Using the Grid to Support Evidence-Based Policy Assessment in Social Science

Peter Edwards, Alison Chorley, Feikje Hielkema, Edoardo Pignotti, Alun Preece, Chris Mellish

Computing Science, University of Aberdeen
{pedwards, achorley, fhielkem, epignott, apreece, cmellish}@abdn.ac.uk

John Farrington

Geography & Environment, University of Aberdeen
j.farrington@abdn.ac.uk

Abstract

The PolicyGrid project is exploring the role of Semantic Grid technologies to support eScience for the social sciences, with a particular emphasis on tools to facilitate evidence-based policy making. In this paper we highlight some of the key challenges facing developers of semantic infrastructure and tools for social science researchers. We outline a framework for evidence management, discuss issues surrounding creation and presentation of metadata, describe a Web-based service which utilises a natural language interface to facilitate creation of RDF, and present a desktop qualitative analysis tool which is integrated with our evidence framework.

1. Introduction

The concept of 'evidence-based policy making' (Bullock, Mountford, & Stanley, 2001) came to the fore in the UK policy environment in response to a perception that government needed to improve the quality of its decision-making processes; it has been argued that in the past policy decisions were too often driven by inertia or by short-term political pressures. Evidence can take many forms: research, analysis of stakeholder opinion, simulation modelling, public perceptions and beliefs, anecdotal evidence, cost/benefit analyses; as well as a judgement of the quality of the methods used to gather and analyse the information.

The UK Government uses a range of evaluation methods to ensure that policies are as effective and efficient as possible. For example, guidance contained within *The Green Book* (HM Treasury, 2003) is used to ensure that a policy is not adopted if its objectives could be achieved in a better way, or if its resources could be used in better ways. The *Green Book* presents the techniques and issues that should be considered when carrying out an economic appraisal or evaluation of a policy, project or programme. These activities form part of a broad policy cycle that is sometimes formalised in the acronym *ROAMEF* - Rationale, Objectives, Appraisal, Monitoring, Evaluation and Feedback. Another government publication, *The Magenta Book* (Cabinet Office, 2003) describes the methods used by social researchers when they commission, undertake and manage policy research and evaluation; it includes discussion on how to use statistics from quantitative data, how to do qualitative research, how to test new policies in social experiments (using control groups, counterfactuals, etc).

APAT (Accessibility Policy Appraisal Tool) (Farrington *et al.* 2004) is an example of a specialised policy evaluation methodology that was designed to examine and evaluate the accessibility impact of policies, using a mixed-method approach. It aims to improve understanding by participants of the accessibility implications of a policy through reflection and analysis and also generates and evaluates alternative policy options.

Our work within the PolicyGrid[1] project is investigating how best to support social researchers in their policy assessment activities through the use of Semantic Grid (De Roure, Jennings & Shadbolt, 2005) technologies. The Semantic Grid is often described as an

---

[1] Funded by the Economic and Social Research Council *eSocial Science* programme.
Award Reference: RES-149-25-1027
(http://www.policygrid.org/)

‘extension of the current Grid in which information and services are given well-defined meaning, better enabling computers and people to work in cooperation’; the analogy here being that the Grid and Semantic Grid have a similar relationship to that existing between the Web and Semantic Web (Berners-Lee, Hendler & Lassila, 2001. Semantic Grids thus not only share data and compute resources, but also share and process metadata and knowledge. eScience applications which utilise semantic technologies now exist in areas as diverse as life sciences – myGrid (Stevens, Robinson & Goble, 2003), chemistry – CombeChem (Taylor *et al.* 2006), and earth sciences – SERVOGrid (Aktas, Pierce & Fox, 2004). However, until recently there has been little work exploring the potential of these techniques within the social sciences, arts and humanities.

1.1 PolicyGrid Overview

The aims of PolicyGrid are as follows: to facilitate evidence-based rural, social, and land-use policy-making through integrated analysis of mixed data types; to demonstrate that Semantic Web/Grid solutions can be deployed to support various facets of evidence-based policy-making through the development of appropriate tools; to focus on the authoring of relevant ontologies to support rural, social and land-use policy domains; to investigate issues surrounding communication of semantic metadata to social scientists and policy practitioners; and to promote awareness of the Semantic Grid vision and supporting technologies amongst social scientists.

