The E-Resources Management Handbook – UKSG

E-resource management and the Semantic Web: applications of RDF for e-resource discovery

GEORGE MACGREGOR
Information Strategy Group, Information Management & Systems
Liverpool Business School
Liverpool John Moores University

Semantic Web technologies and specifications are increasingly finding applications within digital libraries and other e-resource contexts. The purpose of this chapter is to provide an introduction to some essential Semantic Web concepts and the resource description framework (RDF), a key enabling language of the Semantic Web. Applications of RDF including Dublin Core, FOAF, SKOS and RDFa will be explored with practical examples, and recent implementations of these specifications within a variety of e-resource discovery contexts will be discussed.

Introduction

Recent developments in the Semantic Web offer digital libraries and repositories the opportunity to better expose valuable e-resources using a suite of interoperable standards and technologies. Such tools hold the potential for innovative approaches to the navigation and retrieval of resources within heterogeneous and distributed e-resource environments. The outputs of Semantic Web activity also present opportunities for resolving or ameliorating common problems relevant to digital libraries, such as semantic interoperability and advanced metadata integration. Although the deployment of Semantic Web approaches within digital libraries and repositories is growing, the use of such techniques generally remains confined to particular communities of practice (e.g. research centres, academia, research libraries, etc.). To some extent this is consistent with the wider computing and information profession; however, it is something that has been changing in recent years.

Developments in the Semantic Web are of increasing significance to information professionals. As well as having useful applications within digital libraries, information professionals have an emerging role to play in the development and maintenance of the structured data comprising the Semantic Web (e.g. metadata, ontologies, etc.). The relevance of the Semantic Web to Library and Information Science (LIS) has been reflected in recent research and dissemination activity by information professionals[1,2,3] and many are actively participating in the development of important W3C Semantic Web specifications[4].

Given the relevance of the Semantic Web to LIS, the purpose of this chapter is to provide an introduction to some essential Semantic Web concepts and resource description framework (RDF) specifications. Recent applications of these concepts within a variety of contexts will also be explored, particularly within digital libraries and e-resource discovery. Since RDF and applications of RDF provide a key enabling technology within the Semantic Web, the chapter will introduce RDF using practical examples.

The Semantic Web

The Semantic Web is a research agenda originally initiated by Tim Berners-Lee in 2001[5]. It is now considered to be an evolving extension of the existing web, and the agenda is one that has been reiterated more recently by Berners-Lee and his colleagues as a 'web of data'[6].

The purpose of the Semantic Web is to make the semantics of information and services available on the web interpretable and understandable to machines so that user requests can be more accurately satisfied. The difficulty with the current web is that it has evolved to consist primarily of documents designed for humans to read, rather than for machines. For example, machines can interpret the *syntax* of the web documents (e.g. XHTML) and display these documents to users, but they have little ability to interpret their meaning (i.e. semantics). The intention of the Semantic Web is therefore to deliver a web of data which will better facilitate the extraction of *semantics* from documents by intelligent software agents. Equipped with this semantic knowledge, computers can then actively support users in their information tasks as opposed to passively displaying or delivering information to users.

One obvious area in which this semantic data can be put to good use is information retrieval[7]. For example, if information retrieval systems can better understand the meaning of items within an e-resource collection then it will be easier to design systems that provide greater retrieval precision during users' information-seeking tasks. Increased precision could be achieved by better understanding user context, disambiguating conceptually similar items, performing some of the functions controlled vocabularies might; but improvements in recall could also be achieved by augmenting the results with conceptually related resources, perhaps spanning a variety of media. Although the deployment of the Semantic Web within LIS is our focus, such semantic technologies assume greater potential and complexity when applied to everyday tasks, such as booking a medical appointment[8] or ordering wine for a social event[9]. In such instances numerous applications may be involved, requiring a high level of systems interoperability and a shared level of meaning (i.e. shared semantics) through the use of ontologies.

For the Semantic Web vision to work and for intelligent software agents to have data to harness, resources on the web have to be expressed in a machine-interpretable format. This entails annotating resources with machine-interpretable metadata and other structured data which attempts to capture the semantics of resources. Since the ethos of the web is distributed and since the intention is that Semantic Web data be available for manipulation or reuse by any number of heterogeneous applications, the interoperability of this structured data is absolutely essential. Structured and interoperable data is so fundamental to the success of the Semantic Web that Tim Berners-Lee recently conceded that the 'data web' would have been a better name for his vision[10]. Although there are a number of emerging technologies underpinning the Semantic Web[11], it is the resource description framework (RDF) and its various applications which provide the majority of the structured data required to make the Semantic Web work.

Resource description framework

The resource description framework (RDF)[12] is a framework for modelling and representing data on the web. In fact, RDF is simply a data model in which statements are made about web resources. Each statement made about a resource comprises a collection of 'triples' consisting of a *subject*, *predicate* and *object*. The *subject* denotes the object the triple is describing, the *predicate* identifies the attribute of the subject within the statement, and the *object* defines the value of the predicate. A set of triples is known as an RDF graph and is diagrammed using a series of nodes connected by labelled arcs (Figure 1).

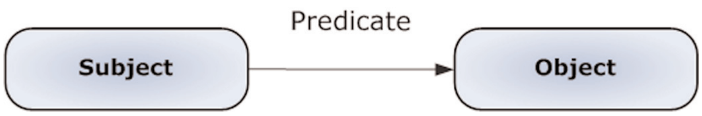

Figure 1. An example of an RDF directed graph

Taken together the subject-predicate-object triple represents a statement of fact about the resource in question and characterizes the nature of the relationship between each node of the directed graph. Consider the following statement as an example:

- 'The title of this chapter is Applications of RDF for e-resource discovery'

Within this statement we can identify the following triple set:

- Subject: `Chapter`
- Predicate: `hasTitle`
- Object: `Applications of RDF for e-resource discovery`

This triple, in turn, could be graphed as in Figure 2.

Figure 2. Identifying triples within an RDF graph

Recall that the purpose of the Semantic Web is to provide machine-interpretable statements about resources on the web in order to derive meaning. For the Semantic Web this entails two things: the use of uniform resource identifiers (URIs) and the way of expressing RDF on the web.

