Abstract

The research project investigated the role of information professionals in geoscience data management. Data management professionals in organisations which generate geoscience data across the government, academic and corporate sectors in Western Australia were surveyed. The results indicate that while information professionals were engaged in managing a range of geoscience data and see opportunities to expand their role, the extent of their involvement and the type of data they manage was impacted by their knowledge base, organisational structures, and stakeholder attitudes toward data management.

Keywords: Information professionals, data management, geoscience, professional development

Introduction

Access to geoscience data is essential for scientific investigation, managing geological hazards, land use evaluation and classification, civil engineering projects, formulating government resource policy and efficient exploitation of natural resources (Cutler & Maples, 2002; Kelly & Phillips, 1986). There is a vast range of geoscience data which requires management, ranging from geological specimens, to non-digital data such as hand drawn maps and field notebooks, through to 3D digital models. A large proportion of modern geoscience data is generated digitally, yet physical data is still produced and access to historical data which predates the digital age remains important.

There is little recent research into the role of information professionals in geoscience data management, and literature relating to the engagement information professionals in the broader field of research data management is focused on digital data, and mainly generated by universities and large scale eScience projects. This study aims to provide insight into the involvement of information professionals in the management of all forms of geoscience data across academic, corporate and government sectors. Through investigating the roles of geoscience information professionals in Western Australia, the study seeks to identify factors which impact on their engagement in geoscience data management, and provide the means to affirm, encourage and promote the work of these professionals for the benefit of all geoscience data stakeholders.

Background and Context

Australia is one of the world’s leading producers of minerals (Grant, Hawkins, & Shaw, 2005) and Western Australia is Australia’s dominant resource state (Western Australia (WA) Department of Mines and Petroleum and Department of State Development, 2011, p. 2). In Australia geoscience data is generated by universities and government agencies, whose data is publicly available, and by corporate entities where access to data is restricted. Statutory reporting requirements ensure that some corporate data is eventually made public, and there is a complex
web of interactions between sectors and agencies engaged in geoscience data generation and management.

Internationally, the creation of the Geoscience Information Society in the United States of America (US) in 1966 saw geoscience information establish a foothold as a separate discipline, gaining international recognition in 1978 with the 1st International Conference on Geological Information in London (Ward & Walker, 1986). The continuing growth of digital data has spurred discussion about how the “data deluge” (Lyon, 2007, p. 5) should be managed, and there is a perception in some sectors that geoscience data management is the responsibility of information technology (IT) professionals (Yacopetti & Mundell, 2010). The focus on digital data has overwhelmed discussion of the management of non-digital geoscience data, and the inter-relationship between different data formats within datasets has been overlooked (Browne & Love, 1997; Lowe, 1995). In the US—and potentially elsewhere—this has culminated in the loss of many valuable collections, with many more identified as in danger of being lost (National Research Council (NRC), 2002).

This study pursues four main areas of inquiry:

1. the geoscience data management environment information professionals operate within, and its influence on their role in geoscience data management;
2. the management activities information professionals are engaged in for the range of geoscience data held by their organisations;
3. the skills and knowledge necessary for geoscience data management, and how information professionals develop them; and
4. the opportunities and challenges identified by information professionals engaged in geoscience data management.

Definitions

There is a wide range of opinion about the definitions of the key terms used in this study—data, data management and information professionals—in part springing from the wide range of professions engaged in the field. In order to clearly set the parameters of this project, the following definitions are applied:

Data are defined as facts, figures and other items which describe or represent an object, idea, condition, or situation, which can then be organized for analysis (NRC, 2002). This project divides geoscience data into three categories:

- geoscience collections: items which originate naturally, including geological samples such as rocks and drill cores;
- non-digital geoscience data: physical items such as hardcopy maps and field notebooks; and
- digital geoscience data: data held in computer files, such as text, tables, maps, and 3D models.

Data management is defined as the range of activities which ensure long-term access to data in a usable form. This includes selection, storage, organisation, preservation and provision of ongoing access to data, and the development and implementation of policies and standards related to these activities (Association of Research Libraries Workshop on New Collaborative Relationships (ARL Workshop), 2006; Lewis, 2010).
The term information professional can embrace a wide range of occupations. This project is interested in the role of information professionals with library and information science (LIS) training, and the definition used is informed by the requirements for professional membership of the Australian Library and Information Association (ALIA) (Australian Library and Information Association (ALIA), 2011). Information professionals are defined as those individuals possessing tertiary qualifications in the fields of library and information science, archival science and/or records management. Correspondingly, the project considers geoscience professionals (referred to as geoscientists) and IT professionals as those with tertiary qualifications in their respective disciplines. Those without tertiary qualifications are referred to as non-professional staff.