Our work involves close collaboration between computer scientists, social scientists and other policy practitioners. These interactions have allowed us to explore a range of issues, including: the extent to which these researchers are comfortable with the Grid as a framework for research practice and collaboration; if ontologies are appropriate (and acceptable) to this community as a way of representing concepts to facilitate evidence-based policy making and assessment; the utility (or otherwise) of existing metadata frameworks in use by the social sciences; and how best to integrate eScience tools and methods into existing working practices. Key facets of our work are:

- Management of heterogeneous evidence types (including how to obtain, store and facilitate the use of evidence in policy assessment).
- Support for creation of metadata and access to resources annotated by metadata.
- Development of software tools which integrate existing working methods with a range of Grid services.

We are developing a service-oriented infrastructure based upon current Grid and Web service middleware support. These services are used as components by Web or stand-alone applications. For example, we have developed a metadata repository service based upon the Sesame open source RDF framework[2] and an object repository using a version of the Fedora[3] digital object repository.

The remainder of this paper is organised as follows: in the next section we discuss the types of evidence which need to be managed to facilitate evidence-based policy assessment, before outlining a proposed solution. A web-based service, used to elicit metadata about a resource using Natural Language techniques, and a desktop tool used to perform qualitative analysis, are then presented. We conclude with a discussion and a roadmap for future work.

2. Managing Evidence

Researchers conducting a policy assessment exercise will employ some methodology to evaluate the policy’s impact (or possible impact) on the community. They may send out questionnaires to members of the public in certain areas of the country to assess public opinion, or organise town meetings and focus groups. They may interview policy makers to gather information about the impact of the policy on the community or on other policies. They may perform a cost-benefit analysis in order to assess the fiscal impact of the policy. Such an approach is termed ‘mixed method’ – as the researcher uses a variety of methods and tools to evaluate the policy. Quantitative techniques use data obtained from questionnaires and surveys and can be analysed statistically to generate numerical evidence. Qualitative methods use data obtained from interviews, town meetings and focus groups and are usually subject to textual analysis against some conceptual ‘coding’ framework. Social simulation methods use data obtained from running a model under a specified set of circumstances and then analyse the outcome.

As we have seen, evidence can take many different forms. However, there is another

[2] http://www.openrdf.org/

[3] http://www.fedora.info/

category of evidence that is essential if one is to allow researchers to assess the quality of the methods used to gather and analyse the information – *provenance*. Groth et al. (2006) define the "provenance of a piece of data as the process that led to that piece of data". A suitable provenance architecture for e-Social Science would allow questions such as 'Where does this evidence come from?' to be answered. Pieces of evidence that form part of a policy appraisal outcome could then be traced back to their source; Figure 1 shows an example of evidence provenance from an APAT policy appraisal case study where a single piece of evidence is derived from two questions in a questionnaire.

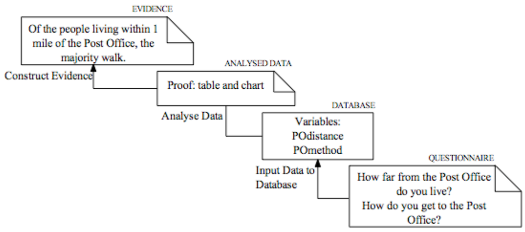

Figure 1 : Example evidence pathway (taken from the APAT rural post offices case study).

Consider the following example, illustrating the steps involved in producing evidence from quantitative data. The initial resource is the questionnaire sent out to the public to gather their responses; data from the completed questionnaires are gathered and stored in some format (database, text file). This raw data is then analysed using some statistical tool to identify patterns, which are treated as evidential statements in some final report (document). At each of these stages information about the processes used should be gathered. For example, when moving from the raw to the analysed data, the raw data may be sent to an external statistical service which performs the analysis and returns results. Knowing how evidence has been derived can prevent problems of misinterpretation and also provides an important audit trail for quality purposes. If a policy maker is exploring the process states 'This evidence seems odd. Show me how it was derived', provenance information could be used to demonstrate that a question used to generate the evidence was misunderstood by subjects completing a survey and they thus gave odd responses, or perhaps a transcription error was introduced.

3. An Evidence Framework for Social Science

Our evidence management framework comprises the following: metadata support for social science resources (qualitative, quantitative, simulation), a model of provenance (process), and argumentation tools. Before we discuss each of these components it is important to elaborate upon the role of ontologies in facilitating semantic e-Social Science.

