Linked Data for Cross-disciplinary Collaboration Cohort Discovery.

Trina Myers[a, b], Jarrod Trevathan[c], Dianna Madden[b], Tristan O'Neill[a, b]

[a] Discipline of IT, School of Business, James Cook University, Townsville, Australia.
[b] e-Research Centre, James Cook University, Townsville, Australia.
trina.myers@jcu.edu.au
[c] School of ICT, Griffith University, Brisbane, Australia.

Abstract. Cross-disciplinary collaborations potentially offer the diversity of understanding required to answer complex problems. However, barriers to cohort discovery exist because content about people is predominantly only in human-readable form on websites and/or in disparate databases. Notably, many cross-disciplinary collaborations never form due to a lack of awareness of cross-boundary synergies. This project applies semantic technologies to automate linkages to reveal hidden connections between people from metadata parameters about data, rather than from publication products. The information in metadata, commonly used for data discovery, can be used to link researchers for potential partnerships. The proposed system combines pre-existing and custom ontologies, populated from a number of accessible repositories, to describe the relationships between researchers based on metadata parameters. The system was tested from the researcher's perspective where significant alignments with potential partners were found based on transitive relationships, similar interests (e.g., research fields) and/or other commonalities (e.g., location/time of research).

Keywords: Semantic Web, ontologies, collaborative research, knowledge systems

1 Introduction

Collaboration is an essential ingredient in modern research efforts directed at answering complex problems. For this reason, funding bodies and policymakers encourage and support two or more disciplines working together. Collaborations have the potential to bridge disciplines, and apply the rich perspectives, diversity of understanding and collective intelligence required to solve the significant issues of our times [1], [2].

The barriers to discovery of both data sets and related research or researchers are acknowledged by the academic publishers who are working towards various solutions. While these collaborator discovery systems are useful, they are neither comprehensive nor widely adopted and focus on linkage through publication authorship only. There are now alternative ways to link people (cohorts) with the open data movement [3] and the standardisation of scientific data citation [4]. For the purpose of this paper, a cohort is a group of subjects that share an event during a time span, a common geographical location and/or other common denominators.

This paper focuses on the challenges of cohort discovery as these challenges possibly reduce the number of emergent collaborations. In fact, many relationships never form because researchers are not aware of the cross-boundary synergies their research may have with other researchers [1], [5]. Software systems can be used to automate data discoveries but barriers exist where content about people are only in human readable form on the Web or stored in disparate data silos [6]. The process of revealing hidden connections between data, people and processes can be automated via semantic technologies and information embedded in metadata for use in the discovery of potential collaborative partners (i.e., people and/or organisations).

Metadata repositories such as *Research Data Australia* (RDA) [7] and data repositories such as the *Tropical Data Hub* (TDH) [8] contain information within metadata to not only link data sources but also the researchers involved. Researchers' profiles can be accurately matched with the intelligent integration of attributed research data and metadata. The use of open linked data formats can therefore be used to identify potential collaboration opportunities between parties that would otherwise not have awareness of each other's existence.

This project aimed to explore the value of open data for cohort discovery. A *Semantic Knowledge Base* (SKB) was created that automates linkages between internal and external data and metadata sources to align researchers with potential partners from inter-disciplinary and/or cross-disciplinary intersections (Figure 1). The SKBs functionality works as a semantic layer on the TDH. Here, we have focused on the use of technology to enable partner discovery by automatically inferring sets of researchers based on the metadata parameters of their data. The linkages automatically reveal hidden connections between related people, data and processes through transi-