Figure 2 illustrates the concept of RDF and triples admirably; however, the English-language text strings used for our triples are more conducive to human interpretation than machine processing. RDF therefore takes advantage of URIs[13] as the principal means of identifying subjects, predicates and objects within RDF triples. Although similar to URLs which *locate* resources, URIs can be far more abstract and can *identify* anything. They can refer to resources available over a network much like a URL but can also refer to non-networked resources (e.g. people, physical documents, places, etc.) and abstract concepts or names which have no physical manifestation (e.g. title, creator, subject). By using URIs within RDF it is therefore possible to describe anything and any type of relationship between these things. The importance of URIs will assume more relevance shortly.

Since RDF is a data model, it remains syntax independent. It is therefore possible to express (or 'serialize') RDF on the web in a variety of ways, including RDF/XML[14], Notation 3 (N3)[15] and Turtle[16]. While the latter two are increasingly popular, RDF/XML continues to be used extensively. The popularity of RDF/XML is attributable to its use of XML[17] to serialize an RDF graph as an XML document. It is used in much of the W3C Semantic Web documentation and continues to be the only serialization recommended by the W3C Semantic Web Activity team[18]. RDF/XML will therefore be the serialization used in examples throughout this chapter.

The importance of RDF/XML and URIs in expressing RDF graphs has been noted and it is now possible to provide an example.

Basic example

In Figure 2 the subject of the RDF graph was `Chapter`. At time of writing, this present chapter lacks an electronic location; however, when it is officially published it will have a URL incorporating the UKSG / MetaPress domain. The URL therefore could be said to be `http://uksg.metapress.com/someURL`.

Dublin Core (DC)[19] metadata allows us to formalize the `hasTitle` predicate from Figure 2 since DC includes a title element fulfilling that purpose. Dublin Core can be expressed as RDF[20] and is defined by an RDF Schema at `http://purl.org/dc/terms`. This allows us to assign a proper predicate for `hasTitle` based not only on a recognized metadata schema, but defined using a URI instead of a text string. In this case `hasTitle` becomes `http://purl.org/dc/terms/title`.

Finally, the object of the RDF graph in Figure 2 is `Applications of RDF for e-resource discovery`. Since this is the value of our object this will remain as a literal (i.e. a text string).

These amendments to the RDF graph allow us to update it accordingly (Figure 3). By doing so we note that the graph now consists of the following triple set:

- Subject: `http://uksg.metapress.com/someURL`
- Predicate: `http://purl.org/dc/terms/title`
- Object: `Applications of RDF for e-resource discovery`

Note also that because the object node in Figure 3 is a literal it is diagrammed as a box.

Figure 3. A simple RDF statement using Dublin Core

Since providing RDF graphs in a machine-interpretable data format is essential for the Semantic Web to operate, it is possible to express the graph in Figure 3 as RDF/XML. Such a graph would be expressed as follows:

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dcterms="http://purl.org/dc/terms/">
<rdf:Description rdf:about="http://uksg.metapress.com/someURL">
 <dcterms:title>Applications of RDF for e-resource discovery</dcterms:title>
</rdf:Description>
</rdf:RDF>
```

The subject and the predicate must always be referenced using a URI. The object is the only component of an RDF triple which is permitted to use literals; but as we have noted in the above example, there are circumstances in which the object must be a literal, often because a URI is inappropriate or unavailable. In the above example the literal was `Applications of RDF for e-resource discovery` and such literals are common when metadata is used. However, the preference in RDF is to use URIs wherever possible to identify triples within an RDF graph so as to aid machine processing and, in many cases, an object URI will be available. Consider the following statement as an example:

- 'The creator of this chapter is George Macgregor'

Within this particular statement we can identify the following triple set:

- Subject: `Chapter`
- Predicate: `Creator`
- Object: `George Macgregor`

With our knowledge of the chapter's URL, of the Dublin Core element set, and of the author's personal homepage (where detailed RDF creator information can be extracted by intelligent software agents), it is possible for us to formalize the triple set using URIs as follows:

- Subject: `http://uksg.metapress.com/someURL`
- Predicate: `http://purl.org/dc/terms/creator`
- Object: `http://www.staff.ljmu.ac.uk/bsngmacg`

Rather than use a literal to describe the creator (i.e. George Macgregor) it is possible for us to reference the creator using a URI. This RDF graph can then be integrated with our previous graph (as in Figure 4) and expressed in RDF/XML as follows:

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dcterms="http://purl.org/dc/terms/">
```

```
<rdf:Description rdf:about="http://uksg.metapress.com/someURL">
 <dcterms:title>Applications of RDF for e-resource discovery</dcterms:title>
 <dcterms:creator rdf:resource="http://www.staff.ljmu.ac.uk/bsngmacg/" />
</rdf:Description>

</rdf:RDF>
```

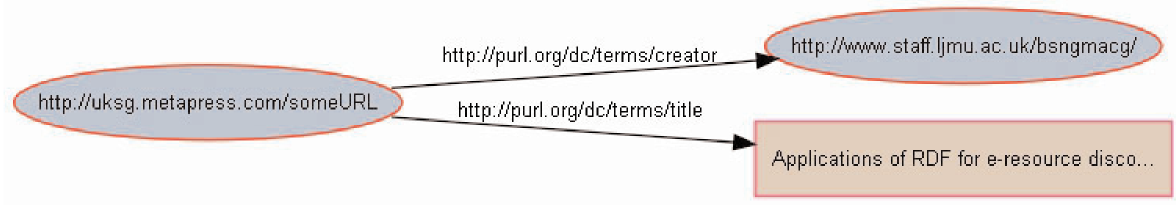

Figure 4. Simple RDF graph demonstrating the use of URIs in RDF

If desired, this simple RDF statement could easily be augmented with further Dublin Core metadata elements. For example, publisher information could be included along with Library of Congress Subject Heading (LCSH) descriptor charactering the aboutness of the resource in question, and the rights could be referred to by a Creative Commons licence[21], all of which could be referenced by URI, thus generating the RDF graph in Figure 5 and providing the following RDF/XML:

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dcterms="http://purl.org/dc/terms/">

<rdf:Description rdf:about="http://uksg.metapress.com/someURL"> <dcterms:title>Applications
of RDF for e-resource discovery</dcterms:title>
 <dcterms:creator rdf:resource="http://www.staff.ljmu.ac.uk/bsngmacg/"/> <dcterms:publisher
rdf:resource="http://www.uksg.org/"/> <dcterms:subject
rdf:resource="http://id.loc.gov/authorities/sh2002000569"/> <dcterms:rights
rdf:resource="http://creativecommons.org/licenses/by-nc-sa/2.0/uk/"/></rdf:Description>