3.1 Ontologies for Social Science

From the beginning of our work, user scientists expressed a fear of 'being trapped in the ontology' due to the contested nature of many concepts within the social sciences. Other researchers (Edwards, Aldridge & Clarke, 2006) have noted that as social science concepts emerge from debate and are open to indefinite modification through debate, vocabularies also tend to be imprecise (e.g. there is no precise definition of 'anti-social behaviour') and mutable (vocabularies tend to change over time to reflect shifts in understanding of social reality). Edwards *et al* describe a case study in which several drug use ontologies were constructed, some representing the use of concepts in real-world applications (so-called *in vivo* concepts), and some reflecting top-down classification knowledge of the same domain. These 'organising concepts' are used to create mappings between the *in vivo* concepts; for example:

```
      [in vivo]: DopeSmoking
isatypeof [organising]:CannabisUse
```

Within PolicyGrid we are adopting a different approach which supports dynamic, community-driven evolution of metadata (Guy & Tonkin, 2006) within a framework provided by a series of utility ontologies. It has recently been argued (Gruber, 2006) that the Semantic Web should act as a 'substrate for collective intelligence' - in other words that the community-driven approach to creation and management of content now increasingly popular on the Web should be integrated with the Semantic Web. Our approach is similar in form to Gruber's suggestion of integrating unstructured user contributions (tags) into a structured framework (ontology); we supply every ontology property with its own folksonomy (i.e. collection of user tags). Use of folksonomies in this way stimulates the emergence of a community set of tags (Guy & Tonkin, 2006), prompting the user to use the same values as other users, or to adopt a similar style. We believe that this provides social science researchers interested in the Grid with a flexible and open-ended means of describing resources, whilst at the same time providing a context for those assertions through more structured concepts.

3.2 Resource Metadata

The resource metadata framework uses an OWL-lite ontology derived from the UK Social Science Data Archive[4] schema (itself based upon the Data Documentation Initiative - an international effort to establish a standard for technical documentation describing social science data). The ontology[5] defines a number of concepts including `document`, `questionnaire`, `dataset` and a range of object and datatype properties. Permitted values for many of the datatype properties are of type 'string' and it is here that users are permitted to enter tags; as users describe their resources, an underlying folksonomy is constructed which can be used to guide others towards popular tag choices.

3.3 Provenance for Social Science

We are developing an evidence-based model of provenance which will go beyond the service-oriented approach of the PASOA (Provenance Aware Service Oriented Architecture) project (Groth *et al.* 2006) as it will also include human-centred activities. Within PolicyGrid we define provenance as the information recorded about the association between two or more digital resources; in other words it describes *how* a resource was produced.

There are three types of association in our framework: *Type1* - associations representing provenance-aware services which will use PASOA. *Type2* are the associations representing the use of provenance-aware software by a user; such software will generate provenance metadata as the user is working. *Type3* associations represent human oriented activities. This type of provenance metadata will have to be almost entirely supplied by the user.

The provenance architecture comprises a generic core of associations which users from any discipline can use; these include reference, feedback and versioning associations. There are also domain specific associations which the user can plug into the architecture as required. In this way a user could use the associations they need from different disciplines such as social simulation or qualitative analysis.

3.4 Supporting Argumentation

To facilitate argumentation within policy appraisal we are developing tools which will aid policy stakeholders (researchers, policy makers, others) in using evidence to construct arguments for and against policies. These arguments will be incorporated into an argumentation framework (Dung, 1995) which consists of a set of arguments and the relations between them. The argumentation framework will allow policy makers to explore a system of conflicting arguments (derived from conflicting evidence) and determine how particular arguments are attacked and defeated (or not) by other arguments.

4. Facilitating Metadata Creation

Our approach to evidence management requires that the various artefacts associated with a policy assessment exercise have metadata associated with them. Unfortunately, not all metadata can be generated automatically; some types can only be created by the user scientist. A tool is therefore needed that facilitates easy creation of RDF metadata by non-experts, to enable researchers to deposit and describe their own social science resources.

