Chapter 61
Virtual Communities as Contributors for Digital Objects Metadata Generation

Joana Sócrates Dantas
Escola Politécnica da Universidade de São Paulo, Brazil

Regina Melo Silveira
Escola Politécnica da Universidade de São Paulo, Brazil

ABSTRACT

Description of online digital content is currently extremely necessary to facilitate a diverse amount of resource sharing over the internet. Many times, content is shared and reused within a virtual community. Virtual communities tend to have their own specific needs of resources, and tend to use a specific vocabulary to describe content. Members of virtual communities also tend to have specific motivations for participating and sharing information and knowledge with other members. In this chapter, the authors discuss the benefits of community members generating content description by analyzing the current literature on the matter. Then, the authors present two studies they have held where they assess the metadata generated by users of an IPTV system and by members of two different virtual communities.

1. INTRODUCTION

One of the most important revolutions the internet underwent in recent years was the emergence of Web 2.0. The term Web 2.0 was coined by Dale Dougherty and popularized by O'Reilly Media in 2004 and, in the words of O'Reilly himself, one of the insights on what was becoming the Web 2.0 back then was the network serving as a platform, where applications would learn and get better with people's use and contribution (O'Reilly & Battelle, 2009). Web 2.0 is both a usage and a technology paradigm that has consolidated the Web in a more collaborative and interactive manner. Being more dynamic it allows users to both access a web site and contribute to it. It can be understood as a collection of technologies, business strategies and social trends, which were ignited by social applications such as MySpace[1], Flickr[2], and YouTube[3] (Murugesan, 2007).

Web 2.0 has had a large impact on the way people use the Internet. This phenomenon can

DOI: 10.4018/978-1-4666-5942-1.ch061

be appreciated in the vast and diverse multimedia content produced, exchanged and suggested by users online. Web sites supporting online collections of digital multimedia contents are very common nowadays, where people may exchange resources, opinions and information on content like music, video, books, periodicals and so on. Web 2.0 interactivity tools permit users to describe content for others or their own personal re-use.

Some examples of multimedia content services where users may participate at different levels are IPTV, online photo albums, bookshops, online radio and so on. Internet protocol television (IPTV) is the name given to a service that provides digital television over Internet Protocol (IP) for residential and business users (Xiao et al., 2007) with the aid of a set-top box connected to a TV set. As another type of service is video over the Internet, that can be viewed on various devices such as TVs, PCs and mobile phones. For this reason IPTV and Internet TV present a convergence of communication, computing and content (Jain, 2005). Some online radio services such as Last FM rely greatly on users' contribution and interaction, where other users are cited based on common tastes. Similarly with Amazon, users may criticize and evaluate books and other products.

Web 2.0 has also promoted the development of virtual communities, which are, currently, large producers and consumers of digital multimedia content on the Web. In virtual communities members not only debate and exchange ideas, but also share information in the form of videos, articles, music etc, and suggest content for each other.

The vast availability of content on the web and the sharing and exchanging of this content within communities created the scenario where the Semantic Web became necessary. In Semantic Web, information is given well-defined meaning, better enabling computers and people to work in cooperation (Lee et al., 2001). The Semantic Web gives the traditional web far greater utility as users, within an area of interest, can use common terms to represent information. With semantic web tools, terms may be linked to each other and be understood automatically by different communities' web software (Feigenbaum et al., 2007).

Recently there has been intense research on user generated metadata alongside the social aspect involved in this activity. With all these new technologies and advances on the Internet, the need for metadata and the new active roles of users make us curious about how virtual communities relate to digital content online. Are virtual community members more active in generating content description? Do they generate higher quality description of content?

In the next section we will briefly elucidate on what a virtual community is and what motivates those involved in actively participating in a virtual community. In section 3 we will discuss the definition of metadata, different types of metadata promoted by the Web 2.0 technology and the quality of metadata generated by these new means. Section 4 will introduce the subject of virtual communities as metadata generators, while sections 5 and 6 will present the two studies we developed on the matter, their results and analysis. In section 7 we draw some conclusions on the subject and, finally, in section 7 some outlines for future work will be presented.

2. VIRTUAL COMMUNITIES

Howard Rheingold is one of the first enthusiasts and researchers of online communities. In 1985, when the internet was not yet broadly open and public, he became part of the WELL, a computer conferencing system, where people from all over the world would gather to have online conversations and discussions. In 1993, Rheingold wrote the book *Virtual Communities*, *Homesteading on the Electronic Frontier,* (re-edited and re-published by MIT Press in 2000), and was credited with inventing the term "Virtual Community". In the