
I. 1.1 LA QUESTION DU SENS EN RECHERCHE D'INFORMATION

La question du sens en recherche d'information est liée à une approche de la langue sur le plan des idées, autrement dit une approche notionnelle d'un champ de reformulations possibles d'une requête émise par un utilisateur humain. Lorsqu'il s'agit d'analyser le sens d'un texte renvoyé par un moteur de recherche à la suite d'une requête en langue naturelle, on se trouve au cœur de la problématique de la constitution de connaissance au moyen d'un véhicule humain qu'est la pensée, qui se donne les moyens d'atteindre le champ du savoir, par la médiation d'un processus automatique mis en œuvre par un ordinateur.

Les Sciences de l'Information et de la Communication sont à la fois transversales et situées au carrefour de plusieurs disciplines. Au carrefour de la linguistique et de l'informatique, le Traitement Automatique des Langues Naturelles (TALN) est une discipline qui est née dans les années 50 et dont l'idée première était de modéliser la langue naturelle, afin de la plier aux exigences d'un traitement par des processus informatiques. Or les chercheurs se sont rapidement aperçus, notamment dans les cas de décalage notionnel en traduction automatique, que la langue ne pouvait guère être considérée comme une série de chaînes de caractères reliées à un dictionnaire, mais que chaque mot au contraire, faisait sens. En traduction automatique, on parle de décalage notionnel lorsque, pour décrire un état de choses ou un objet, la langue-source et la langue-cible ne structurent pas la réalité extralinguistique de la même façon.

Ce problème de décalage, F. de Saussure le mentionnait déjà en 1916 dans son exposé consacré à la valeur du signe linguistique dans son "Cours de linguistique générale" : "le signe linguistique unit non une chose et un nom, mais un concept et une image acoustique" ¹ ainsi, "si les mots étaient chargés de représenter des concepts donnés d'avance, ils auraient chacun, d'une langue à l'autre, des correspondants exacts pour le sens ; or il n'en est pas ainsi.(...) Le français mouton peut avoir la même signification que l'anglais *sheep*, mais non la même valeur, et cela pour plusieurs raisons, en particulier parce qu'en parlant d'une pièce de viande apprêtée et servie sur la table, l'anglais dit *mutton* et non *sheep*. La différence de valeur entre *sheep* et *mutton* tient à ce que le premier a à côté de lui un second terme, ce qui n'est pas le cas pour le mot français".²

1 Ferdinand de SAUSSURE - Cours de linguistique générale - Payot (1972) p 98

2 Idem p 160/161

Très vite, le traitement automatique des langues, s'est heurté à la question du sens. A l'heure du multimédia, la question du sens est d'autant plus prégnante qu'elle concerne le traitement d'importantes masses de données, par les entreprises comme par le grand public. Le traitement de l'information, qu'elle soit monolingue ou multilingue est en effet l'un des défis cruciaux auxquels doit faire face la société de l'information. L'évaluation de la pertinence des moteurs de recherche, qui s'effectue grâce à des technologies de traitement automatique des langues, mais qui ne doit cependant pas être confondue avec celui-ci, est au cœur de l'analyse des notions de bruit et de silence. Or le bruit, comme le silence ne constituent en aucun cas une information utile et ne concourent en aucune façon à la constitution d'un référentiel de connaissances sur un sujet donné.

Tout comme le silence plonge l'utilisateur d'un moteur de recherche dans une espèce de désarroi du néant de l'information, le bruit va l'immerger dans un univers où règne la confusion des idées par absence de pertinence, comme de structuration conceptuelle. Le traitement de l'information par un moteur de recherche doit, en cas de succès d'une recherche, permettre à l'utilisateur d'achever le processus d'interrogation du moteur en ayant les idées claires sur un sujet donné. Car il s'agit bien de traiter des idées et des thématiques et la confusion est toujours possible, comme aux premières heures de la traduction automatique, avec à la clé, l'absence de résultat, ou le risque de ne pas avoir les idées claires sur un sujet donné. Or les idées ne peuvent s'avérer claires que lorsqu'elles sont structurées.