</rdf:RDF>
```

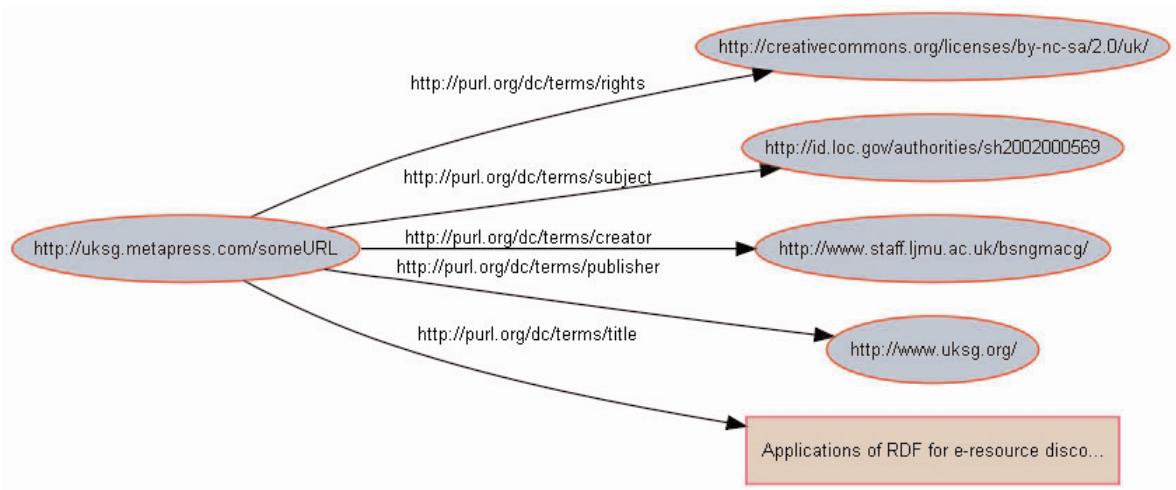

Figure 5. Augmenting our simple RDF graph from Figure 4

In Figure 5 we have been able to augment the RDF graph by making greater use of URIs. The decision to use a URI for subject indexing was based on the increasing use of controlled vocabularies on the Semantic Web expressed in RDF. One such example of this is LCSH[22]. The URI of `http://id.loc.gov/authorities/sh2002000569` denotes the LCSH descriptor, 'Semantic Web'. This URI not only defines the concept of the Semantic Web, but at the end of the URI we discover rich terminological data expressed in a variety of Semantic Web-friendly serializations. Referring to controlled vocabularies in this way will be discussed in more detail in the 'Simple Knowledge Organization System (SKOS)' section of this chapter. Of course, in many circumstances literals will suffice and the subject heading used above, for example, could easily be a literal taken from LCSH rather than a URI.

The use of Dublin Core in the Semantic Web is a useful introduction to the basic concepts of RDF and RDF/XML. Additionally, the ability to integrate RDF data on the web means that DC is often used in conjunction with numerous other RDF applications. Note that the RDF/XML examples and the resulting RDF graphs in this section were created using specialist software[23,24]; however, the validity of the RDF/XML examples (and all others in this chapter) can easily be verified by using the W3C RDF Validation Service[25]. This allows the RDF/XML document to be checked and graphed.

The basic concepts and principles of RDF have now been introduced. The remainder of the chapter will now consider some other applications of RDF.

Friend-of-a-friend (FOAF)

Friend-of-a-friend (FOAF)[26] was one of the first applications of RDF and was originally designed as a Semantic Web version of a personal homepage[27]. FOAF is therefore designed to capture metadata about people. The FOAF vocabulary specification[28] provides a rich vocabulary to describe personal information (e.g. name, mailbox addresses, homepage URLs, blogs, etc.), as well as relationships with other people, groups, projects, and other affiliations.

The FOAF vocabulary defines classes (e.g. `foaf:Person`) and numerous properties (i.e. predicates), such as `foaf:name`, `foaf:knows`, `foaf:interests`, `foaf:depiction`, `foaf:weblog`, etc. Once published on the web (e.g. as RDF/XML), FOAF files can be processed by machines to establish relationships between people or organizations and the nature of these relationships. This data can then be used by computers to locate people or groups with similar interests, allow new entrants to a community to understand its structure, manage online personal identities via URIs, and a variety of other uses too numerous to list here[29]. FOAF's ability to characterize social relationships has also led to its use within online social network applications[30].

For example, we might want to state that there exists a person (`foaf:Person`) with the name 'George Macgregor' (`foaf:name`), who has:

- An e-mail address (`foaf:mbox`)
- A homepage (`foaf:homepage`)
- A blog (`foaf:weblog`)
- And, who knows (`foaf:knows`) another person (`foaf:Person`) with the name 'Emma McCulloch', who also has a homepage (`foaf:homepage`).

Such a 'social graph' could be expressed in FOAF RDF/XML as follows:

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">

<foaf:Person rdf:about="http://www.staff.ljmu.ac.uk/bsngmacg/#me">
 <foaf:name>Macgregor, George</foaf:name>
 <foaf:mbox rdf:resource="mailto:g.r.macgregor@ljmu.ac.uk"/>
 <foaf:homepage rdf:resource="http://www.staff.ljmu.ac.uk/bsngmacg/"/>
 <foaf:weblog rdf:resource="http://ljmuinfostrategy.blogspot.com/"/>
 <foaf:knows>
  <foaf:Person>
   <foaf:name>McCulloch, Emma</foaf:name>
   <foaf:homepage rdf:resource="http://cdlr.strath.ac.uk/people/mcculloche.htm"/>
  </foaf:Person>
```

```
 </foaf:knows>
</foaf:Person>

</rdf:RDF>
```

Recall that URIs can identify anything, even people. A URI has therefore been used in the above example to identify `foaf:Person` (i.e. George Macgregor). By assigning a URI we eliminate any ambiguity about which 'George Macgregor' is being referred to. Not only that, we enable others in the Semantic Web to refer unambiguously to this 'George Macgregor' rather than others with the same name. This URI could also be used to merge all other RDF data available on the web which happens to reference 'George Macgregor'. Where such a URI is missing, other mechanisms could be used (e.g. e-mail address).

