Protégé Knowledge Base Coordinator

Noah Zimmerman, Stephen Meehan and Leonore Herzenberg
Department of Genetics
Stanford University School of Medicine

Introduction

One of the major accomplishments of the knowledge-representation community in recent years has been the creation of formally represented reusable knowledge-sources. Ontologies like the Gene Ontology (GO), Suggested Upper Merged Ontology (SUMO) and the Foundational Model of Anatomy (FMA) [3] are curated resources publicly available for developers to "plug-in" to their domain specific knowledge-based application. This growing repository of resources supports the ultimate vision of ontologies as reusable knowledge components and services that can be invoked across networks [2].

While the resources are emerging to create such a library, there is still an inadequate infrastructure for integrating these disparate sources. Tools such as PROMPT [1] have emerged to address the problem of mapping, merging and extracting semantically independent components of these ontologies; however, the tools are not meant to safeguard the semantics, security, integrity and accessibility of a multi-ontology system. Our knowledge-base coordinator (KBC) addresses these issues by providing a seamless integration of multiple ontologies based on modeled knowledge in a secure, traceable and location transparent design.

The KBC improves upon the ontology inclusion mechanism in Protégé by providing extensions to support multi-ontology hierarchical systems. The KBC provides an integrated model which supports multiple ontology types, rules for ontology inclusion, access control security and ontology versioning. We utilize the Protégé plug-in architecture to interpret this model at runtime to monitor and authorize opening, creating, modifying, including and saving an ontology. The KBC also provides a client-server architecture for coordinating the versioning and offline caching of multiple ontologies.

Background

The KBC emerged out of a need for knowledge-base management in the FacsXpert project. FacsXpert is a Protégé based tool developed at the Herzenberg Laboratory for supporting Fluorescence Activated Cell Sorting (FACS) protocol development. It utilizes a four-tiered knowledge structure (figure 1) to provide the user with context sensitive decision support when designing a FACS experiment. Each of these levels of knowledge maps to distinct Protégé ontologies.

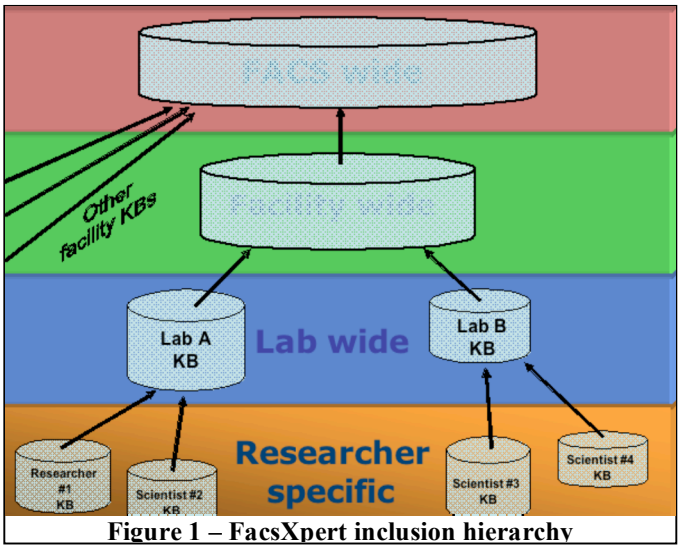

Figure 1 – FacsXpert inclusion hierarchy

At the FACS wide level, we represent concrete knowledge shared by all FACS researchers. This includes knowledge of fluorescent elements, their fluorescence excitation and emission, lasers, and FACS instruments. The facility level provides knowledge about the facility, such as specific instrument configurations. At the lab wide level, we represent shared laboratory information such as inventories and shared protocols. Finally, the researcher specific level contains knowledge about the individual's protocols.

Each level includes its parent ontology so that the user is presented with a single coherent view of a multi-ontology system. These inclusions are constrained by the KBC model (also a Protégé ontology) to prevent invalid inclusions such as a lab wide ontology including a researcher specific ontology. The KBC also provides a model to restrict users' read/write privileges at varying levels of the hierarchy. For instance, a lab administrator has permission to modify her own researcher specific knowledge-base as well as the lab wide knowledge-base, but not the facility knowledge-base or FACS wide knowledge-base. In each case, however, the user need not be aware of the knowledge-hierarchy – the instances are simply presented to them as editable if they have the requisite privileges. Should the user (or more likely, the engineer) be interested in the underlying structure, the KBC extends Protégé's inclusion mechanism to trace the ontology inclusion path for each instance.

The client-server architecture of the KBC manages the versioning and offline caching of ontologies. Versioning is imposed on the ontologies to track changes by multiple users. Currently, we have not implemented a system for merging these changes, resulting in a competitve system. We plan to integrate PROMPT to manage the merging process as part of our upcoming development.

[1] N. F. Noy and M. A. Musen. The PROMPT suite: Interactive tools for ontology merging and mapping. International Journal of Human-Computer Studies, 59(6):983–1024, 2003.

[2] Gruber, T. R. A Translation Approach to Portable Ontology Specifications. 1993.

[3] Rosse C, Mejino JL, Jr. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. J Biomed Inform 2003;36(6):478-500.