L'explosion des médias de masse et du volume des données requêtées confronte aujourd'hui le chercheur à la problématique de l'efficacité de la recherche d'information. Il est aujourd'hui nécessaire dans le cadre de recherches effectuées par des cellules de veille, d'aller droit au but, de disposer de la bonne information au bon moment. Et ces nouveaux impératifs ont des implications aussi bien sur le plan de la constitution de connaissances que sur le plan économique. La recherche d'information rejoint alors la problématique de l'intelligence économique qui lui est transversale, mais demeure néanmoins en lien très étroit avec elle.

Il s'agit dans ce cas de capitaliser la bonne information en fonction d'un certain nombre d'indicateurs. Or ces indicateurs sont des indicateurs qui renvoient vers le sens, c'est-à-dire l'alimentation d'une pensée comme d'un champ de connaissance. Pour vérifier la pertinence d'une information il est nécessaire de l'évaluer, ce qui suppose la mise en place de protocoles. Ces protocoles évaluent la validité d'une information, autrement dit, son sens. L'aspect de validité de l'information recueillie est capital dans une perspective de constitution de connaissance et demeure indissociable de celle-ci. C'est ainsi que l'évaluation, bien qu'elle mette en place des protocoles scientifiquement vérifiables, est susceptible de revendiquer son appartenance au champ des Sciences de l'Information et

de la Communication, par son appartenance à l'univers du sens, d'une part et sa transversalité entre plusieurs disciplines d'autre part. Il s'agit en fait d'évaluer la valeur d'une information et la mise en exergue d'un sens, qui permet non seulement de naviguer à l'intérieur d'un corpus mais encore de constituer des référentiels de connaissances.

I. 1.2 RECHERCHE DU SENS

Afin de mieux cerner l'importance des contextes, des usages et de l'appropriation dans la recherche de l'information, nous allons approfondir notre analyse à partir d'un champ caractéristique et relativement précis de pratiques, d'outils, de relations et d'intentionnalités : celui du Knowledge Management. La constitution d'un champ de connaissances par capitalisation, la création de connaissances, le partage des connaissances, constituent un contexte particulier, très avancé, de la recherche d'information, et un ensemble d'exigences particulières pour son évaluation puisque la valeur de la connaissance ne s'apprécie pas de la même façon que la valeur de l'information.

Le Knowledge Management est à la fois une discipline récente, et une pratique incontournable des TIC, au moins dans les grands groupes et réseaux d'acteurs. Il se caractérise, de notre point de vue, par une composante double dont les éléments sont indissociables : humaine et technologique.

La centralité du sujet humain est un postulat de base. En effet, la connaissance est par nature : humaine, relationnelle, sociale.

[Polanyi, 1967]³, puis [Nonaka, 1994]⁴ identifient la dimension tacite de la connaissance, et lui reconnaissent une importance essentielle. Pour [Polanyi, 1967], qui se concentre sur la connaissance scientifique, aucun savoir ne peut être complètement explicite. Chacun comporte une dimension inexprimée ou inexprimable, des convictions personnelles et une tendance à adopter certains présupposés de base, que nous ne sommes pas toujours en mesure d'expliquer. Dans tel ou tel contexte, par exemple celui d'une communauté scientifique, des convictions personnelles sont partagées et deviennent par accord non formulé la trame sous-jacente d'un savoir tacite. Ces connaissances tacites participent activement à la créativité générative de connaissances sociales et scientifiques nouvelles.

³ POLANYI, M (1967) "Tacit Knowing", THE TACIT DIMENSION. Londres: Routledge & Kegan Paul.

⁴ Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1), 14-37.