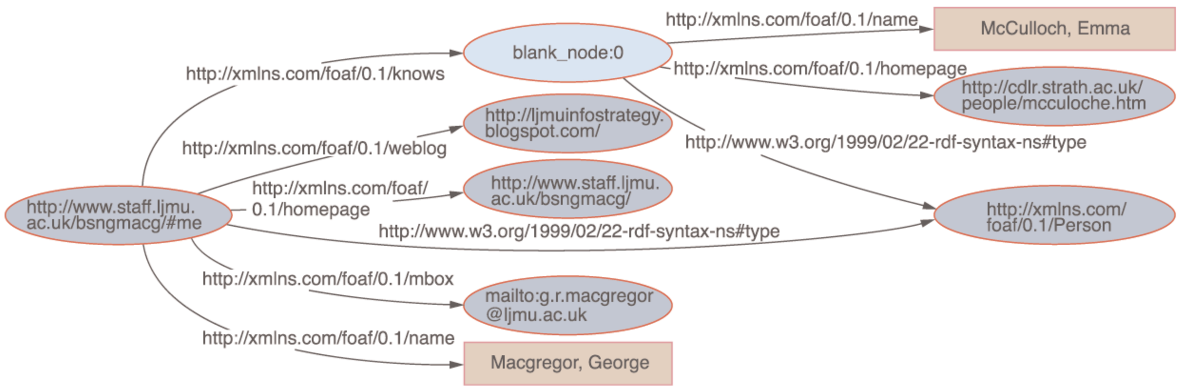

Figure 6. RDF graph of a FOAF file about 'George Macgregor'

Although the above FOAF RDF/XML example is relatively simple, we can observe from Figure 6 that the resulting RDF graph is already more complex than those featured earlier. A 'blank node' can also be observed in Figure 6. Blank nodes are common in RDF and are often unavoidable. Blank nodes essentially represent nodes which do not have a URI or literal (i.e. they are 'blank'). Such nodes therefore do not contain any data; instead they are used as parent nodes to group data together. For example, in the above example the FOAF RDF/XML essentially states that 'George Macgregor' knows a person whose name is 'Emma McCulloch' and who has a homepage. The `foaf:Person` of 'Emma McCulloch' is not uniquely identified by a URI. Since `foaf:Person` does not have its own URI, properties about 'Emma McCulloch' are grouped together using a blank node. This blank node mimics a URI and provides the necessary linkages between nodes within the RDF graph for it to make sense. In the absence of a URI, the software used to generate the RDF graph in Figure 6 has assigned a blank node identifier (`blank_node:0`). Blank node identifiers have no real meaning within RDF graphs other than allowing us to distinguish between other blank nodes within the same graph, thus most dedicated software applications (including the W3C RDF Validation Service) will assign identifiers automatically. Note that blank node identifiers only identify nodes within the same graph. If there is a need to merge multiple RDF graphs, or if others want to reference a blank node from outside the graph, then URIs have to be used instead.

Of course, it is possible to further augment this FOAF example with properties such as `foaf:gender`, `foaf:depiction`, `foaf:pastProject`, and so forth. More relationships can also be established (`foaf:knows`), as well as personal interests (`foaf:interest`) thus increasing the links within the social graph. The following example augments our FOAF RDF/XML with numerous properties and extra classes. Note also that the blank nodes resulting from the previous example have been resolved by assigning URIs to all instances of `foaf:Person`. The resulting RDF graph is too large to reproduce here but can be verified using the W3C RDF Validation Service:

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
```

```
<foaf:Person rdf:about="http://www.staff.ljmu.ac.uk/bsngmacg/#me">
 <foaf:name xml:lang="en">Macgregor, George</foaf:name>
 <foaf:firstName xml:lang="en">George</foaf:firstName>
 <foaf:surname xml:lang="en">Macgregor</foaf:surname>
 <foaf:gender>male</foaf:gender>
 <foaf:mbox rdf:resource="mailto:g.r.macgregor@ljmu.ac.uk"/>
 <foaf:homepage rdf:resource="http://www.staff.ljmu.ac.uk/bsngmacg/"/>
 <foaf:depiction
rdf:resource="http://www.staff.ljmu.ac.uk/bsngmacg/img/georgedepiction.jpg"/>
 <foaf:workplaceHomepage rdf:resource="http://www.ljmu.ac.uk/LBS/92624.htm"/>
 <foaf:publications rdf:resource="http://www.staff.ljmu.ac.uk/bsngmacg/pubs.html"/>
 <foaf:weblog rdf:resource="http://ljmuinfostrategy.blogspot.com/" dc:title="Information
Strategy Group, LJMU - Blog"/>
 <foaf:interest rdf:resource="http://www.w3.org/2004/02/skos/"/>
 <foaf:interest rdf:resource="http://hilt.cdlr.strath.ac.uk/"/>
 <rdfs:seeAlso rdf:resource="http://cdlr.strath.ac.uk/foaf/cdlr.rdf"/>
 <foaf:pastProject>
  <rdf:Description rdf:about="http://hilt.cdlr.strath.ac.uk/" dc:title="HILT: High-level
Thesaurus project phase IV">
  </rdf:Description>
 </foaf:pastProject>
 <foaf:knows>
  <foaf:Person rdf:about="http://www.staff.ljmu.ac.uk/bsnpkell/#me">
   <foaf:name>Kelly, Phil</foaf:name>
   <foaf:title>Dr</foaf:title>
   <foaf:homepage rdf:resource="http://www.ljmu.ac.uk/LBS/92623.htm"/>
  </foaf:Person>
 </foaf:knows>
 <foaf:knows>
 <foaf:Person rdf:about="http://cdlr.strath.ac.uk/people/mcculloche.htm#me">
  <foaf:name>McCulloch, Emma</foaf:name>
  <foaf:mbox rdf:resource="mailto:e.mcculloch@strath.ac.uk"/>
  <foaf:homepage rdf:resource="http://cdlr.strath.ac.uk/people/mcculloche.htm"/>
  <foaf:depiction rdf:resource="http://cdlr.strath.ac.uk/people/mcculloche.jpg"/>
 </foaf:Person>
 </foaf:knows>
 </foaf:Person>

