

Integrating File Systems and the Semantic Web

Bernhard Schandl¹, Arash Amiri², Stefan Pomajbik¹, and Diman Todorov¹

¹ University of Vienna, Department of Distributed and Multimedia Systems
{bernhard.schandl|stefan.pomajbik|diman.todorov}@univie.ac.at

² Research Studio Digital Memory Engineering, Vienna, Austria
arash.amiri@researchstudio.at

Abstract. In this demo, we present an approach how Semantic Web technology can be applied to file systems in order to enrich the storage of unstructured content with metadata. Our system provides a variety of organizational means for digital objects and integrates them into a graphical user interface that is similar to file management utilities, which lowers the entry barriers to semantic systems for end users.

1 Introduction

File systems are the backbone of every data storage system. In addition to situations where the file system is hidden from the user (e.g. in web applications, database-based CRM systems, etc.), a plentitude of unstructured data is stored directly in the file system, organized by the mechanisms typical file systems provide, like directories and file names.

The Semantic Web aims to leverage the World Wide Web to a web of data, where machines are able to process annotations and relations between resources, and where implicit information can be derived from utilizing ontologies and shared vocabularies. There has been effort to consolidate the Semantic Web idea with file systems, especially in the research field of the *Semantic Desktop* [1]. However, to the best of our knowledge, there exists no approach that aims to entirely replace the pseudo-semantics of hierarchical file systems with a metadata-centric storage paradigm.

In this demo, we present the Semporer, a graphical user interfaces for repositories that are able to hold unstructured data objects with associated metadata represented in RDF. The Semporer allows to access such repositories in a way similar to common file management utilities, thus bringing semantics to the user without the need to learn completely new interaction paradigms.

2 System Description

In this demo, we mainly concentrate on the user interface aspects of such an integrated system. The Semporer, as described above, provides a view on a RDF-based repository that is similar to the environment users know from file management utilities (e.g. Windows Explorer). Since we disband certain constraints of classical file systems (such as strict directory hierarchies or the need

for unique file names), and instead add strong semantic support to the system, we mutated the word *file* to *sile* (*semantic file*). Siles consist of content that is not interpreted by the system, but can be enriched with RDF metadata. The view on siles resembles the view on file systems: Siles are represented as icons with labels, and additional information about siles is displayed on demand in dedicated GUI areas.

Instead of organizing files into hierarchical directories, users can use a variety of mechanisms to organize siles: An arbitrary number of *tags* can be attached to siles. We believe that tagging can help users who do not want to cope with the details of ontologies to easily annotate and recover data objects. Similar to traditional directories, siles can be put into an arbitrary number of overlapping *collections*. Instead of being restricted to a file name, users can annotate siles with *attributes* in the form of name/value pairs. These can be defined ad-hoc, or be selected from predefined ontologies. Finally, siles can be brought into context by using *relations*, which allows to bring siles into context.

The GUI allows to search for siles using all the criteria mentioned above. To simplify handling of metadata and search queries, we implemented a drag-and-drop interface for all operations. At any time, users can see the current search context and add or remove search conditions with simple mouse interactions. Files from the user's local system may be imported by dragging them into the application, and siles are opened using the correct applications by a simple double click.

The Semplore is built on top of an API to the SemDAV protocol[2], which is currently developed in a research project carried out by the authors. This protocol provides calls for all operations mentioned above, and for search and retrieval within SemDAV-conformant repositories. In a long-term perspective, this approach could replace the file system as it is used today: By using SemDAV, it is possible to implement cross-application metadata exchange, since the full set of metadata and annotations is exposed by the API to every application and is made available to the user via the Semplore interface.

In the near future, we plan to implement more sophisticated support for user-centric handling of tags and ontologies: We will analyze the possibility of user-defined ad-hoc matchmaking between these, as well as algorithms that trace user interactions and automatically infer relations and annotations out of this data.

References

1. Leo Sauermann, Ansgar Bernardi, and Andreas Dengel. Overview and Outlook on the Semantic Desktop. In Stefan Decker, Jack Park, Dennis Quan, and Leo Sauermann, editors, *Proceedings of the 1st Semantic Desktop Workshop*, volume 175, Galway, Ireland, November 2005. CEUR Workshop Proceedings.
2. Bernhard Schandl. SemDAV: A File Exchange Protocol for the Semantic Desktop. In *Proceedings of the Semantic Desktop and Social Semantic Collaboration Workshop*, volume 202, Athens, GA, USA, November 2006. CEUR Workshop Proceedings.