Nonaka et Takeuchi reprendront la notion de connaissances tacites dans un contexte élargi aux organisations et à leurs processus sociaux, et développeront un modèle connu : pour [Nonaka & al. 1996]⁵, identifiés comme pères fondateurs du KM, celui-ci s'articule autour du concept des savoirs tacites, des "schémas mentaux, croyances, points de vue si profondément ancrés que nous les prenons pour acquis", et qui imprègnent complètement l'action. Les dynamiques possibles de partage du savoir, dans ce cadre, sont résumées par ce schéma

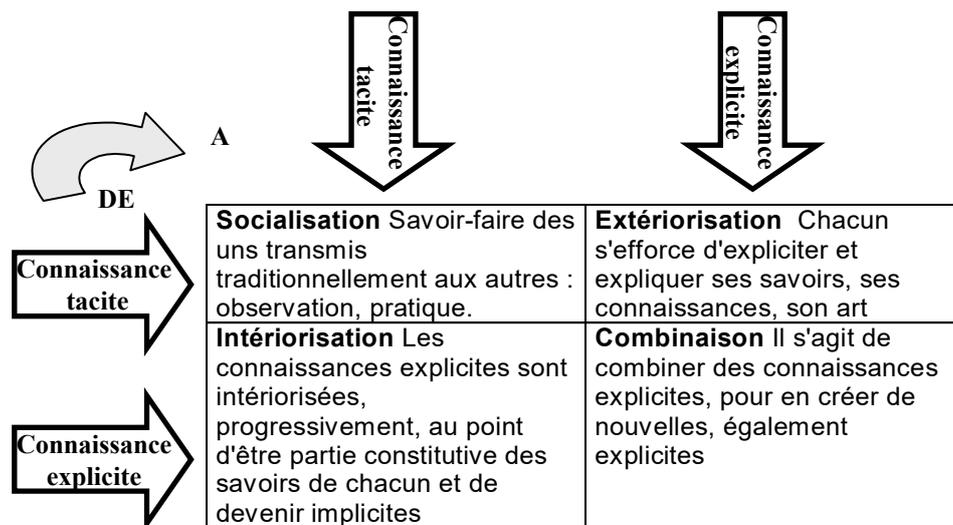


FIGURE 1 - SCHEMA DE LA DYNAMIQUE DU KM D'APRES NONAKA ET HIROTAKA TAKEUCHI 1996

L'accent mis sur le savoir tacite et son rôle, donne une distance face aux technologies et traduit la recherche d'une "spirale du savoir".

Toutefois, la composante humaine du Knowledge Management n'est pas suffisante : elle s'articule en permanence au développement des TIC et de leurs usages. La relation entre facteur humain et facteur technologique est au cœur de la composition du Knowledge Management, qu'on le considère comme une pratique ou comme une discipline émergente des sciences humaines. Il s'agit, au reste, d'une interrelation dynamique et toujours en mouvement, sous ses deux aspects : celui de la connaissance humaine, dont la vocation est de créer en permanence de nouvelles connaissances, de conquérir de nouveaux territoires ; celui des TIC, toujours en prise avec l'innovation, d'une part, et à l'évolution de leurs usages et socialisation d'autre part.

⁵ Nonaka, I., Takeuchi, H. et Umemoto, K. (1996). A theory of organizational knowledge creation. International Journal of Technology Management, 11(7/8), 833-845

Les concepts de la sociologie de l'innovation, particulièrement le « modèle de la traduction » proposé par [Akrich & al. 1988]⁶, permettent d'approfondir cette notion d'interrelation dynamique connaissance humaine – usages des TIC, constitutive du champ du KM.

Ce modèle s'appuie sur la théorie de l'acteur-réseau pour caractériser l'innovation comme un processus non linéaire, interactif et itératif avec l'idée d'une transformation et d'un déplacement des acteurs et des objets. L'objet technique résulte d'une négociation qui évolue en fonction des traductions et des processus d'intéressement des acteurs impliqués. Le modèle, appliqué par ces chercheurs aux raisons du succès d'une innovation, peut être transposé dans la problématique de l'évolution conjuguée de l'innovation TIC, de ses usages, et des nouveaux besoins en connaissances dans le cadre de pratiques et dispositifs de Knowledge Management.

