

# Mémoire organisationnelle adaptative en vue d'une classification automatique pour la capitalisation d'informations

Jérémy Bascans<sup>2,3</sup> — Max Chevalier<sup>2</sup> — Patrice Gennero<sup>3</sup> —  
Chantal Soule-Dupuy<sup>1</sup>

1. IRIT, Université Toulouse 1 Capitole, Faculté d'Informatique

2 rue du Doyen-Gabriel-Marty, 31042 Toulouse, France

2. IRIT, Université Toulouse 3 Paul Sabatier

118 route de Narbonne, F-31062 Toulouse Cedex 9, France

3. SmartKiwi, Parc Technologique du Canal

10 avenue de l'Europe, 31520 Ramonville-Saint-Agne

{jeremy.bascans, max.chevalier, chantal.soule-dupuy}@irit.fr,  
patrice.gennero@smartkiwi.net

*RÉSUMÉ. Une mémoire organisationnelle est un espace de capitalisation d'informations diverses circulant au sein d'une entreprise. L'objectif de nos travaux est d'assister les usagers dans le classement de ces informations au sein de cette mémoire organisationnelle tout en limitant leur contribution. Pour ce faire, nous proposons un méta-modèle de mémoire organisationnelle sur lequel nous souhaitons faire reposer un processus de classification automatique.*

*ABSTRACT. An organizational memory is a capitalization space of diverse information circulating in a company. The aim of our work is to assist users in the information classification in this organizational memory while limiting their contribution. To do that, we propose a meta-model of organizational memory on which an automatic classification process will be based.*

*MOTS-CLÉS : classification automatique, mémoire organisationnelle, organisation de l'information, capitalisation de l'information.*

*KEYWORDS: automatic classification, organizational memory, information organization, information capitalization.*

## 1. Contexte

L'information représente aujourd'hui un capital non négligeable pour les entreprises. De ce fait, une attention particulière est portée aux systèmes manipulant l'information, et par là aux systèmes d'information (SI). Afin d'intégrer un objectif de capitalisation et de partage des connaissances au sein de ces SI (Basaruddin *et al.*,

2011) entre autres ont constaté une évolution vers une intégration de systèmes de mémoire organisationnelle (SMO). Cependant, (Hodge, 2000) souligne qu'une connaissance des métiers, de la structure et des objectifs de l'entreprise est nécessaire pour une implantation de SMO réussie. Par ailleurs, l'implantation d'un SMO nécessite parfois des modifications des processus de l'organisation impactant directement la tâche des usagers (Zacklad *et al.*, 2011, Doria et Zacklad, 2010), et demandant parfois une forte implication de leur part (Mas *et al.*, 2008). Or, (Ackerman et McDonald, 1996) a constaté dans ce contexte que les systèmes basés sur une action communautaire ne perdurent dans le temps que lorsque plusieurs personnes l'alimentent régulièrement. À cause des efforts et de l'implication demandés, de tels systèmes tombent souvent à l'abandon. Dans ce contexte, nous souhaitons réaliser un SMO ayant pour objectif de limiter la contribution des usagers, tout en maintenant une organisation cohérente des informations. Pour ce faire, nous proposons un méta-modèle de mémoire organisationnelle décrivant les différentes dimensions de la structure de cette mémoire. Ce méta-modèle permet d'instancier un SMO qui intègre un processus automatique de classification des informations introduites directement par les usagers dans cette mémoire.

## **2. Etat de l'art**

### **2.1. Système de Mémoire Organisationnelle et problématiques**

Les mémoires organisationnelles ont pour but de conserver, structurer et partager des informations en apparence hétérogènes, dont l'origine n'est pas toujours connue, afin d'en offrir une représentation cohérente (Hodge, 2000). En plus de cela, il est important que cette représentation soit au maximum partagée. En effet, d'après (Pesqueux, 2005), plus une représentation est partagée entre des individus, plus il y aura de connaissances communes déduites de cette représentation.

Or au sein d'une entreprise, chaque collaborateur, ou équipe de collaborateurs, possède une partie de la connaissance de l'entreprise. Ainsi, un problème de fuite de connaissance se pose à chaque départ d'un collaborateur (« turn over », retraite, démission, etc...). L'enjeu pour l'entreprise est donc de capitaliser de façon centralisée les connaissances existantes afin de les rendre durables. Cette capitalisation repose sur une organisation intégrant la notion de contexte des informations liées.

Pour faire face à cet enjeu actuel, les SMO ont pour objectif de gérer les connaissances au niveau des entreprises, notamment pour répondre au besoin de capitalisation et de pérennisation des connaissances (Basaruddin *et al.*, 2011). Dans cet article, nous nous concentrons sur les SMO dédiés aux entreprises. Dans la section suivante, nous les définissons plus précisément et fournissons également une typologie de ces systèmes.