**Literature Review**

While the focus of this review is on literature relating to geoscience data management, it is necessary to investigate the broader field of research data management to explore some issues. The (mostly grey) literature of geoscience information stakeholders is mainly task focused, often based on accounts of personal experience, and rarely dwells on the roles of individual professions. In recent times surveys, case studies and reports from academic libraries and scientific research bodies have provided insight into the role of information professionals in research data management.

**Data management environment**

The literature indicates that those engaged in research data management are commonly drawn from three spheres of expertise: domain experts, IT professionals and information professionals (Corrall, 2008; Swan & Brown, 2008). Reports emanating from the US and the United Kingdom (UK) attempt to define the roles of those engaged in research data management—grouping them variously as data creators, data managers, data scientists, and data librarians—yet there is acknowledgement that these roles are not clearly defined and can overlap (National Science Board, 2005; Swan & Brown, 2008). Hyams (2008) and Pryor and Donnelly (2009) each reflect that in practice data management roles vary widely and are influenced by the aims and resources of organisations.

Milne, Thompson and De Vine (2011, Conclusion, para. 3) state that “identifying, describing and making research data more visible is an enormous undertaking which cannot be undertaken by any one professional area in isolation”. To encourage collaboration organisational structures which encourage cooperation between professions have developed in academic institutions such as Purdue University, the University of California San Diego and Queensland University of Technology (Milne et al., 2011; Soehner, Steeves, & Ward, 2010; Witt, 2008).

Outside the academic realm cooperation between professions is not as evident: Yacopetti and Mundell (2010) assert that there is a perception in the resources sector that IT professionals are responsible for data management issues, and Foote’s (2010, p. 63) survey of geological surveys in the US reveals that only one-third employed an information professional—“among geologists, hiring a professional librarian remains an option”. Studies of geoscience collection management in the US (NRC, 2002) and Australia (Simpson, 2003) both identify a shortage of the resources necessary to employ appropriately qualified staff, placing geoscience datasets at risk. Yacopetti and Mundell (2010) call for resource companies to increase investment in data management, and believe people with information management skills are the most important component of such investment.


**Data management activities**

Libraries and information professionals are identified as being familiar with many of the activities associated with scientific data management and Gold (2007) provides a comprehensive list of these. Downstream (post-publication) activities include many tasks which are an extension of those familiar to information professionals: selection, acquisition and licencing of datasets, creating metadata, preservation services, developing data publication standards and intellectual property policies, and repository services (Gold, 2007). Upstream (pre-publication) activities include developing data management frameworks and standards, and require information professionals to “forge embedded working relationships with research teams” (Gold, 2007, sec. 2.4, para. 2).

While much of the academic sector’s discussion of data management roles for information professionals is aspirational, there is evidence that information professionals have a long tradition of undertaking a range of geoscience data management activities. There are detailed accounts of geoscience information professionals developing databases (Bruce, 1991; Tellis, 1986), bibliographic standards (Gerdes & Smith, 1990), and metadata structures (Major, 2004), and providing education in data management planning (Miller, 2010). Dorn (1995) and Radke (1994) discuss preservation issues associated with non-digital geoscience data, whilst recent literature focuses on digital data which presents challenges borne of the rapid pace of technological change: quality control, bibliographic control, and maintaining format migration programs for large volumes of data (Joseph, 2008; Kuny, 1998; Lowe, 1995; Scott, 1992).

**Skills and knowledge**

Hallmark’s (1998) survey of geology librarians provides an extensive list of skills required by geoscience information professionals: a strong science background, extensive computer skills, expertise with maps and GIS applications, awareness of legal issues, marketing, contract negotiation, strategic planning, financial acumen, and effective communication.

Gabridge (2009) and Soehner et al. (2010) report scepticism that information professionals are equipped to help scientists with data management. Stuart and Drake (1992) note the shortage of information professionals with skills in both scientific and LIS domains, and more recently the ALIA Education and Workforce Summit recognised there is a dearth of information professionals with the IT skills needed for research data management (Wells, 2008). There is widespread agreement that information professionals need to undertake ongoing professional development to build data management expertise (Corrall, 2010; Gold, 2007; Lewis, 2010).

