

Compréhension de recettes de cuisine utilisateurs par extraction de connaissances intrinsèques

Damien Leprovost, Thierry Despeyroux, Yves Lechevallier

*Inria – Rocquencourt, Équipe-projet AxIS, BP 105
78153 Le Chesnay Cedex, France
{prénom.nom}@inria.fr

Résumé. Sur les sites Web communautaires, les utilisateurs échangent des connaissances, en étant à la fois auteurs et lecteurs. Nous présentons une méthode pour construire notre propre compréhension de la sémantique de la communauté, sans recours à une base de connaissances externe. Nous effectuons une extraction de la connaissance présente dans les contributions analysées. Nous proposons une évaluation de la confiance imputable à cette compréhension déduite, afin d'évaluer la qualité du contenu, avec application à un site Web de partage de recettes de cuisine.

1 Introduction

Le Web 2.0 favorise le développement des sites collaboratifs, où les utilisateurs échangent des connaissances, se structurent en communautés, développent des codes, des usages et une sémantique qui leurs sont propres. Dans le cadre de la recherche traditionnelle de fouille de la connaissance, cette évolution en relative autonomie peut se révéler problématique : il n'existe aucune garantie que cette dernière se structure autour d'une sémantique qui soit en adéquation avec les bases de connaissance de référence traditionnelles. La pertinence des conclusions peut alors ne pas ou peu refléter l'évolution réelle du comportement des utilisateurs et de la sémantique de leurs échanges.

Nous proposons une méthode pour construire notre propre compréhension des contributions des utilisateurs, basée uniquement sur les données de celles-ci, afin d'extraire la sémantique des utilisateurs. Nous évaluons cette approche par la mesure d'une valeur de confiance. Nous effectuons notre analyse dans le contexte des recettes de cuisine, dont les sites de partage communautaires sont nombreux et très populaires sur le Web français comme mondial.

2 État de l'art

La recette de cuisine est un type de données particulier, composé d'un ensemble d'ingrédients et de procédures d'exécution. Ce type de données est exploité par de nombreux systèmes de recommandation. Le *Cooking Assistant* (Sobecki et al., 2006) définit un système de recommandation démographique basé sur une inférence à logique floue, efficace pour fournir une réponse globale à un besoin général. Mais la généralisation des caractéristiques conduit à une