

Chapter 7

TagClusters: Enhancing Semantic Understanding of Collaborative Tags

Ya-Xi Chen

Media Informatics, University of Munich, Germany

Rodrigo Santamaría

University of Salamanca, Spain

Andreas Butz

Media Informatics, University of Munich, Germany

Roberto Therón

University of Salamanca, Spain

ABSTRACT

Many online communities use TagClouds, an aesthetic and easy to understand visualization, to represent popular tags collaboratively generated by their users. However, due to the free nature of tagging, such collaborative tags have linguistic problems and limitations, such as high semantic density. Moreover, the alphabetical order of TagClouds poorly supports a hierarchical exploration among tags. This paper presents an exploration to support semantic understanding of collaborative tags beyond TagClouds. Based on the results of the authors' survey of practical usages of collaborative tags, they developed a visualization named TagClusters, in which tags are clustered into different groups, with font size representing tag popularity and the spatial distance indicating the semantic similarity between tags. The subgroups in each group and the overlap between groups are highlighted, illustrating the underlying hierarchical structure and semantic relations between groups. The authors conducted a comparative evaluation with TagClouds and TagClusters based on the same tag set. The results confirmed the advantage of TagClusters in facilitating browsing, comparing and comprehending semantic relations between tags.

DOI: 10.4018/978-1-4666-0285-4.ch007

INTRODUCTION

The advent of the next-generation Web has tremendously influenced our ways of dealing with online digital items. In today's online communities, people are allowed to sharing their content and access others' in turn. They are also encouraged to make contributions by tagging these digital items. The most representative websites are YouTube (<http://www.youtube.com>) for video sharing, Flickr (<http://www.flickr.com/>) for photo sharing, and the online music community Last.fm (<http://www.last.fm>). The general goal of making online search and browsing easier is the main motivation for tagging. Besides facilitating personal organizing, browsing and searching online items, such collaborative tags also help to discover new items, acquire knowledge and derive insights users may be not aware of before.

Currently, there are two main usages of collaborative tags. Users can use tags to conduct a keyword-based search to seek information on the web. Comparing tags generated by single users, large quantities of collaborative tags convey public interests and opinions in a general level. Therefore, TagClouds, a visualization for the most popular tags, has become a fashion in many websites. Although TagClouds is widely used, it still has some intrinsic limitations. Many researchers have already noticed this and dedicate to improve its effectiveness by proposing more aesthetical representations or providing a better semantic understanding. In this paper, we take Last.fm as our experiment platform and explore how to use an improved visualization to support the semantic understanding of collaborative tags. We are specifically interested in the hierarchical structures and relationships among tags, which might lead to more successful tag recommendation and visualization-based music retrieval.

In the remaining parts of this paper, TagClouds and their related work will first be discussed. Based on a survey among Last.fm users, some

possibilities to improve semantic understanding beyond TagClouds will be explored. Finally, a prototype named TagClusters and its evaluation will be presented. The discussions on the evaluation results and future work will conclude this paper.

TAGCLOUDS

TagClouds is a visualization of popular tags (see Figure 1). In this visualization, tags are normally ordered alphabetically, with font size representing their popularity. Additional information, such as recency can be illustrated with color lightness. Comparing with tags generated by single users, these popular tags in TagClouds have a higher degree of generality and accuracy. Such visualization facilitates quickly foraging an overall impression of the most popular items, and thus conveys the general interests among a large audience (Viégas & Wattenberg, 2008), (Hearst 2008), (Hearst & Rosner, 2008). TagClouds can be also used for keyword-based search by selecting one or multiple tags as input.

LIMITATIONS OF TAGCLOUDS

In most online communities, users are allowed to generate tags freely, without any restriction or quality control. This low usage barrier has attracted hundreds of millions users. However, due to the free nature of tagging, some problems are inevitable with these collaborative tags (Li, Bao, Yu, Fei & Su, 2007; Hassan & Herrero, 2006).

Linguistic Problems

Nielsen (2007) and Begelman, Keller and Smadja (2006) discovered educational and cultural background influence people's understanding of tags, which is one of the reasons for tag inconsistency among different users. With no input restrictions,

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/tagclusters-enhancing-semantic-understanding-collaborative/64048

Related Content

Revealing Passwords: Using Social Engineering Methods to Monitor Scammer Email Communication

Andreas Zingerle (2015). *Analyzing Art, Culture, and Design in the Digital Age* (pp. 173-186).
www.irma-international.org/chapter/revealing-passwords/138539

The Use of Media Convergence in the Preservation and Dissemination of Cultural Assets: Case of Mozambicans Timbila

José João Hogueane, Mirian N. Tavares and Sandra Boto (2020). *International Journal of Creative Interfaces and Computer Graphics* (pp. 16-26).
www.irma-international.org/article/the-use-of-media-convergence-in-the-preservation-and-dissemination-of-cultural-assets/264199

Senses of "Selfie" Around the World From Web Search Patterns Over Extended Time

Shalin Hai-Jew (2018). *Selfies as a Mode of Social Media and Work Space Research* (pp. 249-295).
www.irma-international.org/chapter/senses-of-selfie-around-the-world-from-web-search-patterns-over-extended-time/191378

From Zero to Infinity: A Story of Everything

Clayton S. Spada and Victor Raphael (2012). *Biologically-Inspired Computing for the Arts: Scientific Data through Graphics* (pp. 356-377).
www.irma-international.org/chapter/zero-infinity-story-everything/65038

The Relief-Perspectives of Bitonti and Borromini: Design and Representation of the Illusory Space

Giuseppe Amoroso (2016). *Handbook of Research on Visual Computing and Emerging Geometrical Design Tools* (pp. 420-455).
www.irma-international.org/chapter/the-relief-perspectives-of-bitonti-and-borromini/149315