

Chapter 107

An Assessment of Several Taxonomies of Volunteered Geographic Information

Antony K Cooper

CSIR, South Africa & University of Pretoria, South Africa

Serena Coetzee

University of Pretoria, South Africa

Derrick G Kourie

University of Pretoria, South Africa

ABSTRACT

User-Generated Content (UGC) in general, and Volunteered Geographical Information (VGI) in particular, are becoming more important as sources for official data bases, such as those used in national Spatial Data Infrastructures (SDIs). Discovering and assessing VGI as suitable geospatial resources for one's purposes is hence becoming more important, but can be difficult. One way of assessing VGI resources is by classifying them into different types of resources, i.e. a taxonomy of resources. The question is whether such taxonomies can accurately identify suitable VGI resources. We assess five taxonomies both subjectively and using formal concept analysis to determine their discrimination adequacy, that is, how well the taxonomies discriminate between repositories containing UGC in general, or VGI in particular.

INTRODUCTION

User Generated Content

The World Wide Web has evolved from being a network of resources on the Internet developed and delivered by professional programmers and administrators, with users as passive receivers,

to one of social networking and content sharing (sometimes known as *Web 2.0*), facilitated by tools making it easy to publish content in different languages, and fast and cheap Internet access in many parts of the world. It has spawned the development of *virtual communities* or *virtual social networks*, which share data with one another, and with the public at large. This *user generated content* is most obvious in the repositories available through

DOI: 10.4018/978-1-4666-2038-4.ch107

web sites such as Wikipedia (Wikimedia, 2011), the free, online encyclopedia in many languages, consisting of contributions mainly from the public at large, rather than from domain experts (though it does also include much content from encyclopedias that are out of copyright and other expert sources). Similarly, virtual communities have also facilitated *folksonomies* or *collaborative tagging*, which are the classification and identification of content by the general public, rather than by domain experts (Cooper, et al., 2011).

There is no widely accepted definition of User-Generated Content (UGC), and maybe there never will be. As with many concepts in information technology, UGC is interpreted in different ways, and one woman's user generated content could be another man's professionally generated content (Cooper, et al., 2010b). To support a study by the Organisation for Economic Cooperation and Development (OECD) on the participative web and the rapid growth of user-created content, and because they could not find a suitable definition or taxonomy, Wunsch-Vincent and Vickery (2007) defined *user-created content* (their term for user-generated content) as being:

- Content made publicly available over the Internet,
- Which reflects a "certain amount of creative effort," and
- Which is "created outside of professional routines and practices" (Wunsch-Vincent & Vickery, 2007).

Their second criteria could be considered to be controversial, as much content contributed by the public might be done so without any creative effort, such as the material on file sharing sites. Further, it appears to exclude content where the person uploading the content is not the creator of the content but is doing so legitimately, which would be the case of a tribute site, such as for the late Andries Naude (2009), who established the site that was later populated by his wife and

friends. The third criteria is nominally useful for differentiating user-created content from professionally generated content, though they do acknowledge that it is getting harder to maintain this distinction as some amateur content providers obtain sufficient status to then get paid for providing the same content for a media web site. This also excludes the content that the likes of De Longueville *et al.* (2009) consider to be user generated, namely where the data are collected, synthesized and posted by a professional research team, derived from interviews with stakeholders (Cooper, et al., 2010a).

Of course, UGC is not confined to the Internet and was not invented on the Internet—though the Internet brings UGC to a much wider audience and much more quickly, than would otherwise be the case. People generate content whenever they document something or tell someone something. Much of the content is discarded quickly, because the other person was not listening or the document (e.g.: scrap of paper with a shopping list) is used and thrown away. There are no minimum criteria for value, availability or use for considering whether or not content can be deemed UGC. Of particular interest here is the UGC that is made widely available, such as through the Internet (Cooper, et al., 2011).

Pervasive, cheap (or free), easy-to-use and intelligent web services empower users to develop, rate, combine (e.g.: mashups) and distribute content on the Internet; collaborate with peers (known and unknown, with common interests or not); and customise Internet applications. This is the basis of the participative web (Wunsch-Vincent & Vickery, 2007). Gervais (2009) feels that even as a mere conceptual cloud, the term UGC is useful for considering the societal shifts in content creation due to the participative web. A web site can make more than one repository available (e.g.: Wikipedia currently has 269 active editions [Wikimedia, 2011]) and a repository can be made available through more than one web site (e.g.: OpenStreetMap data used by other web sites).

13 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/assessment-several-taxonomies-volunteered-geographic/70535

Related Content

Satellite and Ground Estimates of Surface and Canopy-Layer Urban Heat Island: Comparison and Caveats

Bakul Budhiraja, Prasad Pathak, Girish Agarwal and Raja Sengupta (2021). *International Journal of Applied Geospatial Research* (pp. 1-21).

www.irma-international.org/article/satellite-and-ground-estimates-of-surface-and-canopy-layer-urban-heat-island/289374

CulturalNature Arga #2

Tiago Cruz, Fernando Faria Paulino and Mirian Tavares (2019). *Geospatial Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 804-812).

www.irma-international.org/chapter/culturalnature-arga-2/222926

Modeling Species Distribution

Yongyut Trisurat and Albertus G. Toxopeus (2013). *Geographic Information Systems: Concepts, Methodologies, Tools, and Applications* (pp. 2102-2127).

www.irma-international.org/chapter/modeling-species-distribution/70553

An Evolving Residential Landscape in Post-Katrina New Orleans: Racial Segregation Among Racial and Ethnic Groups, 2000-2010

John Byron Strait and Gang Gong (2012). *International Journal of Applied Geospatial Research* (pp. 1-19).

www.irma-international.org/article/evolving-residential-landscape-post-katrina/70656

Geospatial Mashups in Web GIS for Tourism Infrastructure: Internet-Based Channel Perspective Promotional Measures

Somnath Chaudhuri and Nilanjan Ray (2019). *Geospatial Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 1334-1352).

www.irma-international.org/chapter/geospatial-mashups-in-web-gis-for-tourism-infrastructure/222950