

Les treillis de Galois pour l'organisation et la gestion des connaissances

Laszlo Szathmary, Amedeo Napoli

LORIA/Inria-Lorraine, 615 rue du Jardin Botanique, BP 101
F-54600, Vandœuvre-lès-Nancy, France
{szathmar,napoli}@loria.fr

Résumé. Dans cet article, nous étudions la construction de treillis de Galois à partir de différentes sources d'information (par exemple des documents du Web ou des notices bibliographiques) afin d'organiser les connaissances qui peuvent en être extraites. L'organisation en treillis qui en résulte peut alors être utilisée pour satisfaire un certain nombre de buts, comme par exemple la gestion de la connaissance dans une organisation, la recherche documentaire sur le Web, etc. En outre, la classification par treillis peut également tirer parti d'une ontologie des propriétés du domaine considéré. Notre objectif global est de mettre en place un processus de classification par treillis pour enrichir une ontologie qui à son tour permet de guider le processus de découverte de connaissances dans les données.

1 Introduction

Dans cet article¹, nous cherchons à analyser d'un point de vue global le travail effectué au sein d'une organisation ou d'une entreprise. Dans notre cas, nous avons considéré une équipe de recherche comme une petite entreprise. L'expérience montre que fréquemment les chercheurs, pouvant appartenir à la même équipe ou au même laboratoire, ne savent pas toujours exactement sur quoi travaillent exactement les autres chercheurs. Notre but est de faire émerger des relations entre les travaux des différents chercheurs, pour comprendre les orientations de recherche principales et/ou marginales d'une équipe de recherche et ainsi fournir des explications sur le travail de recherche.

Dans une équipe de recherche, les publications sont une bonne façon de décrire les centres d'intérêts d'un chercheur. C'est pourquoi nous avons choisi d'analyser les notices bibliographiques de l'équipe. Nous avons travaillé sur les descriptions BibTeX des articles, qui nous fournissent les méta-données d'un article, comme le titre, les auteurs, les mots-clés, le résumé, l'année de publication, etc. Une entrée BibTeX s'appuie globalement sur le standard du Dublin Core². La norme Dublin Core comprend un ensemble d'éléments simples et représentatifs décrivant les caractéristiques des ressources sur le réseau, articles scientifiques en particulier. Cette norme est généralement utilisée pour la gestion des méta-données dans les pages HTML. Les descriptions BibTeX ayant un

1. Ce travail de recherche est réalisé en partie dans le cadre d'un programme de recherche franco-hongrois Balaton (Balaton F-23/03).

2. <http://www.dublincore.org>