à son étendue, sa destination, son lieu et sa durée.

Le Code de la propriété intellectuelle énumère en outre un certain nombre d'autres conditions devant être respectées dans le cadre de la réalisation d'une cession de droits d'auteur. Notamment :

- la cession doit, en principe, comporter au profit de l'auteur la participation proportionnelle aux recettes provenant de la vente ou de l'exploitation ; un certain nombre d'exceptions à cette règle est toutefois prévu par le Code de la propriété intellectuelle ;
- la cession globale des œuvres futures est nulle. En vertu de cette disposition, il n'est donc pas possible, pour un auteur, de céder l'ensemble de ses droits sur les œuvres qu'il pourrait être amené à créer, sauf à délimiter cette cession par des critères objectifs.

6.2.1.5 Sanctions

Le non-respect des droits de l'auteur sur son oeuvre constitue une contrefaçon et, à ce titre, un délit sanctionné pénalement et civilement. La contrefaçon est punie de peines pouvant aller jusqu'à deux ans d'emprisonnement et 150 000 euros d'amende pour les personnes physiques, les personnes morales étant susceptibles d'être condamnées à des amendes pouvant aller jusqu'au quintuple de l'amende prévue pour les personnes physiques et de peines complémentaires telles que la dissolution, l'interdiction ou encore l'affichage de l'interdiction.

Il faut préciser que toute utilisation d'une œuvre non autorisée par l'auteur, même si des droits proches de l'acte contrefaisant ont été acquis auprès de l'auteur, constitue une contrefaçon.

¹ L. 131-3 du CPI.

6.2.1.6 Les autres types de protection

Des œuvres protégées par un droit autre que le droit d'auteur sont susceptibles d'être intégrées à l'outil de gestion des connaissances.

Il conviendra que le producteur et/ou l'utilisateur respecte ces divers autres droits soit en obtenant des autorisations de leur auteur/propriétaire, soit en prenant des précautions visant à respecter lesdits droits.

- La protection par le droit de la propriété industrielle

Les marques, les brevets, les dessins et modèles font l'objet de protection par le droit spécifique de la propriété industrielle.

A la différence du droit d'auteur, les droits sur ces éléments de propriété industrielle s'acquièrent par une formalité de dépôt.

Le déposant est en principe titulaire d'un monopole d'exploitation sur ces éléments.

Par conséquent, l'utilisation de ceux-ci sans son autorisation constitue une contrefaçon et est, à ce titre, passible des peines du délit de contrefaçon.

Il a ainsi été jugé le 4 août 1997¹ qu'une société qui utilisait comme mot clé pour que son site soit sélectionné par les moteurs de recherche, des dénominations constituant des marques déposées par une autre société se rendait coupable de reproduction illicite de marque et, à ce titre, engageait sa responsabilité.

Le producteur devra également prendre garde à ne pas reproduire au sein de l'outil de gestion des connaissances des dessins et modèles déposés comme par exemple une charte graphique.

- La protection par l'interdiction de la concurrence déloyale ou parasitaire

La concurrence déloyale est une notion qui a été élaborée par la jurisprudence et qui est fondée sur l'article 1382 du Code civil. Elle permet d'engager la responsabilité civile d'un concurrent dont le comportement, en raison de son caractère particulièrement indélicat ou contraire aux pratiques habituellement pratiquées dans le commerce, caractérise l'usage abusif de la liberté de la concurrence.

La notion de concurrence parasitaire permet, quant à elle, de prendre en compte la pratique qui consiste à profiter des investissements d'autrui pour bénéficier de sa réputation et de son image aux yeux du public, par exemple en se "glissant dans son sillage" et en faisant l'économie des investissements correspondants.

¹ Tribunal de grande instance de Paris, ordonnance de référé rendue le 4 août 1997, REF 97/59360-97/59600.

La société à l'origine du comportement parasitaire n'est pas forcément concurrente de la société "parasitée".

La mise en cause de la responsabilité civile d'un concurrent déloyal ou parasitaire étant fondée sur les règles du droit commun de la responsabilité civile, il appartient au demandeur à une telle action de prouver :

- la faute, c'est-à-dire le comportement déloyal ou parasitaire ;
- le préjudice en résultant, c'est-à-dire la confusion et le détournement de clientèle ainsi engendré et les difficultés commerciales pouvant se manifester par une baisse du chiffre d'affaires ou par un manque à gagner ;
- le lien de causalité entre le comportement invoqué et le préjudice subi.

L'action en concurrence déloyale ou parasitaire est souvent engagée lorsque les conditions de la contrefaçon ne semblent par réunies, par exemple, lorsque seule l'idée à la base de l'œuvre est reprise mais pas l'œuvre elle-même.

Dans ce cas là, les idées n'étant pas protégées par le droit d'auteur, l'action en contrefaçon ne peut aboutir.
L'action en concurrence déloyale peut en revanche être envisagée si les conditions de cette action sont réunies.

La Cour de cassation a ainsi jugé que "l'action en concurrence déloyale peut être intentée pour la protection d'une personne pouvant se prévaloir d'un droit privatif"¹.

Il existe donc un nombre important de moyens juridiques mis à la disposition de l'auteur/créateur/propriétaire qui s'estime lésé dans la jouissance paisible de ses droits.

Le producteur de l'outil devra donc prendre garde à respecter les différents droits existant sur les éléments qui seront intégrés dans l'outil de gestion des connaissances.

6.2.2 Les droits sur le contenu de l'outil : la protection des bases de données

Le contenu de l'outil de gestion des connaissances sera probablement constitué d'un ensemble de données, faits, informations ou documents de nature ou d'origine différente.

¹ Cour de cassation, chambre commerciale, 6 décembre 1984.

En réunissant ces différents documents, le producteur constitue une base de données susceptible d'être protégée, cumulativement ou alternativement, par :

- le droit d'auteur ;
- le droit du producteur, tel que visé dans la loi du 1^{er} juillet 1998.

6.2.2.1 La protection par le droit d'auteur

La protection de la base de données peut porter sur :

- les données elles-mêmes ;
- l'ensemble des données "disposé de manière systématique ou méthodique" appelé "base de données".
- Les données

Le droit d'auteur réserve sa protection aux créations intellectuelles originales, c'est-à-dire celles qui portent l'empreinte de la personnalité de leur auteur. Sont donc protégées par le droit d'auteur, les données qui, par elles-mêmes, satisfont à la condition d'originalité.

En revanche, les informations et données factuelles, dont la nature ne permet guère à la personnalité d'un auteur de s'exprimer, échappent à la protection par le droit d'auteur.

Ainsi la jurisprudence a considéré que les cartes et atlas pouvaient être protégés par le droit d'auteur.

Le Tribunal de grande instance de Paris a en effet jugé qu'une présentation, tout à fait inhabituelle, peut fonder la protection d'une carte par le droit d'auteur, même si celle-ci consiste en une projection qui repose sur des formules mathématiques au moyen desquelles deux paramètres sont soumis à des variations arbitraires¹.

- La base de données

La réunion d'un ensemble complet et ordonné de données peut être protégée par le droit d'auteur, indépendamment des données qu'elles contiennent.

En effet, aux termes de l'article L. 112-3 du Code de la propriété intellectuelle, les bases de données sont protégeables dès lors que "par le choix et la disposition des matières" elles constituent des créations intellectuelles.

¹ TGI Paris, 27 avril 1989, PIBD 1989, n° 463, III, page 489.



La jurisprudence a ainsi admis la protection par le droit d'auteur des recueils d'informations, à condition que ceux-ci présentent une certaine originalité.

Le Conseil d'Etat, par arrêt du 10 juillet 1996¹, a jugé que le répertoire Siren, créé et exploité par l'Insee, constitue non une simple collection de données mais un ensemble organisé et structuré d'informations protégées par le droit d'auteur.

Néanmoins, la preuve de l'originalité dans le choix ou l'ordonnement des informations n'est pas toujours aisée à rapporter.

6.2.2.2 La protection par le droit du producteur

Le Code de la propriété intellectuelle octroie aux producteurs de bases de données "une protection du contenu de la base"², indépendamment de tout droit d'auteur sur cette base.

Le producteur bénéficie d'une protection sur le contenu de cette base lorsque "la constitution, la vérification ou la présentation de celle-ci atteste d'un investissement financier, matériel ou humain substantiel"³.

La durée des droits ainsi conférés au producteur d'une base de données est de 15 ans à compter du 1^{er} janvier de l'année qui suit celle de son achèvement ou celle de sa première mise à disposition du public⁴.

Cependant, lorsque la base de données fait l'objet d'un "nouvel investissement substantiel", sa protection expire 15 ans après le 1^{er} janvier de l'année civile suivant celle de ce nouvel investissement⁵.

Dès lors, les mises à jour de la base de données susceptibles d'être diffusées au sein de l'outil pourront être considérées comme des nouveaux investissements substantiels, allongeant d'autant la durée de protection de la base de données.

¹ Conseil d'Etat, 10 juillet 1996, Expertises 1997, n° 204, page 148.

² Article L. 341-1 du CPI.

³ Article L. 341-1 du CPI.

⁴ Article L. 342-5 du CPI.

⁵ Article L. 342-5 du CPI.

6.2.2.3 Le droit des utilisateurs sur la base de données

Aux termes des articles L. 342-1 et L. 342-2 du Code de la propriété intellectuelle, le producteur d'une base de données peut, que les données soient protégées ou non par le droit d'auteur :

- interdire l'extraction, c'est-à-dire le transfert permanent ou temporaire, du contenu de la base de données sur un autre support, quels que soient le moyen et la forme, de la totalité ou d'une partie substantielle du contenu, la notion de substantialité s'appréciant de façon qualitative ou quantitative ;
- interdire la réutilisation, c'est-à-dire la mise à disposition du public, de la totalité ou d'une partie qualitativement ou quantitativement substantielle du contenu de la base ;
- transférer, céder ou donner en licence les droits visés ci-dessus ;
- interdire l'extraction ou la réutilisation répétée systématique d'une partie qualitativement ou quantitativement non substantielle du contenu de la base lorsque ces opérations excèdent manifestement les conditions d'utilisation normales de la base de données.

Ainsi, le producteur est en droit d'interdire l'extraction substantielle du contenu de sa base ou l'extraction répétée non substantielle du contenu de sa base.

La loi a d'ores et déjà fait preuve de son efficacité.

