

## Bourses doctorales 2015

Candidature présentée par l'équipe Wimmics (Inria-I3S) et le Département Communications Multimédia (Eurecom)

### TITRE

#### ***Conception d'une méthode d'évaluation des moteurs sémantiques de recherche exploratoire***

*SOUS-TITRE : Evaluation des moteurs sémantiques de recherche exploratoire dans des données liées multimédia : conception d'une méthode centrée utilisateurs, basée sur une modélisation du processus de recherche exploratoire.*

### RÉSUMÉ

Les moteurs de recherche exploratoire (MRE) sont des applications dont la fonction est d'aider les utilisateurs à explorer un domaine de recherche méconnu pour y faire des découvertes. D'une certaine manière, elles apportent une visualisation et des possibilités d'interaction au-dessus de techniques d'analyse de données (*data mining*). Les méthodes d'évaluation de ces systèmes visent à vérifier que ces derniers aident effectivement les utilisateurs à réaliser et tirer les bénéfices de leurs tâches d'exploration. On constate toutefois, que les méthodes existantes ne permettent pas véritablement cette vérification.

La raison principale est qu'on manque encore de connaissances sur le processus de recherche exploratoire ainsi que sur les comportements qu'adoptent les utilisateurs lors de telles tâches. Un préalable à la conception d'une méthode d'évaluation de MRE est de mieux comprendre la relation tenue entre l'utilisateur et le système. Cela inclut entre autres les besoins en information, les processus cognitifs et stratégies de recherches des utilisateurs.

L'objectif de cette thèse est d'élaborer une méthode d'évaluation des MRE basée sur une modélisation du processus de recherche exploratoire issue d'études empiriques d'utilisation de tels systèmes. Cette méthode sera applicable dès les premières étapes de la conception des MRE. De plus, cette thèse produira un certain nombre de briques logicielles, configurables et ré-utilisables, instrumentalisant les préconisations de la méthode proposée dans cette thèse pour faciliter la conception et le développement de MRE.

### POSITIONNEMENT PAR RAPPORT AUX AXES SCIENTIFIQUES DU LABEX

- *Axe scientifique : Réseaux orientés contenu.* – Ce projet de thèse se propose de contribuer à : 1) « développer les possibilités de recherche de contenus intelligents sur le réseau » (contenus

multimédia), 2) répondre à la question de « la difficulté de trouver une information pertinente ou fiable » (par un travail en particulier sur les méthodes d'évaluation des qualités de pertinence et de fiabilité, du point de vue de l'utilisateur), 3) « l'élaboration d'outils (...) offr[ant] la possibilité de recherches de contenu intelligent sur le réseau » (par l'évaluation et la reconception de MRE).

- *Axe applicatif : Interaction avec les SHS.* – L'UCN ayant « pour but de placer l'utilisateur au cœur du réseau », ce projet de thèse se propose de contribuer à ce but, par l'accent mis sur l'utilisateur des réseaux et ses processus d'utilisation des données présents sur les réseaux, et par la mise en œuvre de modèles et de méthodes issues des SHS.

## DIRECTION DE THÈSE

- Directeur de thèse : Fabien Gandon, DR, responsable scientifique de l'équipe Inria-I3S Wimmics<sup>1</sup>
- Co-directeur portant le projet : Alain Giboin, CR (Equipe Inria-I3S Wimmics)
- Co-directeur : Raphaël Troncy, MCF (Eurecom, Département des Communications Multimédia)

Le doctorant sera embauché par l'Inria.

