Ontology Management and Selection In Re-Use Scenarios

Kim Finney (B.Sc, M.Sc)

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

University of Tasmania

School of Computing and Information Systems

October 2012

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

The research associated with this thesis abides by the international and Australian codes on human and animal experimentation, the guidelines by the Australian Government's Office of the Gene Technology Regulator and the rulings of the Safety, Ethics and Institutional Biosafety Committees of the University.

This thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Acknowledgements

Several people require special mention who have supported me to deliver this thesis.

First, is my partner, Leanne Wilkes who has lived through this episode vicariously and who probably deserves the award more than I do. Without her critical eye, robust comments and stamina beyond belief in reading and re-reading this dissertation I would never have finished.

Second are my parents. My father, Bob Finney a man of prodigious talent whose enquiring mind inspired me to undertake this research in the first place, offered much useful criticism and advice. My mother, Brenda Finney as always practical and supportive, helped me push on when the end wasn't clearly in sight.

Last, but not least are my supervisors, Simon Milton, Chris Keen and Peter Marshall whose reviews, useful insights, experience and comments helped me to make a thesis.

I would also like to dedicate this dissertation to Elizabeth Bell (1916 – 2009), who didn't understand the Web, had never heard of an ontology and who disliked computers, but encouraged me none-the-less, just as she had done for everything else I've ever attempted in my life.

Abstract

One of the main impediments to realising the Semantic Web vision is that most scientific data, even those data deployed on the web, are not generally expressed or encoded in an unambiguously defined, machine-interpretable manner. This is particularly the case for Antarctic-themed data. Ontologies that are linked to datasets via semantic annotation are required to achieve semantic-enablement of scientific data infrastructure. In scientific communities that adhere to the Open Geospatial Consortium Service-Oriented-Architecture (Web services) paradigm, Feature Catalogues are the repositories intended to manage and publish descriptions of dataset concepts. This thesis explores how Feature Catalogues can be ontologically-grounded to facilitate semantic annotation and in doing so addresses the lack of guidance in current standards about how to configure an ontologically grounded Feature Catalogue and how best to access the resources it contains for the semantic annotation of Web services. Also investigated is how ontology selection and evaluation is currently taking place in practise because ontology evaluation methodologies mentioned in the literature are resource intensive to apply, often requiring a high level of ontological expertise. Both contributions seek to lower barriers for ontology uptake and reuse within scientific communities.

To address these issues, two scientific communities of practise (i.e., AODN and SCAR) were used as case studies within a Design Science research method to ground-truth the design and to prototype an ontologically grounded, service-enabled Feature Catalogue. To address research questions pertaining to ontology selection and evaluation practise, fourteen experts (from outside of the AODN and SCAR communities) with experience in building semantically-enabled scientific infrastructure, were surveyed and interviewed to ascertain what ontology evaluation methods and criteria are being used in practise. A hierarchical evaluation model was established from analysed expert data using Template Analysis (Crabtree and Miller, 1992; King, 2004). The Analytical Hierarchical Processing (AHP) technique (Saaty, 1980), was then harnessed to establish the relative importance given by experts to each of the model elements.

The contributions arising consisted of an enhanced ISO 19110 Feature Catalogue model which accommodated additional concepts necessary to describe the observation-centric dataset paradigms of the two case study communities. The extended conceptual model was semantically grounded using the DOLCE (upper ontology) and expressed in both OWL and SKOS. Demonstration REST-based service interfaces (and REST query patterns) were created for serving Catalogue content to requesting Web clients. To the author's knowledge, no other Feature Catalogue implementation, founded on the ISO 19110 conceptual model, has attempted to model the Catalogue as an ontology, or permits access to Catalogue content via REST-based service interfaces. This thesis also delivers a

“practical” framework for evaluating and then selecting reusable ontological content which encompasses weighted model elements (indicating relative levels of importance), coupled with expert-derived evaluation metrics. Although the evaluation criteria listed in the framework are not novel in themselves, identifying which criteria are of most utility to experts who are operating in real-world scenarios, is an important contribution to practise.

Table of Contents