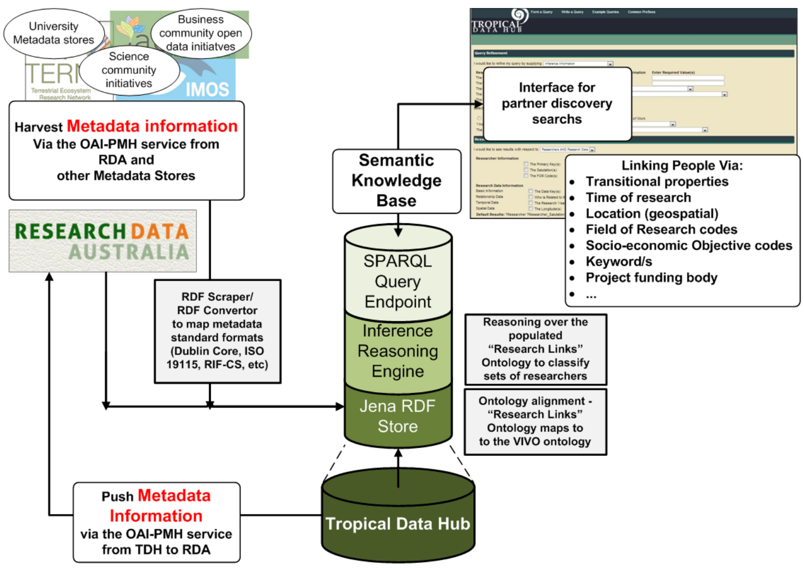

Fig. 1. – The Semantic Knowledge Base workflow for partner discovery

tive relationships, similar fields of study and interests and/or with diverse interests but other underlying commonalties (i.e., methods, time or location of research, etc). The initial results show cohort discovery across disciplines can be enabled with automated inference over metadata attributes pertaining to data and project information.

This paper is structured as follows: Section 2 provides background details and related work. Section 3 outlines the development of the SKB ontologies and the user interface to view inference outputs. Section 4 provides a summary of the results and section 5 concludes with a discussion of outcomes and implications for the future.

2 Background

There is a need for new approaches to partner discovery to assist in coordinating institutional or team-based collaborations. Although there are many factors that contribute to a successful collaboration, for example, formation, size and duration, organisation bureaucracy, technological practices, and participant experiences, this study explores a potential technique for the initial cohort discovery phase of team formation.

Academic publishers such as Thomson Reuters ISI researcherID [9], Elsevier's SciVal Experts [10], CiteSeerX's CollabSeer [11], are related initiatives that work towards various solutions in partner discovery.

Thomson Reuters ISI researcherID [9] provides a solution to the author ambiguity problem by assigning a unique identifier to enable researchers to manage their publication lists, track citations, identify potential collaborators and avoid author misidentification. Researchers can search the registry to find collaborators based on publication lists in the Web of Knowledge.

Elsevier's SciVal Experts [10], is an expertise profiling and research networking tool that simplifies the process of finding experts for collaboration within their institution and across organizations. Similar to researcherID, SciVal Experts creates a directory of research expertise using information found in publication lists and individual researcher updates. SciVal Experts applies semantic technologies by generating "Fingerprints" at the researcher and department level, and links researchers across common concepts and expertise to find connections among authors.

CollabSeer [11] is based on, and draws solely from, the CiteSeerX scientific literature digital library to build a co-author network. CollabSeer discovers potential collaborators by analysing the structure of a co-author network and the user's research interests and analyses the network structure with similarity algorithms.

In contrast to related works, this project aims to infer connections among researchers based on information extracted from metadata standards such as ISO 19115, Darwin Core, etc., instead of forming links from publication lists. The open data movement and the publication of research data offers alternative ways to link people to find potential partners rather than only connections found in authorship. The above systems focus on linkage through information from printed publications only. Open data initiatives, such as the *Australian National Data Service* (ANDS), aim to enable more researchers to re-use research data. Data citation refers to the practice of providing a reference to research data in the same way as researchers routinely provide a bibliographic reference to printed resources. Further, this project links information sourced from metadata stored in disparate Web available repositories from unrelated institutions, as opposed to singular or proprietary source boundaries.

2.1 The Semantic Web technologies

Software can use the contextual definitions written in ontologies to find hidden links, answer questions and infer conclusions from open data currently available on the Web [12]. Herein, a hierarchy of pre-existing and newly developed domain-specific ontologies has been combined to describe to a computer the relationship between researchers and their research data. The ontologies form a flexible, dynamic SKB where disparate data from a diverse range of sources are ingested so reasoning and inferences can be applied to extract latent connections among the researchers within.

The Vivo Ontology and the Research Links Ontology are the foundation of the SKB. VIVO is an extensive research-focused semantic application that manages an ontology populated with linked data representing scholarly activity [13]. It is a discovery tool to browse or search information on people, departments, courses, grants, and publications and enables collaboration among researchers across all disciplines.