## PRÉSENTATION DU SUJET DE THÈSE

### Contexte

#### *Contexte institutionnel*

- Cette thèse est l'une des suites données aux travaux de doctorat de Nicolas Marie sur Discovery Hub [8], le moteur sémantique de recherche exploratoire dans les données liées, travaux menés dans l'équipe Wimmics sous la direction de Fabien Gandon. Une partie non négligeable de ces travaux a porté sur l'évaluation utilisateurs de Discovery Hub. Ces travaux ont révélé la nécessité d'améliorer les méthodes existantes d'évaluation de ce type de systèmes et ont conduit à mettre au point une première série d'améliorations.
- Cette thèse fait aussi suite aux travaux de doctorat de Houda Khrouf, menés sous la direction de Raphaël Troncy, sur EventMedia<sup>2</sup> [3], un autre moteur sémantique de recherche exploratoire dédié à des données liées décrivant des événements (par exemple des événements culturels) et des médias les illustrant. Un système de recommandation ainsi que de nouvelles interfaces de navigation par facettes parallèles ont été conçus pour faciliter l'aide à la décision. La navigation par facettes<sup>3</sup> parallèles est aussi le point central de l'application 3cixty<sup>4</sup>, co-développée par EURECOM, qui sera une application officielle de l'exposition universelle organisée à Milan en 2015 (EXPO 2015) et qui sera susceptible d'être utilisée par 20 millions de visiteurs attendus.
- Le candidat à la bourse de thèse est déjà identifié. Il s'agit d'une candidate : Emilie Palagi. Emilie est aujourd'hui diplômée du M2 Sociologie et Ergonomie des Technologies numériques de l'UNS, avec une spécialisation en Ergonomie des Interactions Humain-Machine (IHM). Emilie a effectué son stage de

---

<sup>1</sup> Wimmics (Web-Instrumented Man-Machine Interactions, Communities, and Semantics):

<https://wimmics.inria.fr/>

<sup>2</sup> EventMedia: <http://eventmedia.eurecom.fr/>

<sup>3</sup> La navigation par facettes est une technique de recherche d'information permettant à l'utilisateur de filtrer des informations à l'aide de dimensions (*facettes*) explicites de filtrage (mots-clés, métadonnées).

<sup>4</sup> 3cixty ExplorMI 360 application, une application officielle de EXPO 2015: <https://www.3cixty.com/>

M2 au sein de l'équipe Wimmics sous la direction d'Alain Giboin, sur précisément le projet Discovery Hub. Durant tout son stage (mars-septembre 2014), Emilie a collaboré étroitement avec Nicolas Marie. Cette collaboration s'est poursuivie jusqu'à la soutenance de thèse de Nicolas Marie (décembre 2014). Emilie Palagi est engagée actuellement dans une étude sur la recherche exploratoire d'images dans le cadre d'une collaboration Wimmics-Mind<sup>5</sup> au sein du projet ANR VISIIR. Durée de l'étude : 6 mois (janvier-juin 2015).

- Ce projet de thèse présenté au labex est motivé par le souhait de mettre davantage en avant l'aspect centré utilisateurs du labex UCN (<https://www.ucnlab.eu/>). Faire participer un doctorant ergonomiste au labex UCN justifierait davantage l'emploi de l'expression « user-centric » dans le titre du laboratoire (*User-Centric Network*) ou dans le slogan UCN <*The "user at the heart of the network" paradigm*> (« L'utilisateur au cœur du réseau »). Aujourd'hui, il est plus juste de dire, comme il est écrit dans le descriptif de l'UCN, que ce labex est « centré sur les services que l'internet de demain pourra offrir aux utilisateurs » ou « centré sur les données ».
- Ce projet de thèse est en rapport avec l'objectif de la série de workshops « Human-Semantic Web Interaction » (HSWI)<sup>6</sup>, d'exploiter les travaux pertinents de la communauté IHM (HCI) pour relever les défis technologiques, socio-techniques et scientifiques posés par la recherche en Web sémantique et, plus généralement, par la Science du Web. Raphaël Troncy est membre du « Semantic Web Interfaces Community Group » à l'origine de HSWI.
- Ce projet de thèse présente également un rapport certain avec le projet Metroscope d'Observatoire de l'Internet envisagé un moment par l'Inria<sup>7</sup>. L'objectif cible de cet observatoire était en effet de pouvoir mener des observations « en tout point » (petits et grands sites, liens spécifiques, utilisateurs finaux, etc.) et « de toute nature » (mesures techniques, d'usages, etc.). Des chercheurs en sciences humaines étaient impliqués dans ce projet, pour étudier les usages d'Internet, des services, des protocoles, et leurs détournements éventuels.