</rdf:RDF>
```

Merging of RDF data is where FOAF is potentially of most use to digital libraries. For example, Dublin Core metadata (in RDF) about this chapter could be merged with FOAF metadata (in RDF), thus providing an enhanced metadata record containing rich authorship information. Malmsten[31] describes the use of a series of Semantic Web specifications to build a semantic digital library, in particular the use of FOAF to structure name authority files. A similar approach is demonstrated by Kruk et al.[32]. Their semantic digital library ('JeromeDL'[33]) uses FOAF to manage an authority file of authors, editors and publishers, but also uses FOAF to connect users and manage user profiles within their system[34]. JeromeDL deploys FOAFRealm[35], a FOAF-based technology developed by members of the same research team, to establish user identities[36]. FOAF is also used to offer novel resource discovery mechanisms described as 'social semantic collaborative filtering'[37]. For example, two colleagues will often share similar academic interests such that one might be able to find resources relevant to their information need within the profile of the other (e.g. resources held within virtual bookshelves, bookmarks, etc.).

Even less formal tools, such as those optimized for personal information management, increasingly deploy FOAF. BibSonomy[38], the social bookmark and publication management tool, exposes user profiles and interests via publicly available FOAF files, each providing personal information and subject interests which can be discovered by Semantic Web applications wishing to reuse bookmarks or publications stored and tagged by users. BibSonomy also exposes bookmarks in a variety of formats, including RDF/XML, XML, RSS and BibTeX.

Although RDF is optimized for machine processing, an increasing number of freely available tools can be used to explore FOAF files on the web[39, 40, 41]. Browser plug-ins for Mozilla Firefox[42] are also available[43], enabling the automatic extraction of FOAF data (and other RDF data) from web pages and their interrogation using a number of technologies.

Simple Knowledge Organization System (SKOS)

It was noted earlier that an important aim of the Semantic Web is to improve information retrieval and information organization on the web. SKOS[44] is an application of RDF designed to provide a data model for Knowledge Organization Systems (KOS) and is currently under active development by the W3C Semantic Web Deployment Working Group[45]. KOS – also referred to as controlled vocabularies or terminologies, and as 'concept schemes' by the SKOS specification – includes tools such as information retrieval thesauri, taxonomies, classification schemes, subject heading lists, and other forms of authority list or knowledge structure. It is therefore immediately understandable why SKOS will contribute to improvements in resource discovery, and practical examples of this will be discussed later.

SKOS is primarily designed to enable the publication of controlled vocabularies for use in the Semantic Web, thus enabling their machine interpretation to facilitate the retrieval and organization of resources. SKOS also enables KOS interoperability, data sharing, linking and data merging. The ability to merge and link SKOS with other data sources is consistent with RDF generally and enables SKOS data to be linked or merged by Semantic Web applications with other controlled vocabularies or subject indexes. This can be useful for a number of reasons, but particularly in retrieval circumstances where multiple collections have to be queried as it avoids the need for complex database integration[46].

An important Semantic Web specification in the area of knowledge modelling and representation is the W3C Web Ontology Language (OWL)[47]. Discussion of OWL can be complex and is therefore outside the scope of this chapter. Nevertheless, OWL assumes an important role in enabling intelligent software agents to infer and reason over knowledge captured in ontologies[48]; however, it is generally acknowledged that OWL is insufficient to fulfil the Semantic Web vision on its own and the "construction of detailed 'maps' of particular domains of knowledge"[49] are necessary, along with metadata. SKOS is therefore about harnessing LIS expertise in the area of knowledge organization to create these 'maps'. The large number of well-developed vocabularies already in use and under continual revision are well suited to achieving this. Additionally, SKOS enables the easy creation and publication of new vocabularies to fulfil emerging knowledge domains.

SKOS is very flexible and can accommodate most forms of KOS, with special provisions made for modelling arrays, notation and other features peculiar to controlled vocabularies. SKOS essentially consists of a series of classes and properties to express the structural characteristics of KOS. For example, a thesaurus would be a `skos:ConceptScheme` containing a series of `skos:Concepts`, each of which might have properties such as `skos:broader`, `skos:narrower`, `skos:related` and `skos:altLabel` (i.e. BT, NT, RT and UF respectively). Consider the following example taken from the *UNESCO* Thesaurus[50] for the concept, 'Information scientists':

Information scientists

SN A person who works on the theory or application of informatics or information science, i.e. analyses, designs, implements, etc. information systems

UF Information officers
BT Information/library personnel
RT Archive personnel
RT Information science education

Such a thesaurus concept could be expressed in SKOS RDF/XML as follows:

```
<?xml version="1.0" encoding="UTF-8"?>

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:skos="http://www.w3.org/2004/02/skos/core#">

<skos:Concept rdf:about="http://.../mt5.20/Informationscientists#concept">
 <skos:prefLabel xml:lang="en">Information scientists</skos:prefLabel>
 <skos:scopeNote xml:lang="en">A person who works on the theory or application of
informatics or information science, i.e. analyses, designs, implements etc. information
systems.</skos:scopeNote>
 <skos:altLabel xml:lang="en">Information officers</skos:altLabel>
 <skos:broader rdf:resource="http://.../mt5.20/Informationlibrarypersonnel#concept"/>
 <skos:related rdf:resource="http://.../mt5.20/Archivepersonnel#concept"/>
 <skos:related rdf:resource="http://.../mt1.50/Informationscienceeducation#concept"/>
</skos:Concept>

