

Preface

WORKSHOP ON CROSS-LINGUAL AND CROSS-CULTURAL ISSUES OF THE SEMANTIC WEB (C3LSW-2010)

1. Motivation

Ontologies, are the conceptual framework of choice for the Semantic Web. In theory, ontologies would allow cross-lingual querying and presentation of query results across multiple languages, provided that appropriate lexical mappings of the source and target languages have been implemented. In reality, lexical mappings are just one part of the equation, as other non-lexical issues, more specifically; cultural-dependency issues also play an important role. Addressing these issues, has become a particularly urgent necessity for the realization of the Semantic Web as over 76% of the estimated 1.8 billion total number of Web users, speak a language other than English, so far the lingua-franca of the Web. We desperately need novel approaches to address the theoretical and practical issues related not only to lexical-to-conceptual mappings of natural language to ontologies, but also to a variety of non-linguistic issues. Just as internationalization (i18n) and localization (L10n) do not simply consist of translation of interface components, but also of careful cross-cultural and cross-functional consideration to the cultural sensitivities of the intended source and target languages, similar considerations should be given to those issues in the realization of the multilingual Semantic Web. This workshop is intended to gather researchers who are performing innovative research that directly or indirectly address these Issues.

2. WORKSHOP DESCRIPTION

This workshop is intended to provide a forum to share and discuss innovative ideas approaches in ontology, agent and Semantic Web research in areas related to the cross-lingual and cross-cultural issues of the multilingual Semantic Web. The issues discussed will not only be relevant and timely to ISWC2010 participants, but also of much urgency, because the exponential growth non-English-speaking users of the Web and thus the Semantic Web. This urgency becomes apparent as most known written languages become part of the Web and as the rapid penetration rate of the Web users to over 25% of the World population, most of which speak a language other than English, so far the lingua franca of the Web, requires that we pay close attention to the potential issues that might arise from a multilingual and multicultural Semantic Web. In particular, this is true for languages spoken in Asia as it is interesting to note that just for a period of 9 years, from end of the year 2000 to the end of the year 2009, Asia enjoyed a 568.8% increase in Internet usage penetration and currently holds over 42% of the total world Internet users, this with a penetration rate of just over 20% of the total Asian population, which leaves it much room to grow compared to the rest of the World. This is interesting, because we have barely scratched the surface of the Semantic Web with languages such as Chinese, Japanese, Korean and Russian and Arabic, which together account for about 646.8 million users or about 35.8% of the total 1.8 billion current users[1]. As the

penetration rate for these languages increases, new approaches to enable the realization of the multilingual Semantic Web become even more important.

INTENDED WORKSHOP AUDIENCE

One of the important topics of the workshop, is how to map language-dependent lexica to language-independent ontologies, to enable cross-lingual and cross-cultural querying and presentation of query results across multiple and diverse languages. The increasing use of nonnative languages on the Web can be insightfully compared with the development of interlanguages by second language learners. Moreover, the successful implementation of the multilingual Semantic Web across multiple languages and cultures requires novel approaches to address the theoretical and practical issues related, not only to the mappings of natural language lexica to conceptual ontologies, but also to a variety of other factors peculiar to cultures and their spoken languages. Consider for example, that internationalization (i18n) and localization (L10n) do not simply consist of the relexification of interface components (e.g., phrase-for-phrase translation), but also of careful cross-cultural considerations of the sensitivities of the intended speakers of the source as well as the target languages. Conversely, although ontologies are fundamentally language-independent, when ontologies are used as an interlanguage framework to translate lexical information from one language to another, this usage is subject to the cross-cultural influences which have been historically identified by i18n and L10n efforts. Thus, just as in the case of traditional i18n and L10n approaches, we must address these issues, not only from the theoretical point of view, but also from the engineering, methodology and work-flow points of view.

Moreover, these issues transcend linguistic and cultural aspects as they affect functional implementation of software components, specifically software agents, which must inevitably cross language and cultural boundaries when querying and performing reasoning across ontologies developed under the many linguistic and cultural biases possible in a multilingual Semantic Web. Some questions that come to mind are: Do we need to adapt agent designs, to accommodate for the multi-linguistic and/or multi-cultural boundaries, which agents must inevitably cross to gather knowledge on which to perform reasoning? If so, do we also need to adapt agent communication and negotiation policies for agents collaborating from across these boundaries? Can we learn from and therefore generalize these policies from the inter-linguistic and multi-cultural negotiations that take place in multi-cultural societies?

Among other topics we anticipate that authors and participants will be interested in one or more of the following topics:

- Design principles of cross-lingual/cultural ontologies.
- Cross-lingual/cultural semantic annotation.
- Cross-lingual/cultural information extraction.
- Cross-lingual/cultural frameworks, workbenches or platforms.
- Cross-lingual/cultural applications.
- Cross-lingual/cultural ontologies for web services
- Cross-lingual/cultural policies for collaborative agent interaction.
- Agent-based cross-lingual annotation, mapping and extraction.

2. LIST OF PROGRAM COMMITTEE MEMBERS

Jeffrey M. Bradshaw, IHMC, Pensacola, Florida (confirmed)

Andrzej Uszok, IHMC, Pensacola, Florida (confirmed)

David W. Embley, Brigham Young University, Provo, Utah (confirmed)

Stephen W. Liddle, Brigham Young University, Provo, Utah (confirmed)

Yasuhiko Kitamura, Kwansei Gakuin University, Sanda, Japan (confirmed)

Koji Kozaki, Osaka University, Osaka, Japan (confirmed)

Takahira Yamaguchi, Keio University, Tokyo, Japan

Youngim Jung, KISTI, Daejeon, Korea

Yuri Tijerino, Kwansei Gakuin University, Japan

Tarek Helmy, King Fahd University of Petroleum and Minerals, Saudi Arabia

3. ORGANIZING COMMITTEE

Chair:

Yuri A. Tijerino, Associate Professor

Department of Applied Informatics,

School of Policy Studies,

Kwansei Gakuin University

2-1 Gakuen, Sanda, Japan 669-1337

Co-chair:

Tarek Helmy, Associate Professor College of Computer Science and Engineering, Department of Information and Computer Science, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.