Ainsi, le Tribunal de commerce de Paris, dans une décision en date du 18 juin 1999¹, a jugé que France Télécom, qui exploitait un service d'annuaire inversé sous forme de base de données bénéficiait d'une protection, en vertu des dispositions de la loi du 1^{er} juillet 1998, sur ladite base de données. France Télécom a ainsi été autorisée à interdire l'extraction non-autorisée d'une partie de sa base de données.

Il a également été jugé que l'extraction de données d'une base de données de notice bibliographique n'était pas autorisée dès lors que l'utilisateur avait signé un contrat excluant la diffusion des bases de données à des tiers².

Le producteur de l'outil de gestion des connaissances, s'il bénéficie de la protection de la loi, pourra ainsi interdire la réutilisation de tout ou partie de sa base de données, dans les conditions susvisées.

¹ TC Paris, 18 juin 1999, Expertises, octobre 1999, page 308.

² TC Paris, 7 mai 1999, RG 99008160.

6.2.2.4 Les modalités de protection des droits : le dépôt

- Le dépôt de marques

Afin de renforcer la protection des bases de données développées par une entreprise, il est conseillé à celle-ci de déposer les noms de ces bases de données en tant que marque auprès de l'Institut national de la propriété industrielle (INPI).

A cet égard, il est impératif de procéder avant tout dépôt de marque à des recherches d'antériorités et de similitudes afin de vérifier la disponibilité du signe retenu.

En effet, l'enregistrement d'une marque auprès de l'INPI confère au déposant un monopole d'exploitation pour l'utilisation de cette marque dans la désignation des produits et services qu'il aura réservés dans le cadre du dépôt.

Il s'agit d'une protection potentielle perpétuelle puisqu'elle est accordée pour une durée de dix années, renouvelables sans limitation.

En contrepartie de ce monopole d'exploitation, le titulaire de la marque a l'obligation d'exploiter celle-ci directement ou indirectement.

- Le dépôt de dessins et modèles

La charte graphique des bases de données commercialisées peut également faire l'objet d'un dépôt auprès de l'INPI au titre des dessins et modèles.

Une société peut ainsi déposer l'environnement graphique de ses bases de données (icônes, menu déroulant, fenêtres...).

La forme ainsi déposée doit être nouvelle, c'est-à-dire ne pas avoir déjà été appropriée par un tiers.

La protection est acquise pour une période de 25 ans à compter de la date du dépôt et peut être prorogée pour une nouvelle période identique à la demande du titulaire.

- L'enveloppe Soleau

Le dépôt auprès de l'INPI d'une enveloppe Soleau permet de donner date certaine aux éléments déposés dans l'enveloppe et d'établir ainsi la preuve de la date de création de leur contenu.

Par contre le dépôt d'une enveloppe Soleau ne confère aucun droit ou titre de propriété au déposant : il permet simplement de se préconstituer la preuve de l'existence d'une création intellectuelle à une date donnée.

La durée de conservation de l'enveloppe Soleau par l'INPI est de 5 années à compter de la date du dépôt. Elle peut être renouvelée une fois.

- Le dépôt auprès de l'APP

Afin d'améliorer la protection d'une société sur ces bases de données, il peut être utile de procéder au dépôt du logiciel de navigation auprès de l'Agence pour la Protection des Programmes (APP).

Le dépôt à l'APP permet à l'auteur déposant de préconstituer la preuve du contenu exact de son œuvre, de la date de création et d'achèvement.

Ce dépôt permettra de démontrer l'antériorité de l'œuvre, antériorité pouvant être opposable aux tiers.

Le dépôt est effectué pour la durée de protection des droits d'auteur.

Cependant, seule une reconduction annuelle de l'adhésion à l'APP permet de procéder à des mises à jour des programmes déposés.

- Le dépôt à la Bibliothèque Nationale de France

Le décret n°93-1429 du 31 décembre 1993 prévoit le dépôt des bases de données à la Bibliothèque Nationale de France dès qu'elles sont mises à la disposition du public, à titre onéreux ou gratuit, par diffusion en nombre d'un support matériel de quelque nature que ce soit.

Les bases de données ne sont pas soumises à l'obligation de dépôt lorsqu'elles sont importées à moins de 100 exemplaires.

Il convient également de souligner que les bases de données accessibles en ligne ne sont pas soumises à l'obligation de dépôt légal.

Ainsi, une société ne doit procéder qu'au seul dépôt de ses bases de données commercialisées sur support matériel.

Le dépôt légal est une formalité administrative obligatoire. Par conséquent, il ne confère aucun droit au déposant.

Le dépôt légal peut cependant s'avérer utile pour une société dans la mesure où il facilite la preuve de l'acte de création de la base de données.

6.2.2.5 Les autres moyens permettant d'améliorer la protection des bases de données

Il peut être procédé à l'apposition sur l'exemplaire qui reproduit l'œuvre de la mention "Copyright", matérialisée par le symbole ©, suivi du nom de celui qui revendique la titularité des droits d'auteur ainsi que de l'année de la revendication.

Bien que cette mention ne soit pas exigée pour la protection des œuvres en France et qu'elle ne soit, en aucun cas, constitutive de droits, elle peut néanmoins être utile, afin d'informer les utilisateurs.

6.2.3 Les droits sur l'outil : la protection des logiciels

6.2.3.1 Une protection de plein droit par le droit d'auteur

En matière de logiciel, la protection ne porte pas sur les fondements abstraits des logiciels, tels que les principes mathématiques, les méthodes de programmation et les règles de calcul qui sont utilisés lors de l'élaboration d'un programme. Il s'agit là des règles ou des méthodes qui ne constituent pas l'objet direct du programme, et qui relèvent du domaine des idées dans la mesure où leur matérialisation peut donner lieu à des formes différentes.

Or, seule la forme originale est protégée.

La notion de forme vise tant les spécifications externes du logiciel (analyse fonctionnelle) que les spécifications internes (suite logique d'instructions).

Dans ces conditions, le principe est que la réutilisation des idées est toujours possible. Seule la composition de l'expression est protégée. Dès lors, seul l'emprunt d'une forme identique ou similaire est répréhensible.

La jurisprudence donne un exemple de ce que les tribunaux considèrent comme constitutif d'un logiciel original.

Un très célèbre arrêt rendu par l'assemblée plénière de la Cour de cassation le 7 mars 1986 a donné une définition de cette notion clé qu'est l'originalité¹.

Pour la Cour de cassation, il faut que :

- l'auteur "fasse preuve d'un effort personnalisé allant au-delà de la simple mise en œuvre d'une logique automatique et contraignante" ;
- "la matérialisation de cet effort réside dans une structure individualisée ;
- le logiciel porte "la marque de l'apport intellectuel" de son auteur.

Aucun formalisme (dépôt, enregistrement, déclaration...) n'est nécessaire. Cette absence de formalisme présente une limite lorsque, dans le cadre d'une procédure, il est nécessaire de faire la preuve de sa titularité sur le logiciel, l'absence de formalisme pouvant rendre la constitution de la preuve difficile.

¹ Cour de cassation, Assemblée plénière, 7 mars 1986, Dalloz 1986, page 405.

Dès lors un dépôt auprès de l'APP notamment permet l'établissement d'une présomption de propriété.

Par ailleurs, il permet d'organiser un accès des utilisateurs aux codes sources du logiciel.

Cette procédure peut s'avérer nécessaire pour remédier à la défaillance du créateur (cessation d'activité, perte du programme source) ou pour répondre à certaines obligations légales et notamment fiscales.

Si cet accès aux codes sources se révèle être un avantage au profit des utilisateurs, il peut également constituer un argument commercial pour le créateur ou l'éditeur du logiciel.

Ainsi, de nombreux éditeurs de logiciels de comptabilité axent aujourd'hui leur communication sur le fait que leurs programmes sont conformes à la réglementation et que les codes sources en sont déposés en vue d'être mis à la disposition des utilisateurs.

6.2.3.2 Les logiciels "libres"

L'utilisation des logiciels libres par les professionnels se développe de manière importante.

Or, l'utilisation d'un logiciel libre ne traduit pas seulement un choix d'ordre technique mais également l'adhésion à un système de distribution originale pour des éléments susceptibles d'être protégés par le droit d'auteur.

Pour autant, il semble que c'est par abus de langage que l'on peut qualifier de "libres" les logiciels.

Les logiciels libres sont des œuvres de l'esprit. En conséquence, leur seule création (même inachevée) est constitutive de droit, à la différence des brevets ou des marques pour lesquels c'est le dépôt qui confère un monopole d'exploitation.

Au regard de ces dispositions légales tant sur le plan national, que sur le plan international, les logiciels ne "naissent" pas libres mais asservis à la personne de leur auteur.

En conséquence, un logiciel ne pourrait être libre qu'à la condition que son auteur ait expressément renoncé à son monopole d'exploitation sur la création dont il est l'auteur.

La gratuité d'un logiciel ou la mise à disposition du public par son auteur du logiciel n'entraîne pas automatiquement renonciation de son auteur à ses droits sur le logiciel.

Au regard du droit français, une telle renonciation ne pourra jamais être que partielle compte tenu de l'existence des droits moraux auxquels l'auteur lui-même ne peut valablement renoncer puisqu'ils sont perpétuels, inaliénables, imprescriptibles et intransmissibles.

Dès lors, un logiciel, quel qu'il soit, ne peut être utilisé au préjudice du droit moral de son auteur, lequel consiste notamment, en un droit au respect de son nom, de sa qualité et un droit de s'opposer à toute atteinte au respect de l'œuvre logicielle susceptible de nuire à son honneur ou à sa réputation ce qui peut, théoriquement, constituer une entrave à la réalisation de certaines œuvres dérivées qui ne respecteraient pas les droits moraux de l'auteur sur l'œuvre initiale.

S'agissant des droits patrimoniaux de l'auteur, c'est-à-dire les droits de reproduction, représentation, transformation, adaptation, traduction, ce dernier ne pourra y renoncer tacitement.

En effet, la cession des droits d'auteur ne se présume pas et doit respecter un ensemble de conditions précises définies à l'article L.131-3 du Code de la propriété intellectuelle, dont la destination des droits cédés, la durée et le territoire de la cession.

L'utilisateur d'un logiciel dit libre devra donc s'assurer de l'étendue de cette liberté et disposer d'un moyen de preuve opposable pour justifier du libre usage de ce produit.