Existing tools are often graphical (Handschuh, Staab & Maedche, 2001). Natural language approaches include GINO (Bernstein & Kaufmann, 2006), an ontology editor with an approach reminiscent of natural language menus (Tennant *et al.* 1983), and controlled languages such as PENG-D (Schwitter & Tilbrook, 2004). We believe that, for most social scientists, natural language is the best medium to use, as the way they conduct their research and the structure of their documents and data indicate that they are more oriented towards text than graphics. We also require a tool that is open-ended and flexible. Natural language applications are often domain specific and not very flexible. This makes the open-endedness we need a great challenge. Existing elicitation approaches, such as using controlled languages, restrict in great measure what the user can and cannot say. We believe that to achieve the desired open-endedness and flexibility, the best approach is not based on natural language processing, as it is as yet beyond the state of the art to reliably parse all user utterances, but based on natural language generation. We have chosen an approach (Power, Scott & Evans,

---

[4] http://www.data-archive.ac.uk/

[5] http://www.csd.abdn.ac.uk/policygrid/ontologies/UKDA/UKDA.owl

1998) in which the user can specify information by editing a feedback text that is generated by the system, based on a semantic representation of the information that the user has already specified. As the text is generated by the system and does not have to be parsed, we do not have to restrict what can and cannot be said, so the language can retain its expressivity and the user does not need to learn what is acceptable as input. The system is guided by an underlying data-structure, in our case a lightweight ontology plus a series of supporting folksonomies used to suggest feasible tags to influence user behaviour, without restricting the user to a pre-defined set of concepts.

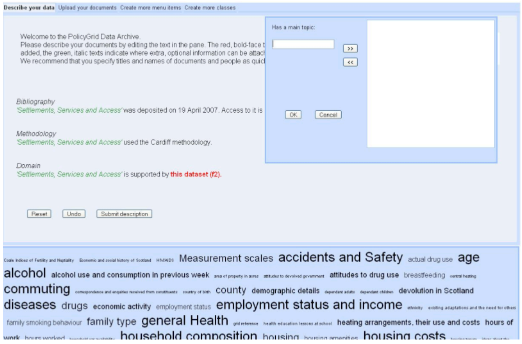

Figure 2 : Natural language metadata interface, including tag-cloud.

Figure 2 shows a feedback text (generated by the current system) for the APAT scenario. We are building a tool that elicits metadata from the user, by presenting them with a text containing an expansion point (anchor) for each object that is mentioned, which has a menu with possible properties associated with that object. These objects and properties are defined by an underlying OWL-lite ontology (e.g. the resource metadata ontology mentioned above). Each 'string' datatype within the ontology has an associated folksonomy, which is used to generate a tag-cloud that is shown to the user when he/she has to enter a value for that property (see Figure 2); the tag cloud shows the tags in the folksonomy and the frequency with which they have been used (reflected in the relative font size). This (in part) protects the system from mistakes such as spelling errors, increases the chance that users use the same terms instead of synonyms, and, when queried, increases the likelihood of a search term being associated with more than one resource. However, the user still has complete freedom, as he/she does not have to use the folksonomy values but can still use free text; and every entry the user makes is immediately added to the folksonomy. The folksonomy, then, allows us to subtly guide user behaviour, while being completely unrestrictive.

The current system consists of five components: the semantic graph, the ontology reader, the RDF-creator, the natural language generator (consisting of a text planner and a surface realiser) and the interface.

The **interface** shows the feedback text with anchors indicating expansion points, which contain menus with types of information that can be added. Google Web Toolkit[6] was used to create the prototype interface.

The **semantic graph** stores the information the user is adding. Each node corresponds to an ontology class, each edge to an ontology property. Initially a generic graph is created, so an initial feedback text can be produced; the graph is updated each time the user adds information (see Figure 3).

The **ontology reader** creates a model of a given OWL-lite ontology, which is consulted throughout the building of the semantic graph and extended with all new properties or objects that the user adds. The ontology specifies the range and domain of the properties; i.e. the properties in each anchor menu, and the (type(s) of) resource that can be selected or added as the range of a selected menu item.

The semantic graph is translated to a list of RDF triples by the **RDF-creator**. These triples are stored, with the relevant resource(s), in a Grid-enabled repository of social science resources.

The natural **language generator** maps the semantic graph to (HTML) text that contains anchors. The text is divided in paragraphs to give a clear overview. To keep the text concise some sentences are aggregated (combining two sentences into one, e.g. 'Mary sees Bill' and 'Claire sees Bill' become 'Mary and Claire see Bill') and where possible pronouns (he, she, it) are used.