### **Contexte scientifique : Les méthodes d'évaluation des MRE existantes et leurs limites**

L'évaluation des moteurs de recherche exploratoire<sup>8</sup> est reconnue par la communauté scientifique comme « difficile et subtile, car elle implique une analyse qualitative et quantitative à la fois du comportement de l'utilisateur et des résultats de la recherche » [1]. Les systèmes eux-mêmes sont des ensembles complexes qui regroupent différentes fonctionnalités et interactions formant ainsi une alchimie difficile à évaluer ; à quoi il faut ajouter le nombre important de données liées. Cette complexité et cette difficulté se traduisent dans les limites des méthodes existantes d'évaluation (liste non exhaustive) :

- *Limites liées aux métriques d'évaluation.* Les métriques utilisées dans les évaluations existantes ne sont pas appropriées à une tâche de recherche exploratoire. Certaines métriques standards comme la précision ou encore le rappel sont insuffisantes, en particulier pour les données multimédia ([12]), pour mesurer l'engagement des utilisateurs lors de cette tâche ou encore la charge cognitive qu'elle induit. Même si, dans la littérature scientifique, les évaluations tendent à prendre de plus en plus en compte les utilisateurs à l'occasion de tests utilisateurs, les processus cognitifs et les comportements des utilisateurs lors d'une tâche de recherche exploratoire sont encore méconnus, mal définis ou tout simplement pas encore pris en compte.

---

<sup>5</sup> MinD (Knowledge Extraction, Integration and Algorithms) : <http://mind.i3s.unice.fr/>

<sup>6</sup> [http://hswi.referata.com/wiki/Main\\_Page](http://hswi.referata.com/wiki/Main_Page)

<sup>7</sup> <http://owni.fr/2012/01/19/mesure-internet-inria-metroscope-qos/>

<sup>8</sup> Pour plus de détails sur la recherche exploratoire et les moteurs de recherche exploratoire, voir l'annexe.

- *Limites liées au focus de l'évaluation (résultat de la recherche vs processus de recherche).* Dans les méthodes existantes, ce qui est le plus souvent évalué, ce sont les résultats de la recherche d'information et beaucoup moins les processus mis en œuvre par l'utilisateur pour obtenir ces résultats (cf. [9][13]).
- *Limites liées à la nature de la tâche de test proposée aux utilisateurs.* La majorité des tests utilisateurs organisés proposent aux participants des tâches artificielles et imposées, qui ne correspondent pas ou peu aux objectifs et intérêts propres des utilisateurs. Afin d'impliquer davantage les participants dans la tâche de recherche exploratoire, des tests basés sur des tâches plus « écologiques » de RE seraient nécessaires ; par exemple, la tâche de recherche d'informations nouvelles dans le but d'écrire un article (pour des étudiants en journalisme [5]) ou la tâche de recherche d'œuvres originales pour une exposition (pour un organisateur d'expositions).
- *Limites liées au modèle du processus de recherche sous-tendant la méthode d'évaluation.* Lorsque les méthodes s'appuient sur un modèle du processus de recherche décrit du point de vue de l'utilisateur, ce modèle ne reflète pas exactement l'activité de recherche exploratoire. C'est le cas par exemple de la méthode utilisée dans [1], qui repose sur le modèle ISP (*Information Search Process*) de Kuhlthau (1991)[4].
- *Limites liées aux interfaces de recherche dans les bases multimédia et à la quantité de données liées.* De nombreuses interfaces, utilisant la spatialisation (2 dimensions) ou même la profondeur (3 dimensions) ont été proposées pour naviguer dans des grandes bases d'images ou de vidéos. L'idée centrale est que la proximité des éléments induit une similarité visuelle ou sémantique entre les médias. Ces interfaces ont par exemple été développées dans le cadre de la compétition CIVR Video Olympics ou plus récemment Video Search Showcase<sup>9</sup>.