6.3 DEPLOYER L'OUTIL DANS LE RESPECT DE LA LOI

Le déploiement de l'outil de gestion des connaissances au sein de l'entreprise utilisatrice nécessite le respect des dispositions issues notamment :

- de la loi "Informatique et libertés" ;
- du Code du travail.

6.3.1 Le respect de la loi "Informatique et libertés"

6.3.1.1 La déclaration auprès de la CNIL

Des traitements automatisés d'informations nominatives¹ pourront être mis en œuvre dans le cadre de l'outil de gestion des connaissances.

¹ Conformément aux dispositions de l'article 5 de la loi Informatique et libertés, constitue un traitement automatisé d'informations nominatives : tout ensemble d'opérations réalisées par les moyens informatiques, relatifs à la collecte, l'enregistrement, l'élaboration, la modification, la conservation et la destruction d'informations nominatives, ainsi que tout ensemble d'opérations de même nature se rapportant à l'exploitation de fichiers ou base de données et notamment les interconnexions ou rapprochements, consultations ou communications d'informations nominatives. Aux termes de l'article 4 sont considérées comme nominatives toutes informations "qui permettent, sous quelque forme que ce soit, directement ou non, l'identification des personnes physiques auxquelles elles s'appliquent, que le traitement soit effectué par une personne physique ou par une personne morale".

Or, aux termes de l'article 16 de la loi du 6 janvier 1978 relative à l'Informatique, aux fichiers et aux libertés:

- "Les traitements automatisés d'informations nominatives effectués pour le compte [du secteur privé] doivent préalablement à leur mise en œuvre, faire l'objet d'une déclaration auprès de la CNIL".

L'article 226-16 du Code pénal sanctionne pénalement le délit de mise en œuvre d'un traitement automatisé d'informations nominatives sans déclaration préalable auprès de la CNIL. Les peines prévues par cet article sont de 45 000 euros d'amende et 3 ans d'emprisonnement (avec quintuplement de la peine d'amende encourue en cas de reconnaissance de la responsabilité pénale de l'entreprise en tant que personne morale).

6.3.1.2 Les autres obligations du responsable de traitement

- L'obligation de sécurité et de confidentialité

Conformément aux dispositions de l'article 29 de la loi Informatique et libertés, le maître du fichier (ou responsable du traitement) est responsable de la sécurité et de la confidentialité des informations qu'il traite.

Il s'engage de ce fait à "prendre toutes précautions utiles afin de préserver la sécurité des informations et notamment d'empêcher qu'elles soient déformées, endommagées ou communiquées à des tiers non autorisés".

Cette sécurité doit être entendue aussi bien de la sécurité physique des traitements mis en œuvre par le déclarant que de leur protection logique, ou des mesures prises pour former les personnes à leurs obligations (sensibilisation des utilisateurs).

Les mesures de sécurité physiques et logiques prises pour garantir la confidentialité des informations, ainsi que pour sensibiliser les utilisateurs des systèmes, doivent être décrites dans la déclaration auprès de la CNIL.

- L'interdiction de traiter des données sensibles

En vertu de l'article 31 de la loi Informatique et libertés, il est interdit de procéder, sans l'accord des intéressés, au traitement de certaines données considérées comme sensibles.

Ces données, énumérées par l'article 31, sont celles qui font référence directement ou indirectement :

- aux origines raciales ;
- aux opinions politiques, philosophiques ou religieuses ;
- aux appartenances syndicales ;
- aux mœurs des personnes.

La violation de cette interdiction est sanctionnée pénalement.

En effet, la collecte, l'enregistrement et la conservation de données sensibles sont punis de 300 000 euros d'amende et 5 ans d'emprisonnement (avec quintuplement de la peine d'amende encourue en cas de reconnaissance de la responsabilité pénale de l'entreprise en tant que personne morale par l'article 226-19 du Code pénal).

Il convient par conséquent d'informer, par exemple au moyen d'une charte, les utilisateurs amenés à produire du contenu au sein de l'outil de gestion des connaissances de cette interdiction.

Le projet de loi Informatique et libertés, destiné à transposer en droit interne la directive communautaire du 24 octobre 1995 rajoute à cette liste de données sensibles interdites, les données relatives à la santé.

- L'obligation d'information des personnes

L'obligation d'information des personnes est prévue par l'article 27 de la loi Informatique et libertés qui dispose que :

"Les personnes auprès desquelles sont recueillies des informations nominatives doivent être informées :

- du caractère obligatoire ou facultatif des réponses ;
- des conséquences à leur égard d'un défaut de réponse ;
- des personnes physiques ou morales destinataires des informations ;
- de l'existence d'un droit d'accès et de rectification ;

Lorsque de telles informations sont recueillies par voie de questionnaires, ceux-ci doivent porter mention de ces prescriptions.

Le non-respect des dispositions de l'article 27 de la loi Informatique et libertés est passible de 1500 euros par infraction constatée.

Par ailleurs, le projet de loi Informatique et libertés susvisé prévoit quatre mentions supplémentaires, à savoir :

- l'identité du responsable du traitement ;
- les finalités du traitement auquel les données sont destinées ;
- si les données sont destinées à être communiquées à des tiers ou utilisées pour le compte de tiers à des fins de prospection, une information sur ce point accompagnée d'une possibilité pour les personnes de s'opposer (au moyen d'une case à cocher ou à cliquer notamment) ;
- si les données sont destinées à être communiquées à l'international à des pays tiers à l'Union européenne, une information sur ce point.

6.3.2 Le respect du droit du travail

6.3.2.1 L'information et la consultation préalable du comité d'entreprise

L'article L.432-2 du Code du travail dispose que :

- "Le comité d'entreprise est informé et consulté, préalablement à tout projet important d'introduction de nouvelles technologies, lorsque celles-ci sont susceptibles d'avoir des conséquences sur l'emploi, la qualification, la rémunération, la formation ou les conditions de travail du personnel".

Une circulaire ministérielle du 30 novembre 1984 définit la notion de "nouvelles technologies".

Selon l'administration, cette notion doit être entendue de manière particulièrement large.

Ainsi, ont été considérées comme des nouvelles technologies l'automatisation, l'informatisation, la robotisation, même si cette technologie introduite dans l'entreprise ou dans l'établissement est largement répandue dans les secteurs d'activité concernés.

La circulaire précise néanmoins qu'une entreprise déjà équipée d'un matériel informatique n'est pas soumise aux dispositions de l'article L.432-2 précité du Code du travail, si elle acquiert pour le remplacer, un matériel de technologie identique qui n'apporte pas de modification à la situation du personnel.

Au vu des textes, il apparaît que la notion de "nouvelles technologies" doit être examinée au regard des conséquences concrètes de son introduction dans l'entreprise, et plus particulièrement au regard des modes de travail de chacun des salariés.

C'est pourquoi, la jurisprudence procède à une évaluation au cas par cas en fonction notamment du degré d'informatisation de l'entreprise et en fonction de l'impact concret de la technologie mise en place sur les processus de travail.

Par ailleurs, l'article L.432-2 du Code du travail ajoute que :

- "Les membres du comité reçoivent, un mois avant la réunion, des éléments d'informations sur ces projets et leurs conséquences quant aux points mentionnés ci-dessus".

Les textes imposent donc la réalisation d'une étude d'impact qui permettra :

- de déterminer s'il est obligatoire de consulter préalablement le comité d'entreprise ;
- dans l'affirmative, de communiquer certains éléments du résultat de l'étude un mois avant la réunion du comité d'entreprise.

En ce qui concerne la notion de "projet important", par un arrêt du 3 mai 1994¹, la Chambre criminelle de la Cour de cassation a précisé que cette notion s'apprécie :

- "Nécessairement en fonction des modifications que le projet apporte dans l'organisation et la marche de l'entreprise et les conséquences qu'il peut avoir (...) sur l'emploi, la qualification, la rémunération et la formation, les conditions de travail du personnel".

Ainsi la mise en place d'un système de gestion des connaissances pourrait être qualifiée de "projet important de nouvelles technologies".

Pour s'en assurer, il est conseillé de mener un audit et ce, d'autant que le défaut de consultation du comité d'entreprise sur l'introduction de nouvelles technologies dans le réseau informatique, constitue le délit d'entrave² engageant la responsabilité pénale du chef d'entreprise.

6.3.2.2 La formation des salariés

Au regard du principe d'exécution de bonne foi des contrats, la jurisprudence considère que l'employeur a l'obligation "d'assurer l'adaptation des salariés à l'évolution de leurs emplois".

Ce principe est rappelé dans un arrêt du 6 avril 1994³ de la Chambre sociale de la Cour de cassation qui a considéré qu'en raison des mutations technologiques modifiant les postes de travail, les salariés auraient pu être maintenus dans l'entreprise sous réserve d'une formation complémentaire et que leur licenciement était en conséquence dénué de cause réelle et sérieuse.

En d'autres termes, dans l'hypothèse où l'introduction d'un outil de gestion des connaissances ne ferait pas l'objet d'une formation particulière des salariés et pourrait amener certains d'entre eux à ne plus être en mesure d'exécuter leur contrat de travail, un licenciement pour ce motif pourrait être considéré comme dénué de toute cause réelle et sérieuse.

¹ Cour de cassation, 3 mai 1994, Bulletin criminel n° 164.

² Article L. 483-1 du Code du travail.

³ Cour de cassation, 6 avril 1994, Gazette du Palais 1995, n° du 22 avril, page 18.



A cet égard, l'inadaptation d'un salarié suite à l'informatisation de l'entreprise et la rupture de son contrat de travail par l'employeur en raison de cette inadaptation ont, d'ores et déjà, été qualifiées par la jurisprudence de licenciement pour motif économique.

Toutefois, il convient de s'assurer que l'inadaptation invoquée ne découle pas d'un manquement de l'employeur à l'une de ces obligations et, en particulier, à celle d'assurer l'adaptation des salariés à l'évolution de leurs emplois.

C'est pourquoi, afin d'éviter ce risque, des formations spécifiques peuvent apparaître nécessaires au sein de l'entreprise au titre de la mise en œuvre de l'outil de gestion des connaissances.

6.3.2.3 L'information des salariés : charte/guide/livret de procédures

L'information des salariés peut être diffusée au moyen de différents outils propres à l'entreprise, à savoir :

- une charte utilisateur ;
- un guide ;
- un livret de procédures.

