Web Semantics: Science, Services and Agents on the World Wide Web 27–28 (2014) 70–77

Contents lists available at ScienceDirect

Web Semantics: Science, Services and Agents on the World Wide Web

journal homepage: www.elsevier.com/locate/websem

Lessons learnt from the deployment of a semantic virtual research environment

Peter Edwards [a,*], Edoardo Pignotti [a], Chris Mellish [a], Alan Eckhardt [a], Kapila Ponnamperuma [a], Thomas Bouttaz [a], Lorna Philip [b], Kate Pangbourne [b], Gary Polhill [c], Nick Gotts [c]

[a] Computing Science, University of Aberdeen, United Kingdom
[b] School of Geosciences, University of Aberdeen, United Kingdom
[c] The James Hutton Institute, Aberdeen, United Kingdom

ARTICLE INFO

Article history:
Received 28 September 2013
Received in revised form
6 May 2014
Accepted 24 July 2014
Available online 23 August 2014

Keywords:
Provenance
Policies
Natural language
OPM
PROV-O

ABSTRACT

The *ourSpaces* Virtual Research Environment makes use of Semantic Web technologies to create a platform to support multi-disciplinary research groups. This paper introduces the main semantic components of the system: a framework to capture the provenance of the research process, a collection of services to create and visualise metadata and a policy reasoning service. We also describe different approaches to authoring and accessing metadata within the VRE. Using evidence gathered from data provided by the users of the system we discuss the lessons learnt from deployment with three case study groups.

1. Introduction

Many of the contemporary challenges facing society such as climate change, require researchers from a range of disciplines to work together. Underpinning the scientific process is the transfer of ideas, knowledge and resources, and in recent years, the Web has drastically altered both the nature and speed of this exchange.

Web-based Virtual Research Environments (VREs) [1] have been proposed as one way to help researchers across all disciplines to manage the increasingly complex range of tasks involved in carrying out research. In the UK, the Joint Information Systems Committee (JISC) VRE programme[1] explored the virtual research environment collaborative landscape. JISC recognised that a major shift in research practice will occur through the formation of common taxonomies, data standards and metadata as researchers collaborate with others across disciplinary, institutional and national boundaries [1]. Semantic web technologies [2] are seen as crucial in this context in order to provide a common framework to allow the creation of intelligent applications and services which can be integrated with data resources, people and other objects in a VRE.

Some of these issues have been explored by the PolicyGrid[2] project, a collaboration between human geographers and computer scientists as part of the UK Digital Social Research initiative. As part of the project we developed *ourSpaces*,[3] a virtual research environment that allows researchers to collaborate online. A screenshot of the *ourSpaces* web interface is presented in Fig. 1.

The system was co-developed with three interdisciplinary case study groups: a research team investigating *E. coli* O157 risk in communities, members and affiliates of the Aberdeen Centre for Environmental Sustainability, and a group of agent based social simulation modellers. Based upon interactions with these groups

* Corresponding author. Tel.: +44 1224274065.
E-mail addresses: p.edwards@abdn.ac.uk (P. Edwards), e.pignotti@abdn.ac.uk (E. Pignotti), c.mellish@abdn.ac.uk (C. Mellish), a.e@centrum.cz (A. Eckhardt), k.ponnamperuma@abdn.ac.uk (K. Ponnamperuma), thomas.bouttaz@googlemail.com (T. Bouttaz), l.philip@abdn.ac.uk (L. Philip), k.pangbourne@abdn.ac.uk (K. Pangbourne), Gary.Polhill@hutton.ac.uk (G. Polhill), ngotts@gn.acp.org (N. Gotts).

[1] http://www.jisc.ac.uk/whatwedo/programmes/vre.aspx.

[2] http://www.purl.org/policygrid.

[3] http://www.purl.org/policygrid/ourspaces.

http://dx.doi.org/10.1016/j.websem.2014.07.008

Fig. 1. A screenshot of the *ourSpaces* VRE showing a user's personal page, a natural language representation of project metadata, and a provenance graph associated with a report.

the core requirements for the *ourSpaces* VRE were identified as follows: (a) It should be possible to describe and uniquely identify a range of entities: artefacts (digital and physical); processes (both computational services and human activities); people; organisational structures and membership; social networks; (b) The system should incorporate online communication (e.g. instant messaging, blog entries, email) into the research record; (c) It should be possible to define relationships (e.g. causal, social, organisational) between entities; (d) It should be possible to define access control and documentation policies.

To satisfy these requirements the *ourSpaces* architecture implements a number of core and Web services for creating, editing and querying data, metadata and digital artefacts. These include a service used to upload and access digital artefacts, a natural language service to support browsing and querying data, and a policy reasoning service [3]. The diagram in Fig. 2 illustrates the basic components of the system architecture. While *ourSpaces* is currently accessible online, it is no longer supported (as the PolicyGrid project ended in Summer 2012). The source code is available for download via GitHub[4] under the GNU LGPL version 2.1. The system requires a linux-based environment capable of running Java EE 1.6.0, apache tomcat 6.18 or above, MySQL version 14.2 or above and openrdf-sesame 2.0.

The system has been designed in order to encourage users to share their digital artefacts, download and comment on each other's work and form cohesive groups with other researchers. In order to support the formation of interdisciplinary groups in *ourSpaces*, users are presented with various means of establishing their social presence, e.g. tagging, blogging, personal status updates, instant messaging.

In a collaborative environment such as *ourSpaces*, understanding the provenance of scientific data and other research artefacts is crucial in order to understand and verify their authenticity and completeness [4]. *ourSpaces* is thus underpinned by a provenance framework capable of capturing the derivation history of research artefacts, including the original sources, intermediate products and processes involved.

In this paper we revisit the design, implementation and deployment of the *ourSpaces* VRE introduced in a previous publication [5]. We introduce a new Personal Lexicon Service (PLS) designed to create a set of vocabulary mappings and to utilise them during search and online communication. We also introduce a possible solution for adapting the provenance framework used within *ourSpaces* to the recent W3C provenance recommendations.

The remainder of this paper is organised as follows. In Section 2 we describe the design of the ontological framework required to support provenance in the VRE; Section 3 presents a number of tools that we have developed in order to support interaction with semantic metadata. In Section 4 we discuss lessons learnt from deployment of *ourSpaces* with our case-study communities. In Section 5 we introduce a set of mappings and rules to convert an *ourSpaces* provenance record to PROV-O. Finally, we discuss related work, our conclusions and future directions.

2. The *ourSpaces* ontological framework

We have developed an extensible ontological framework for capturing the provenance of the research process based on the requirements highlighted in Section 1. In order to describe and uniquely identify entities (such as artefacts, people, locations) and to make explicit relations between entities we follow the linked data principles [6]. At the heart of *ourSpaces* (and thus, our provenance framework) is an OWL representation of the Open

[4] https://github.com/policygrid/ourSpaces.