

# **Veille Scientifique et Technique :**

## **Modèle d'observation et d'analyse issu de l'intelligence collective en nanosciences**

**Conférencier invité :**

**Sahbi SIDHOM**  
Université de Lorraine & LORIA Lab., France

**Dans la journée d'études sur :**

**La veille : plus qu'une démarche, une stratégie au service des entreprises, des citoyens et  
des élus. Université Jean Moulin Lyon 3. France**

**Organisée par :**

**Université Jean Moulin Lyon 3. France. EL-HACHANI Mabrouka (Eds.)**

**Résumé long :**

De nos jours, la diversité des outils et applications en « veille scientifique et technique » recouvre plusieurs problématiques et ouvre de nouvelles pistes de réflexion. Ce travail de recherche s'articule autour des « traitements automatiques de la langue naturelle » (En. NLP : naturallanguageprocessing), l'organisation des connaissances (En. KO : knowledgeorganization) et la veille en IST : informations scientifique et technique. Dans cette conjoncture, l'apport de la ré-indexation sociale a été mis en valeur : l'évolution des pratiques documentaires avec le web et ses usages actuels. Particulièrement, nous montrons un intérêt qui marque son importance par la valorisation des contenus <sup>[6]</sup> dans les réseaux sociaux professionnels. Pour nous, il s'agit d'une nouvelle orientation de recherche qui donne un apport au domaine d'étude sur la ré-indexation des ressources et qui réactualise la réflexion sur l'apport et l'action de l'utilisateur <sup>[11], [12]</sup>.

Prenons l'exemple d'un veilleur ou un professionnel de l'information qui est plongé dans une activité de condensation/ de synthèse/ d'écriture, sans contraintes de style ou de rédaction, comme le cas d'un usager sur les réseaux sociaux. Son action préconise au moins l'existence d'une idée mentale développée. Dans ce contexte, le résultat (ou l'objet "écrit" de cette action) nous offre la réflexion première pour étudier les structures syntaxiques inhérentes. Ainsi, l'objet écrit est vu comme un ensemble cohérent d'unités linguistiques plus ou moins complexes. Chaque unité s'articule avec les autres et contribue à la réalisation d'un équilibre structurel : à la fois morphologique, syntaxique et sémantique.

Dans une autre considération, l'action d'écriture sans contraintes préconise le relâchement des règles et styles dans la langue de l'auteur : il s'agira bien de la réalisation d'un objet écrit propre à son auteur avec ses structures conceptuelles, morpho-syntaxiques et sémantiques.

Dans une dernière considération, l'étude nous amène à examiner des corpus encapsulant les objets écrits de divers acteurs, de les analyser et de découvrir dans cette observation cognitive : « quelle est la diversité des règles employées (aussi bien les nouvelles règles véhiculées) ? » Il s'agit d'examiner les règles dans l'ordre conceptuel, morpho-syntaxique et sémantique<sup>[14]</sup>.

Dans cette réflexion à problématiser, le choix porté à un modèle morpho-syntaxique est essentiel et ne doit pas, d'une part, perdre de vue la qualité des résultats d'un analyseur (ou parseur) placé dans un système ouvert à l'usage<sup>[13]</sup>; Et d'autre part, la qualité de conception (ou le formalisme d'organisation des connaissances) doit rapprocher les concepts théoriques et pratiques sur l'objet d'étude : objets ou traces écrites dans les contenus (documents, IST, analyses de veille, annotations, avis, tags sociaux, etc.) et qu'on pourra étendre aux enquêtes d'opinion (ressources ouvertes sur le web, web des usages, opinions sociaux-professionnels, etc.)<sup>[16]</sup>.

Le corpus utilisé récemment (après celui de l'INA comme corpus d'études et d'implémentation) pour l'observation et la manipulation est fondé sur une enquête d'opinion associant des questions fermées et ouvertes dans le cadre de la constitution d'un Club thématique en nanosciences et nanotechnologies (nanoMétrologie). Le club nanoMétrologie est coordonné par des grands acteurs du domaine en France, comme le Laboratoire National d'Essai et de Métrologie (LNE), le Centre de Compétences NanoSciences France (C'NANO) qui rassemble l'ensemble des laboratoires du domaine. Une hypothèse faite sur la nature du texte libre véhiculé dans les questions ouvertes est qu'il ne s'agit pas de contraintes de style ou de rédaction proposées ou imposées (aux répondants de l'enquête). Pour nous, il s'agit d'un texte libre à préparer pour l'analyse automatique et l'extraction des connaissances<sup>[7]</sup>.

L'objectif de ce travail est d'apporter un modèle adaptatif à l'analyse du langage dans les ressources ouvertes (corpus documentaires, corpus web, corpus web-usages et corpus d'opinions) pour la ré-indexation sociale.

Au terme de ce travail de recherche, nous prendrons appui sur l'analyse de L. Bloomfield<sup>1</sup> (1933) qui soulignait particulièrement que : "*Ce qui concerne le sens est le point faible des études sur le langage, et le restera jusqu'à ce que nos connaissances aient avancé bien loin de leur état actuel*".

Cet article s'articule autour de trois idées pour répondre à la réflexion problématisée, à savoir : « comment valoriser les analyses de contenus et les processus inhérents pour fonder les actions de l'utilisateur et ses apports à l'information et à la connaissance ? ». Pour se faire, nous présentons le processus d'indexation et les apports cognitifs à l'élaboration d'une grammaire d'analyse. Puis, nous présentons le processus de réindexation pour la valorisation des contenus par les usages. En dernier, les expérimentations qui ont permis de développer la méthodologie pour diagnostiquer la structure interne dans un réseau professionnel (celui du club nanoMétrologie) et pour proposer les recommandations adaptées.

