



Avis de Soutenance

Monsieur Smail SELLAH

Présentera ses travaux en soutenance

Soutenance prévue le **mardi 17 décembre 2019** à 14h30
Lieu : UTBM site de Belfort Rue Thierry Mieg 90000 BELFORT
Salle : A204

Titre des travaux : approche automatisée d'assistance à la structuration des connaissances

Ecole doctorale : SPIM - Sciences Physiques pour l'Ingénieur et Microtechniques

Section CNU : 27

Unité de recherche : CIAD - Connaissance et Intelligence Artificielle Distribuées

Directeur de thèse : Vincent HILAIRE

Codirecteur de thèse : HDR NON HDR

Soutenance : Publique A huis clos

Membres du jury :

<u>Nom</u>	<u>Qualité</u>	<u>Etablissement</u>	<u>Rôle</u>
M. Vincent HILAIRE	Professeur des Universités	Université de Technologie de Belfort-Montbéliard	Directeur de thèse
Mme Marie-Hélène ABEL	Professeur des Universités	Université de Technologie de Compiègne	Rapporteur
Mme Nada MATTA	Professeur des Universités	Université de Technologie de Troyes	Rapporteur
M. Christophe NICOLLE	Professeur des Universités	Université de Bourgogne	Examineur
M. Patrice WIRA	Professeur des Universités	Université de Haute Alsace	Examineur

Mots-clés : Gestion des connaissances, Ontologies, Regroupement de documents.

Résumé de la thèse (en français) :

Dans un contexte globalisé, les entreprises doivent être innovantes pour augmenter leur productivité et continuer d'exister dans un marché de plus en plus concurrentiel. Les innovations, sources de bénéfices potentiels pour une entreprise, peuvent se situer au niveau d'un processus, d'un nouveau produit ou d'un service, etc. Une entreprise innovante est une entreprise qui capitalise sur ses connaissances. La gestion de connaissance (GC) constitue un ensemble d'approches qui permettent de répondre à un ensemble de problématiques liées aux connaissances et notamment la capitalisation de ces connaissances. Cependant, malgré les bénéfices et l'impact positif que peuvent avoir de telles pratiques sur une organisation, celles-ci sont très peu mises en place. Dans la thèse défendue dans ce manuscrit, nous nous intéressons à améliorer la capitalisation des connaissances et en particulier la structuration des informations afin de proposer des connaissances candidates. Notre objectif est de rendre plus efficace la mise à disposition des connaissances aux acteurs métiers. Pour cela, il faut réduire le nombre de résultats non pertinents et identifier les connaissances qui peuvent aider les acteurs métiers dans leurs problématiques quotidiennes. Par cette approche, on peut ainsi aider l'organisation à optimiser les retours d'expériences et le temps passé dans les différents procédés mis en place. Afin de relever ces challenges, nous nous intéressons à mettre en place un ensemble de briques élémentaires, chacune de ces briques ayant un rôle spécifique. Ces briques sont organisées sous forme d'un cycle interactif. Chaque brique sera en interaction avec les autres, l'idée sous-jacente est qu'une brique améliore ses résultats en apprenant des résultats des autres briques. Les collaborateurs métiers interagissent directement avec ces briques de manière transparente. Pour rechercher les connaissances, le cycle scrute et analyse le comportement des collaborateurs métiers pour mieux comprendre leurs attentes. Ainsi, le cycle est capable d'apprendre et d'améliorer pour mieux capturer et rechercher les connaissances de l'entreprise. La première brique consiste en l'identification et représentation de la connaissance, cette brique a pour rôle d'exploiter un corpus de documents afin d'extraire les connaissances au sein de ce corpus. La seconde brique a pour but d'organiser ce corpus de documents en utilisant les connaissances extraites par la première brique. La dernière brique s'appuie sur les résultats fournis par les précédentes briques, le rôle de cette brique est de permettre à l'utilisateur de pouvoir faire une recherche sémantique en exploitant le modèle de connaissances construit par la première brique et l'organisation des documents qu'offre la deuxième brique. Cette dernière brique aura pour rôle de partager la connaissance et de la diffuser, cette brique ne se restreint pas uniquement à une recherche, elle intègre aussi un mécanisme de suggestions qui assiste l'utilisateur dans sa recherche en lui proposant des documents similaires, etc. L'approche est globale est testée et validée avec un corpus de documents issus d'articles du journal Reuters. Les résultats de l'analyse automatique sont comparés aux tags produits par des lecteurs humains.

Abstract (in English):

In a globalized context, companies must be innovative to increase their productivity and continue to exist in an increasingly competitive market. Innovations, potential sources of profit for a company, can be at the level of a process, a new product or a service, etc. An innovative company is a company which capitalizes on its knowledge. Knowledge management (KM) is a set of approaches that can address a range of issues related to knowledge including capitalization of knowledge. However, despite the benefits and the positive impact which can have such practices on an organization, these are very little implemented. In the thesis defended in this manuscript, we are interested in improving the capitalization of knowledge and in particular the structuring of information in order to propose candidate knowledge. Our goal is to make the access to knowledge more effective to business. To do this, we must reduce the number of irrelevant results and identify the knowledge that can help business in their daily problems. By this approach, we can help an organization to optimize its feedbacks and the time spent in the different processes put in place. In order to meet these challenges, we are interested in setting up a set of elementary components, each of these components having a specific role. These components are organized as an interactive cycle. Each component will interact with others, the underlying idea is that a component improves its results by learning results from other components. Users interact directly with these components in a transparent way. To search for knowledge, the cycle scrutinizes and analyzes the behavior of users to better understand their expectations. Thus, the cycle is able to learn and improve to better capture and seek knowledge of the company. The first component is named «identification

and representation of knowledge», this component has the role of exploiting a set of documents in order to extract the knowledge within this corpus. The second component aims to organize this set of documents using the knowledge extracted by the first component. The last component builds on the results provided by the previous components, the role of this component is to allow the users to be able to do a semantic search by exploiting the knowledge model built by the first component and document organization which the second component offers. This last component will aim to share knowledge, this component is not restricted to only a search, it also includes a mechanism of suggestions that assists the users in their search by offering similar documents, etc. The global approach is tested and validated with a set of documents from Reuters newspaper articles. The results of the automatic analysis are compared to the tags produced by human readers.