

# Réflexions sur l'apport de l'exploration des traces d'usage pour améliorer le tri des résultats des moteurs de recherche sur le Web

Rushed Kanawati

LIPN – CNRS UMR 7030  
99 Av. J.B. Clément 93430 Villetaneuse  
rushed.kanawati@lipn.univ-paris13.fr  
<http://www-lipn.univ-paris13.fr/~kanawati>

**Résumé.** Nous présentons dans ce papier un système de fouille coopérative de données d'usage de moteurs de recherche sur le Web dont l'objectif est d'améliorer le tri des résultats rendus par un moteur de recherche. Le système est construit selon une architecture multi-agents où chaque utilisateur est assisté par un agent personnel. Les agents coopèrent entre-eux et utilisent la méthodologie du raisonnement à partir de cas pour re-trier les résultats rendus par un moteur de recherche. Nous nous servons de ce système pour 1) présenter notre analyse des choix de conception d'un système d'exploration coopérative de données d'usage du Web et 2) montrer les problèmes qui restent à résoudre et l'apport attendu des techniques de fouille de données d'usage pour les résoudre.

## 1 Introduction

La présentation et le tri des résultats des moteurs de recherche est un problème important dans le domaine de la recherche d'information sur le Web. En soumettant une requête à un moteur de recherche l'utilisateur attend en retour un ensemble de documents *triés* en fonction de leur pertinence par rapport à ses besoins informationnels. La contre performance des moteurs existant est le meilleur témoin de la nécessité de nouvelles approches de tri de résultats. Différents travaux se sont intéressés récemment au problème du tri de résultats de recherche. Nous les classifions selon les trois axes suivants :

1. Approches fondées sur l'exploration de la structure du Web (Brin & Page, 1998) (Ding, 2002).
2. Approches fondées sur l'exploration des données d'usage (Chen & Meng, 2000) (Arezki et al., 2004).
3. Approches coopératives ou approches orientées communauté d'utilisateurs (Chidloveski et. al., 2000).

Nous nous intéressons à une approche hybride qu'on qualifie de fouille coopérative de données d'usage. L'idée de base est de permettre à un groupe d'utilisateurs de partager implicitement leurs expériences en recherche d'information (Trousse et. al., 1999), (Kanawati, 2003), (Freyne, et. al, 2004). Un premier prototype d'un système d'agents assistants d'aide au tri des résultat est proposé. Ce système est construit selon une architecture égal-à-égal (Peer to Peer). Les agents assistants utilisent la méthodologie du