On Community-based Ontology Evolution

Pieter De Leenheer

Semantics Technology and Applications Research Laboratory (STARLab)
Vrije Universiteit Brussel
Pleinlaan 2, B-1050 BRUSSELS 5, Belgium

Abstract

The second-generation Web (2.0) is a complex socio-technical system of unprecedented growth and dynamics. On-line communities emerge and interact all around a usually self-organising manner, supported by *interactive applications* including bookmarking, tagging, blogging, and wikis, being developed and shared at little or no cost. The emerging range of Semantic Web and other open technologies promises an increase in scale and maturity of knowledge sharing, achieved through collaboration and integration within and between diverse communities.

Considering the current pace of social and technological development, it seems that the transformation of the Web from a network of separately siloed applications and content repositories to a more seamless and *semantically interoperable* ecosystem is at hand, opening a wide range of scientific challenges and opportunities. However, while simple, the idea of the Semantic Web remains largely unrealised.

The ontologies that will furnish the Semantic Web are lacking, and those few that have been published are usually not based on consensus, and hence unreliable and not reusable beyond individual purposes. Our key observations are that in current ontology engineering practices, the underlying *methodological* principles are mostly ignored. Furthermore, they systemically disregard the subtle *gap* that looms between knowledge sharing among people at the community/social level on the one hand; and information exchange between computer systems at the operational/technical level.

Architecting community-driven internet systems will require a paradigm shift that goes beyond mere technological fits. In order to bridge the gap between the social and technical part of the community, one must put into practice the necessary activities to identify common needs from socialisation activities and bring the stakeholders together to find and ontological agreement to support these needs.

> *Community-based ontology evolution* establishes the co-evolution of (A) *social interactions* enabled by the community's design; (B) the information systems that support them; and (C) the *semantic patterns* to fulfil *semantic interoperability* between these systems, as a fundamental property of *knowledge-intensive communities*.

In order to enact this co-evolution, in this dissertation, we introduce a *community semantics management* method and system, as illustrated below. The approach is based on the following seven principles:

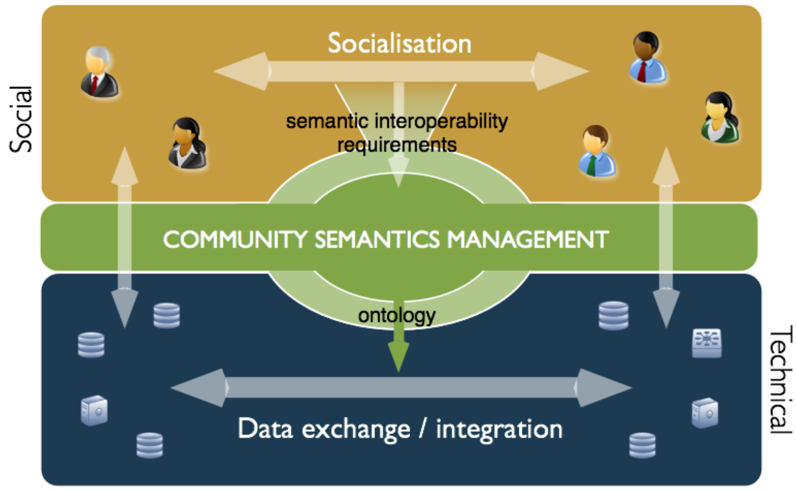

Fig. 1. In order to bridge the gap between the social and technical part of the community, community semantics management must enact the necessary activities to identify common needs from socialisation activities and bring the stakeholders together to find and ontological agreement to support these needs.

1. **ICT Democracy** An ontology should be defined by its community, and not by a single developer.
2. **Autonomy** Semantic interoperability requirements emerge autonomously from community evolution processes.
3. **Co-evolution** Ontology evolution processes are driven by the changing semantic interoperability requirements.
4. **Perspective Rendering** Ontology evolution processes must reflect the various stakeholders perspectives.
5. **Perspective Unification** In building the common ontology, relevant parts of the various stakeholder perspectives serve as input for the unified perspective.
6. **Validation** The explicit rendering of stakeholders perspectives allows us to capture the ontology evolution process completely, and validate the ontology against these perspectives respectively.
7. **Satisfaction** Ultimately, co-evolving communities with their ontology will increase overall stakeholder satisfaction.

From these principles we deduced the following three key requirements for the design of our approach.

1. We *formally describe community evolution* processes in terms of Nonaka's knowledge conversion modes that define the *cyclic* interaction between tacit conceptions and explicit representations in a typical community. We illustrate how these conversion modes conceptually bridge the *gap* that looms between the computerised subsystem and the social (human) subsystem of the knowledge-intensive community,

and by doing so engender an *upward knowledge spiral*, where individual knowledge opinions become commonly accepted, through an iterative interplay between externalisation and internalisation, and socialisation and combination respectively. This interplay illustrated how knowledge artefacts such as facts and fact types, being key building blocks for ontologies, autonomously co-evolve with their communities of use.

2. Linking these (socio-technical) community evolution processes with the ins and outs of the (technical) ontology evolution activities and vice versa is not straightforward. Therefore, we outline a *community semantics management methodology* consisting of seven activities that enacts community-based ontology evolution, consisting of the following constructs.
   - a two-tier *community semantics* platform to store and version community semantics patterns;
   - a typology of *perspective policies* to interlink and control the evolution of the patterns;
   - a general community semantics *architecture* in terms of patterns, perspectives on these patterns, and policies.

   In order to implement the community semantics storage and formally describe the seven community semantics management activities, and their link with the community evolution deliverables, we adopt and extend the DOGMA ontology framework. In that context we also define a perspective management extension to control the divergence and convergence of the community architecture using graph transformation.
3. Finally, we establish ICT democracy via the explicit identification of *affordances and roles* of knowledge workers contributing to the ontology. This presumes an explicit *metamodel for professional communities* including micro-scale semantic patterns such as actor, action, goal, etc., as underlying semiotic framework for community semantics management.

Finally, we demonstrate and validate our approach and tool in a realistic case study that was carried out within the framework of the European Codrive project. The vision of Codrive was to contribute to more effective and meaningful competency-centric human resource management. Key challenges were the *uniform* publication of *unambiguous* competency information and "time-to-competency" agility. To this end, we developed and evolved a shared and formal knowledge representation of competence domains. Community stakeholders include educational institutes, public employment organisations, and industry partners from different European countries. The resulting Vocational Competency Ontology wanted to provide a candidate best practice for engineering a community-shared and reusable semantic pattern base that can be applied by all stakeholders to semantically reconcile their competency models.