La théorie de l'acteur-réseau introduit une perspective dynamique dans le dispositif de management de la connaissance, ou tout au moins dans cette partie précise du dispositif qui nous intéresse : la recherche des informations qui vont nourrir les processus de création-partage des connaissances, la recherche d'informations qui se transforment en connaissances, qui émergent comme connaissances tacites, ou qui sont capitalisées sous forme de connaissances explicites. Dans le schéma suivant, nous appelons cette recherche spéciale « contenus d'informations vs connaissances »

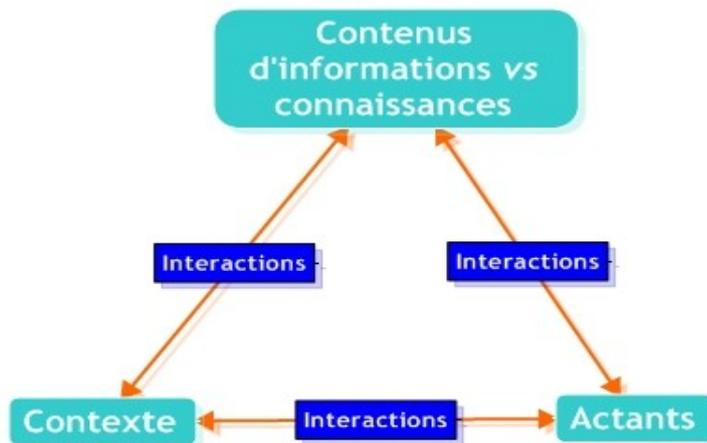


FIGURE 2 - « CONTENUS D'INFORMATIONS VS CONNAISSANCES »

⁶ AKRICH, Madeleine, LATOUR, Bruno, CALLON, Michel. A quoi tient le succès des innovations. Annales des mines, n°11, juin 1988, p. 16

L'un des pôles de ce triangle concerne les « actants ». Il apparaît en effet qu'un autre apport de la théorie de l'acteur-réseau est de considérer comme entité agissante non seulement le sujet humain, mais aussi des sujets non-humains. Les actants ne sont pas considérés comme des sujets isolés, mais bien comme les parties prenantes du processus, de la création, de l'usage. Nous nous éloignons, comme souhaité, d'un individu, face à l'océan informationnel, armé de son seul moteur de recherche et/ou instrumentalisé par celui-ci. Le réseau – humain et outils – opère les transformations, génère le changement, au sein d'un ensemble d'interactions que l'on peut identifier.

La connaissance peut se construire à partir d'informations, mais elle « n'est pas plus « déposée » dans le psychisme individuel que dans un document. Les individus sont les dépositaires de schèmes ou de compétences psychiques et physiques mais c'est uniquement dans l'action, c'est-à-dire dans le cadre de la réalisation des transactions, qu'ils rendent manifestes les connaissances qui les caractérisent en tant qu'acteurs à travers leurs productions sémiotiques ou physiques. De plus, la plupart du temps, le déroulement des transactions nécessite le recours à divers instruments qui peuvent être eux-mêmes de nature principalement sémiotique, un discours, un document, ou physique, un outil »⁷.

Le contexte, est celui des domaines d'activités dans lesquels les acteurs travaillent et conduisent des projets. Les actants sont aussi bien les sujets humains que les outils technologiques qu'ils emploient. L'ensemble est en mouvement au travers d'interactions multiples. Ce modèle permet de rendre compte de situations concrètes au cours desquelles nous pouvons tester notre approche des moteurs de recherche et de leur évaluation.

⁷ Zacklad, M. (2004), Transférabilité des connaissances : une reconceptualisation de la distinction tacite /explicite, in Actes du colloque « En route vers Lisbonne » : 1er Colloque luxembourgeois sur l'économie de la connaissance dans une perspective européenne, 12-13 Octobre, Luxembourg