## 2.2. Typologie des SMO spécialisés pour les entreprises

Les SMO spécialisés pour des entreprises sont également appelés « mémoire d'entreprise ». Plusieurs typologies de mémoires d'entreprise ont été identifiées (Dieng *et al.*, 1999). Cependant, la littérature remonte quatre principaux types de mémoires d'entreprise :

- une mémoire **métier** (ou technique) qui capitalise tous les référentiels, documents, outils et méthodes d'un domaine métier ;
- mémoire **société** qui est fortement liée à l'entreprise, ses activités, ses produits et ses partenaires ou interlocuteurs (fournisseurs, clients, etc...) ;
- mémoire **individuelle** qui capitalise des éléments propres à un individu comme son statut, ses compétences, son savoir-faire et ses activités ;
- mémoire **projet** qui capitalise toutes les informations liées à un projet comme sa définition, ses activités, son historique et ses résultats.

Plusieurs SMO peuvent être identifiés dans la littérature (Gandon *et al.*, 2000, Zacklad *et al.*, 2011). Chacun d'entre eux intègre un processus de classification visant à assister l'utilisateur dans la tâche d'organisation manuelle des informations au sein de la mémoire. Cependant, le processus proposé est dédié à certains types de mémoire d'entreprise (particulièrement la mémoire métier) et est essentiellement manuel. Dans notre approche nous souhaitons limiter cette dimension manuelle en intégrant un processus guidé de classification. Ce processus guidé permet également de favoriser une organisation cohérente et partagée des informations.

## 3. Méta-modèle d'un SMO

Nos travaux ont pour objectif de proposer une organisation intuitive et pérenne des informations introduites dans la mémoire organisationnelle par les usagers. Nous souhaitons que cette approche soit indépendante du type de mémoire implantée (individuelle, métier...). Pour cela nous avons défini un méta-modèle représentant une structuration « générique » des informations au sein d'une telle mémoire. Nous souhaitons ensuite étudier les caractéristiques des informations pouvant être capitalisées dans cette mémoire. Ces caractéristiques permettront par la suite d'identifier les mesures ou techniques adaptées et à utiliser dans un processus de classification de ces informations dans la mémoire.

Le méta-modèle de mémoire organisationnelle que nous proposons repose sur 4 éléments : *Sujet, Document, Catégorie* et *Caractéristique* (figure 1)

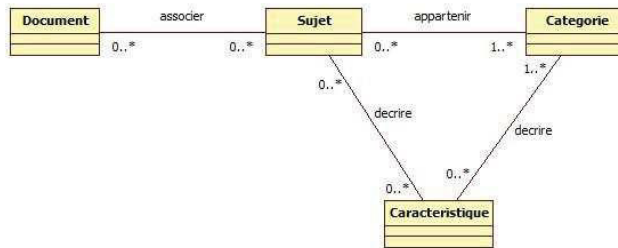


Figure 1. Meta-modèle de l'organisation des sujets

En premier lieu, pour permettre une organisation cohérente des informations collectées dans la mémoire et, plus particulièrement établir un contexte de ces informations, nous proposons d'organiser le méta-modèle autour de la notion de *Sujet*. Un *Sujet* est un « thème » (objet du discours ou de la connaissance dont il est question) autour duquel vont être organisées des informations liées entre elles afin d'en offrir une représentation cohérente. Ainsi, ces sujets permettent aux entreprises d'obtenir une représentation partagée des informations qu'elles souhaitent collecter. En d'autres termes ces sujets permettent de rassembler toutes les informations sur une thématique, un projet, une personne qui pourrait être un contact particulier de l'entreprise... Ceci permet, dans le même temps, pour les usagers et pour chaque sujet, de proposer un accès « centralisé » à l'ensemble des informations liées. Chaque sujet peut être décrit par ensemble de caractéristiques '(éléments de type *Caractéristique*) propres à chaque entreprise (titre, description, historique d'activités...). Ces caractéristiques peuvent être partagées entre différents sujets.

La notion de *Catégorie* associée aux sujets décrit ce que le sujet représente dans l'entreprise. Par exemple, le sujet « Jeremy Bascans » représente un contact de l'entreprise et sera donc associé à la catégorie « Contact ». Grâce à la notion de catégorie, on peut regrouper les objets de même nature. Chaque catégorie est également décrite par un ensemble de caractéristiques (éléments de type *Caractéristique* spécifiques qui seront identiques et valorisées pour chaque sujet de cette catégorie. Par exemple, la catégorie « Contact » peut être décrite par des caractéristiques comme nom, prénom, adresse, entreprise... Certaines de ces caractéristiques permettent d'établir des liens entre différents sujets de catégories différentes. Par exemple, le fait de valoriser, la caractéristique « Entreprise » du sujet « Jérémy Bascans » (catégorie « Contact »), avec la valeur "SmartKiwi" permettra d'établir un lien implicite entre le sujet « Jérémy Bascans » et le sujet « SmartKiwi » (catégorie « Entreprise ») également enregistré dans la mémoire.

Enfin, le méta-modèle fait apparaître la notion de *Document* qui représente toute information explicitement introduite dans la mémoire. Ainsi, toute « granule informationnelle », tel qu'un fichier texte, un mail, etc..., est considérée comme un document.