---

Mots-clés : réindexation sociale, observations cognitives, organisation de la connaissance, intelligence collective, nanosciences, corpus d'enquêtes d'opinion.

## Biographie

- [1] Bachimont B. (1999). La documentation au coeur du processus de production. in Dossier de l'Audiovisuel, Janvier-Février 1999, n°83, INA-Publications, p.38-39.
- [2] Bachimont B. (2004). Signes formels et computation numérique : entre intuition et formalisme. In H. Schramm, L. Schwarte & J. Lazardzig (Eds.), *Instrumente in Kunst und Wissenschaft - Zur Architektonikkultureller Grenzen im 17. Jahrhundert*. Berlin: Walter de Gruyter Verlag.
- [3] Browne G. (1996). Automatic indexing and abstracting. in *Indexing in Electronic Age Conference*, Robertson, NSW 20-21 April 1996, Australian Society of Indexers, 8p.
- [4] Carbonell J.G., alii. (1997). Translingual Information Retrieval: a comparative evaluation. in *Proceedings IJCAI-97*, Nagoya, Japan, Morgan Kaufmann, San Mateo, CA (1997).
- [5] Guimier-Sorbets A-M. (1993). Des textes aux images : accès aux informations multimédias par le langage naturel. *Documentaliste – Sciences de l'information*, 1993, vol.30, n°3, p.127-134.
- [6] Régimbeau G. (1998). Accès thématiques aux d'art contemporaines dans les banques de données. in *Documentaliste Sciences de l'Information*, Volume 35, n°1, janvier 1998, p.15-23.
- [7] Lambert P., Sidhom S. (2011). Problématique de la veille informationnelle en contexte interculturel : étude de cas d'un processus d'identification d'experts vietnamiens". in *Proceedings : ISKO-Maghreb'11 – Concept and Tools for Knowledge Management (KM)*. ESCE-University of la Manouba Edition. Hammamet (Tunisia) May. 2011.
- [8] Maniez J. (1993). L'évolution des langages documentaires. *Documentaliste et Sciences de l'information*, 1993, vol.30, n°4-5, p.254-259.
- [9] VAN SLYPE G. (1987). Les langages d'indexation : conception, construction et utilisation. in *dans les systèmes documentaires Paris : Editions d'organisation*, 1987. 277 p. - (Systèmes d'information et de documentation).
- [10] Maret P., Pinon J-M., Martin D. (1994). Capitalisation of consultants' experience in document drafting. *Conference Proceedings RIAO 1994*, Printed by CID Paris France, p.113-118.
- [11] Calmet J., Maret P. (2013). Toward a trust model for knowledge-based communities. *WIMS 2013*: 47.
- [12] Mseddi R., Sidhom S., Ghenima M., Ben Ghezala H. (2011). From information to decision: information management methodology in decisional process. in *Proceedings SIIE'2011 : Information Systems and Economic Intelligence (SIIE'2011) vol.1 (2011) pp.219-226*, IGA Edition. Marrakech (Morocco) Feb. 2011.
- [13] B. Bertin, V. Scuturici, J.M. Pinon, E. Risler. (2012). CarbonDB : a Semantic Life Cycle Inventory Database. in *Conference on Information and Knowledge Management (CIKM) 2012*, Maui, Hawaiï.2012.
- [14] Sidhom S. (2002). Plateforme d'analyse morpho-syntaxique pour l'indexation automatique et la recherche d'information: de l'écrit vers la gestion des connaissances. Thèse de doctorat de l'Université Claude Bernard Lyon1. France. Mars 2002. p.247.
- [15] Sidhom S., and Lambert P. (2011). "Information Design for Weak Signal detection and processing in Economic Intelligence: case study on Health resources". in *Proceedings SIIE'11: Information Systems and Economic Intelligence*, IGA Edition. Marrakech (Morocco) Feb. 2011.

[16] Sidhom S. (2013). Conjoncture des processus d'indexation et de gestion des connaissances : vers la réindexation par les usages. in *Didactiques et métiers de l'humain et de la relation : nouveaux espaces et dispositifs en question.* (direction de Frisch M.), ID Collection L'Harmattan. pp.85-125. Paris, 2013.

[17] SIDHOM S. et LAMBERT, P. (2014). « Project Management in Economic Intelligence: NooJ as diagnostic Tool for Nanometrology cluster ». (chapter) in Cambridge Scholars Publishing (in 2014) :NooJ for NLP. "à paraître en 2014", (sous la direction de) SvetlaKoeva, Slim Mesfar, Max Silberztein (Eds.)

[18] Donabédian A., Khaskarian V., Silberztein M. (2013). NooJ Computational Devices. In *Formalising Natural Languages with NooJ.* Eds. Cambridge Scholars Publishing: Cambridge.

[19] Chaudiron S., Ihadjadene M. (2010), « Electronic Information Access Devices : Crossed Approaches and New Boundaries », in *Information Science*, London, ISTE, p. 167-189

[20] COUZINET Viviane, 2009. Complexité et document : l'hybridation des médiations dans les zones en rupture, RECIIS, *Electronic journal of communication information and innovation in Health*, vol.3, n°3, p. 10-16.