

Ontologies for folksonomies and for profile-based online social networks

Ing. David Zejda, KIT, FIM, UHK, david@zejda.net

Abstrakt

S přerodem webu podle paradigmat Webu 2.0 směrem k prostředí otevřenému nejrozličnějším sociálním aktivitám přibývá i uživatelů komunitních portálů, které tak zažívají nebývalý rozvoj. Nicméně stále vlastně nevíme, jak vhodně popsat uživatele komunitního webu, jak co nejlépe zachytit jeho zájmy a touhy a jak mu vhodně pomoci při hledání přátel či navazování vztahů. Jak se co nejlépe dozvídat o přáních uživatelů portálu? Měla by komunita vedle samotného obsahu sama více určovat strukturu a celkové zaměření webu, který jí slouží? S růstem možností a významu sociálních sítí nabývají na aktuálnosti otázky spolupráce mezi jednotlivými sítěmi a jejich integrace, otázky určování identity a formování důvěry napříč sítěmi a bezpečnosti. Příspěvek zkoumá, jak by mohly k řešení přispět ontologie. Zvláštní pozornost přitom zaměřuje na folksonomie a na sociální sítě založené na uživatelských profilech.

Klíčová slova

ontologie, sociální síť, folksonomie

Abstract

Current web takes more and more out of Web 2.0 concepts, becomes more central to the lives of people, millions of users register in many community portals, web becomes annotated and driven by a community. The sites are doing well, but they may do even better. Many issues remain open, like how to describe a person, how to capture his interests and passions, how to help him find the relations according to his wishes. How to get to know about the requirements of users or even better, how to allow users to enrich the portal in a systematic way? With proliferation of rich social networks the importance of interconnection, integration, identity, trust and security challenges grows too. The article examines possible role of ontologies in search for answers. We particularly focus on folksonomies and on social networks based on semi-structured user profiles.

Keywords

ontology, social network, folksonomy

Lead-in

By definition an ontology as an explicit specification of a conceptualization [3] represents the common understanding of a group of people about a certain domain [2]. Moreover, it does not only reflect the reality in form of a formal model, but also describes the result of a process of negotiation [12]. An ontology typically provides a vocabulary that describes a domain of interest and a specification of the meaning of terms used in the vocabulary. Depending on the precision of this specification, the notion of ontology encompasses several data/conceptual models - classifications, database schemas or axiomatized theories. Ontologies tend to be put everywhere, in many applications, such as peer-to-peer systems, electronic commerce, semantic web services and, of course, social networks [8].

Social network as a term of social science is based on the graph theory. Actors (individuals, groups) are viewed as nodes, and the relationship of some kind between them as edges. Networks can have few or many actors (nodes) and one or more kinds of relations between pairs of actors (Hannemann, 2001). We are focused on social networks somehow supported by the means of web rather than in social networks in general. Today, it is more and more apparent, that web becomes semantically richer and driven by the community. Mika [13] points out, that the field of knowledge representation and reasoning has existed long before, but what is new is its application to a large-scale, open, distributed web environment. Web 2.0 is not so much about a new technology but rather about the perception of Web as user-oriented social space for collective action and creativity.

Sites such as del.icio.us, 43Places, 43Things, Flickr, YouTube, CiteULike, Last.FM, base.google.com (most general of the mentioned) models social relationships in an indirect manner – through common interests reflected in a form of tags. Besides milieu for implicit socialization [15], web provides social sites, community portals or dating sites, such as Friendster, Orkut, Ecademy, Facebook, LinkedIn, MySpace, people not only specify their friends and acquaintances, but they also maintain an explicit, self-crafted run-down of their interests and passions [5]. Maybe also interest discussion lists and groups, like Yahoo Groups should be mentioned. Many of those mentioned are toys for their users. But there are many even a professional areas, where they can prove as a highly useful and powerful means for cooperation and collaboration.

Challenges

The question now is – should we make an effort to accommodate data structures at a background of the social network sites somehow to push them forward and maybe overcome

some of their weaknesses? And the other question – what kind of methods are most applicable for representing and reasoning with the emerging structures. Mika suggests methods from graph theory and statistics as opposed to formal logics [13].

There are the two opposing requirements – the profile should be as semantically rich as possible, but the user should not feel bored, filling the profile. Maybe, the second requirement is even more important. Cantador et al. notes, that users are generally not willing to spend the precious time describing the detailed preferences to the system, even less to assign weights to them, especially if they do not have a clear understanding of the effects and results of this input [1]. Web ontologies need to carry minimal commitment in order to be adopted in a wide scope [17]. Further, the dynamic environment of the Web means dictates low commitment, as an ontology carrying minimal commitment is more likely to persist in time despite changes in the environment [11].