- Charte utilisateur

La charte a pour objectif de présenter les règles et principes généraux qui commandent l'utilisation de l'outil de gestion des connaissances.

Ce document répond à la nécessité d'informer les salariés sur les spécificités de l'outil de gestion des connaissances et de porter à leur connaissance les règles générales de son utilisation.

La charte rappellera également la nécessité pour les salariés de respecter, dans le cadre de l'utilisation de l'outil, les droits des tiers et notamment les droits de propriété de ceux-ci sur les différents documents et/ou programmes utilisés.

Enfin et compte tenu de la nécessaire évolution des règles d'utilisation d'un tel outil au sein d'une entreprise, il conviendra d'insérer un article "évolution de la charte" offrant la possibilité à tout utilisateur et à la direction d'émettre des propositions quant à son évolution et des conditions d'utilisation de l'outil de gestion des connaissances.

- Le Guide

Le guide a, par rapport à la charte utilisateur et au livret des procédures, une finalité plus juridique.

En effet, il définit les dispositions légales en vigueur qui s'appliquent dans le cadre de l'utilisation de l'outil de gestion des connaissances.

Les différents principes juridiques applicables en matière de libertés individuelles, de propriété, de respect des droits d'auteur, de traitement de données nominatives et de responsabilité pourront être ainsi rappelés dans ce guide.

La mise en place d'un tel guide permet de s'assurer que les salariés auront pris connaissance des conséquences juridiques de l'utilisation qui pourrait être faite de l'outil de gestion des connaissances.

- Le Livret des Procédures

Le livret des procédures a pour objet de définir les procédures d'utilisation de l'outil de gestion des connaissances dans l'entreprise.

Le livret des procédures pourra comprendre une partie générique ainsi qu'une partie spécifique décrivant :

- le système de gestion des connaissances ;
- les conditions d'attribution de ses composants ;
- la gestion des accès et leur sécurisation ;
- la gestion électronique des échanges électroniques ;
- la gestion des procédures de sauvegarde et d'archivage,

et plus généralement les conditions d'emploi de cet outil dans l'activité quotidienne des utilisateurs.

En outre, comme pour la charte et le guide, il pourra être prévu une évolution possible du livret qui pourrait notamment résulter de propositions de modifications ou d'ajouts des utilisateurs.

- Déploiement de la Charte utilisateur, du Guide et du Livret des procédures

Ces documents doivent être portés à la connaissance des utilisateurs. Il en va de même pour leurs évolutions.

A cet effet, un déploiement doit être organisé.

Il est précisé qu'un déploiement électronique peut être réalisé et que l'acceptation des procédures et documents par les utilisateurs pourrait intervenir sur la base d'un double clic d'acceptation ("contrat clic").

6.4 ASSURER LA SECURITE DE L'OUTIL

6.4.1 Le contrôle de l'utilisation de l'outil

Pour que l'outil de gestion des connaissances soit créateur de valeur, il faut en protéger son contenu qui constitue le patrimoine intellectuel de l'entreprise.

A ce titre, le projet de loi relatif au droit d'auteur et aux droits voisins dans la société de l'information prévoit, en son article 7 :

- la mise en place de "mesures techniques efficaces destinées à empêcher ou limiter les utilisations non autorisées par le titulaire d'un droit d'auteur ou d'un droit voisin du droit d'auteur(...)" ;

- et précise que "(...) ces mesures techniques sont réputées efficaces lorsqu'une utilisation est contrôlée grâce à l'application d'un code d'accès, d'un procédé de protection tel que le cryptage, le brouillage ou toute autre transformation de l'objet de la protection, ou d'un mécanisme de contrôle de la copie qui atteint cet objectif de protection".

Des systèmes de tatouage des contenus ou d'audit afin de savoir qui fait quoi, qui imprime quoi, pour quelle raison, afin également d'identifier les failles du systèmes pourront également être mis en place dans le respect du droit du travail car ce type de système ne saurait être réalisé à l'insu des salariés.

Aucun contrôle de l'activité des salariés ne peut se faire par l'utilisation de moyen technique déployé à l'insu du personnel¹.

Ce principe ressort des dispositions de l'article L.120-2 du Code du travail qui prévoit que :

- "Nul ne peut apporter au droit des personnes et aux libertés individuelles et collectives de restrictions qui ne seraient pas justifiées par la nature de la tâche à accomplir ni proportionnées au but recherché".

Ce principe ressort également des dispositions de l'article L.121-8 du Code du travail qui prévoit que :

- "Aucune information concernant personnellement un salarié ou un candidat à un emploi ne peut être collectée par un dispositif qui n'a pas été porté préalablement à la connaissance du salarié ou du candidat à un emploi".

¹ Cour de cassation, chambre sociale, 15 mai 2001, première espèce, pourvoi n°R99-42.937, arrêt n°2084FS-PB.

En ce sens également, la CNIL, par une délibération du 18 septembre 1984, portant adoption d'une recommandation concernant l'usage des autocommutateurs téléphoniques sur les lieux de travail, a précisé que la mise en œuvre d'un autocommutateur téléphonique sur les lieux de travail nécessitait l'existence de garantie minimale afin d'éviter que de telles installations ne portent atteinte aux libertés individuelles.

Elle a plus particulièrement recommandé la publicité du système et la formation du salarié sur la nature et la périodicité des contrôles pratiqués.

De leur côté, les tribunaux ont eu à juger d'atteintes aux libertés résultant de la mise en œuvre de tels moyens de contrôle de l'activité des salariés.

Ainsi et à titre d'exemple, par un arrêt du 7 mars 1997, la Cour d'appel de Paris a estimé que la production d'un listing de relevé de communications téléphoniques émanant du poste d'un salarié ne faisant apparaître que les premiers chiffres des numéros appelés, constituait un moyen de preuve illicite¹.

Dans cette affaire, la salariée, chef du personnel, avait été licenciée pour avoir violé son obligation de loyauté et de confidentialité dans la mesure où la production du listing des relevés téléphoniques où figuraient les numéros de téléphones d'anciens collègues de l'entreprise permettait d'établir une correspondance entre la date des appels effectués par la salariée et la date des réclamations faites par les ex-salariés.

Sur cette base, la cour a estimé que l'employeur n'apportait pas la preuve que la salariée avait dévoilé des informations confidentielles sur l'entreprise et a écarté ce moyen de preuve, le jugeant illicite.

En outre, la loi confie au comité d'entreprise un rôle important sur le contrôle de l'activité des salariés dans la mesure où une procédure d'information ou de consultation préalable de cette institution représentative du personnel doit être respectée avant de mettre en œuvre tout moyen ou technique permettant de contrôler l'activité des salariés.

L'article L.432-2-1 alinéa 3 du Code du travail prévoit, en effet, que le comité d'entreprise est informé et consulté préalablement à la décision de mise en œuvre dans l'entreprise des moyens ou des techniques "permettant un contrôle de l'activité des salariés".

On relèvera la difficulté d'interprétation du terme "permettant".

En effet, faut-il considérer que la procédure d'information et de consultation concerne exclusivement les méthodes et techniques dont la fonctionnalité première est d'organiser un contrôle ?

¹ Cour d'appel de Paris, 7 mars 1997, Gazette du Palais, 1999, I, Jur. page 30, note Joëlle Berenguer Guillon et Laetitia Guignot.

Dans le cas d'un outil dont la vocation n'est pas la surveillance, permettre le contrôle suffit-il à générer cette obligation vis-à-vis du comité d'entreprise ? Il semble que oui.

En tout état de cause, si un outil de travail permet d'obtenir des informations sur l'activité d'un salarié et que ce salarié et le comité n'en n'ont pas été informés, il ne pourra pas servir de preuve quant à la légitimité d'une éventuelle sanction.

Par ailleurs, l'objectif de l'article L.432-2-1 du Code du travail est, selon la circulaire du 15 mars 1993, d'assurer une plus grande transparence à l'égard du comité d'entreprise sur le contrôle des salariés.

Il ressort également de cette circulaire que le comité d'entreprise peut, par l'intermédiaire de cette procédure d'information et de consultation, donner son avis sur la pertinence et sur la proportionnalité entre les moyens techniques utilisés et le but recherché par l'entreprise.

Il s'agit, en pratique, d'une procédure de consultation tout à fait ponctuelle qui doit être mise en œuvre dès lors que l'employeur envisage de mettre en place un tel outil.

6.4.2 Le vol d'informations

La notion de vol d'informations n'est pas prévue par le Code pénal.

Toutefois, une partie de la doctrine a admis cette possibilité, appuyée par certaines décisions de jurisprudence.

Ainsi, le Tribunal correctionnel de Montbéliard a, le 26 mai 1978, consacré le principe du vol d'informations en condamnant le salarié, auteur d'un programme informatique qui, après son départ de l'entreprise, avait recopié le programme dont il était l'auteur.

La Cour d'appel de Reims a également expressément fait référence au vol d'informations dans une affaire où d'anciens salariés d'une imprimerie "avaient frauduleusement soustrait le contenu informationnel de 47 disquettes"¹.

Cet arrêt a été confirmé par la Cour de cassation² ou, pour la première fois, la Cour suprême se prononce clairement pour l'extension de l'article 379 du Code pénal au vol du contenu informationnel d'un support documentaire indépendamment de ce dernier.

¹ Cour d'appel de Reims, 27 février 1987, Expertises 1987, n° 94, page 148.

² Cour de cassation, 12 janvier 1989, Bulletin criminel 1989, n° 14, page 38.



Cependant, la reconnaissance par la jurisprudence de la notion de vol d'informations n'est pas unanime et la Cour d'appel de Paris a ainsi jugé par un arrêt du 25 novembre 1992¹ que :

- "Il résulte des éléments de la cause que les vols allégués par la partie civile s'analysent en des transferts de disques durs à disquettes (faits concernant le logiciel Turbo C) et de disquettes à disques durs (faits concernant Sprint 1.5) de données d'instructions informatiques qui sont de la nature même des logiciels de programmation Turbo C et de traitements de textes Sprint 1.5 (...);

- considérant qu'il apparaît que ces transferts qui portent exclusivement sur des données immatérielles, quelle qu'en soit la valeur intellectuelle, ne sauraient entrer dans le champ d'application de l'article 379 du Code pénal qui exige que la soustraction frauduleuse porte sur une chose matérielle ou corporelle";

- en conséquence, et en l'absence de toute "soustraction frauduleuse" de supports informatiques (disquettes, cédéroms), les transferts de données ne sauraient constituer le délit de vol prévu et sanctionné par l'article 425 du Code pénal".

