

Collaborative Wiki Tagging Platform for Multilanguage Glossaries

Position Paper

Milorad Tasic, University of Nis, Faculty of Electronic Engineering, Nis, Serbia,
mbtotic@yahoo.com, <http://infosys1.elfak.ni.ac.yu>

Olivera Tasic, Virtuona d.o.o., Serbia,
olivera.totic@virtuonasoft.com, <http://www.virtuonasoft.com>

Background

We have reach experience in using as well as developing wiki applications in a number of different application domains, such as blended learning approach to higher education (<http://infosys3.elfak.ni.ac.yu/nastava/>), collaborative wiki tagging (<http://www.tagleen.com>), content management (<http://infosys1.elfak.ni.ac.yu>, <http://www.virtuona.co.yu>), etc. In this position paper, we address the important role that a glossary has in a wiki translation process. We propose the collaborative wiki tagging platform for developing wiki glossaries (based on our first practical experiences <http://glossary.tagleen.com>). Also, we point out the open challenges related to the social protocols design and heterogeneous community interaction. The paper's proposition is that the emergent semantic of the proposed wiki glossary is naturally complementing the role of the glossary in the wiki translation process.

Problem Statement

Recently, we have been involved in a distributed, multi-project effort for localization of several Open Source applications. The localization was focused on user interface translation on the local language but also included translation of the help system, documentation, FAQs, etc. The localization teams were primarily consisted of IT people with some support of local language experts. The main target audience was computer-illiterate, non-English speaking people that are going to use computers on their daily jobs in the near future. Language barrier is one of the biggest obstacles for that people. However, the package of the localized applications would be available to the general public also. The most important and the most visible issue on the project as a whole was a consensus over a common glossary. The glossary would be uniformly used across user interfaces of different applications by different localization teams. In this way, different localized applications would have the same look-and-feel that is of crucial importance for lowering the overall adoption barrier. However, several problems immediately appeared making the development of the common glossary extremely challenging problem. Two of the most important problems are: 1) How to support distributed, collaborative authoring of the Multilanguage glossary, and 2) How to achieve sufficient adoption across different communities of users.

Collaborative Wiki Tagging and Semantic Model of the Wiki Glossary

Semantic model of the proposed wiki glossary is based on semantics of the collaborative wiki tagging systems [1], [2]. Each term is considered as a resource that is uniquely distinguished by it's identity. The identity of the resource is reified by it's URL. In the

wiki glossary, the term is represented by a wiki page with the page name equal to the term. Links on the term's wiki page are considered as a semantic attributes of the term. For example, there are two links [Term] and [English] on the "Account" page meaning that the "Account" is a term in English language. Translation of a term is implemented as a relation between two terms. The relation is a resource class. We can make an arbitrary number of instances from the class, so we can instantiate different translation resources. The wiki page of the translation resource contains links [Translation], [English], and [Srpski] (meaning "This is translation from English to Serbian language") but also links [Account] and [Racun] (meaning "Racun is a term obtained by translation of the term Account"). This approach, that may be referred to as a "end-user meta programming", proves to be very powerful. Namely, using the same technique we implemented user comments on arbitrary page. Also, it is easy to scale the number of supported languages by simply allowing end users to create new instances of the "Language" resource class.

Social Protocols and Cross-Community Glossary Development

Meaning negotiation for a multilingual glossary is very difficult. One of the main sources of the complexity is different roles that different communities play in the process. Meaning of a term is heavily dependent on the specific context that the term has in the specific community. Also, the translation with the lowest adoption resistance in a specific community is very dependent on the vocabulary used within the community. For example, adoption resistance for computer related terms that are translated on a local language is very high in IT community since they are accustomed the English terminology. Some people are involved as active contributors, some as occasional contributors while most of the people are involved as end users only. People from different communities (like "linguistic community" vs. "IT community" vs. "academic community" vs. ...) bring in different background knowledge, different communication norms, and different vocabularies, among other. Such a cross-community software-intensive socially-empowered system has a specific architecture [3] [4]. Social protocols in such systems are very important success factor and we should pay special attention on their design, selection and adoption. This is particularly challenging task because of the participative and open nature of the wiki glossary system.

Further Reading

- [1] [Valentina Milicevic, Milorad Tosic, "Collaborative Wiki Tagging"](#), T. Pellegrini, S. Schaffert (Eds.): Proceedings of I-MEDIA'07 and I-SEMANTICS'07, International Conferences on New Media Technology and Semantic Systems, as part of [TRIPLE-I 2007](#), Graz, Austria, September 5-7, 2007.
- [2] [Milorad Tosic, Valentina Milicevic, "Semantics of The Collaborative Tagging Systems", "3th European Semantic Web Conference", "2nd Workshop on Scripting for the Semantic Web"](#), June 11-14, 2006, Budva, Serbia & Montenegro.
- [3] [Milorad Tosic, "Meta-Architecture for Intelligent Information Systems", Workshop on Designing for Reflective Practitioners: Sharing and Assessing Progress by Diverse Communities, CHI2004.](#)
- [4] [Milorad Tosic, Ladders of Reflection in the Community Infrastructure, <http://infosys1.elfak.ni.ac.yu/InfosysWiki/Wiki.jsp?page=ReflectiveCommunity>](#)