En résumé, les méthodes existantes ne prennent pas assez en compte les utilisateurs lors des évaluations : que ce soit leurs besoins, leurs intérêts, etc. Il est alors nécessaire, pour intégrer davantage les utilisateurs dans les évaluations de ces systèmes, de comprendre leurs comportements, leurs processus cognitifs et de les modéliser. L'idée ici est d'élaborer une méthodologie centrée utilisateurs qui prendrait en compte les comportements exploratoires des utilisateurs et vérifier si le système évalué soutient et facilite la recherche exploratoire. En un mot, il s'agit d'*aller vers une méthode d'évaluation des MRE plus ergonomique.*

## Objectif de la thèse et questions de recherche

Le présent projet de thèse a ainsi pour objectif d'*élaborer une méthode d'évaluation centrée utilisateurs des moteurs sémantiques de recherche exploratoire* dans des grandes masses de données multimédia. La méthode sera basée sur une *modélisation du processus de recherche exploratoire* issue d'études empiriques d'utilisateurs effectuant des recherches exploratoires. La méthode sera *testée sur des systèmes existants ou en cours de conception.*

Pour dépasser les limites évoquées plus haut, la question générale de recherche qui se pose est la suivante :

**Quelle méthode ergonomique élaborer pour évaluer la recherche exploratoire sur le web sémantique multimédia et comment l'élaborer ?**

Question générale que l'on peut décomposer en sous-questions telles que :

- *Sur quel modèle de recherche exploratoire (RE) faire reposer la méthode : un modèle existant ou un nouveau modèle ? Comment représenter le processus réel d'une tâche de recherche exploratoire ?*
- *Comment faire en sorte que la méthode élaborée s'appuie effectivement sur ce modèle ?*
- *Quelles métriques/procédures d'évaluation utiliser pour une évaluation centrée utilisateur ?*

<sup>9</sup> <http://www.videobrowsershowdown.org/>

On notera que répondre à ces questions sur la méthodologie d'évaluation des MRE pourra aider à répondre à des questions de conception des MRE telles que : Comment capturer les intérêts explicites et implicites des utilisateurs ? Comment détecter la progression d'un utilisateur sur un sujet qu'il ne connaissait pas au départ ? Comment accompagner et s'adapter à la progression ? Comment assister le repérage de l'utilisateur dans un domaine de recherche ?

## Défis

- *Défi majeur* : Passer des méthodes d'évaluation de moteurs de recherche ciblée (*look up*) à des méthodes d'évaluation de MRE (la recherche exploratoire étant une activité plus complexe, impliquant des interactions plus nombreuses avec le moteur de recherche). Ce défi, exprimé en 2008 (cf. [13]), reste toujours d'actualité. Un défi similaire, qui reste aussi toujours d'actualité, avait été exprimé en 2006 par [6] à propos de la recherche d'information multimédia basée sur le contenu, car avait été identifié le besoin d'utilisateurs comme les décisionnaires d'explorer un domaine de recherche pour prendre des décisions pertinentes ; ce défi était associé au défi de concevoir des « systèmes d'exploration multimédia par l'expérience permettant aux utilisateurs de mieux comprendre et d'explorer des collections [multi]média ».
- *Défi plus spécifique* : Trouver automatiquement les facettes (critères) sur lesquelles un utilisateur peut pivoter dans une tâche d'exploration de contenus multimédia (ou même de collections). Cette tâche est non triviale (voir par exemple les travaux de [10] sur la visualisation de graphes de réseaux multiplexes ; voir également [11].)
- *Défi relatif aux relations STIC-SHS* : Amener la communauté Réseaux à prendre davantage au sérieux les validations de leurs solutions par leurs utilisateurs et à réaliser effectivement ces validations. Ces études sont aujourd'hui le plus souvent mentionnées comme perspectives et plus rarement réalisées. Voir par exemple [2, p. 15] : « *A user-oriented evaluation would be necessary to validate our approach [to summarization of Life-Science e-experiments from provenance traces] and study the possible usage of experiment summaries. It would also bring valuable inputs on how e-scientists search for their simulated data, and if the proposed approach fosters sharing of simulation data/models.* »