La jurisprudence relative au vol d'informations ne paraissant pas définitivement fixée, il conviendra d'envisager avec prudence toute action pénale reposant sur ce moyen.

6.4.3 La mise en place d'une charte de sécurité

De plus en plus d'entreprises précisent les règles de sécurité et d'utilisation des systèmes d'information.

Elles sont dénommées chartes ou codes d'éthique, selon la culture de l'entreprise.

De manière générale, ces chartes prévoient que l'utilisation des outils informatiques doit rester exclusivement professionnelle et indiquer que des contrôles sont effectués. Cette position ne correspond pas à la position de la jurisprudence actuelle qui reconnaît à chacun des membres de l'entreprise un droit à la vie privée résiduelle.

Dans son rapport sur la cybersurveillance des salariés dans l'entreprise, la CNIL recommande l'élaboration de charte ou code de conduite négociés.

Sur la question de l'utilisation de la messagerie électronique à titre privé, elle considère que l'interdiction pure et simple devrait être écartée.

¹ Cour de cassation, 25 novembre 1992, Gazette du Palais 1993 du 23 octobre, page 27.

Toutefois, il convient de reconnaître qu'un contrôle minimum peut être salvateur et devrait, en conséquence, être assuré compte tenu des risques encourus à défaut de surveillance (notamment pour les images répréhensibles aux yeux de la loi et de la morale – comme des images pédophiles – ou la divulgation d'information sensible interdite).

La valeur et la justification juridique de ces chartes sont importantes.

Si la charte a valeur de règlement intérieur, l'article L.122-39 du Code du travail trouve matière à s'appliquer.

Une fois que la procédure de consultation du comité d'entreprise pour l'introduction de la charte à valeur de règlement intérieur est respectée, la charte est opposable à tous les salariés, sans que l'entreprise n'ait à recueillir l'acceptation individuelle des salariés.

En tout état de cause, l'article L.122-35 du Code du travail impose que le règlement intérieur n'apporte pas des restrictions aux droits et libertés individuelles et collectives qui ne seraient justifiées ni par la tâche à accomplir ni proportionnées au but recherché.

L'article L.122-39 du Code du travail impose le même régime à toutes les notes de services ou tout autre document portant prescription générale et permanente relevant du règlement intérieur, c'est-à-dire l'avis du comité d'entreprise ou, à défaut, des délégués du personnel et la communication à l'inspecteur du travail.

La charte aura valeur de règlement intérieur, lorsqu'elle édicte des prescriptions générales et permanentes.

7. RECOMMANDATIONS

Pour le groupe de travail, la mise en place d'un outil de gestion des connaissances au sein de l'entreprise est source d'enrichissement en ce qu'il permet de développer le capital intellectuel de l'entreprise, source de compétitivité, de fertiliser les savoirs individuels en croisant les connaissances et les meilleures pratiques et de réaliser des gains de productivité.

Le groupe de travail relève toutefois l'existence d'obstacles au développement d'un outil de gestion des connaissances au sein de l'entreprise. Ces obstacles sont principalement dus à la résistance des membres de l'entreprise de partager leurs connaissances.

La réussite d'un programme de gestion des connaissances nécessite donc de créer des situations de confiance permettant l'adhésion des membres de l'entreprise au programme et de les amener à adopter une culture de partage au sein de l'entreprise.

L'élaboration d'une charte comportementale dont l'objet est de sensibiliser les membres de l'entreprise à l'intérêt d'utiliser l'outil et à produire du contenu, sans pour autant s'en faire une obligation est donc recommandée.

De plus, l'incitation – en donnant aux membres de l'entreprise du pouvoir dans le système – et la reconnaissance – en valorisant l'individu à l'origine du document diffusé au sein de l'outil par l'indication de son nom sur le document, par exemple – font partie des moyens que le groupe de travail considère comme pouvant contribuer à créer une culture de partage.

En outre, le groupe de travail constate que l'échec étant également source de connaissances, il paraît utile de diffuser des informations, facilement accessibles, par exemple sous forme de fiches, contenant non seulement les meilleures pratiques, mais également les pièges à éviter (les pratiques à bannir et à proscrire).

Enfin, le groupe de travail estime que l'outil de gestion des connaissances doit faire l'objet d'un encadrement juridique fort et recommande à ce titre :

- de disposer de contrats protecteurs (contrats de conseil, de formation, de licence d'utilisation de l'outil, de réalisation de développements spécifiques, de maintenance ...)
- de gérer les droits sur l'outil de gestion des connaissances et son contenu (droit d'auteur, protection des logiciels, droits sur la base de données ...)

- de déployer l'outil au sein de l'entreprise utilisatrice dans le respect des dispositions issues de la loi "Informatique et libertés" (déclaration auprès de la CNIL ...) et du Code du travail (information et consultation préalable du comité d'entreprise ...);
- d'informer les salariés par le biais d'une charte présentant les règles et principes généraux qui commandent l'utilisation de l'outil de gestion des connaissances, d'un guide juridique définissant les dispositions légales en vigueur qui s'appliquent dans le cadre de cette utilisation et d'un livret des procédures ayant pour objet de définir les procédures d'accès à l'outil et d'utilisation ;
- de mettre en place une charte de sécurité permettant de protéger le contenu de l'outil, base de la compétitivité de l'entreprise.

8. ANNEXES

- annexe 1 : liste des participants ;
- annexe 2 : présentation de Alain Bensoussan Selas ;
- annexe 3 : présentation de LexisNexis ;
- annexe 4 : présentation de WoltersKluwer ;
- annexe 5 : présentation de l'ACFCI ;
- annexe 6 : présentation de Factiva ;
- annexe 7 : présentation de Sensoria Technology Limited ;
- annexe 8 : présentation de l'INPI ;
- annexe 9 : présentation du GFII ;
- annexe 10 : rappel ou précision, rappel et précision ? Question de pertinence ;
- annexe 11 : bibliographie.

ANNEXE 1

LISTE DES PARTICIPANTS

Alain Bensoussan, Alain Bensoussan Selas

Alain Bensoussan, avocat à la Cour d'appel de Paris, spécialiste en droit de l'informatique et des technologies avancées, a fondé, en 1978, un cabinet dédié à ces domaines. Il a rédigé et publié en 1985, aux Editions Berger-Levrault, le premier traité de droit de l'informatique. Il est arbitre auprès de l'Organisation mondiale de la propriété intellectuelle (OMPI), enseigne le droit de l'informatique à l'Ecole Centrale de Paris. Il est Président d'honneur de l'Association française de l'informatique et de la télécommunication (AFDIT).

Denis Berthault, Juris-Classeur

Denis Berthault intervient depuis 20 ans dans l'information juridique en ligne. Il a occupé respectivement des fonctions commerciales à Européennes de Données (1985-1994), puis directoriales chez ORT où il a réalisé notamment les sites Jurifrance et Legifrance (1^{ère} version) ainsi que les CD-Rom "50 années de Journal officiel", "les révisions de la Constitution 1995 et 1996" et "Infocodes". Entré en 2000 chez Juris-Classeur, il a bâti le site Jurisclasseur.com sorti en juillet 2002. Depuis avril 2003, il définit la politique d'acquisition de contenus de presse et d'information sur les entreprises pour le compte de LexisNexis.

Titulaire d'un DEA de droit public et d'un DESS Droit des collectivités locales (Paris X), il est Vice-président du Groupement français de l'industrie de l'information (GFII) depuis 2003 et administrateur de l'ADIJ (Association pour le développement de l'informatique juridique) depuis 1999. Il a participé à plusieurs groupes de travail notamment pour le rapport Lorentz sur le commerce électronique en 1998 et celui sur les Services en ligne : modèles économiques et systèmes de paiement réalisé par l'ACSEL en février 2004.

Michèle Côme, Wolters Kluwer

Michèle Côme, titulaire d'une maîtrise de Droit des affaires et de brevets de traducteur juridique anglais, allemand et italien a exercé de 1979 à 1993 au sein du cabinet Berlioz, cabinet d'avocat international, où elle a développé un centre de documentation juridique, une base de données de doctrine et de contrats. Elle a participé à la constitution de l'association Juriconnexion qui regroupe les utilisateurs de banques de données juridiques. Elle en a été présidente de 1992 à 1994, avant de rejoindre les Editions Lamy en qualité de responsable de l'édition électronique.

Depuis 2002, elle est "business program manager" chez Wolters Kluwer Europe, maison mère des Editions Lamy, et exerce une activité d'échange de bonnes pratiques et de projets transversaux visant à créer une synergie plus forte parmi les filiales du groupe en Europe, dans les segments de marchés juridiques, fiscaux, ressources humaines, santé, sécurité et environnement.

Olivier Duval, ACFCI

Diplômé DESS "Applications des réseaux et de la télématique" à Paris VII Jussieu, Olivier Duval intègre une petite structure spécialisée en télématique, en tant que développeur C, administrateur Unix et réseau Transpac. Il a rejoint l'ACFCI en 1998 en qualité de responsable technique NTIC et intègre, par la suite, le projet CCINET comme coordinateur technique. Il exerce les fonctions d'administrateur système, d'interface avec les entités techniques et d'intégrateur de progiciels. Développeur en langage objet, il est l'auteur du chapitre sur le polymorphisme du présent livre blanc.

Patrick Kervern, Factiva

Spécialiste de l'Internet BTOB et BTOC depuis 1995 (Responsable de programmes senior chez Infonie, Directeur de production chez Excite France), Patrick Kervern a intégré en 1998 Reuters comme Chef de produit nouveaux médias et business information products. Il a rejoint Factiva en 1999 en qualité de Marketing Manager Europe du Sud et y exerce, depuis 2001, la fonction de Directeur Marketing Europe Continentale et Moyen-Orient.

Arnaud Le Canu, Influo / Sensoria Technology

Arnaud Le Canu vient du secteur publicitaire et documentaire. Il bénéficie d'une large expérience en création d'entreprise consolidée par une formation à HEC Management. Il est en charge du lancement de la division de Sensoria dédiée au traitement de l'information appelée Influo.