According to Kim [4], there are many examples of highly successful sites with profiles inwards primitive, but funny to be filled e.g. just by tagging items in collaborative tagging systems, folksonomies. According to Vander Wal, folksonomy means the user-generated and distributed classification system, emerging through a bottom-up consensus. Commonly cited advantages of folksonomies are their flexibility, rapid adaptability, free for all collaborative customization and the high level of serendipity [10]. People can use any term as a tag without exactly understanding it's meaning. The power of folksonomies stands in the aggregation of tagged information that one is interested in. This improves social value by enabling social connections and by providing social search and navigation [16].

But on the other hand, Kim [4] states, that even though the simplicity and ease of use of tagging is desirable, it leads to troubles caused by the low semantic value. Zhdanova [18] points out that also profile-driven community portals, related to business or leisure are rather inflexible when it comes to the specification of user profiles, the content of the portals, the ways in which this content is organized, and to the search options. Moreover, social networks are generally separated with each other, though there are many online communities, whose interests and ideas overlaps with one another [7] and the low semantic value of the data brings further barriers for possible cooperation or integration.

Synergy of ontologies in social networks

One of the central aspects of social networks is a term of relationship. We can consider social relations either explicit, stated by users themselves, or implicit, inferred, learned somehow. There are many possible sources for relationship maps mining – either relatively direct ones -

address books, databases of citations, or the more decent ones – lists of shared interests stated in profiles or even raw text of article. As soon as we have relations harvested, ontology may offer the appropriate structure, to realize them and perform desired operations on top of them.

As a part of social network analysis, ontology based social network models help explicating relationships, that may not be immediately obvious [14]. Relations should be inferred neither solely out of keywords used, nor of general profile match. Rather, the hidden links based on the similarity of common preferences should be established. For example, the opinions of users on music could be valuable for each other, but many collaborative recommender or matching systems ignore it, because a global relatedness between the users seems to be low. Ontologies allow inferring more generic concepts from the more specific and vice versa – book implies both its author and the literary genre. It is apparent, that user related to a narrower term is also associated with the broader term, so we are talking about effort to extract a hierarchy based on sub-community relationships. There are interesting initiatives on the field of clustering user profiles based on common preferences. Cantador et al. proposes multi layered semantic social network model, which would help to find deeper similarities and relations among individuals. They suggest three steps: semantic preference extension, concept clustering (to identify cohesive interests) and finally the clustering of users [1].

People are linked in many ways - through events, family relations, professional relations, common interests. Liu et al. did research in ways, how we can build models of people outside of narrow application domains, by capturing the traces they leave, and inferring their everyday interests from this. They harvested thousands of user profiles in order to apply natural language processing to capture interests into ontologies of books, music, movies, etc. They mined out also passions - pivotal interests, which are more central to one's own self-identification. It resulted into the map of social and cultural identities (e.g. “Book Lover”, “Rock Musician”). Further, they analysed patterns of how these interests and identities co-occur, and generated a network-style map of affinities between different interests and identities. [5]

If we talk about folksonomies, Mika [11] suggests to take users more seriously as an integral part of ontology model. He represents networks of folksonomies at an abstract level as a tripartite graph with hyper edges, utilizing the set of actors (users), the set of concepts (tags, keywords) and the set of objects annotated (bookmarks, photos etc.), so he extended the traditional bipartite model of ontologies (concepts and instances) by incorporating actors in the model. Out of the model, several interesting graphs may be generated, such a actor/concept (called affiliation network also), concept/object, actor/instance. Experiments

leded Mika to the conclusion, that the actor/concept network better reflects and more closely match the conceptualizations of those involved in particular community.

Zhdanova [18] goes a step forward in semantic richness of primary profile data, providing portal members with means to cooperate on direct creation of shared ontology. Users should be able to introduce new concepts, available for utilization by others. Benefits are obvious. Whereas ontology management is an expensive process, in community ontology management, the expenses are shifted from the portal maintainers to the community. This shift results in an adequate investment distribution among the ontology items, such as classes and properties, so the ontology items of higher importance to the community gain more support. The ontologies which are being constructed, aligned and operated by their community represent the domain more comprehensibly than ontologies designed and maintained by an external knowledge engineer. The approach also provides a higher dynamics and up-to-date relevance to the outside world's changes in time. The most obvious risk of the approach is the possible redundancy, maybe also inconsistency of the ontology. Zhdanova believes, that these troubles could be diminished by means of general consensus expressed on a well designed architecture.