</rdf:RDF>
```

The above example produces the RDF graph given in Figure 7. Note that URIs have been used to identify the concepts within the KOS. At time of writing, the *UNESCO Thesaurus* remains unpublished for the Semantic Web so the URIs in the above example are merely illustrative. Increasingly, vocabularies published in SKOS infer their structure or use their notation within URI. The micro-thesaurus notation from the *UNESCO Thesaurus* has therefore been incorporated into the URI. This approach to 'minting' URIs is consistent with the 'Cool URI' trend within the Semantic Web community[51]; an attempt to maintain the purpose of a URI in uniquely identifying resources (in their various permutations) whilst simultaneously making them more meaningful than simply a random sequence of characters. The significance of minting Cool URIs has recently attracted wider discussion and research by SKOS researchers. For example, Panzer discusses the minting of URIs for publishing DDC for the Semantic Web[52], whilst Summers et al. discuss URIs in their conversion of LCSH from MARCXML to SKOS[53].

Figure 7. RDF graph for the UNESCO Thesaurus concept (skos:Concept), 'Information scientists'

Recall that in our DC RDF/XML example (Figure 5), the subject of our resource (`dcterms:subject`) was indicated by the LCSH descriptor, 'Semantic Web'; however, rather than identify this descriptor by using a literal we elected to identify the concept by URI (`http://id.loc.gov/authorities/sh2002000569`). This 'concept URI' not only defines the concept of the 'Semantic Web' and the preferred lexical label, but points to rich terminological data (e.g. the PT, BT, RT, SN, etc.) expressed in SKOS by the Library of Congress Authorities & Vocabularies service[54], thus enabling information retrieval which is less dependent on free-text searching and more concerned with the representation of *concepts*. Indeed, it is possible for concept definitions (i.e. URIs) to be reused with alternative lexical labels. This ethos is central to SKOS (and the Semantic Web generally) and forms part of the 'linked data' principle[55, 56]: exposing and reusing RDF data and URIs to maximize data connections and relationships in a manner which is useful to both humans and machines. In essence then, linked data is about creating connections between data which previously may not have existed and exposing this data for sharing on the Semantic Web by using URIs and RDF. Tim Berners-Lee has noted that linked data is essential to connect the components of the Semantic Web[57]. The more connections there are between data, the greater the value and usefulness of that data, thus allowing humans and machines to follow semantic threads across disparate data sources (using URIs). The linked data approach holds great potential for SKOS as it allows "concepts from different concept schemes [to be] connected together […] to form a distributed, heterogeneous global concept scheme. A web of concept schemes can serve as the foundation for new applications that allow meaningful navigation between KOSs"[58].

More generally, the use of SKOS makes it easier to design distributed information retrieval systems because the identification of concepts is based on concept URIs and structured according to KOS rules in RDF. For example, upon retrieving a resource via subject searching, a system could be designed to retrieve other resources on the Semantic Web identified in the same way, thus improving recall whilst maintaining a level of precision. This can be a particularly useful mechanism given the distributed and decentralized nature of resource publication on the web. Since a concept URI links to a detailed description of the concept (e.g. its preferred label, BT, NT, RT, etc.), it is also possible to reuse this data to provide extra retrieval aids for the user. For example, broader and related terms could be used to deliver query expansion search techniques[59], or the terms could be displayed to assist the user in refining their search query, perhaps allowing the user to browse the KOS hierarchically. Visual search interfaces could be created showing the relationships between concepts (e.g. based on the RDF graph), for example see the Library of Congress Authorities & Vocabularies[60]. Software could also be designed to enable users to browse concept schemes and retrieve resources identified using its concept URIs.

Some of the aforementioned techniques have been demonstrated by the Explicator project[61]. Gray et al. demonstrate a web service for searching and exploring concepts within SKOS-encoded astronomical vocabularies[62]. Their 'Vocabulary Explorer' web application enables users to traverse astronomical concepts and formal scientific definitions, their relationships to other concepts, and their relationships with similar concepts in alternative vocabularies. Further work undertaken by the same research team demonstrates how rich semantic relationships within SKOS can be exploited to improve retrieval precision and deliver a variety of searching aids for users[63].

Another interesting feature of SKOS is its ability to capture mappings between concepts in different concept schemes. This can be useful where problems of semantic heterogeneity exist (i.e. a collection is using more than one vocabulary to index resources). To accommodate such scenarios, SKOS provides properties such as `skos:closeMatch`, `skos:exactMatch`, `skos:broadMatch`, etc. These properties can be used to state a conceptual link between SKOS concepts in different concept schemes, thus ameliorating the vocabulary mis-match difficulties which often arise in distributed contexts, or where several heterogeneous collections are merged. For example, Isaac et al. report on the use of SKOS to resolve semantic heterogeneity within digitized cultural heritage collections[64]. They use their methods of 'semantic alignment' to create mappings between different concept schemes, thereby providing users with integrated access to resources which have been indexed using a number of different vocabularies.

An increased need to deliver KOS data (with mappings) in a web services context has emerged in recent years. Such web services are considered necessary to effect improvements in digital library searching functionality and/or to offer users the option of searching multiple third-party repositories indexed using disparate vocabularies. Use of SKOS within a web services context has unsurprisingly attracted attention. For example, the STAR project[65] has created a series of pilot Semantic Web services for KOS data based on SKOS, providing term look-up functionality, browsing and semantic concept expansion[66]. Macgregor et al.[67] demonstrate the use of SKOS in a web services context as part of the High-Level Thesaurus (HILT) project[68]. Their 'terminology mapping server' uses SKOS to structure terminological data (including mappings via a DDC spine) when responding to SRW/U requests from digital libraries. Similar work is also being conducted by the Deutsche Nationalbibliothek[69].

Of course, almost all of the aforementioned is entirely dependent upon KOS being published for the Semantic Web in SKOS. Although SKOS is currently a W3C 'candidate recommendation', several well-known vocabularies have already been made officially available in SKOS for use on the Semantic Web, such as LCSH[70], *STW Thesaurus for Economics*[71], AGROVOC[72], and GEMET[73]. Many others have been temporarily published in SKOS, but these lack provenance and stability owing to their use within research experiments.

RDFa

RDF specifications such as FOAF, SKOS, OWL, and even Dublin Core RDF, necessitate understanding of the underlying RDF data model, as well as knowledge of the various RDF serializations. Such applications of RDF are typically made available independently of the resource(s) they are describing or associated with (i.e. as a separate file).

More recently the W3C has introduced RDFa (Resource Description Framework in attributes)[74]. RDFa provides a series of XHTML[75] extensions which can be used to annotate web pages with semantic data. As the official RDFaWiki[76] and RDFa Primer[77] indicate, RDFa is a simple way of embedding RDF statements within XHTML and an attempt to encourage publishers, bloggers, web developers and the like to participate in the development of the Semantic Web. RDFa enables simple semantic data to be encoded without detailed knowledge of RDF or the need for separate RDF files containing detailed RDF/XML or other RDF serializations. In fact, knowledge of XHTML is the only prerequisite to deploying RDFa in practice, although more detailed applications of RDFa would obviously benefit from a wider knowledge of RDF.

Consider the following snippet of 'vanilla' XHTML. This example represents what typical XHTML might look like in a fictional web page publishing this chapter at the MetaPress domain (`http://uksg.metapress.com/someURL`):

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
"http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en">

<head>
<title>E-Resource management and the Semantic Web: applications for RDF for e-resource
discovery</title>
<meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1"/>
</head>

<body>
<h1>E-Resource management and the Semantic Web: applications for RDF for e-resource
discovery</h1>
<p>George Macgregor</p>
<p>16 April 2009</p>
<p>Keywords: Semantic Web, digital libraries</p>

<h2>Abstract</h2>
<p>Semantic Web technologies and specifications are increasingly finding applications
```

```
within digital libraries and other e-resource contexts. The purpose of this chapter is to
... within a variety of e-resource discovery contexts will be discussed.</p>
<h2>About the author</h2>
<p>George Macgregor is currently a Lecturer in Information Management and a <a
href="http://www.ljmu.ac.uk/LBS/92624.htm">member</a> of the Information Strategy Group at
<a href="http://www.ljmu.ac.uk/LBS/Index.htm">Liverpool Business School</a>, <a
href="http://www.ljmu.ac.uk/">Liverpool John Moores University</a>. George helps maintain
the <a href="http://ljmuinfostrategy.blogspot.com/">Information Strategy Group
blog.</a></p>

</body>
</html>
```

