

Molecular Biotechnology Programme

Uppsala University School of Engineering

UPTEC X 10 014

Date of issue 2010-07

Samuel Lampa

Title (English)

SWI-Prolog as a Semantic Web Tool for semantic querying in Bioclipse: Integration and performance benchmarking

Abstract

Author

The huge amounts of data produced in high-throughput techniques in the life sciences and the need for integration of heterogeneous data from disparate sources in new fields such as Systems Biology and translational drug development, require better approaches to data integration. The semantic web is anticipated to provide solutions through new formats for knowledge representation and management. Software libraries for semantic web formats are becoming mature, but there exist multiple tools based on foundationally different technologies. SWI-Prolog, a tool with semantic web support, was integrated into the Bioclipse bio- and cheminformatics workbench software and evaluated in terms of performance against non-Prolog-based semantic web tools in Bioclipse, Jena and Pellet, for querying a data set consisting of mostly numerical, NMR shift values, in the semantic web format RDF. The integration has given access to the convenience of the Prolog language for working with semantic data and defining data management workflows in Bioclipse. The performance comparison shows that SWI-Prolog is superior in terms of performance over Jena and Pellet for this specific dataset and suggests Prolog-based tools as interesting for further evaluations.

Keywords

Semantic Web, Prolog, Bioclipse, RDF, SPARQL, NMR shift, Eclipse, Java

Supervisors	
Dr. Egon Willighagen Uppsala University	
Scientific reviewer	
Prof. Mats Gustafsson Uppsala University	
Project name	Sponsors
Language	Security
English	
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages
	37
Biology Education Centre Biomedical Center Husargatan 3 Uppsala	
Box 592 S-75124 Uppsala Tel +46 (0	$)18\ 4710000 \text{Fax } + \overline{46}\ (0)18\ 471\ 4687$