**Summary**

From the literature reviewed it is difficult to construct a clear picture of the role information professionals play in geoscience data management. Literature generated by geoscience information groups indicates that information professionals have a long history of involvement in the management of geoscience data, but the coverage of this involvement is fragmented. The academic and research library sector provides a more holistic view of the role of information professionals in research data management, although much of this is based on discussions, and recent research is focused on digital data. Research into the geoscience data management activities of information professionals in small information services—which are common in
geoscience organisations—is largely absent from recent literature generated by information professionals.

The review identifies an opportunity to capture an overview of the role of information professionals in geoscience data management, for both the full spectrum of geoscience data and the range of organisations which are engaged in geoscience data management.

Research Process
In order to contain the size of the project, and also for logistical reasons, the study was limited to Western Australia, which is Australia’s dominant resource state and home to a large range of organisations which generate geoscience data (WA Department of Mines and Petroleum and Department of State Development, 2011).

To identify information professionals engaged in geoscience data management the researcher consulted a number of sources:

- a list of information services developed for the Australian Geoscience Information Association (AGIA) research project Access to Australian geoscience information: a tragedy of the commons (Smith, 2010);
- members of AGIA;
- the Geological Society of Australia’s listing of institutions engaged in tertiary geoscience education (Geological Society of Australia, 2011);
- geoscience organisations’ websites; and
- professional networks.

The consultation process revealed that some organisations had multiple work units engaged in geoscience data management. Consequently, individuals across a range of geoscience data management work units were recruited in an effort to capture as many information professionals as possible. A non-probability sampling method was used to select an expert sample, and individuals in 30 work units across 22 organisations were identified.

The research project used a mixed methods approach, combining qualitative and quantitative research approaches through a descriptive survey. The survey instrument was a self-administered questionnaire designed by the researcher. The questionnaire comprised 37 questions which were a mix of closed questions and open-ended questions across four broad categories relating to geoscience data management:

1. Environment – what structural and professional data management frameworks exist within participants’ organisations?
2. Activities – what geoscience data do participants manage and how do they manage it?
3. Knowledge and skills – what qualifications do participants hold? What skills do they consider necessary for geoscience data management and how are they developed?
4. Opportunities and challenges – are there opportunities for information professionals to improve geoscience data management, and what challenges do they face?

After pilot-testing, a hard copy of the questionnaire was mailed to the recruits, with a reminder email (including a digital version of the questionnaire) sent to non-respondents after the due date had lapsed. The responses were de-identified and transcribed into Microsoft Word. An iterative
analysis of qualitative data gathered via the open-ended questions was conducted to identify categories and themes, which were then logged in a Microsoft Excel spreadsheet, along with the quantitative data from the closed questions. A cross-case analysis was conducted to identify common themes and patterns within the data.

**Findings**

1. **Response**

The response rate was 37%, with 16 valid responses received from 11 work units (across nine organisations).

A summary of questionnaires sent and responses received can be seen below in Table 1.

Table 1. Questionnaire Distribution and Response

<table>
<thead>
<tr>
<th>Sector of organisation</th>
<th>Number of work units sent questionnaires</th>
<th>Questionnaires received</th>
<th>Other responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic</td>
<td>6</td>
<td>1</td>
<td>No involvement (1)</td>
</tr>
<tr>
<td>Government</td>
<td>Commonweal th</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>State</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Corporate</td>
<td>Minerals</td>
<td>8</td>
<td>2*</td>
</tr>
<tr>
<td></td>
<td>Petroleum</td>
<td>10</td>
<td>10*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>On long leave (1)</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td></td>
<td>No consent form &amp; no involvement</td>
</tr>
<tr>
<td>Total</td>
<td>31</td>
<td>17</td>
<td>4</td>
</tr>
</tbody>
</table>

16 valid responses

Note: * multiple questionnaires received from some work units.

Eight participants were information professionals, four were geoscientists, two were IT professionals and two were non-professional staff. One participant felt that the study’s definition of information professionals was too narrow: “we have 10 staff in Data and Information Management only two of which are ‘Information Professionals’ by your definition. The other eight are also info[rmation] professionals but with other qualifications – not library, records or archives”. This participant also felt that IT professionals should have been included in the study as “they are the key to successful management of geoscience data. The role of Info[rmation] Professionals (by your definition) are [sic] more towards managing geoscience information”. 

