



Ontology-driven collaborative annotation in shared workspaces



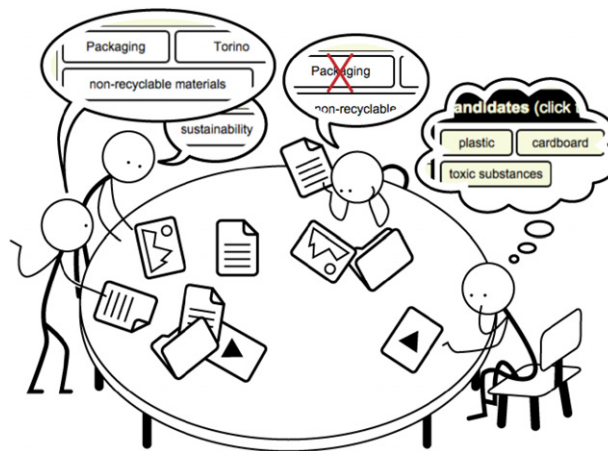
Anna Goy*, Diego Magro, Giovanna Petrone, Claudia Picardi, Marino Segnan

Dipartimento di Informatica, Università di Torino, C. Svizzera, 185, 10149 Torino, Italy

HIGHLIGHTS

- Sem T++ is a framework for the collaborative management of shared online resources.
- Sem T++ supports collaborative semantic annotation, thanks to a formal semantic model.
- Implementation of collaborative annotation is based on results of a user study.
- User study analyzes three collaboration policies: consensual, authored, supervised.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 19 December 2014
 Received in revised form
 29 March 2015
 Accepted 24 April 2015
 Available online 6 May 2015

Keywords:

Online collaboration
 Collaborative semantic annotation
 Ontologies
 Knowledge representation
 Knowledge sharing

ABSTRACT

In the last decade, collaboration and sharing on the Web have become mainstream. Digital, remote interaction happens on a daily basis, not only to share digital resources, but also to create, manage and discuss them, in every possible situation where collaboration is required: from work teams to groups of friends, from community committees to no-profit organizations. In this paper we address the task of collaborative management of digital resources within a team, with a special focus on the task of semantic annotation, where team members, possibly supported by automated reasoning, enrich resources with properties that help in organizing, retrieving and creating connections between contents of different types. We focus in particular on the problem of reaching an agreement on the annotation itself among the participants. The paper presents a qualitative user study aimed at observing users behavior when faced with this task. The results of the study are then analyzed in order to draw guidelines, which are then implemented in a tool for collaborative annotation. This study is carried out in the context of the Semantic Table Plus Plus (Sem T++) Project, a framework supporting collaboration over thematic workspaces, whose goal is to enhance cooperation through awareness, enhanced communication and easy sharing of digital content.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The recent evolution of the web brought along new perspectives for its users. In particular, we embrace the idea that the current Web, sometimes referred to as *Web 3.0*, can be considered as

* Corresponding author. Tel.: +39 0116706786; fax: +39 011751603.

E-mail addresses: annamaria.goy@unito.it (A. Goy), diego.magro@unito.it (D. Magro), giovanna.petrone@unito.it (G. Petrone), claudia.picardi@unito.it (C. Picardi), marino.segnan@unito.it (M. Segnan).

<http://dx.doi.org/10.1016/j.future.2015.04.013>
 0167-739X/© 2015 Elsevier B.V. All rights reserved.

the evolution of “participative”, service-oriented *Web 2.0* when it meets the *Semantic Web* (see [1] for an early and effective discussion about this point of view). Moreover, the convergence of social aspects and semantic technologies also led to a new vision which is referred to as *Social Semantic Web* [2]. In this perspective, the following aspects are important for the approach presented in this paper: both software applications and web sites tend to turn into interactive online services; emerging web technologies encourage participation and collaboration among users; the usage of semantically-based metadata about web resources is exploited in order to make them processable by “intelligent” software agents.

This scenario is experienced in our everyday life: in recent years, almost everybody has faced the growing need of interacting with online services and managing digital resources, in order to carry on many daily tasks, including work and business (e.g., managing projects, organizing and sharing digital documents, and so on), interaction with e-government applications, family management and leisure (e.g., online shopping or reservations for travels, hotels, restaurants, etc.). Moreover, users can take advantage from the enhanced possibility of sharing and collaboratively handling such digital resources, for example in social tagging systems, project management applications, collaborative editing tools, online conferences, file sharing tools, just to mention a few within a very large set.