Although our system is driven by an ontology, we have kept this lightweight (OWL-lite, using only domain and range of properties and cardinality restrictions), and plan to give the user the power to adapt this ontology to his/her own needs, or even substitute another ontology entirely. Such new ontology resources should be well-formed and clear about which objects are permitted in the domain and range of properties.

---

[6] http://code.google.com/webtoolkit/

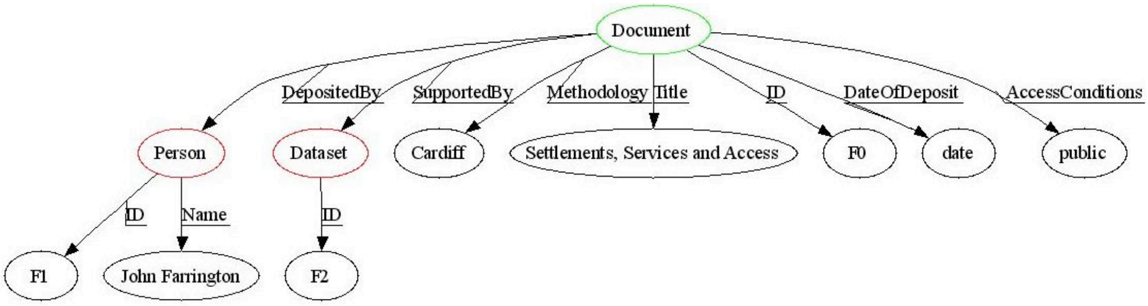

Figure 3 : Semantic graph underlying the feedback text shown in Figure 2.

Our challenge is that they also need a lexical specification that dictates how they should be mapped to natural language. In our system, classes are represented by a noun phrase (e.g. 'this person', 'Thomas'), which can easily be supplied by the user. However, properties are represented by natural language sentences, and therefore need linguistic trees that represent such sentences. These trees are complex structures, containing a wealth of linguistic information and specifying where the source and target of the property should be inserted. A straightforward way to obtain these linguistic specifications is to let a system administrator create them when needed. However, this would cause considerable delays for the user and would appear almost as restrictive as not allowing new property creation at all. Instead, we are enabling the system to create these linguistic specifications immediately, so the user can use the new property during the same session. Using the property name that the user provides, the system generates possible sentences, based on the common sentence types it has stored in linguistic tree form; the user, who knows what the natural language representation should be, chooses a suitable sentence from the options offered and perfects it by changing individual words, verb tense, adding adjectives, etc. The resulting lexical specification is then stored with the new ontology resource.

5. Integrating Desktop Qualitative Analysis & The Grid

As highlighted above, qualitative analysis (of interview transcripts, focus group discussions, etc.) is a key tool within policy evaluation and researchers already make use of software tools to support such activities, e.g. NVivo[7]. We therefore chose to develop a desktop client that would allow social science researchers to conduct qualitative analysis in an environment of Grid enabled resources. *Squanto* (see Figure 4) supports qualitative analysis, while stimulating collaboration and data sharing. It incorporates the basic functionality of existing tools, i.e. the mark up of documents with annotations, memos and codes. However, *Squanto* encourages the user to code their text with terms from ontologies, referred to as 'structured codes'. These can be shared with other users and are used by the system to automatically locate (and present to the user) resources tagged with the same codes.

In keeping with the grounded approach of coding in social science, *Squanto* also enables the user to create and use 'free' codes, which are text tags; these free codes afford the users all the freedom they want. Within *Squanto* we allow the user to create relations between free and structured codes: from the detailed 'subclass of' to the generic 'is related to'. When such a relation is created, the system uses the corresponding structured code to find related resources for the free code. In this manner, a document coded from a grounded approach, entirely with free codes, can still be connected to, and made accessible for, the wider community.

Squanto is designed to make use of available Semantic Grid services to perform collaboration-oriented tasks, such as:

- locate related resources;
- upload and download documents;
- share annotations, memos and codes;
- load and display a shared ontology.