*LIBRES ISSN 1058-6768 Volume 22, Issue 2, September 2012*
To simplify the reporting of results, each work unit in the study was allocated a unique name as follows:

- Academic Library;
- Computing Unit;
- Corporate Library;
- Data Management;
- Government Library;
- Government Minerals;
- Government Petroleum;
- Information Management;
- Information Services;
- Minerals Team; and
- Petroleum Team.

2. Geoscience data management environment
The study revealed five types of work units engaged in geoscience data management:

1. Specialist: small specialist information services predominantly staffed by information professionals. Four work units fitted this profile and were found in the corporate and government sectors;
2. Team: multi-disciplinary data/information management teams with workers from a range of professions, including information professionals. The two work units that fitted this profile were in the corporate sector;
3. Geoscience: geoscience units where most staff were geoscientists. Two work units fitted this profile and were both in the government sector;
4. Data: work units specialising in data management with no information professionals. Two work units fitted this profile and were both located in the corporate sector; and
5. Academic: large academic information services where geoscience is one of many subjects covered. One work unit—a university library—fitted this profile.

The study investigated whether work units managing geoscience data were commonly found within a particular operational area of organisations. The study identified four operational areas:

- Administration;
- Geoscience;
- Information Services; and
- Information Technology.

With the exception of the Academic Library, all work units contained three or fewer information professionals. Geoscience operational areas were the most common location for work units managing geoscience data, and all work units in this area were very involved in data management, regardless of work unit type. Work units in the Geoscience divisions were consistently satisfied or very satisfied with their work unit’s level of involvement with data management, which had either increased or stayed the same. This positive trend was not so consistently reported in other operational areas where levels of satisfaction and engagement were mixed. Information professionals from both Team units believed management strategy had led to their increased involvement.

All of the participating work units indicated that they liaised internally on data management issues, and information professionals indicated that they liaised with a range of stakeholders.
within their organisations including geoscientists, geographic information systems (GIS) professionals, information technology departments, management, libraries, and operational departments.

Participants were asked “do you believe there is one professional group in your organisation which is seen as having a lead role in geoscience data management?” The responses were mixed:

- four participants (two geoscientists and two information professionals) across the range of academic, corporate and government organisations indicated there was no single professional group seen as leaders in geoscience data management;
- five participants nominated their work units as providing leadership in geoscience data management: a geoscientist in Government Petroleum, information professionals in the Corporate Library and both Teams, and an IT professional in Data Management;
- the information professional from Information Management felt geoscientists were leaders; and
- the information professional from the Government Library saw a geoscience division (to which the Government Petroleum and Government Minerals units belonged) as leaders.

Although the structure of the work units and professions which provided leadership varied, all (with the exception of the Petroleum Team which did not specify an operational area) were found within Geoscience operational areas.

3. Geoscience data management activities
The questionnaire asked participants to list the types of geoscience data held by their organisations, the geoscience data management activities carried out within their work unit, and the professional groups responsible for these activities. These questions were asked separately for each of the three data categories: geoscience collections, non-digital geoscience data, and digital geoscience data. The answers of all participants were collated to give an overview of each organisation’s geoscience data and details of this can be found Figures 1, 2 and 3 below.

Figure 1. Geoscience Collections Held by Organisations Showing Proportion Managed by Information Professionals
All of the organisations surveyed (except the Academic Library) held geoscience collections. Figure 1 (above) shows that the engagement of information professionals in the management of geoscience collections was generally low. Responses regarding data management activities for geoscience collections were sparse, and an information professional explained:

*Information Services doesn’t manage the geoscience collections. The Exploration Dept is responsible for managing the samples at Corelab – they know what samples have been collected for new wells, know what analyses the reservoir engineers have ordered, etc, so they are in a better position to manage those items.*

The responses of the three information professionals who specified data management activities were varied. Two information professionals listed maintaining databases or records management systems, archiving, and cataloguing as part of their management of geoscience collections. The information professional from the Corporate Library provided a detailed list of activities, which—in addition to those already listed— included:

- re-boxing of material if required;
- submission of samples to government;
- handover of material to new permit operators; and
- collection management – where a well fell into an area no longer of interest to the company the librarian offered samples to the relevant state or national regulatory body.