VIVO can help institutions highlight researcher expertise and enable collaboration [13]. When installed and populated with researcher interests, activities, and accomplishments, it enables the discovery of research and scholarship across disciplines at that institution and beyond. Organization's data is brought into VIVO in automated ways from local systems of record, such as HR, grants, course, and faculty activity databases, or from database providers such as publication aggregators and funding agencies. Applications can then read organization's data and share researchers' profile data, which is in semantic-web compliant format.

ANDS have extended the VIVO ontology to capture ANDS-compliant descriptions of research data sets and create a metadata store solution on VIVO. The enhancements (called the ANDS VITRO ontology) are being built as a community initiative involving several Australian universities [14]. Although there are other suitable ontologies available that describe activities in research, such as the AKT reference ontology [15], the ANDS VITRO ontology was chosen as it is currently implemented in Australian institutional data repositories as part of the ANDS initiative. Ontology alignment was less complex by using the ANDS VITRO and importing it to the higher-level domain-specific "Research Links" because the need to declare equivalencies between ANDS VITRO and another research-domain vocabulary was eliminated.

As a component of this project, the domain-specific "Research Links" ontology was created to describe the relationships between researchers based on metadata elements, such as location and time of data collection. The Research Links ontology aligns to the ANDS-VITRO ontology to enlist the pre-existing descriptions of researchers, affiliations and projects. The Research Links ontology extends the ANDS-VITRO ontology with classes for transitive relationships, specific location, time, keywords, *Field of Research* (FoR) codes, and *Socio-economic Objective* (SEO) codes written in OWL-DL so reasoning is possible [12].

2.2 Metadata resources

The Tropical Data Hub (TDH) is a knowledge management platform that provides a data-hosting infrastructure to store, aggregate and serve significant tropical data sets from a single virtual location [8]. It provides researchers, managers and decision-makers access to key national and international research data from disparate data sources for a more accurate holistic view of the current state of the tropics. Currently the TDH has implemented functionality for data deposition and retrieval as well as metadata creation via a web portal. A key function of the portal is the amalgamation of data exposed for harvesting metadata so searching across data sources is possible.

A problem currently exists where the data and metadata available within the TDH are stored in a repository that does not enable semantic linkages. The current metadata description is in a HTML format, which does not allow for intuitive searching without embedded information to add context to the metadata fields. Current metadata standards focus predominantly on gathering information about data for human readable presentation that makes mapping between the multitudes of metadata standards non-trivial.

The aim of this project was to incorporate semantic technologies within the TDH to provide data integration and knowledge discovery across the "vertical" data silos (Figure 1). Then, semantic technologies will link different terminologies to bridge across the data stored in the TDH to other repositories and incorporate these linkages to the Linked Data cloud [16]. Further, semantic correlation and inference capability can merge data, metadata and infer linkages between users.

Research Data Australia (RDA), the flagship service of the ANDS, is a metadata repository that provides access to the Australian research data commons [7]. It is an Internet-based discovery service designed to provide connections between data, projects, researchers and institutions, and promote visibility of Australian research data collections in search engines.

Many of Australia's data repositories, such as the TDH, feed into the RDA. The *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) [17] is a low-barrier mechanism for repository interoperability and is initialised to perform this ingestion of metadata from diverse sources. The *Online Research Collections Australia* (ORCA) assessment workflow allows ANDS to incorporate a level of quality assessment and approval within the record publishing process (both manual and harvested). The ORCA-Registry is a PHP/PostgreSQL software utility that enables import, entry and management of collection metadata [18]. The ORCA-Registry is designed to be housed in an instance of the *Collection Services Infrastructure* (COSI)-Framework, which stores information about roles, activities and authorisations to control access to web application functionality. The ANDS COSI/ORCA package is available to institutions to create a local collections registry. This project initialised an instance of the COSI/ORCA framework within the TDH.

3 Approach

The approach involved three stages: (1) The development of the ontologies, (2) the inclusion of the semantic layer in the TDH and (3) the development of an interface so the test participants could view inference outputs.