6. Discussion

To support semantic e-Social Science we are developing an evidence-based framework to manage the heterogeneous resources employed during policy appraisal; a hybrid metadata approach combining both lightweight ontologies and folksonomies forms a part of this framework, in order to give social scientists the

[7] http://www.qsrinternational.com/

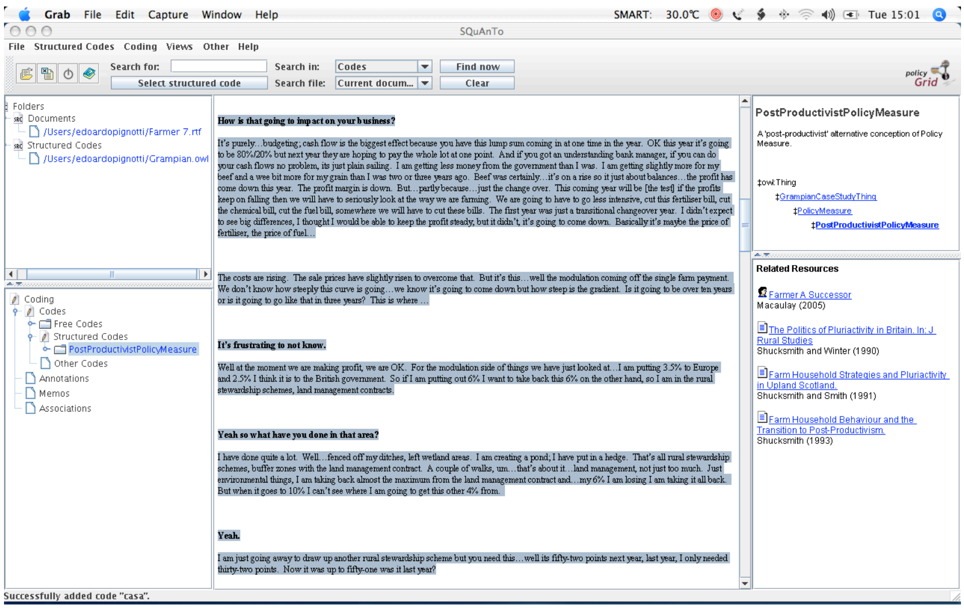

Figure 4 : *Squanto* Qualitative Analysis Tool.

open-ended mechanism they require to annotate resources.

The use of tags helps to fill the acceptance/participation gap by allowing individual researchers to describe resources as they deem appropriate, without being constrained by a centrally imposed conceptualisation. To allow users to deposit their resources on the Grid we have developed a tool for metadata elicitation which stores descriptions using RDF triples; the user has freedom to add values for datatype properties, and is supported by a folksonomy that suggests suitable and popular tags for each property. Although the current version of the tool is driven by one (lightweight) ontology we are exploring ways to enable the user to extend this ontology. As the process used to generate evidence is as important as the evidence itself, we are also embedding support for provenance within our approach. Our abstract provenance model can be instantiated in different ways depending upon the stage in a policy assessment exercise being recorded and whether the activity was performed by a computational service or a human researcher.

At present our hybrid ontology-folksonomy approach is basic, with many outstanding issues still to be resolved. To date we have employed a standard utility ontology (derived from the DDI standard); this ontology does not attempt to represent concepts which might be imprecise, mutable or contested – as it simply defines standard documentation concepts which then act as containers for tags. We still need to determine what other utility ontologies might be appropriate (and acceptable) for use in semantic e-Social science. In our current implementation, tags are simply string tokens with associated frequency counts. What properties (if any) should a tag possess? Should relationships between tags be supported? Can (should?) tags "evolve" to become ontology concepts?

We intend to use the same approach for the querying and information presentation tools as for metadata elicitation. The user should be able to construct a short text describing the type of resource he/she is looking for, which the system should then translate to a SPARQL-query; the answer could be presented in another feedback text in which the anchors would cause related information to be added. We think this approach will be suitable for all three tasks; it also means the user only has to learn to work with one interface which improves usability.

To integrate the various components of our evidence management framework for policy assessment (and potentially other social science research tasks) we are constructing a community Web portal (*ourSpaces*[8]) which will provide the following functionality: submission of resources, searchable archive(s), enhanced collaboration support, integration with client (desktop) tools such as *Squanto*. We intend to let users annotate each other's resources and share those annotations (in the manner of Connotea). This approach would especially suit quantitative social science data sets, as they are

8 http://www.ourspaces.net/

frequently used for secondary research, and are re-analysed from different perspectives. Users will be able to associate existing resources with new ones, add to an existing description of a resource if they feel the description is missing valuable information, mark resources from one discipline as relevant to another discipline, etc.