**Figure 2. Non-Digital Geoscience Data Held by Organisations Showing Proportion Managed by Information Professionals**

All organisations which participated in the survey held non-digital data. Figure 2 (above) shows that information professionals were very involved in the management of the full range of non-digital data held by their organisations. The most common data management activities engaged in by information professionals were organisation of datasets (including database management), metadata (including cataloguing), archiving/storage/preservation and collection management, and these activities were spread across the full range of non-digital geoscience data. Information

*LIBRES ISSN 1058-6768 Volume 22, Issue 2, September 2012*
professionals in three Specialist units supplied detailed accounts of their management of non-digital data, summarised as follows:

**Government Library:**
- acquisition, cataloguing/indexing, circulation and physical management of hardcopy data;
- management of online catalogue for access to collections; and
- database creation, data entry and management of databases and indexes of specific geoscience data collections (e.g. maps).

**Corporate Library:**
- individually catalogue and store technical reports, well logs (film and prints);
- ensure all relevant data submitted to government regulatory bodies to meet statutory requirements;
- handover of data to new operators as required;
- weeding – offer unwanted material to government regulatory bodies; and
- archive working files and destroy them when they are no longer required.

**Information Service:**
- manage offsite data storage;
- catalogue data;
- order data from government and commercial sources;
- organise duplication/transcription of tapes; and
- handle data transfer to/from other companies.

**Figure 3. Digital Geoscience Data Held by Organisations Showing Proportion Managed by Information Professionals**

![Digital Geoscience Data Held by Organisations Showing Proportion Managed by Information Professionals](image)
Figure 3 (above) shows that for digital data information professionals were mainly involved with the management of reports, drilling and well data, seismic data, and maps and plans; the same four areas which attracted the greatest involvement in non-digital data.

The majority of information professionals’ data management activities were centred on the maintenance of databases and electronic records management systems. This included activities such as loading and archiving data, as well as creating metadata. An information professional from the Minerals Team provided details of digital data management activities undertaken by her work unit:

- upkeep of thesauri and metadata standards;
- guidance on information management;
- training on system use;
- create maps and plans (with geoscientists); and
- reports, maps and plans, and point data – catalogue, manage access, manage database, governance and guidance notes for database, publication via web tools.

Comparison of Data Management Activities across Geoscience Data Categories

Data management activities were grouped as follows:

- collection management – including acquisition, weeding, digitisation, physical management, managing access;
- compliance & standards – including legislation and processes;
- education, training and guidance;
- metadata - including cataloguing;
- organisation of datasets – encompassing database & records management;
- preservation – including archiving and storage; and
- publication.

Figure 4 (below) shows a comparison of information professional’s data management activities across the three data categories: geoscience collections, non-digital geoscience data and digital geoscience data.
The most consistent management activity for information professionals across data types was the organisation of datasets. Metadata, collection management, and compliance and standards, were strong areas of engagement for non-digital geoscience data, but lower for geoscience collections and digital geoscience data. Both geoscience collections and digital geoscience data were regularly handled by other work units and professions:

- geoscience collections by geoscience operational areas or outsourced to external providers; and
- digital geoscience data by geoscientists, technicians, GIS staff and IT professionals.

4. Skills and knowledge for geoscience data management

Fourteen participants held tertiary qualifications. Six participants held university science degrees (five majoring in geology) and postgraduate areas of study included geoscience (three), business (one) and LIS (one). Five participants held university LIS degrees, with three proceeding to postgraduate studies in records management, knowledge management, and business respectively. Two participants held university arts degrees and both completed postgraduate LIS studies; one also completed a Masters of Information Management and Systems. The participating IT professional completed a computer programming diploma through the Technical and Further Education system.

Participants were asked to list what knowledge and skills they considered important for geoscience data management. The responses were grouped as shown in Figure 5 (below).

Ten of the 11 respondents to this question believed an understanding of geoscience data was important—only the Academic Librarian omitted this factor. “Organisation” was seen as an important skill by the majority of participants, and this term was used in a general sense rather than as the organisation of any specific item. Other general skills such as attention to detail and
analytical ability were also commonly nominated. Technical skills included competency with computer applications and database knowledge. Compliance included an interest in standards and the ability to follow them and other data management processes, and knowledge of legislation.

When asked “what has contributed most significantly to your geoscience data management knowledge and skills?” the response was almost unanimous: 14 of the 15 responses (including seven information professionals) nominated experience gained through work. Five participants from non-geoscience backgrounds (including three information professionals) found working alongside geoscientists particularly valuable.