## Démarche, Plan d'action prévisionnel et Résultats attendus

Pour répondre aux questions de recherche énoncées plus haut, la démarche et le plan d'action prévisionnel suivants sont proposés.

### *Eléments d'une démarche pour répondre aux questions de recherche*

| Questions de recherche   | Eléments d'une démarche  |
|--|--|
| <i>Sur quel modèle de recherche exploratoire (RE) faire reposer la méthode ?</i> | 1) Recenser les modèles existants du processus de RE. Recenser les modèles existants de recherche d'information présentant des composantes exploratoires<br><br>2) Observer les comportements d'utilisateurs dans des situations réelles de RE<br><br>3) Elaborer ou adapter un modèle du processus de RE à partir de 1 et 2 |
| <i>Comment faire en sorte que la méthode</i>                                     | Recenser les méthodes d'évaluation basées sur des modèles du   |

|  |   |
|--|---|
| <i>élaborée s'appuie effectivement sur ce modèle ?</i>                                 | processus de RE ou de recherche d'information<br><br>Identifier la manière dont ces méthodes sont dirigées par les modèles. Faire ressortir des principes de conception d'une méthode basée sur un modèle                         |
| <i>Quelles métriques/procédures utiliser pour une évaluation centrée utilisateur ?</i> | Recenser les métriques/procédures utilisées dans les méthodes existantes<br><br>Déterminer les points forts et les points faibles de ces métriques/procédures<br><br>Adapter ces métriques/procédures ou en élaborer de nouvelles |

Ces éléments de démarche entreront dans un plan d'action tel que le suivant.

### **Plan d'action prévisionnel et résultats attendus**

#### I. Préalable à l'élaboration d'une méthode d'évaluation des MRE :

##### 1. Etat de l'art

- a. Les moteurs de recherche exploratoire (MRE). Discovery Hub (Wimmics) et EventMedia/3xcity (Eurecom) seront les MRE de référence.
- b. Modèles sous-jacents de la recherche exploratoire.
- c. Recensement des méthodes d'évaluations des MRE utilisées dans la littérature scientifique, en particulier des méthodes basées sur des modèles de l'activité de recherche
- d. Recensement de méthodes similaires mais appliquées à d'autres types de systèmes (ex. : *cognitive walkthrough*<sup>10</sup>)

##### 2. Analyse et modélisation du processus de recherche exploratoire.

- a. Analyse : études empiriques (observation / expérimentation) d'utilisateurs réalisant des activités de recherche exploratoire.
- b. Modélisation : élaboration d'un modèle de la recherche exploratoire s'appuyant sur les données des études empiriques ainsi que sur les données pertinentes de l'existant.