3.1 The ontologies

A hierarchy of ontologies were developed using the *Web Ontology Language* (OWL) [19] that describe the characteristics and information contained in metadata standards. The triplestore links researchers based on the semantic tags derived from metadata standards such as Darwin Core, ISO 19115, etc [20]. The relevant metadata components included the provenance or descriptive details about the actual dataset (e.g., keywords, geospatial details, dates created and published, data formats, etc.), administrative data (e.g., researcher's name, affiliations, institutions, physical location of the dataset, etc.) and process data (methodologies, hardware/software configurations, version information, etc).

The "Researcher Links" ontology is a task-specific heavyweight ontology written in OWL-DL [19] that defines axioms to describe the relationships between the researcher and the data that could link researchers. For example, sets of "like" individuals are linked based on the location of the research data, the time the data was collected or fields of study. Classes were created so the reasoning engine could subsume individual researchers to concurrently belong to a time-period, a location and/or a general field of study (i.e., *Science Technology, Engineering and Medicine* (STEM) and *Health Arts and Social Science* (HASS), based on information about data collection. The Research Links ontology subscribes to the ANDS VITRO ontology, which represents scholarly activity (Figure 2).

This project was designed with the intent on mapping metadata information to the SKB Researcher Links ontology from a variety of harvested data sources. Metadata from repositories using the VIVO ontology and/or the *Registry Interchange Format -*

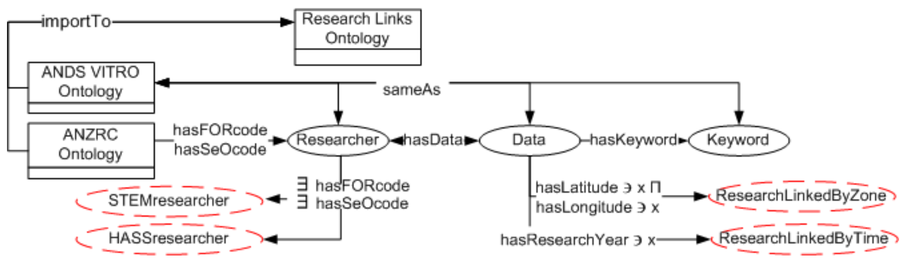

Fig. 2. – Extract of the Research Links ontology to show anonymous classes as subsumed by property restrictions.

Collections and Services (RIF-CS) metadata standard, which describes data collections, are mapped to the Researcher Links ontology. The most relevant elements of the Researcher Links ontology, shown in Figure 2, consist of the following:

- A class structure of geospatial locations by zone where membership is inferred by the longitude and latitude coordinates within the metadata;
- A class structure of temporal location where membership is inferred by year of data collection based on metadata start and end dates;
- Subscription to the *Australian and New Zealand Research Classification* (ANZRC) FoR and SEO codes ontology, which was developed as part of this project. It is written in OWL for import into the Researcher Links ontology for reasoning[1];
- STEM and HASS classes where membership is based on FoR codes;
- Data description;
- Researcher details;
- Project details, which are used to link researchers through transitive relationships based on links with projects; and
- Keywords.

3.2 The triplestore implementation

The aim was to populate the SKB with data available within the TDH combined with researcher information from external sources because the TDH is only concerned with tropic zone data collections. The metadata information was harvested from the TDH and external repositories including RDA, CSIRO, Open Data, university metadata repositories, etc. via OAI-PMH (Figure 1). The information was converted to triples and ingested to the Jena triplestore (TDB) where it was used to populate the "Researcher Links" ontology; the highest level of the SKB. The implementation described here required reasoning over OWL-DL and queries over RDF so the Pellet reasoning engine [21] and the Fuseki SPARQL server [22] were invoked.

Web Scraping - Data was extracted from the RDA COSI/ORCA site by web scraping the records that were displayed through the available OpenSearch functionality. Harvesting from a single source yields a significant number of records, which proved a primary issue. When these records exceed 32,000, PHP (on the server side) was not capable of processing the request. To extract records from individual data sources, the method of harvesting on the TDH COSI/ORCA site was altered so each data source did not exceed this limitation. A custom-built RDF convertor was created in Java to harvest the records for each data source and parse them into RDF triples based on the given XML tags. As each data source was parsed, a flat file was constructed consisting of these triples. The tags map to the pre-defined ontology framework and are ingested directly into the Triplestore to populate the Research Links ontology.