Isabelle Leclercq, INPI

Isabelle Leclercq a rejoint l'INPI en 1995. Responsable du marketing et de la stratégie au sein du Département de la Documentation et de l'Information, elle a mis en place le système de gestion des connaissances clients à partir duquel sont élaborées les stratégies de diffusion de l'information. Sa démarche prospective lui permet d'identifier et d'évaluer les technologies de gestion des connaissances les plus prometteuses. Isabelle Leclercq possède un DESS de gestion (IAE) et une maîtrise de chimie inorganique. Aujourd'hui basée à Paris, elle a travaillé auparavant à Londres et aux Etats-Unis. Elle a près de 15 ans d'expérience professionnelle dans l'industrie de l'information.

Ruth Martinez, GFII

Ruth Martinez est Déléguée générale du GFII depuis 1999. A ce poste, elle assure la gestion administrative, financière et technique de l'association. Elle participe aux travaux des différents groupes de travail et représente le GFII dans les instances administratives et professionnelles. Avant de rejoindre le GFII, Ruth Martinez a été chef de projet dans un cabinet de consultants et chargée de la mise en œuvre de systèmes d'information et de communication interne au sein de grandes entreprises et organismes publics.

Eric Sergent, ACFCI

Diplômé de l'ISG et du CESI, après avoir été responsable des ventes indirectes micro-informatiques au groupe Olivetti, Eric Sergent a mis en œuvre chez Ernst et Young les outils CASE (Computer Aided Software Engineering). Puis, il a intégré l'éditeur Rational Software spécialisé dans les outils Orienté-Objet. Chef de service à l'ACFCI, il dirige depuis 1998 CCINET, l'extranet des 26 000 collaborateurs et 8 000 élus des Chambres de commerce et d'industrie et pilote les relais humains de la gestion des connaissances et du contenu.

Chloé Torres, Alain Bensoussan Selas

Avocat à la Cour d'appel de Paris, Chloé Torres a rejoint Alain Bensoussan en 1999. Elle exerce son activité au sein du pôle Informatique et Télécoms et plus spécifiquement dans les domaines de la gestion des connaissances, de l'e-learning et de l'internet, pour la conduite de projets, en matière contractuelle et dans la gestion des précontentieux et contentieux. Elle intervient en tant que conférencière notamment dans le cadre de KM Forum et E-learn Expo.

Ralph Villoing, Influo / Sensoria Technology

Ralph Villoing est Chief Executive Officer de Sensoria Technology Ltd et Directeur général de Sensoria France. Après sept années dans un grand groupe de presse informatique en tant que journaliste, il s'est investi totalement dans l'aventure Sensoria. Il en est l'un des principaux instigateurs et s'est impliqué dans toutes les facettes de son lancement depuis le premier jour. Il en est devenu le Chairman en 2003. Il possède une formation universitaire de mathématique (Paris VII) et de germaniste (Institut d'allemand d'Asnières).

ANNEXE 2

PRESENTATION DE ALAIN BENSOUSSAN SELAS

Orienté dès sa création en 1978, vers le droit de l'informatique, le cabinet Alain Bensoussan¹ a poursuivi cette orientation et se consacre entièrement au droit des technologies avancées.

Autour de son cœur de métier constitué par l'informatique et les télécoms, il intervient, tant en France qu'à l'étranger, dans les secteurs du multimédia et de la communication, de l'internet, des biotechnologies et de la santé, de l'espace, de l'énergie, de l'industrie, des matériaux, des travaux publics et du bâtiment, et ce, tant en conseil qu'en contentieux.

Les membres du cabinet associent la connaissance de ces secteurs techniques et du droit spécifique qui s'y applique, à celle des grandes catégories du droit, telles que le droit des affaires, le droit de la propriété littéraire et artistique, le droit de la publicité, le droit des assurances, le droit social, le droit pénal, le droit administratif, le droit de l'environnement, le droit des sociétés et le droit fiscal.

La qualité au sein du cabinet est une longue tradition qui l'a amené à être le premier cabinet d'avocats français certifié. Cette certification, obtenue en 1998, a été renouvelée en 2001 puis en 2003 sur la base de la nouvelle norme ISO 9001 : 2000 pour le "conseil et contentieux dans le domaine du droit des technologies avancées, notamment droit de l'informatique et des télécommunications".

¹ <http://www.alain-bensoussan.com>



ANNEXE 3

PRESENTATION DE LEXISNEXIS

Créée en 1973, LexisNexis constitue la plus grande bibliothèque d'information économique, financière, juridique et scientifique sur Internet.

Avec plus de 2 millions d'utilisateurs dans le monde, LexisNexis est reconnue pour l'exhaustivité et la pertinence de son fonds documentaire ; elle offre aujourd'hui accès à 2,5 milliards de documents issus de 35 000 sources telles que la presse quotidienne et spécialisée, les études de marché, les rapports des plus grandes institutions financières et banques d'affaires dans le monde, les principaux fournisseurs de droits étrangers, les dépôts de marques et brevets ...

LexisNexis emploie 12 000 personnes dans le monde et compte 50 bureaux de représentation aux Etats-Unis, mais aussi dans les principales métropoles mondiales comme Amsterdam, Londres, Frankfurt, Paris, Singapour, Hong Kong, Tokyo, Sydney et Toronto. Le bureau de LexisNexis France est basé à Paris (141, rue de Javel 75015 Paris).

A propos du Groupe LexisNexis

La base de données LexisNexis est intégrée au Groupe LexisNexis, Division de Reed-Elsevier, un des leaders mondiaux de l'information professionnelle (36 000 personnes, CA 2002 : 7,9 M€).

Le Groupe LexisNexis réunit de nombreuses marques de référence en matière de sources documentaires et de nouvelles technologies à destination des professionnels du droit, des entreprises, des administrations et des universités. Ces produits sont disponibles via Internet et sous forme d'ouvrages papier et de CD-Rom.

La branche française du groupe LexisNexis est composée de :

- Juris-Classeur et ses marques et filiales (Litec, Légisoft, Infolib, La Documentation Organique) ;
- LexisNexis France.

En savoir plus :

- Groupe Reed-Elsevier : <http://www.reed-elsevier.com/> ;
- LexisNexis France : <http://www.lexisnexus.fr> ;
- Reed Business information (RBI) en France : <http://www.reedbusiness.fr/index.html> ;
- Elsevier en France : <http://www.elsevier.fr/html/index.cfm> ;
- Reed Exposition : <http://www.reedexpo.com>.

ANNEXE 4

PRESENTATION DE WOLTERS KLUWER

Wolters Kluwer¹ est un groupe multinational d'édition et de presse spécialisée, présent dans 26 pays, en particulier en Europe, aux États-Unis et dans la zone Pacifique-Asie.

Le groupe est né en 1985 du mariage de deux maisons d'édition hollandaises, Wolters Samson Group et Kluwer. Son siège est basé à Amsterdam.

Sa mission, apporter à ses clients :

- une information fiable exhaustive à forte valeur ajoutée ;
- des produits éditoriaux adaptés sur supports papier et électronique.

Ses produits couvrent toute la gamme de l'édition traditionnelle et électronique :

- ouvrages, annuaires, etc. ;
- fascicules à feuillets mobiles ;
- revues et journaux ;
- cédéroms ;
- banques de données, services en ligne www.lamylinereflex.fr ;
- programme de formation ;
- e-business.

Ses filiales en France sont :

- Les Editions Lamy ;
- Teleroute ;
- Val Informatique ;
- Ciceron ;
- Heliaste ;
- Liaisons Sociales ;
- Dalian ;
- Annonces et Formalités Légales ;
- Le Point Vétérinaire.

¹ <http://www.wolterskluwer.com>



ANNEXE 5

PRESENTATION DE L'ACFCI

L'Assemblée française des Chambres de commerce et d'industrie (ACFCI)¹ est au service d'un réseau constitué de 159 Chambres de commerce et d'industrie locales et 20 Chambres régionales.

Le réseau des CCI comporte 8 000 élus au suffrage universel des entreprises et 26 000 collaborateurs répartis sur l'ensemble du territoire et reliés par CCINET.

Le budget global des CCI représente 3,39 milliards d'euros.

Parmi leurs activités, les CCI forment chaque année 500 000 étudiants, apprentis, stagiaires, des premiers niveaux de qualification aux écoles supérieures de commerce ou d'ingénieurs.

Elles gèrent 121 aéroports, 180 ports, 36 plates-formes multimodales, 18 complexes routiers, 28 entrepôts ou parcs à vocation logistique, 55 palais des congrès et parcs d'expositions, 2 ponts.

Elles enregistrent plus de 780 000 formalités liées à la création, reprise ou transmission d'entreprise dans leurs Centres de formalités d'entreprises (soit environ 60% des formalités effectuées en France).

¹ <http://www.acfci.cci.fr>

ANNEXE 6

PRESENTATION DE FACTIVA

Factiva¹, une société de Reuters et Dow Jones, fournit des informations de haut niveau, telles les dépêches de Reuters et Dow Jones ainsi que Le Wall Street Journal, Le Monde, Les Echos, La Tribune etc...

Ces informations constituent une combinaison unique et inégalée de sources d'information. Factiva, offre la seule solution de contenu conjuguant de nombreuses interfaces de langues et un contenu multilingue couvrant plus de 8 000 sources d'information.

Les produits et services Factiva aident ses clients à intégrer des informations économiques au sein même de leur entreprise afin d'accroître l'intelligence organisationnelle, enrichir le contenu interne et externe et optimiser la gestion des connaissances (Knowledge Management). Les services Factiva de Content Management et d'intégration sont utilisés par les plus grandes sociétés de par le monde.

Basés sur des standards et une architecture ouverte, les produits Factiva fournissent des solutions flexibles, modulables et personnalisables afin de permettre une intégration facile au sein de l'entreprise. Son département de consulting éditorial et technique, son expérience dans la mise en place de taxonomies et son expertise dans les solutions d'intégration et les programmes d'e-learning reflètent l'approche innovante de Factiva : délivrer des solutions "au-delà du contenu".

Factiva a été sélectionnée comme l'une des sociétés du "Knowledge Management Top 100" et fait partie des "Top 100 sociétés à suivre" selon eContent Magazine.

¹ Pour plus d'information, visitez <http://fr.factiva.com/>.