Developing Semantic Grid solutions for social science presents many challenges, some of which we have outlined in this paper. Based upon our experiences to date we are convinced that a solution integrating Grid services with ontologies and community-driven folksonomies is appropriate and will meet the needs of researchers (and others) interested in evidence-based policy development and assessment.

References

Aktas, M. S.; Pierce, M.; and Fox, G. C. 2004. Designing Ontologies and Distributed Resource Discovery Services for an Earthquake Simulation Grid. In Proceedings of GGF-11 Global Grid Forum Semantic Grid Applications Workshop.

Berners-Lee, T.; Hendler, J.; and Lassila, O. 2001. The Semantic Web. In Scientific American.

Bernstein, A., and Kaufmann, E. 2006. Gino - a guided input natural language ontology editor. In International Semantic Web Conference 2006, 144–157.

Bullock, H.; Mountford, J.; and Stanley, R. 2001. Better Policy-Making Centre for Management and Policy Studies. Technical report, Cabinet Office.

Cabinet Office, 2003. The Magenta Book: A Guide to Policy Evaluation, Cabinet Office, London, Government Chief Social Researcher's Office. http://www.policyhub.gov.uk.

De Roure, D.; Jennings, N.; and Shadbolt, N. 2005. The Semantic Grid: Past, Present, and Future. Proceedings of the IEEE 93(3):669–681.

Dung, P. M. 1995. On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games. Artificial Intelligence 77(2):321–358.

Edwards, P.; Aldridge, J.; and Clarke, K. 2006. A Tree full of Leaves: Description Logic and Data Documentation. In Proceedings of the Second International Conference on e-Social Science.

Farrington, J.; Shaw, J.; Leedal, M.; Maclean, M.; Halden, D.; Richardson, T.; and Bristow, G. 2004. Settlement, Services and Access: The Development of Policies to Promote Accessibility in Rural Areas in Great Britain. H.M. Treasury, The Countryside Agency, Scottish Executive, Welsh Assembly Government.

Groth, P.; Jiang, S.; Miles, S.; Munroe, S.; Tan, V.; Tsasakou, S.; and Moreau, L. 2006. An Architecture for Provenance Systems. ECS, University of Southampton.

Gruber, T. 2006. Where the Social Web Meets the Semantic Web. In Cruz, I.; Decker, S.; Allemang, D.; Preist, C.; Schwabe, D.; Mika, P.; Uschold, M.; and Aroyo, L., eds., Proceedings of the 5th International Semantic Web Conference, ISWC 2006, volume 4273, 994. Athens, GA, USA: Springer Lecture Notes in Computer Science.

Guy, M., and Tonkin, E. 2006. Folksonomies: Tidying up Tags? D-Lib Magazine 12(1).

Handschuh, S.; Staab, S.; and Maedche, A. 2001. Cream: creating relational metadata with a component-based, ontology-driven annotation framework. In K-CAP '01: Proceedings of the 1st international conference on Knowledge capture, 76–83. New York, NY, USA: ACM Press.

HM Treasury, 2003. The Green Book: A Guide to Appraisal and Evaluation, London, HM Treasury.

Power, R.; Scott, D.; and Evans, R. 1998. What You See Is What You Meant: direct knowledge editing with natural language feedback. In Proceedings of the Thirteenth European Conference on Artificial Intelligence.

Schwitter, R., and Tilbrook, M. 2004. Controlled natural language meets the semantic web. In Proceedings of the Australasian Language Technology Workshop 2004.

Stevens, R.; Robinson, A.; and Goble, C. 2003. my-Grid: Personalised Bioinformatics on the Information Grid. Bioinformatics 19(1):302–304.

Taylor, K.; Essex, J. W.; Frey, J. G.; Mills, H. R.; Hughes, G.; and Zaluska, E. J. 2006. The Semantic Grid and Chemistry: Experiences with CombeChem. Journal of Web Semantics 4(2):84–101.

Tennant, H.; Ross, K.; Saenz, R.; Thompson, C.; and Miller, J. 1983. Menu-based Natural Language Understanding. In Proceedings of the Twenty-first Annual Meetings on Association for Computational Linguistics, 151–158.