The ontologies are populated by performing two harvests to manage duplicate entities. The first harvest extracts unique researchers and metadata that exist within the

[1] http://mmisw.org/ont/ANZRC/ANZRC_Codes

given source and generates a Java object for each. The second harvest appends the main aggregate RDF file with the objects generated in the first harvest, which populate the ontologies. The *Permanent Identifier* (PID) or repository identifier that is given to each dataset and researcher from the originating source is maintained. If a unique identifier does not exist, one is created during the first harvest. The identifiers are used in the rdf:about tag so on ingestion into the triplestore, duplicate triples are discarded but researchers that have the same name are maintained as unique. However, a researcher may be given more than one identifier if they had datasets in multiple repository sources, which can cause some inconsistency. To counter this, if there are associated triples (e.g., project name, dataset name, etc) that link two different identifiers as the same researcher, an equivalency is generated.

The Zone/Region class structure is made up of six global geographical zones (i.e., north and south torrid, north and south temperate and north and south frigid zones), which entails 65,341 coordinate pairs on a Cartesian map. These zones are represented as classes in the Research Links ontology, each with region subclasses. Each region covers an approximate area of 111km^2 across the surface of the earth. Class membership is inferred based on the hasLatitude and hasLongitude properties present in the data instances. Data instances can be inferred to multiple regions per pair of matching longitude and latitude.

The Web interface to the SKB was created as a proof of concept to test the system by allowing individual researchers to find potential cohort partners. The interface, shown in Figure 1, connects to a SPARQL endpoint (via Fuseki) so the current state of the Triplestore can be queried. The interface is designed to be researcher-centric rather than metadata-parameter centric where the system will compare all other instances in the SKB with the variables specified from an individual researcher's metadata. Therefore, an individual researcher can search for partners based on specific metadata characteristics from their own data (e.g., the location of their research). Once the reasoning is complete and instances are subsumed to appropriate classes, the connections can be retrieved through predefined searches, which are based on specific metadata parameters (i.e., similar research fields, location, keywords, etc.), customised searches, where the user can choose the variables via check-boxes, or manual searches where the user can create their own queries.

4 Results - Proof of concept

The framework maps the harvested metadata provenance components (e.g., geospatial location, creation date, keywords etc.) and user profile components (e.g., affiliations, publications, etc.) to the ontologies within the SKB so links between the data metadata can be established through reasoning. After reasoning, researcher instances are subsumed to belong to the classes (sets) defined in the Researcher Links ontology. For example, Researcher A, a member of the "Researcher" class, conducted research in 2005 and 2009 in the Daintree Rainforest, North Queensland. The reasoner would classify Researcher A to belong to the "ResearchRegion_38592" class, the "SouthernTropicsResearchZone" class (the super-class of the specific research regions), the "ResearchYear" classes (2005 and 2009) and the "ResearchDecade_2000s" class (the super-class of the specific research years) concurrently. SPARQL queries were also run to find specific sets of researchers. These sets, for example, were based on transitive relationships or commonalities in location or time of data collection, FoR/SEO codes and/or keywords.

The SKB was populated from 55 different metadata sources that extracted information on 3,287 researchers and 1,819 data sets (Table 1). Table 2 show the statistics of the properties populated from the metadata of each imported dataset. The mix of researchers showed individuals from both HASS and STEM disciplines derived from the FoR codes in the metadata.

Table 1. –Instance statistics of the data automatically ingested to the knowledge base used for testing cohort linkage outcomes.

Instance and property description	Instance data
Data Sources	55
Data Individuals	1819
Researcher Individuals	3287

Table 2. –Property statistics of the data automatically ingested to the knowledge base used for testing cohort linkage outcomes.