#### II. Elaboration de la méthode d'évaluation des MRE à partir des éléments issus des étapes précédentes.

1. Opérationnalisation du modèle de recherche exploratoire (pour être utilisé comme cadre d'évaluation)
2. Elaboration de critères et d'heuristiques (d'évaluation/conception de MRE) basés sur le modèle de recherche exploratoire
3. Elaboration d'une procédure d'inspection sans utilisateur (type *cognitive walkthrough*) basée sur le modèle de recherche exploratoire
4. Elaboration d'une technique de tests utilisateurs basée sur le modèle de recherche exploratoire

#### III. Test de la méthode proposée sur des MRE et propositions d'améliorations de ces derniers.

EventMedia/3xcity et Discovery Hub seront deux des MRE évalués à l'aide de la méthode. Cette phase conduira dans un premier temps à l'élaboration de maquettes. Elle contribuera également à la production d'un certain nombre de briques logicielles, configurables et ré-utilisables, dans le but de faciliter la conception et le

<sup>10</sup> Méthode d'analyse de l'utilisabilité d'un système, basée sur un modèle de la séquence des opérations cognitives mises en œuvre par un utilisateur pour apprendre à réaliser une tâche à l'aide du système.

développement de MRE. Ces briques pourront être appliquées à la reconception de EventMedia/3xcity et Discovery Hub.

## RÉFÉRENCES

1. Bozzon, A., Brambilla, M. Ceri, & S. Mazza, D. (2013). Exploratory search framework for Web data sources. *The VLDB Journal*, 22, 641-663.
2. Gagnard, A., Montagnat, J., Gibaud, B., Forestier, & G. Glatard, T. Domain-specific summarization of Life-Science e-experiments from provenance traces. *Web Semantics: Science, Services and Agents on the World Wide Web* 29 (2014) 19-30.
3. Khrouf, H., & Troncy, R. EventMedia: a LOD Dataset of Events Illustrated with Media. In *Semantic Web journal Special Issue on Linked Dataset descriptions, 2012*. IOS Press, ISSN: 1570-0844.
4. Kuhlthau, C.C. (1991). Inside the search process: information seeking from the user's perspective. *Journal of the American Society for Information Science*. 42(5), 361-371.
5. Kules, B., & Shneiderman, B. (2008). Users can change their web search tactics: Design guidelines for categorized overviews. *Information Processing & Management*, 44(2), 463-484.
6. Lew, M.S., Sebe, N. Djeraba, C., & Jain, R. Content-based multimedia information retrieval: State of the art and challenges. *Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, Volume 2 Issue 1, 2006, 1-19.
7. Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, 49(4), 41-46.
8. Marie, N. *Linked Data Based Exploratory Search*, Université de Nice Sophia Antipolis, Décembre 2014.
9. Marie, N., Gandon, F., Giboin, A., & Palagi, É. (2014, September). Exploratory search on topics through different perspectives with DBpedia. In *SEMANTICS 2014. Proceedings of the 10th International Conference on Semantic Systems*, ACM, 45-52.
10. Renoust, B. 2013. *Analysis and Visualisation of Edge Entanglement in Multiplex Networks*. PhD Thesis, Graduate School of Computer Science, University of Bordeaux.
11. Schraefel, M.C., Wilson, M., Russell, A., & Smith, D.A. mSpace: improving information access to multimedia domains with multimodal exploratory search. *Communications of the ACM*, Volume 49 Issue 4, April 2006.
12. Slaney, M. Precision–Recall is Wrong for Multimedia. *Visions and Views Column, IEEE Multimedia Magazine*, Fall 2011, 4-7.
13. White, R.W., Marchionini, G., & Muresan, G. (2008). Evaluating exploratory search systems: Introduction to special topic issue of information processing and management. *Information Processing and Management*, Vol. 44, Issue 2, (2008), pp. 433–436.

