## Enhancing Personal File Retrieval in Semantic File Systems with Tag-Based Context

Hung Ba Ngo<sup>1,2</sup>, Frédérique Silber-Chaussumier<sup>1</sup>, Christian Bac<sup>1</sup>

Institut National des Télécommunications-France<sup>1</sup>, Cantho University-Vietnam<sup>2</sup> {hung.ngo\_ba, frederique.silber-chaussumier, christian.bac}@int-edu.eu

**Abstract.** Recently, tagging systems are widely used on the Internet. On desktops, tags are also supported by some semantic file systems and desktop search tools. In this paper, we focus on personal tag organization to enhance personal file retrieval. Our approach is based on the notion of context. A context is a set of tags assigned to a file by a user. Based on tag popularity and relationships between tags, our proposed algorithm creates a hierarchy of contexts on which a user can navigate to retrieve files in an effective manner.

## **1** Introduction

Nowadays, tagging systems such as (Delicious) are widely used on the Internet. These tagging systems enable users to add keywords (or tags) to Internet resources without relying on a controlled vocabulary. On the desktop, tags are also supported by some semantic file systems and desktop search tools. Users in LFS (Padioleau, 2005), for example, can manually assign tags to a Jpeg file to annotate the names of persons in that photo for later retrieval. With tags, users are flexible in describing their opinions and interests on files (or resources). As a result, users' personal files are classified per tags and each user has a personal vocabulary of tags. Users then can retrieve files using logical expressions of tags. By default, tagging systems are more suitable for file retrieval using querying than browsing. However, experiments in personal information management (Barreau et al., 1995), and (Khoo et al., 2007) show that most users prefer browsing than querying (logical search) as retrieving their files from a desktop. That is the reason why recently tagging systems such as Delicious on web or LFS (Padioleau, 2005) and TagFS (Bloehdorn et al., 2006) on desktop concentrate on tag organization to help users to browse tags for file retrieval. We continue to enhance personal file retrieval in tagging systems with *context-based searching*. A context in our approach is a set of tags assigned to a file (or resource) by a user. Based on tag popularity and relationships between tags, our proposed algorithm creates a hierarchy of contexts on which a user can navigate to retrieve files in an effective manner. In this paper, we first present the interesting techniques for tag organization in section 2; introduce tagbased context and how to enhance tagging systems with context-based searching in section 3. Our algorithm for creating a Directed Acyclic Graph of Tags (DAGoT) based on tag popularity and relationship of tags is in section 4. This DAGoT is used to organize contexts into a hierarchical structure so that we can enhance personal file retrieval with context-based searching. An implementation and experimental results using real data are presented in section 5. Our conclusion and perspectives are in the last section.