Eleven of the 16 participants were engaged in continuing professional development, and this included six information professionals. Professional development activities included:

- conferences;
- training courses;
- membership of professional associations: four participants (three information professionals and one geoscientist) specifically mentioned AGIA; and
- seminars and workshops.

5. Geoscience data management opportunities and challenges

Participants were asked if they believed there were additional data management roles which information professionals could undertake to improve geoscience data management in their organisation. As can be seen in Figure 6 (below), the response to this question could be roughly divided along professional groupings.

Six corporate sector information professionals believed there were additional roles they could undertake, while the librarians from the Academic Library and Government Library were unsure.

Information professionals were able to identify specific areas of involvement which could improve data management, and the researcher could sense the frustration behind some responses. Information professionals from the each of the Team units identified opportunities for a role in
data management leadership within the organisation, including “communicating clear and transparent work flows between professionals managing data”.

Most of the information professionals’ comments related to opportunities for improving the organisation of digital geoscience data, such as the introduction of naming conventions and electronic content management systems. Shared drives seemed to be a particular problem in terms of organisation and duplication, and information professionals saw opportunities to educate geoscience data stakeholders in issues relating to the organisation of datasets. The information professional from Information Services commented:

*Having seen the chaotic state of the Exploration Department’s shared drive, it’s fairly obvious that they would benefit from a more structured approach to organizing their data and more consistency in naming conventions. We are in the process of implementing an electronic content management system and are using this as an opportunity to advise departments on records management issues and nudge them into organizing and cleaning up their data.*

The views of other participants regarding additional geoscience data management roles for information professionals were not as clear. Two geoscientists who responded to this question both discussed ways their organisations could generally improve data management, rather than ways information professionals could contribute. The remaining participants were divided evenly between the “don’t know” and “no response” camps.

Participants were asked to nominate the biggest challenges in managing geoscience data in their organisation. Fifteen participants responded to this question, some very enthusiastically. Participants also discussed geoscience data management challenges at the end of the questionnaire where open comment was invited. Responses were analysed and grouped into nine themes, as shown in Figure 7 (below).

**Figure 7. Challenges of Geoscience Data Management**

![Challenges of Geoscience Data Management](image)

Technological change was a significant challenge for both information professionals and other participants, and pervaded many of the geoscience data management issues identified.
Participants commented on the range and complexity of different geoscience data types which need to be managed, and the struggle to keep up with formats which are constantly evolving.

The Government Librarian was concerned about the best way to physically store digital geoscience data, and the implications of changing technology on the ability to access data which was stored long-term. The information professional from Information Management noted that old formats (such as 9” reels) presented preservation challenges, and these needed to be migrated to new digital formats to maintain accessibility. Information professionals found it difficult to manage access to older data formats when so much attention was focused on digital data management issues, and an information professional in the Minerals Team discussed “keeping a balance between leading change and maintaining legacy data” and “managing change - making sure not too much falls through the cracks”. The information professional in the Corporate Library described the difficulty of ensuring the preservation of data for use beyond the boundaries of the organisation’s self-interest: “in a corporate environment you tread a fine line in managing the data for your organisation which must be your first priority and ensuring valuable data/material has the chance to be offered to government before it is destroyed”.

The biggest challenge for information professionals was the attitude toward data management by other geoscience data stakeholders (especially management), which they believed impacted on the allocation of resources to, and frameworks for, geoscience data management within their organisations. An information professional from the Petroleum Team found “making all participants aware and willing to participate in set processes” a challenge while another from Information Services said “the Information Services department is only 5 years old but the company has existed for over 20 years so there is a lot of history/inertia/lack of awareness of good practice in data and records management”.

Concern was expressed by a number of participants about the lack of resources—both in terms of money and qualified people—to deal with the volume of data being generated. One information professional summarised it as “lack of time, lack of staff, lack of structure”. The link between management attitude and the allocation of resources was discussed by an information professional in the Corporate Library who believed her organisation would benefit from an electronic document management system: “head office […] controls this project and funding, and nothing has happened in more than eight years”.