## ANNEXE

- *La recherche exploratoire.* – Le terme de « recherche exploratoire » a été popularisé par Gary Marchionini en [7]. Sa définition est encore instable et façonnée par les recherches actuelles. Il est cependant possible de trouver certaines caractéristiques récurrentes dans la littérature scientifique. Lors d’une tâche de recherche exploratoire, les utilisateurs méconnaissent en général le domaine de recherche. Leurs besoins en information ne sont pas fortement déterminés ou du moins sont généraux. Cette activité de recherche est donc caractérisée par un besoin en information et des objectifs de recherche qui sont eux-mêmes indéfinis. Elle ne répond pas à un besoin en information précis mais à une volonté de l’utilisateur de découvrir de nouvelles choses ou d’approfondir ses connaissances sur un sujet d’intérêt, par exemple : apprendre l’histoire du free-jazz, découvrir la culture du Sénégal, etc. Dans [7], Gary Marchionini distingue cette activité de recherche des recherches dites *lookup* qui sont caractérisées par un besoin en information nécessitant une réponse précise (par exemple une recherche de numéro de téléphone ou d’horaires de cinéma). A l’inverse, lors d’une tâche de recherche exploratoire l’utilisateur n’a pas d’objectifs définis et adopte un comportement opportuniste en explorant différents sous-ensembles de l’information disponible. Cette activité de recherche exploratoire peut durer dans le temps et être exécutée à l’occasion de différentes sessions de recherche distribuées dans le temps.
- *Les moteurs de recherche exploratoire.* – Depuis leur création les principaux moteurs de recherche (tels Bing, Google et Yahoo) sont optimisés pour les requêtes de type *lookup*, et ceci pour plusieurs raisons. Premièrement, ils ont popularisé la recherche par mots-clés. Ce mode de requêtage favorise la résolution de besoin en information bien définis. En effet, pour requêter efficacement l’utilisateur doit utiliser des mots-clés précis, déjà connus. Or dans le cas d’une recherche exploratoire le flou caractérisant le besoin en information et la méconnaissance du domaine d’information empêchent une formulation précise du besoin en information sous forme de mots-clés. Deuxièmement, les principaux moteurs de recherche ont basé leur succès sur une simplicité d’utilisation maximale. La simplicité de leur interface et des interactions proposées est un atout pour solutionner les besoins en information simples. Cette simplicité constitue en revanche un goulet d’étranglement pour les besoins en informations plus complexes. Ainsi, lors d’une recherche exploratoire, les utilisateurs ont besoin d’être soutenus continuellement dans le requêtage, le tri, la compréhension de l’information, etc. Ceci nécessite une interface plus sophistiquée.

Néanmoins nous avons observé récemment des évolutions intéressantes concernant le support des tâches de recherche exploratoire par les principaux moteurs de recherche. En 2012 Google a déployé son *knowledge panel* permettant d’ « approfondir les recherches »<sup>11</sup> (« *explore your search* » en anglais). En 2013 Bing et Yahoo ont mis en ligne des fonctionnalités similaires. Les approches et technologies utilisées sont propriétaires donc méconnues. En revanche nous savons avec certitude que ces nouvelles fonctionnalités sont basées sur des technologies « sémantiques » ; sur des graphes de connaissances structurés tels que le Google Knowledge Graph, Bing Satori<sup>12</sup> ou le Yahoo Knowledge Graph<sup>13</sup>.

Les technologies sémantiques apparaissent comme étant très prometteuses pour satisfaire les besoins en information et requêtes complexes, comme en témoigne le développement de moteurs de recherche reposant sur ces technologies pour assister explicitement la recherche exploratoire, par exemple Discovery Hub ([www.discoveryhub.co](http://www.discoveryhub.co)) et Seevl ([www.play.seevl.fm](http://www.play.seevl.fm)). Ce caractère prometteur des MRE sémantiques nécessite cependant d’être validé. L’équipe de recherche qui accueillera la candidate est spécialisée dans les domaines du web sémantique. Pour ces deux raisons nous nous intéresseront ici principalement aux MRE basés sur les technologies du web sémantique.

---

<sup>11</sup> <http://www.google.com/insidesearch/features/search/knowledge.html>

<sup>12</sup> <http://blogs.bing.com/search/2013/03/21/understand-your-world-with-bing/>

<sup>13</sup> <http://semtechbizsj2014.semanticweb.com/sessionPop.cfm?confid=82&proposalid=6452>