ANNEXE 7

PRESENTATION DE SENSORIA TECHNOLOGY LIMITED

Sensoria Technology Ltd¹ a développé une technologie adaptée au traitement de l'information de l'ère numérique. Il s'agit d'une technologie radicalement nouvelle de reconnaissance de formes et d'apprentissage non supervisé, dynamique et adaptatif. Le premier domaine d'application de cette technologie est celui des moteurs de recherche intelligents à apprentissage autonome et indépendant de la langue. Sensoria est une société de technologie.

Les produits et les services exploitant cette technologie voient leur commercialisation confiée à d'autres entités. Ainsi, le domaine des moteurs de recherche est exploité sous la marque Influo, propriété de Sensoria. Financée depuis 2000 par des capitaux britanniques et français, Sensoria Technology Ltd est basée à Londres et à Paris. La société originelle, Sensoria SARL, ainsi que la technologie sont d'origine française (CNRS).

¹ <http://www.influo.com> ; <http://www.sensoriatechnology.com>



ANNEXE 8

PRESENTATION DE L'INPI

L'Institut national de la propriété industrielle (INPI)¹ délivre les titres de propriété industrielle (brevets, marques, dessins et modèles), centralise l'information du Registre National du Commerce et des Sociétés et le Répertoire Central des Métiers. En parallèle, l'INPI diffuse l'information et la documentation recueillies dans le cadre de ses missions; il propose des textes législatifs et réglementaires dans le domaine de la propriété industrielle. L'INPI est un établissement public sous tutelle du ministère de l'Economie, des Finances et de l'Industrie.

¹ <http://www.inpi.fr>

ANNEXE 9

PRESENTATION DU GFII

Le GFII¹ est une association qui regroupe les principaux acteurs publics et privés de l'industrie de l'information : producteurs, serveurs, intermédiaires, diffuseurs d'information, prestataires, éditeurs de logiciels, grands comptes acheteurs d'information...

Le GFII est pour ses adhérents un lieu d'information, d'échange d'expériences, de réflexion et d'action.

Son action se construit autour de quatre idées de base :

- dégager des objectifs communs et être une force de proposition pour mieux défendre les intérêts de l'industrie française de l'information ;
- proposer un large savoir-faire en matière d'information ;
- être une structure d'interface et un relais avec les instances gouvernementales et les autres acteurs du secteur de l'information ;
- valoriser l'industrie de l'information en France, en Europe et dans le monde.

Le GFII est coorganisateur d'i-expo, le salon de l'information numérique.

¹ <http://www.gfii.asso.fr>

ANNEXE 10

RAPPEL OU PRECISION, RAPPEL ET PRECISION ? QUESTION DE PERTINENCE

Les spécialistes et les experts, dans de nombreux domaines, ont l'avantage sur l'utilisateur, même professionnel, de souvent savoir jauger la qualité des outils d'un seul coup d'œil ou après une simple prise en main.

Cependant, en ce qui concerne les moteurs de recherche ou les systèmes documentaires, ces mêmes experts ne peuvent réellement s'appuyer dans cette tâche sur les notions classiques de "rappel" et de "précision". Du moins, pas sans se poser au préalable un certain nombre de questions sur la difficulté, voire le paradoxe, de leur mise en œuvre.

- un problème circulaire

La définition des deux notions se formule comme suit :

Rappel = Nb de docs pertinents obtenus / Nb de docs pertinents de la base

Précision = Nb de docs pertinents obtenus / Nb de docs obtenus

On leur ajoute leurs notions plus ou moins inverses de *bruit* et de *silence*.

Silence = 1 / Rappel (*le silence mesure donc le pourcentage de documents pertinents non renvoyés par l'outil pour une requête donnée*).

Bruit = 1 / Précision (*le bruit mesure donc le pourcentage de documents non-pertinents renvoyés par l'outil pour une requête donnée*).

Comment ne pas pointer l'aspect circulaire de ces notions qui, censées permettre de juger peu ou prou de la pertinence ("la qualité") d'un moteur de recherche, font elles-mêmes appel à cette même notion de pertinence ?

On rétorquera que c'est justement la tâche de l'expert humain de faire appel à son propre jugement de la pertinence. Mais alors surgit au premier plan la complexité de la mise en œuvre, voire sa faisabilité :

- combien d'experts doivent se prêter à l'exercice pour en atténuer l'effet subjectif ?

- combien chaque expert doit-il faire de requêtes pour parvenir à un résultat statistiquement valide ?

N.B. : Pour se convaincre de l'aspect déterminant des questions de masse critique et d'intervalle de validité, on songera à la représentativité des sondages ; que vaut un sondage sur une poignée de personnes ?

- que se passe-t-il lorsqu'une requête retourne un grand nombre de documents ?

L'expert étant censé les examiner toutes pour déterminer combien de documents (et lesquels) sont pertinents, veillera-t-il à choisir uniquement celles des requêtes qui retournent une petite dizaine de documents chacune (et donc à éliminer celles qui lui demanderaient une trop longue évaluation) ? Ou des requêtes d'un type prédéterminé ? Ou qui satisfassent tel ou tel critère arrangeant ?

N.B. : On songera à nouveau aux sondages et à l'importance de l'établissement du fameux "panel".

- comment la procédure s'applique-t-elle lorsque la base de documents comprend plusieurs milliers, plusieurs centaines de milliers, voire plusieurs millions de documents ?

Cela étant posé, continuons à nous interroger sur les notions de rappel et de précision, avec l'objectif d'en retirer des enseignements qui nous serviront par la suite.

- la précision, notion imprécise

Prenons un exemple exagérément simple : que vaut de savoir que, sur 100 documents obtenus, 80 sont pertinents, si les 20 premiers de la liste ne le sont pas ?

Autrement dit, non seulement le savoir-faire des experts dépasse de très loin la définition de la précision telle qu'elle a été formulée, mais les besoins réels de l'utilisateur la dépassent eux aussi. Il ne s'agit pas seulement de retourner un certain pourcentage de documents pertinents parmi tous les documents pertinents du fonds d'information, mais bel et bien de faire en sorte que la liste de documents retournés par le moteur soit ordonnée. Faute de quoi le travail est alors tout bonnement laissé à l'utilisateur¹.

La précision en elle-même se montre donc insuffisante, d'autant qu'elle est souvent décrite dans la littérature spécialisée comme quasi antinomique du rappel : l'expérience montrerait que plus elle est élevée, plus le rappel est faible, et inversement.

¹ Pour une réflexion autour de la question de la pertinence, voir aussi le livre blanc Moteurs de recherche : où est la technologie ? / www.influo.com.

Pour clore ce paragraphe, nous citerons le travail de chercheurs de l'Université du Québec à Montréal, qui, bien que datant de quelques années, n'en montre pas moins les limites de la notion de précision¹.

On y apprend notamment que, lors des très nombreuses expériences réalisées, le meilleur taux de précision obtenu pour le meilleur taux de rappel possible est inférieur à 50 %.

Autrement dit, un document sur deux n'est pas en rapport avec la requête. On imagine sans peine le peu d'utilité d'un système capable de telles "performances". On prie en tout cas pour que cette moitié de documents hors sujet figurent tous à la fin de la liste de résultats.

- principe de réalité

Venons en maintenant au "rappel", hors du cadre de sa formulation technique et examinons le plutôt "en situation".

Depuis quelques années, les chercheurs comme les éditeurs de logiciel ont fait porter leurs efforts sur deux axes principaux : d'une part, la reconnaissance (l'extraction) des *unités nominales polylexicales* (UNP), également appelées *synapsies* ; d'autre part, la caractérisation morphologique de manière à traiter les procédés linguistiques d'extension du groupe nominal. Nul doute que savoir identifier les noms propres et des groupes complexes, tels que *performances du système de repérage, comptes de la sécurité sociale, direction de la construction navale*, etc., permettent d'augmenter les performances de rappel et de précision.

On comprend aisément que lier ensemble les constituants de ces groupes puisse réduire le risque d'ambiguïté en soi, que générerait une indexation selon les termes pris isolément (*système*, ou *sociale*, ou *construction*, etc.).

Cependant, les documents "réels" recèlent également des problèmes très concrets.

Un premier problème vient du fait que, malgré des efforts souvent importants de correction-révision, les documents soumis à indexation comportent encore un nombre de "variantes orthographiques" autour de certains termes.

Prenons un exemple : *gouvernement* pourrait se retrouver écrit sous diverses formes, telles que *gouvrenement, gouvrenement*, etc., voire *ugouvernement* et aussi *gou-vernement*. Pour de multiples raisons, qui tiennent de la vie réelle des documents, ce type de variantes existe plus souvent qu'on ne le pense au premier abord (ainsi, l'exemple de *gouvernement* est-il tiré d'un grand quotidien économique).

¹ Analyse d'une méthode d'indexation automatique basée sur une analyse syntaxique du texte, 1996.

Dès lors, la question du rappel devient celle du "couplage" de ces variantes avec la forme "correcte" dont elles sont issues. Sinon, une requête portant sur *gouvernement* (pour reprendre notre exemple) risque de ne jamais retourner les documents où seules les variantes seraient présentes.

On notera que cette sorte de "correction orthographique à la source" prend encore plus d'importance lorsque les documents à indexer proviennent d'une phase de numérisation de documents papier, compte tenu des erreurs de reconnaissance souvent engendrées à cette occasion.

Une autre source d'erreurs résidera parfois dans la conversion, opérée par certains moteurs d'indexation, en tout-minuscules ou en tout-majuscules.

En conséquence, et par exemple, la différence de graphie entre l'*art* et l'organisme de régulation des télécommunications, l'*ART*, se verra supprimée ; autre exemple, par la disparition des accents en français, *cote* se verrait correspondre à *côté*, *côte*, *cote* ou *coté*. Ici encore, le "rappel", tout comme la précision, risque de souffrir de ces variantes orthographiques, le plus souvent indétectables pour une machine.

Intéressons-nous un instant, à présent, aux fameuses listes contrôlées, stop-list, ou autres listes de mots-vides.

Le problème qu'elles posent se comprend aisément. Un mot comme *or*, en français, est-il supprimé, que la locution *or noir* n'indexera plus les documents où elle est présente.