Conclusion

Further integration of ontologies and social networks or community portals seems to be a real challenge. We wish to increase the semantic value, but not at the cost of too much complicating. Maybe, the way indicated by Jung et al. [9] will prove as the vital one - first, portals should lead users to make their profiles a bit more semantic, allowing them to explicit a bit of their knowledge or thoughts, then we have to do our best to process the data provided, in order to deepen or intensify the semantic extent. Mika suggests to take users more seriously as an integral part of ontology model. Then, it is natural to consider only the associations created (explicitly or implicitly) by those involved in a subgroup, relevant to the relation type searched. The similarities to clustering works of Cantador et al. [1] of are obvious. In order to overcome the current ontology engineering bottlenecks, Zhdanova [18] urges to do more on delivery of a domain-independent, easy to use ontological infrastructure in order to support bottom-up community-driven both general and domain-dependent ontology construction where consensus modeling would be vital part of the system. Recent research has shown effectiveness of knowledge acquisition and other techniques in social space of online communities, but the same research also brought understanding that we need to move further towards implementation [6].

In the article we focused on ontologies for profile specification, besides this we may consider ontologies also for integration or for trust formation and propagation just to mention some of opportunities. But as power of our tools increases, we also have to be more careful. Ontology-based profiles will be able to hold a lot of our personalities, step further an integrated social super-network will cover up most of us and our relations. Having this in mind, Mika [13] points out, that there is a latent danger hidden in the answers.

References

- [1] CANTADOR IVAN, CASTELLS PABLO, *Multilayered semantic social network modeling by ontology-based user profiles clustering: Application to collaborative filtering*, Departamento de Ingeniería Informática, Universidad Autónoma de Madrid, Spain, 2006
- [2] FENSEL D., *Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce*, Springer, 2001
- [3] GRUBER T. R., *A Translation Approach to Portable Ontology Specifications.*, Knowledge Acquisition, 1993
- [4] HAK LAE KIM, SIMON SCERRI, JOHN G. BRESLIN, STEFAN DECKER, HONG GEE KIM, *The State of the Art in Tag Ontologies: A Semantic Model for Tagging and Folksonomies*, Digital Enterprise Research Institute, National University of Ireland, Galway and Biomedical Knowledge Engineering Lab, Seoul National University, 2008
- [5] HUGO LIU, PATTIE MAES, *InterestMap: Harvesting Social Network Profiles for Recommendations*, MIT Media Laboratory, 2005
- [6] CHKLOVSKI T., *LEARNER: a system for acquiring commonsense knowledge by analogy*, Proc. of Second International Conference on Knowledge Capture, 2003
- [7] CHUNYING ZHOU, HUAJUN CHEN, TONG YU, *Social network mashup: Ontology-based social network integration*, College of Computer Science, Zhejiang University, 2008
- [8] JEROME EUZENAT, PAVEL SHVAIKO, *Ontology Matching*, Springer-Verlag, Berlin Heidelberg, 2007
- [9] JUNG JASON J., EUZENAT JEROME, *Towards semantic social networks*, Department of Computer and Information Engineering, Inha University Incheon, Republic of Korea, 2007

- [10] MATHES A., *Folksonomy-Cooperative Classification and Communication throughs shared Metadata*, Computer Mediated Communication, Graduate School of Library and Information Science, University of Illinois Urbana, 2004
- [11] MIKA PETER, *Ontologies are us: A unified model of social networks and semantics* , Business Informatics group, Faculty of Sciences, Vrije Universiteit, Amsterdam, 2007
- [12] NICOLAS WEBER, *The Evolution of Social Ontologies*, Graz University of Technology, Knowledge Management Institute, Austria, 2008
- [13] PETER MIKA, *Social tagging and ontologies will co-exist*, Business Informatics group, Faculty of Sciences, Vrije Universiteit, Amsterdam, 2007
- [14] PINAR OEZDEN WENNERBERG, *Ontology Based Knowledge Discovery in Social Networks*, JRC Joint Research Center, European Commission, Institute for the Protection and Security of the Citizen (IPSC), 2005
- [15] PINAR OEZDEN WENNERBERG, TAMARA OELLINGER, *Ontology Based Modelling and Visualization of Social Networks for the Web*, JRC Joint Research Center, European Commission, Institute for the Protection and Security of the Citizen (IPSC), 2006
- [16] QUINTARELLI E., *Folksonomies: Power to the People*, Proceedings of 1' ISKO Italy- UniMIB meeting, <http://www.iskoi.org/doc/folksonomies.htm>, 2005
- [17] VAN ELST L., ABECKER A., *Ontologies for information management: balancing formality, stability, and sharing scope.*, Knowledge Management Department, German Research Center for Artificial Intelligence, 2002
- [18] ZHDANOVA ANNA V., *Community-driven ontology construction in social networking portals*, Web Intelligence and Agent Systems: An International Journal 6, 2008