Property description	Domain	Range	Instance data
isDataOf	Data	Researcher	2165
isDataAssociatedWith	Data	Data	166
hasLongitude	Data		2793
hasLatitude	Data		2563
hasResearchYear	Data		4392
usesFoRcode	Researcher	ANZSRCcodes:FOR_CODE	741
usesSEOcode	Researcher	ANZSRCcodes:SEO_CODE	180
hasData	Researcher	Data	2165

Connections between researchers based on the metadata parameters can trigger obscure research correlations and possibly invoke new hypotheses or lines of enquiry. The parameters available for the identification of comparable researchers or data through inference and search mechanisms include the following:

- Those that have transitive relationships based on affiliation with data collection or projects (i.e., researcher A worked with researcher B, researcher B worked with researcher C, therefore A => C, these relationships extend to researcher *n*);
- Those in a specific field of expertise
  - Researchers Linked By SEO code
  - Researchers Linked By FoR code

- Those with a desired affiliation or agenda to benefit the research; and/or
  - Those with provenance commonalities between data collection characteristics, which include research linked by: Year, decade, region, zone and project and researchers linked by project and location.

Five *early career researchers* (ECR) were chosen to test the outcome of the SKB. Of these five ECRs, three were from disciplines of STEM and two were from disciplines of HASS. For each subject, the outcome of the searches had commonalities with a large range of other researchers in time and/or location and transitive relationships. The transitive relationships proved interesting where discontiguous associations with others were exposed via "friend of a friend" characteristics to link individual researchers.

Table 3 shows the results from one ECR (subject A) whose research is within a STEM discipline (i.e., Earth Sciences). Subject A's research involves the sustainability of a marine species whose refugia is predominantly inshore and adversely effected by anthropogenic activities. This subject's data collections span from 1985 to 2007, which resulted in a large catchment of other data collections that were conducted at the same time.

Table 3. Results from an ECR in STEM discipline Biological Science (FoR category 06)

Linkage parameters for subject A	TDH data repository	All 55 sources
Linked by transitive relationships	2	7
Links to different projects	67	785
Links to researchers potentially unknown to subject	35	340
Same FoR code (Projects)	58	72
Different FoR code (Projects)	9	713
Linked to STEM Projects	63	698
Linked to HASS Projects	4	87
Linked to STEM Researchers	31	211
Linked to HASS Researchers	6	131
Linked by Year	37	667
Linked by location - Southern Tropic Zone	37	342
Linked by location - North QLD Region	32	32

The outputs from the system were analysed to determine inference of cohort linkages for Subject A. Here, quantitative data were gathered on the linkages found within all 55 external sources including the TDH repository (Table 3).

The transitive relationships, which link researchers within the same area of research, exposed at least two possible unknown researchers from the TDH repository and seven from all 55 resources (i.e, both TDH and external resources). These researchers were unknown to Subject A as the other studies occurred at different times with the common linking researcher (Researcher B).

In addition, there were linkages with researchers that were not in the same area of research and were unknown to Subject A that showed potential cross-disciplinary collaborations. For example, there was an environmental monitoring project on herbicide use (STEM) and a social study project on how people in the region use water (HASS). Both projects occurred at similar time-periods but were in different locations from Subject A's data collection. These links show promise for further combined cross-discipline studies within a similar location.

5 Conclusion

This project has explored automated collaborative partner discovery through semantic intelligence over data products rather than publication products. Referencing standards for open data are emerging, which allows for the citation of published data not just publication documents. The partner-discovery tool described in this paper is a significant contribution because it explores linking people based on metadata. The proof of concept was trialled within the infrastructure of the TDH.

The SKB architecture developed is an exemplar of the evolving methods for managing rich data sources in new and unique ways. The architecture, which employs semantic inference, includes methods for modularity, reusability and data integration. The system combines a pre-existing ontology (ANDS-VITRO ontology), a custom FoR and SEO code taxonomy (ANZRC codes ontology) and a domain-specific ontology (Research Links ontology) to describe the relationships between researchers and metadata parameters. Reasoning and inference engines were applied to automatically classify researcher entities to find implicit links and make possible the disclosure or extraction of knowledge in data from disparate sources. Sample metadata of scientific data was imported from a diverse range of data repositories to explore the integration potential for dynamic cohort discovery based on information within metadata.

The system aligned researchers with potential partners at inter-disciplinary and/or cross-disciplinary boundaries with embedded contextual information based on metadata characteristics. The semantic linkages disclosed researchers through transitive relationships (i.e., friend of a friend), those from different fields but with other shared commonalties (i.e., complementary research methods, location of research etc.) and/or those with similar interests (i.e., disciplines, keywords, etc.). The outcome of the automated cohort discovery can potentially lead to new and deeper collaboration across the research sector.