In summary, the current web poses the challenge of managing a very large amount of heterogeneous, and often shared, digital resources (documents, pictures and videos, emails, posts, bookmarks, etc.), but it also suggests the techniques and tools to face such a challenge, by providing the set of technologies and approaches usually grouped under the label of *Semantic Web*. In particular, from the perspective sketched so far, the most important aspect is represented by *semantic annotation*, which enables software applications to access, manipulate, and in some sense “understand” the content of digital resources.

In this paper we present a framework supporting the collaborative management of shared digital resources, designed and developed within the *Semantic Table Plus Plus* (Sem T++) project. A major role, within this framework, is played by formal semantic representations of *information objects*, collaboratively built up by groups of users working together in collaborative thematic workspaces. In particular, in this paper, we will focus on this activity, discussing the results of a qualitative user study aimed at analyzing user requirements and defining the collaboration model supporting collaborative semantic annotation of shared digital resources.

The rest of the paper is organized as follows. Section 2 discusses the most relevant related work, taking into account different research fields, which provide the background of our approach, focusing, in particular, on the fields related to collaborative semantic annotation. Section 3 presents our previous work on Sem T++ by describing its main features. Section 4, which represents the original contribution of the paper, describes the role of semantic annotations of information objects in Sem T++. This section also contains the description of a qualitative user study about collaborative semantic annotation, a discussion of its results, and an explanation of how the model based on such a study has been implemented in Sem T++. Section 5 concludes the paper by summarizing its findings and open issues.

2. Related work

2.1. Sem T++ general background

A survey and a discussion of existing Web-based applications supporting collaboration, including groupware and project management tools or suites, can be found in [3], a previous paper introducing our approach.

Sem T++ is grounded into the critique to the so-called *desktop metaphor*, and in the approaches trying to replace it. A good overview of this issue can be found in [4], where an articulate discussion of the problem and a presentation of the proposals going beyond it can be found. The major critique to systems based on the desktop metaphor outlined by editors and contributors of the mentioned book is that such a metaphor, based on application-centered and file-centered models, fails in providing a support to the new needs emerging from the current ICT scenario, and in particular the need for an effective support to user collaboration, heterogeneous objects and multiple contexts management. Some of the alternatives to the desktop metaphor have influenced the design of Sem T++ more than others. In particular, Sem T++ proposes the metaphor of “table”, as opposed to the “desktop” – where a person usually works alone – since tables are places where many people can sit in group, discuss, share resources, and work together. In this perspective the most relevant alternative approaches are Haystack [5] and the proposals grounded into Activity-Based Computing [6,7]. Haystack [5] provides users with a flexible and personalized control over resource properties which help to organize them into coherent workspaces, referring to specific *tasks*, by providing support for uniform annotation, links to other resources, and retrieved. Activity-Based Computing [6,7] is based on the assumption that the most important principle exploited to organize is not *application* or *file*, but *user activity*, which enables users to build a set of applications and documents related by a common context. A more detailed account of this literature is provided in [8].

An interesting work, which aims at integrating desktop-based user interfaces and technologies from the Semantic Web, is the *Semantic Desktop* initiative [9], mainly developed within the NEPOMUK project (nepomuk.semanticdesktop.org). The main goal of this project was to support collaboration among knowledge workers, through the integration of existing applications, achieved by the definition of an open-source framework for implementing semantic desktops, based on a set of ontologies. A proposal to connect the Semantic Desktop to the Web of Data, which inspired the future enhancement of Sem T++ with a connection to open datasets (mentioned in Section 4.1), is discussed in [10].

The exploitation of semantic knowledge to support users in collaborative resources management underlies many approaches. For example, in [11] the authors present an ontology to support media data management within the CineGrid Exchange network. [12] presents an extensible and domain-independent ontology-based architecture for data management systems, aimed at enabling the creation, storage, validation, query, and search of large amounts of data (and metadata) in heterogeneous formats.

In a slightly different perspective, with more emphasis on the formal description of interactions, instead of resources themselves, [13] presents a tool supporting the analysis of trends and patterns in collaboration activities taking place within multi-disciplinary design teams and relying on web-based heterogeneous collaborative applications. The described tool, TCN (Team Collaboration Networks), exploits an ontology providing the vocabulary for describing interactions between persons and/or information objects.

2.2. Sem T++ annotation background

Approaches and systems supporting resources annotation have been designed and developed within different research communities, with different purposes and characteristics. In this section, we will provide a non-exhaustive survey, focusing on the characteristics which enable us to define the notion of “annotation” suited to the approach presented in this paper.

In NLP-oriented annotation tools, “annotations” are typically labels associated to phrases within a document and usually rely

Download English Version:

<https://daneshyari.com/en/article/424554>

Download Persian Version:

<https://daneshyari.com/article/424554>

[Daneshyari.com](https://daneshyari.com)