The above example is an instance of how the web has evolved to provide a series of documents conducive to human interpretation, but has failed to capture the semantics of these documents for machine interpretation. This web page does little to assist machines in interpreting who the creator of the chapter is, or even what its title is. Humans know who the creator is and what the title is, but only because this is loosely inferred by the page structure when the file is viewed in a web browser.

To embed semantics we could use RDFa to annotate the XHTML (XHTML+RDFa) by embedding the necessary RDF triples. For example, we could annotate the previous example by extending the XHTML to include Dublin Core and FOAF. The relevant RDFa extensions are visible in bold font:

```
<html
  xmlns="http://www.w3.org/1999/xhtml" version="XHTML+RDFa 1.0"
  xml:lang="en"
  xmlns:dcterms=http://purl.org/dc/elements/1.1/
  xmlns:foaf="http://xmlns.com/foaf/0.1/">

<head>
<title>E-Resource management and the Semantic Web: applications for RDF for e-resource
discovery</title>
<base href="http://uksg.metapress.com/someURL" />
</head>

<body>
<h1 property="dcterms:title">E-Resource management and the Semantic Web: applications for
RDF for e-resource discovery</h1>

<p><span rel="dcterms:creator"><span about="http://www.staff.ljmu.ac.uk/bsngmacg/#me"
typeof="foaf:Person">George Macgregor</span></span></p>
<p><span property="dcterms:date" content="2009-04-16">16 April 2009</span></p>
<p>Keywords: <span rel="dcterms:subject"
resource="http://id.loc.gov/authorities/sh2002000569">Semantic Web</span>, <span
rel="dcterms:subject" resource="http://id.loc.gov/authorities/sh95008857">digital
libraries</span></p>

<h2>Abstract</h2>
<p property="dcterms:abstract">Semantic Web technologies and specifications are
increasingly finding applications within digital libraries and other e-resource contexts.
The purpose of this chapter is to ... within a variety of e-resource discovery contexts
will be discussed.</p>

<h2>About the author</h2>
<p about="http://www.staff.ljmu.ac.uk/bsngmacg/#me" typeof="foaf:Person"><span
property="foaf:name">George Macgregor</span> is currently a Lecturer in Information
Management and a <a rel="foaf:workPlaceHomePage"
```

```
href=”http://www.ljmu.ac.uk/LBS/92624.htm”>member</a> of the Information Strategy Group at
<a rel=”foaf:workInfoHomePage” href=”http://www.ljmu.ac.uk/LBS/Index.htm”><span
property=”dc:title”>Liverpool Business School</span></a>, <a
href=”http://www.ljmu.ac.uk/”><span property=”dc:title”>Liverpool John Moores
University</span></a>. George helps maintain the <a rel=”foaf:weblog”
href=”http://ljmuinfostrategy.blogspot.com/”><span property=”dc:title”>Information Strategy
Group blog</span>.</a> </p>

</body>
</html>
```

One of the advantages of XHTML+RDFa is that it allows semantics to be embedded within running text. This is clearly demonstrated in the paragraph providing biographical information about the author. `foaf:Person` has been used to identify the author and other FOAF and Dublin Core properties have been used.

RDFa Distiller[78] is a W3C tool for ‘scraping’ RDF triples from XHTML+RDFa web pages and for outputting them in standalone RDF serializations (e.g. RDF/XML). By using RDFa Distiller on the above XHTML+RDFa we can observe in the example below that the relevant triples have been extracted and structured in RDF/XML, and in a manner not dissimilar to examples earlier in this chapter. This example generates the RDF graph in Figure 8 and is easier to decipher:

```
<?xml version=”1.0” encoding=”utf-8”?>
<rdf:RDF
 xmlns:dcterms=”http://purl.org/dc/elements/1.1/”
 xmlns:foaf=”http://xmlns.com/foaf/0.1/”
 xmlns:rdf=”http://www.w3.org/1999/02/22-rdf-syntax-ns#”
 xmlns:rdfs=”http://www.w3.org/2000/01/rdf-schema#”
 xmlns:xhv=”http://www.w3.org/1999/xhtml/vocab#”
 xmlns:xml=”http://www.w3.org/XML/1998/namespace”
>
 <rdf:Description rdf:about=”http://uksg.metapress.com/someURL”>
  <dcterms:abstract xml:lang=”en”>Semantic Web technologies and specifications are
increasingly finding applications within digital libraries and other e-resource contexts.
The purpose of this chapter is to ... within a variety of e-resource discovery contexts
will be discussed.</dcterms:abstract>
  <dcterms:creator>
   <foaf:Person rdf:about=”http://www.staff.ljmu.ac.uk/bsngmacg/#me”>
    <foaf:workPlaceHomePage rdf:resource=”http://www.ljmu.ac.uk/LBS/92624.htm”/>
    <foaf:workInfoHomePage rdf:resource=”http://www.ljmu.ac.uk/LBS/Index.htm”/>
    <foaf:name xml:lang=”en”>George Macgregor</foaf:name>
    <foaf:weblog rdf:resource=”http://ljmuinfostrategy.blogspot.com/”/>
   </foaf:Person>
  </dcterms:creator>
  <dcterms:subject rdf:resource=”http://id.loc.gov/authorities/sh2002000569”/>
  <dcterms:subject rdf:resource=”http://id.loc.gov/authorities/sh95008857”/>
  <dcterms:date xml:lang=”en”>2009-04-16</dcterms:date>
  <dcterms:title xml:lang=”en”>E-Resource management and the Semantic Web: applications for
RDF for e-resource discovery</dcterms:title>
 </rdf:Description>
</rdf:RDF>
```