Some information professionals felt there was a shortage of external support systems to help with geoscience data management. The corporate information professional from Information Services found it difficult to locate standards, guidelines, advice and recommendations on managing geoscience data and felt that there was “very little literature that is relevant and current, and tools, standards, guidelines etc. that may be useful tend to have to be adapted”. The same participant also commented that “it’s hard to find relevant courses/seminars that are pitched at the right level” and “geoscience data is quite technical and it is hard to learn how to manage it well”. An information professional from the Minerals Team stated “the legal framework of digital information management is poorly understood and therefore a minefield”.

Participants identified challenges around geoscience data visibility and access. The Academic Librarian found locating appropriate materials and making academics aware of the geoscience data holdings a challenge, while the geoscientist from Government Petroleum found locating legacy data and complete data sets difficult.
Discussion

Data management environment
The structure of geoscience data management in the participating organisations was bespoke – there was no off-the-rack solution and the mix of data management roles, work units and their location within organisations was varied. These results correlate with Hyams’s (2008) and Pryor and Donnelly’s (2009) observations that data management roles vary widely between institutions as a consequence of differing data management aims and resources.

The literature suggests that information professionals wanting to engage in scientific data management need to work closely with data creators (Gold, 2007), and the results support this observation:
- all work units in geoscience operational areas were very involved in geoscience data management;
- all work units identified as providing leadership in geoscience data management (except the Petroleum Team, which did not specify an operational area) were located in geoscience operational areas; and
- all participants in work units in the geoscience operational areas were either satisfied or very satisfied with their work unit’s role in geoscience data management.

Data management activities
The study revealed that information professionals were engaged in all of the data management activities identified in the literature review. Traditional downstream activities as identified by Gold (2007) were common across all three categories of geoscience data, namely organisation of datasets, metadata, collection management, preservation, and standards and policies. Within the category of digital geoscience data there was small amount of involvement in the upstream data management activities identified in the literature (Gold, 2007).

There was a great deal variation in data management activities by information professionals across the three data categories. Information professionals were very engaged in the management of non-digital geoscience data, but were less involved with geoscience collections and digital geoscience data, especially where the data was close to its raw form. Browne and Love (1997) and Lowe (1995) identify that data managers must be aware of the links between different data categories, and the results show information professionals trying to balance the maintenance of legacy collections against keeping up with new data formats.

Skills and knowledge
The results concur with Stuart and Drake’s (1992) observation that there is a scarcity of information professionals with dual science and LIS qualifications. Many of the skills which the participants listed as necessary for geoscience data management correlated with those outlined in Hallmark’s (1998) investigation of ideal skills for geoscience information specialists: geoscience knowledge, extensive computer skills, and an awareness of legal issues.

Understanding the data was identified as essential for geoscience data management, and for most information professionals this knowledge was gained through work experience, especially working with geoscientists. The literature identified the need for information professionals to undertake ongoing professional development to maintain skills for data management (Corrall, 2010; Gold, 2007; Lewis, 2010) and while most of the information professionals in the study...
participated in professional development, a desire for further training in geoscience, data management, technology, and compliance with statutory reporting requirements was identified.

**Geoscience data management opportunities and challenges**

Historically information professionals have been involved in the development of standards and guidelines for the management of geoscience data (Gerdes & Smith, 1990); and Gold (2007) identified this as an area information professionals could increase their engagement. Several participants identified opportunities for information professionals to improve the organisation of data. Implementing electronic records management systems was seen as one way to improve data management, as was the use of naming conventions and establishing workflow procedures.

The education of geoscience data stakeholders in data management was another opportunity identified by information professionals. Miller (2010) describes the challenges of establishing a geoscience data literacy course at Purdue University, and the results suggest information professionals may also be facing barriers in this area with only two information professionals engaged in training and guidance in geoscience data management.

The information professionals surveyed believed they could improve geoscience data management in their organisations by providing leadership. The results demonstrated that there was no single professional group which had assumed leadership in geoscience data management and the role seems ripe for the plucking. Certainly the study showed information professionals liaise with geoscientists, technology professionals, and management, so these relationships do exist and there is potential for them to be built upon.

The attitude of both management and other data stakeholders toward the management of geoscience data was the challenge most commonly cited by information professionals. Yacopetti & Mundell (2010) discuss the difficulty of securing resources for data management because of management attitudes, and this issue was evident in the results. The literature suggests IT professionals and scientific colleagues are often seen as the most likely source of data management solutions (Gabridge, 2009; Soehner et al., 2010; Yacopetti & Mundell, 2010), and the results demonstrated that other professional groups have a low awareness of how information professionals might contribute to geoscience data management. The literature suggests the most effective weapon to change attitudes toward data management roles is advocacy, supported by professional development to equip information professionals with the full range of competencies—in geoscience, data management, compliance and technology— required to take on these additional roles (Gold, 2007; Lewis, 2010).