Sur la base de ce méta-modèle, chaque entreprise pourra structurer l'information comme elle le souhaite : définition des sujets, des caractéristiques et des catégories. De ce fait, les différentes informations introduites dans la mémoire par les usagers devront être associées aux différents sujets auxquels elles se réfèrent. Pour ce faire, nous souhaitons valider ce méta-modèle afin de définir un processus de classification de ces informations dans les différents sujets existants dans la mémoire.

### **3. Perspectives futures**

Dans cet article, nous nous intéressons à l'organisation des informations dans un SMO. Cette organisation a pour objectif d'être générique et adaptative. Il est impératif que, pour répondre aux pratiques et aux usages de toute entreprise, il soit possible d'ajouter de nouveaux processus d'intégration automatiques des informations. Dans le SMO proposé, via le méta-modèle que nous définissons, les informations sont organisées en sujets décrits par un ensemble de caractéristiques. Ces sujets sont eux-mêmes organisés en catégories. Ces éléments permettent de fournir un contexte d'interprétation des informations du SMO en fournissant des liens entre ces informations.

Ces travaux ont posé les principes de base pour la définition d'un SMO et de nombreuses perspectives restent ouvertes. En ce qui nous concerne, il convient dans un premier temps de valider ce méta-modèle. Nous souhaitons vérifier qu'il répond à tous les verrous issus des objectifs imposés par un SMO, notamment en termes d'adaptabilité et de généricité, en l'expérimentant via différents scénarios conçus avec plusieurs entreprises partenaires de SmartKiwi.

Ensuite, il faudra proposer un processus générique de classification automatique, ainsi que les moyens pour une entreprise de définir ses propres processus. Ces processus seront validés via différents scénarios d'expérimentations, (avec mise en situation réelle via une ou plusieurs entreprises partenaires de SmartKiwi). Même si le processus est automatique, il faudra toutefois pouvoir proposer aux utilisateurs de créer de nouveaux sujets lorsque, pour un document donné, le processus ne trouve pas de sujet associé dans le SMO, ou lorsque le contenu du document n'est pas suffisamment couvert par les sujets présents dans le SMO.

Enfin, un des verrous majeurs pour la mise en place d'un SMO, et surtout garantir sa pérennité, est de pouvoir gérer son évolution dans le temps de façon automatique, et la moins intrusive possible. Pour ce faire, nous envisageons de mettre en place un système d'apprentissage basé sur des interactions avec les utilisateurs qui pourront s'exprimer sur la classification proposée et faire des propositions. Ces retours permettront au système de mettre en place un processus d'apprentissage concernant à la fois : (1) l'utilisation des différentes caractéristiques dans la prise de décision finale, afin de coller au mieux aux documents circulant dans l'entreprise ; (2) l'identification automatique des mesures de similarité optimales en

fonction des textes à classer, ainsi que des contenus des caractéristiques des sujets et des catégories ; (3) la définition automatique de descripteurs de sujets (dictionnaires spécifiques, ontologies par exemple) en fonction des vocabulaires métiers utilisés au fil du temps et détectés dans le contenu des informations ajoutées, prenant ainsi en compte l'évolution de ces vocabulaires.

## 7. Bibliographie

- Ackerman M. S., McDonald D. W., « Answer Garden 2: merging organizational memory with collaborative help », *Proceedings of the 1996 ACM conference on Computer supported cooperative work*, 1996, p. 97-105.
- Basaruddin S., Haron H., Noordin, S. A., « Understanding Organizational Memory System for Managing Knowledge », *International Conference on Advancements in Information Technology ICAIT 2011*, IPCSIT vol. 20, 2011, Singapour, IACSIT Press.
- Dieng R., Corby O., Giboin A., Ribiere M., « Methods and tools for corporate knowledge management », *International journal of human-computer studie*, vol. 51, n° 3, 1999, p. 567-598.
- Doria O. D., Zacklad M., « Améliorer la recherche d'information à l'aide de thésaurus « ad hoc » », *Document numérique*, vol. 13, n° 2, 2010, p. 13-40.
- Gandon F., Dieng R., Giboin, A., « CoMMA: Une approche distribuée de la mémoire organisationnelle », *Séminaire Systèmes distribués et Connaissances*, INRIA Sophia-Antipolis, France, Novembre 2000, p. 27-28.
- Hodge, G., « Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files », *Digital Library Federation*, Council on Library and Information Resources, 1755 Massachusetts Ave., NW, Suite 500, Washington, DC 20036, 2000.
- Mas S., Bénel A., Cahier J. P., Zacklad M., « Classification à facettes et modèles à base de points de vue : Différences et complémentarité », *Actes du 36e congrès annuel de l'Association canadienne des sciences de l'information (ACSI)*, 2008, University of British Columbia, Vancouver, p. 5-7.
- Pesqueux, Y., « Management de la connaissance: un modèle organisationnel ? », *26ème congrès de l'AFC*, 2005.
- Zacklad M., Desfriches-Doria O., Bertin G., Mahe S., Ricard B., Musnik, N., *et al.*, « Miipa-Doc: Gestion de l'hétérogénéité des classifications documentaires en entreprise », *Actes de la onzième édition de la conférence internationale H2PTM (Hypertextes et Hypermédias)*, 2011, Hermès, p. 323-333.