On rétorquera que c'est un cas bien connu du français, et que des règles et des "moulinettes" informatiques y pourvoiront efficacement. Mais tous les cas peuvent-ils toujours être recensés à l'avance, ou de manière suffisamment réactive ? Par exemple, lorsque l'expression *ni-ni* (*ni ni, ni / ni*, etc.) apparaîtra soudainement dans les documents liés à la politique de nationalisation / privatisation du nouveau gouvernement du président François Mitterrand (sans parler de l'orthographe à géométrie variable du nom de ce dernier, rencontrée aussi bien sous la forme "mitterrand" que "miterrand" dans un grand quotidien, pourtant lu et relu avant parution) ?

- rappel, précision et langage

Il n'y a guère de doute que l'imprévisibilité de la langue est à l'origine de ces difficultés pour la machine.

On en profitera pour ajouter que si les exemples cités se rapportent volontairement au français, le problème de la diversité du langage est, lui, général. Par conséquent, les réponses techniques devront être indépendantes des langues, et donc "autonomes".

C'est le cas de la technologie neuromimétique de Sensoria. Ce n'est pas le cas de la linguistique. La statistique, pour sa part, n'est pas adaptée par nature à la variabilité : elle exige au contraire une stabilité de l'échantillon¹.

On le voit, les questions de rappel et de précision, liées entre elles, sont donc définitivement dépendantes d'un traitement étendu du langage, tant la forme et le sens sont en réalité intimement imbriqués. Deux derniers exemples achèveront d'en convaincre.

Dans le but de favoriser le rappel, la logique suggère en première analyse d'établir un lien entre les différentes formes, par exemple selon le genre (tout au moins en français). La machine créera donc une connexion entre *femme* et *femmes*, *pieu* et *pieux*, etc. Ainsi, les documents comprenant l'une ou l'autre seront retournés quelle que soit le "nombre" demandé (pluriel ou singulier). Un automatisme de ce type fait malheureusement s'effondrer la précision dans beaucoup de cas, mais selon le contexte sémantique des documents, rendant de ce fait caduques les listes établies à l'avance.

A contrario, un français n'éprouve aucune difficulté à différencier les constructions telles que *rugueux* / *rugosité*, *animiste* / *animosité*, *animal* / *animaux* / *animation*, *gain* / *gainage*, *labour* / *labeur* / *labial*, etc., qui jouent pourtant à cache-cache avec leur "racine" telle que l'entend la linguistique. Ce mécanisme de formation des mots, connu entre autres sous le nom de lemmatisation, ne laisse apparaître aucun fondement statistique particulier, ni aucun automatisme évident. Sans aucun doute, il fait partie, avec d'autres, de la complexité de la langue française pour un étranger... et pour une machine. Ce nouveau problème, ainsi que tous les exemples déjà cités, mettent définitivement l'accent sur la nécessité, pour les systèmes d'indexation, de "savoir lire" les documents censés pouvoir être retrouvés par la suite².

- la connaissance au pluriel

C'est finalement une question de bon sens : rappel et précision ne sont qu'une autre façon de définir une "intelligence minimum" pour un moteur de recherche, à condition toutefois que les deux notions ne s'excluent pas mutuellement.

Toute la difficulté est de pouvoir s'adapter en permanence à la réalité des documents soumis au système.

Or ces documents ont été écrits par des êtres humains, qui ne sont pas infallibles, et ont subi parfois des transformations apportant elles aussi leur lot d'ambiguïtés par-dessus celles dues à la langue.

¹ Voir aussi un traitement du langage non-supervisé, Sensoria Technology, 2003 / www.influo.com

² cf. Au cœur de l'apprentissage et de la désambiguïsation : la Tête de lecture neuromimétique / www.influo.com.

Cette adaptation permanente nécessite une capacité d'apprentissage en temps réel sans présupposition au départ, laquelle devra surmonter les difficultés qui se posent face à telle langue ou telle autre, sachant que lorsque l'une paraît plus simple que l'autre sur un point précis, il n'est pas rare qu'elle présente elle-même des difficultés (d'autres ambiguïtés) sur un autre plan.

Apprentissage rime avec connaissance. On l'a vu, cette connaissance porte justement sur des plans différents. Elle est plurielle. Elle conditionne par exemple la capacité à "coupler" les variantes orthographiques telles qu'elles existent *réellement* dans les documents. Plus généralement, elle porte sur les liaisons entre forme et sens (le couplage devrait également concerner *sécurité_sociale* et *sécu*, *club_méditerranée* et *club_med*, etc.). Elle attache également à chaque terme un environnement sémantique. Et elle organise cet environnement en contextes hiérarchisés. C'est donc une connaissance multi-aspects et totalement dynamique, issue d'un traitement du langage aussi difficilement linguistique que statistique. L'apprentissage dynamique et adaptatif auquel doit se livrer un système d'indexation offrant rappel et précision rappelle, en l'occurrence, celui de l'enfant, souvent incapable de fournir la règle de formation de *anima-lité / anima-tion*, mais sans hésitation sur leur divergence fondamentale. C'est un apprentissage dit "non-supervisé", ou "sans professeur", et donc adaptatif par nature. Influo iSearch et sa tête de lecture neuromimétique respecte ce principe¹.

- conclusion

Les notions de rappel et de précision nous semblaient circulaires dans leur formulation. Et ce n'est pas la littérature sur le sujet, qui pointe leur comportement antinomique obligeant à un compromis insatisfaisant dans son principe même, qui nous aidera à les voir différemment. Devoir privilégier l'une au détriment de l'autre paraît en effet un choix douloureux. Le bon sens laisserait penser que rappel et précision, c'est bien une question de pertinence, et donc de connaissance. Mais une connaissance obligatoirement non figée et forcément en rapport direct et constant avec les documents à examiner.

"Pour utiliser ces deux mesures, il faut pouvoir déterminer les documents pertinents pour chaque requête. Il faut alors disposer de jugements de pertinence pour les requêtes qui donnent pour chaque requête la liste de documents pertinents. Ces jugements de pertinence sont généralement donnés par des personnes [ou par des systèmes] qui connaissent la collection de documents."². *

¹ Cf. iSearch et apprentissage : l'enfance de l'art / www.influo.com.

² Analyse d'une méthode d'indexation automatique basée sur une analyse syntaxique du texte, 1996. Le texte mis entre [] a été rajouté par l'auteur.

* Pour se convaincre de la difficulté, pour l'expert intéressé, à mettre en œuvre les notions de rappel et de précision, on prendra connaissance des chiffres suivants, cités par les auteurs de la thèse.

Caractéristiques de la collection, des requêtes et des jugements de pertinence.

Documents

Nombre 177

Taille minimum (en mots) 46

Taille maximum (en mots) 1472

Taille moyenne (en mots) 271,8

Requêtes

Nombre 40

Taille minimum (en mots) 5

Taille maximum (en mots) 25

Taille moyenne (en mots) 12,1

Jugements de pertinence

Nombre minimum de documents par requête 1

Nombre maximum de documents par requête 35

Nombre moyen de documents par requête 6,7

ANNEXE 11

BIBLIOGRAPHIE

- J. Aitchison, A. Gilchrist, "Thesaurus construction : a practical manual", ASLIB, 1972, distribué par Chicorel Library Publishing Group, NY.
- D. Amidon, "Le knowledge management pour une entreprise innovante", Editions d'organisation ;
- "Outils de management. Capitalisation d'expérience", Afnor, FD X 50-190, septembre 2000 ;
- "Archivage électronique, aspects technique et juridique", Afnor 2000, Pratique Collection ;
- "Prestations de veille. Prestations de veille et prestations de mise en place d'un système de veille", Afnor, XP X50-053 ;
- J.F. Ballay, "Capitaliser et transmettre les savoir-faire de l'entreprise", Eyrolles, Collection de la Direction des Etudes et Recherches d'Electricité de France, 1997 ;
- A. Bensoussan, "Informatique, Télécoms, Internet", Editions Francis Lefebvre 2001 (2ème édition n° 195 et suivants) ;
- J.Y. Bück, "Le management des connaissances", Editions d'organisation, 1999 ;
- W. Bukowitz, Ruth Williams, "Gestion des connaissances en action", Editions Les Echos ;
- Dieng, "Méthodes et outils pour la gestion des connaissances", Editions Interéditions ;
- C. Donzel, "L'art du management de l'information", Editions Village Mondial ;
- L. Edvinsson, M. Malone, "Le capital immatériel de l'entreprise", Editions Maxima, 1999 ;
- J.L. Ermine, "Les systèmes de connaissance", Editions Hermès ;
- GFIL, groupe de travail présidé par Alain Bensoussan et Danièle Delaval, "Bases de données : Guide juridique pratique pour les producteurs, les serveurs et les utilisateurs", 1995 ;

- F. Guillet, H. Briand, C. Donzel, "Extraction et gestion des connaissances", Editions Hermès ;
- F. Harvard, Jacques Chaize, "Le knowledge management", Editions d'organisation, Collection Harvard business review ;
- J. Honeycutt, "Knowledge Management Stratégies", Microsoft Press ;
- F.W. Lancaster, "Vocabulary Control for Information Retrieval", Second Edition, Information Resources Press, Arlington, Virginia, 1986 ;
- P. Lévy, M. Authier "Les arbres de connaissance", Editions La Découverte ;
- P. Maret, J.M. Pinon, "Ingénierie des savoir-faire", Hermès Collection Informatique et Gestion, 1997 ;
- E. Morin, "Relier les connaissances", Editions Le Seuil.
- OCDE, "Société du savoir et gestion des connaissances" ;
- J.Y. Prax, "Manager la connaissance dans l'entreprise", INSEP Editions, 1997 ;
- J.Y. Prax, "Le Guide du Knowledge Management", Editions Dunod ;
- J.Y. Prax, "Le manuel du knowledge management : une approche de 2^{ème} génération", Editions Dunod ;
- L. Rosenfeld, P. Morville, "Information Architecture for the World Wide Web", O'Reilly & Assoc., Sebastopol, CA, 1998 ;
- K. E. Sveiby, "Knowledge Management : la nouvelle recherche des entreprises", Editions Maxima ;
- J.C. Tarondeau, "Le management des savoirs", PUF, collection Que Sais-je, 1998 ;
- D. Thévenot, "Le partage des connaissances", Lavoisier, 1999 ;
- R.C. Tisseyre, "Knowledge Management : théorie et pratique de la gestion des connaissances", Editions Hermès ;
- B. Weinberg, "Complexity in Indexing Systems – Abandonment And Failure: Implications for Organizing The Internet", ASIS, 1996 Annual Conference Proceedings.