Semantic Kalevala —Accessing Cultural Content Through Semantically Annotated Stories

Eero Hyvönen, Joeli Takala, Olli Alm, Tuukka Ruotsalo, Eetu Mäkelä

Semantic Computing Research Group (SeCo)
Helsinki University of Technology (TKK) and University of Helsinki
http://www.seco.tkk.fi/, first.last@tkk.fi

Abstract. An event-based approach is presented for annotating events and narrative structures underlying texts and stories semantically. The idea is applied to using the Finnish national epic Kalevala for accessing related cultural contents, such as artifacts, paintings etc. in the semantic portal "CultureSampo—Finnish Culture on the Semantic Web".

The idea of representing knowledge in terms of perduring events (e.g., a man *riding* a horse in a forest) and related thematic roles [1] (e.g., man as an agent, horse an instrument, and a forest as a place) has been successfully applied in many fields, such as natural language processing [2] and image content description [3]. We use events as a basis to represent semantic knowledge in the portal "CultureSampo—Finnish Culture on the Semantic Web"[1] [4]. In this paper we present how narrative texts can be annotated as sequences of events and be used as views to organizing and accessing cultural contents of different kinds, such as artifacts, photos, paintings, cultural buildings, stories, etc.

In CultureSampo the text to be rendered is annotated using ontologies and an event-based annotation scheme [4, 5]. The basic use-case from the end-user's viewpoint is that she selects runes or their parts for viewing using a table of contents on the screen. The selected text is then rendered, and the system automatically creates semantic recommendation links to related cultural content items in the portal for the user. Annotated text parts, called *scenes*, can be seen highlighted. By moving the mouse on top of them, the recommendations are changed on the fly. Furthermore, additional metadata concerning the scenes, such as literal interpretations or explanations of the text, are shown as tool tips.

This approach is related to e.g. the "semantic filter" Magpie [6] used for browsing web pages enriched with semantic links, and to various RDF browsers. However, semantic browsing in CultureSampo is more advanced w.r.t. these approaches in the following ways: 1) Recommendation links are based on events and narrative structures [7], not only on extracting and identifying ontological concepts in the text, allowing creation of more "intelligent" linking. 2) Semantic rules are used for determining complex associations and recommendation

[1] http://www.kulttuurisampo.fi/

links between content items. 3) The system provides the end-user with literal explanations of why the recommendation links should be of interest to her.

One of the texts being incorporated in the system is (a part of) the national Finnish epic Kalevala[2] published originally in 1835. Kalevala has been a continuous source of inspiration in Finnish fine arts, music, sculpture, literature, and other branches of culture. The mythical world of Kalevala also nicely relates to the original agrarian Finnish life and artifacts that are available in Finnish museums, and that can be associated with the texts.

Four runes, each about 1500 words long, were used in our case study. Each scene was annotated by a set of instances of perduring concepts (e.g., running) of the general Finnish upper ontology YSO[3] of 20,000 concepts, with a set of thematic roles filled with other resources of YSO and two Kalevala-specific ontologies: places and actors. Scenes (132 in total with 383 events) could be related with each other by "larger scene", "next scene", and "consequence scene" relations.

Acknowledgements

Jouni Hyvönen annotated Kalevala, and Heli Kautonen contributed with discussions. This research is part of the National Finnish Ontology Project (FinnONTO) 2003-2007[4], funded mainly by The National Technology Agency (Tekes) and a consortium of 36 companies and public organizations.

References

1. Sowa, J.: Knowledge Representation: Logical, Philosophical and Computational Foundations. Brooks/Cole, Pacific Grove, CA (2000)
2. Zarri, G.P.: NKRL, a knowledge representation tool for encoding the meaning of complex narrative texts. Natural Language Engineering **3**(2) (1988)
3. Schreiber, A.T., Dubbeldam, B., Wielemaker, J., Wielinga, B.: Ontology-based photo annotation. IEEE Intelligent Systems **16**(3) (2001)
4. Hyvönen, E., Ruotsalo, T., Häggström, T., Salminen, M., Junnila, M., Virkkilä, M., Haaramo, M., Kauppinen, T., Mäkelä, E., Viljanen, K.: CultureSampo—Finnish culture on the semantic web. The vision and first results. In Robering, K., ed.: Information Technology for the Virtual Museum, LIT Verlag (2007)
5. Ruotsalo, T., Hyvönen, E.: An event-based approach for semantic metadata interoperability. In: Proceedings of the 6th International Semantic Web Conference (ISWC/ASWC 2007), Busan, Korea, Springer–Verlag, Berlin (2007)
6. Dzbor, M., Motta, E., Domingue, J.: Magpie: Experiences with supporting semantic web browsing. Journal of Web Semantics **3**(2) (2007) 224–241
7. Junnila, M., Hyvönen, E., Salminen, M.: Describing and linking cultural semantic content by using situations and actions. In Robering, K., ed.: Information Technology for the Virtual Museum, LIT Verlag (2007)

[2] http://www.finlit.fi/kalevala/index.php?m=163&l=2
[3] http://www.yso.fi/
[4] http://www.seco.tkk.fi/projects/finnonto/