This project aimed to explore the cross-disciplinary synergies that could emerge from studies across varied disciplines and locations. To trial the outcome of the system, both homogenous and heterogeneous combinations of metadata were examined to offer the highest probability of generating novel pairings. A full implementation of such a system would benefit from examining all angles of juxtaposition of metadata, as well as interviewing potential partners to see if these pairings were of interest. However, this extension of the project is left as future work along with enhancing capacity through a greater range of geospatial coordinates and performance improvements to minimise harvesting, reasoning and inference process times.

Acknowledgements. The authors wish to thank the Australian National Data Service (ANDS) and the eResearch Centre, James Cook University for their assistance.

References

1. Lee, S., Bozeman, B.: The Impact of Research Collaboration on Scientific Productivity. Social Studies of Science 35, 673-702 (2005).
2. Shrum, W., Genuth, J., Chompalov, I.: Structures of scientific collaboration. The MIT Press, Cambridge, MA, USA (2007)
3. The Linking Open Data Project, www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData
4. Green, T.: We need publishing standards for datasets and data tables, White Paper. OECD Publishing (2009)
5. Haines, V., Godley, J., Hawe, P.: Understanding Interdisciplinary Collaborations as Social Networks. American Journal of Community Psychology 47, 1-11 (2010).
6. Hunter, J., Cole, T., Sanderson, R., Van de Sompel, H., The open annotation collaboration: A data model to support sharing and interoperability of scholarly annotations. In: Digital Humanities 2010, London, United Kingdom, 2010.
7. Research Data Australia, http://researchdata.ands.org.au/
8. Myers, T., Trevathan, J., Atkinson, I.: The Tropical Data Hub: A Virtual Research Environment for tropical science knowledge and discovery. International Journal of Sustainability Education 8, 11-27 (2013).
9. Thompson Reuters ResearcherID, http://www.researcherid.com/
10. Scopus® SciVal Experts, http://info.scival.com/
11. Chen, H.-H., Gou, L., Zhang, X., Giles, C.L., CollabSeer: a search engine for collaboration discovery. In: Proceedings of the 11th annual international ACM/IEEE Joint Conference on Digital Libraries (JCDL2011), Ottawa, Ontario, Canada, 2011.
12. Allemang, D., Hendler, J.: Semantic Web for the working ontologist, 2nd Edition: effective modeling in RDFS and OWL. Morgan Kaufmann, Burlington, MA, USA (2011)
13. Börner, K., Conlon, M., Corson-Rikert, J., Ding, Y.: VIVO: A Semantic Approach to Scholarly Networking and Discovery, Vol. 2. Morgan & Claypool Publishers (2012)
14. Metadata Stores Solutions, http://ands.org.au/guides/metadata-stores-solutions.html#VITRO
15. The AKT Reference Ontology http://www.aktors.org/publications/ontology/
16. Heath, T., Bizer, C.: Linked Data: Evolving the Web into a Global Data Space. In: Hendler, J., vanHarmelen, F. (eds.) Synthesis Lectures on the Semantic Web: Theory and Technology, vol. 1, pp. 1-136. Morgan & Claypool Publishers (2011).
17. Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), http://www.openarchives.org/pmh/
18. ORCA-Registry Software, http://www.globalregistries.org/orca.html
19. OWL: Web Ontology Language overview http://www.w3.org/TR/owl-features/
20. Bikakis, N., Tsinaraki, C., Gioldasis, N., Stavrakantonakis, I., Christodoulakis, S.: The XML and Semantic Web Worlds: Technologies, Interoperability and Integration: A Survey of the State of the Art. In: Anagnostopoulos, I.E., Bieliková, M.r., Mylonas, P., Tsapatsoulis, N. (eds.) Semantic Hyper/Multimedia Adaptation, vol. 418, pp. 319-360. Springer, Berlin Heidelberg (2013).
21. Pellet: the open source OWL DL reasoner, http://clarkparsia.com/pellet
22. SPARQL query language for RDF, http://www.w3.org/TR/rdf-sparql-query/