The literature frequently discusses the challenges facing information professionals as a result of the pace of developments in technology and the resultant data deluge (Joseph, 2008; Kuny, 1998), and these challenges were also identified in the results of this project. The range and complexity of data, visibility and access, data volume, and preservation were all challenges which had significant technological components.

**Research Limitations**

Identifying the population of information professionals working in geoscience data management was difficult, and the sample surveyed had several limitations:

- the sample was small and conclusions could not be confidently generalised;
• information professionals in organisations without an information service fell outside the reach of the study; and
• the sampling process captured recruits who were not information professionals (by the study’s definition) as it was not possible to ascertain the tertiary qualifications of recruits prior to the execution of the questionnaire.

While the last point was initially seen as a disadvantage, the diversity of outlooks provided insight into the perceived role of information professionals within geoscience data management, and illustrated the common challenges faced by all geoscience data stakeholders.

The data gathering process faced many of the limitations typical of questionnaires (Williamson, 2002):
• the response rate was low at 37%;
• the responses were unrepresentative of the cross-section of the sample: only one organisation in each of the academic and government sectors responded;
• ambiguous results: the results which concerned the engagement of information professionals in the management of digital geoscience data were ambiguous. The quantitative data indicated information professionals had a low level of involvement with the management of digital data in terms of the range of data handled and the diversity of activities, yet qualitative feedback suggested a great deal of information professionals’ concern is focused on digital data management;
• variation in the quality of responses: there was a wide variety of detail in the reporting of data management activities.

Implications and further research
The results suggest that there is a strong link between leadership and engagement in geoscience data management and alignment with geoscience operational areas within organisations, although a wider study of geoscience data management is necessary to validate this conclusion. Information professionals see opportunities to expand their role in data management and may benefit from focusing their efforts on building alliances with geoscience divisions within their organisations. Further investigation of the advocacy undertaken by information professionals involved in work units identified as leaders in geoscience data management may provide insights into successful strategies to raise awareness of the potential of information professionals and the importance of data management.

Information professionals are engaged in the management of all geoscience data categories, yet they are less involved with the management of geoscience collections and digital geoscience data. This reflects the fragmentation of the management of different data categories within organisations, and raises concern about the longevity, visibility, integrity, and efficiency of access to complete datasets. Given the increasing focus on the management of digital data, a lower level of engagement in digital data management also threatens the future of information professionals in geoscience data management. Further qualitative research is required to understand the practices around the management of geoscience collections and digital geoscience data, together with the extent and impact of disaggregated geoscience data management.

Most of the information professionals participating in the study did not come from a geoscience background— which may in part explain their low level of engagement with raw data commonly found within geoscience collections and digital data—and developed their knowledge of
geoscience through work experience. The results identified a need for—and shortage of—continuing professional development opportunities to meet the challenges of geoscience data management, especially geoscience, technology, data management, and statutory reporting requirements. The most effective means of delivering both workplace-based training and structured professional development in these four areas of need requires further investigation.

Conclusion
The research project provides a view of information professionals’ engagement in geoscience data management across a breadth of data categories which is missing from current investigations into scientific data management. The study also provides insight into the geoscience data management environment of information professionals working in the corporate sector—commonly in small specialist information services or embedded within multidisciplinary teams—and the challenges associated with a highly competitive and often secretive environment.

A range of groups can extract value from the results of this research project. Information professionals seeking to expand their engagement in scientific data management can learn from the long experience of geoscience information professionals in this field. The value of working closely with geoscientists—both in terms of structuring scientific data management units and gaining domain knowledge—is readily transferrable to other forms of scientific data management. Geoscience organisations should consider the impact of their departmental structures on efficient data management, especially the costs of disaggregation. Professional associations and government agencies who are geoscience stakeholders need to consider ways to improve opportunities for professional development to improve geoscience data management. Geoscience information professionals need to employ advocacy backed-up by continuing professional development in order to secure their place at the data management table.
References


Western Australia (WA) Department of Mines and Petroleum and Department of State Development. (2011, April-May). WA exports top $100 billion. Prospect, April - May 2011.