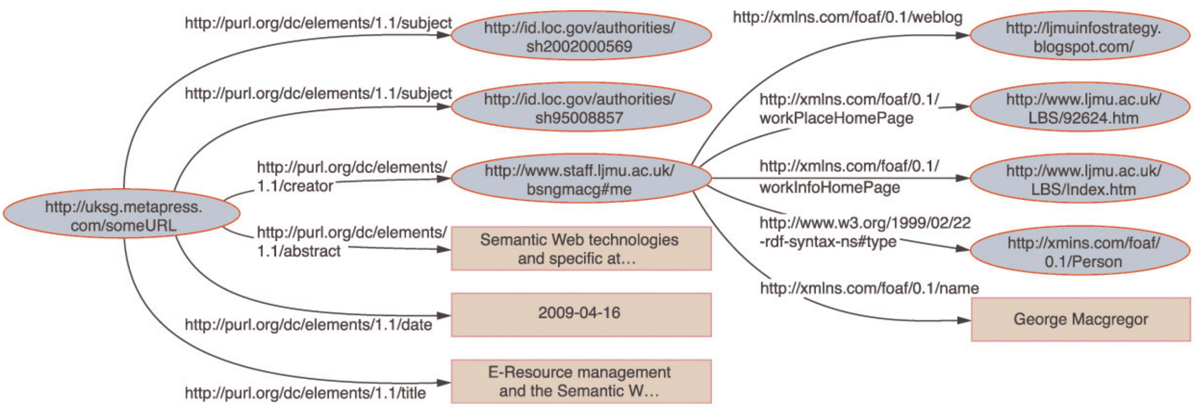

Figure 8. RDF graph of fictional XHTML+RDFa, based on triples extracted using RDFa Distiller

RDFa remains a relatively new Semantic Web standard and only received W3C recommendation status in late 2008. Implementations have therefore been predominantly confined to those offered by the W3C. Nevertheless, large scale implementations within digital libraries are already visible. Neubert[79] describes the publication of the *STW Thesaurus for Economics*[80] for the Semantic Web. *STW* is a richly interconnected multilingual thesaurus (English and German) accommodating subjects within the economics and business-related disciplines. It provides 'topical entry points' to the German National Library of Economics (ZBW)[81] digital library and aims to provide an economics and business hub within the web of linked data. *STW* is delivered as XHTML+RDFa pages (using Dublin Core, SKOS, OWL and others), with searching and concept tree browsing functionality offered in the interface. A standalone SKOS RDF/XML dump version can also be downloaded.

Conclusion

This chapter has attempted to introduce the key Semantic Web concepts and its principal enabling language using a series of practical examples. As we have noted, applications of RDF, such as Dublin Core, FOAF and SKOS, have clear applications within e-resource discovery contexts and their increased deployment can effect improvements in information retrieval and enable the delivery of other information tools for users. They also enable a level of improved data sharing, linking, merging and interoperability which can enrich the structured data already managed by digital libraries, thus contributing to the web of 'linked data' and better exposing invaluable e-resources. The benefits of interacting, contributing and maintaining the structured data required to support the Semantic Web have been recognized by information professionals and the increased deployment of Semantic Web techniques within digital libraries has proliferated. Fulfilling the vision of the Semantic Web for those outside the information profession is an immense task owing to the lack of structured data available with which to work. It is therefore appropriate that digital libraries and repositories assume an increased responsibility in bringing the Semantic Web vision to fruition.

References

1. Ed. Greenberg, J and Méndez, E, *Knitting the Semantic Web*, 2007, Binghamton, The Haworth Press.
2. Ed. Macgregor, G, Digital libraries and the Semantic Web: context, applications and research, *Library Review (Special issue)*, 2008, 57(3).
3. International Conferences on Digital Libraries and the Semantic Web (ICDC): http://www.icsd-conference.org/ (Accessed 16 April 2009)
4. Ed. Miles, A and Bechhofer, S, *SKOS Simple Knowledge Organization System Reference. W3C Candidate Recommendation 17 March 2009*, 2009, Cambridge, MA, W3C. http://www.w3.org/TR/2009/CR-skos-reference-20090317/ (Accessed 16 April 2009)

5. Berners-Lee, T, Hendler, J and Lassila, O, The Semantic Web, *Scientific American*, 2001, May.
   http://www.sciam.com/article.cfm?id=the-semantic-web (Accessed 16 April 2009)
6. Shadbolt, N, Berners-Lee, T and Hall, W, The Semantic Web Revisited, *IEEE Intelligent Systems*, 2006, 21(3), 96–101.
7. Guha, R, McCool, R and Miller, E, Semantic search. In: *Proceedings of the 12th International Conference on the World Wide Web, 20–24 May, Budapest, Hungary*, Ed. Hencsey, G and White, B, 2003, New York, ACM, 700–709.
8. Berners-Lee, T et al., ref. 5.
9. Ed. Smith, M K, Welty, C and McGuinness, D L, *OWL Web Ontology Language Guide*, 2004, MA, W3C.
   http://www.w3.org/TR/owl-guide/ (Accessed 16 April 2009)
10. Berners-Lee, T, Q&A with Tim Berners-Lee, *Business Week*, 2007, 9 April.
    http://www.businessweek.com/technology/content/apr2007/tc20070409_961951.htm (Accessed 16 April 2009)
11. W3C Semantic Web Activity:
    http://www.w3.org/2001/sw/ (Accessed 3 August 2009)
12. Ed. Klyne, G and Carroll, J J, *Resource Description Framework (RDF): concepts and abstract syntax*, 2004, Cambridge, MA, W3C.
    http://www.w3.org/TR/rdf-concepts/ (Accessed 16 April 2009)
13. Thompson, H S, *What's a URI and why does it matter?*, 2008, Edinburgh, University of Edinburgh / W3C.
    http://www.ltg.ed.ac.uk/~ht/WhatAreURIs/ (Accessed 16 April 2009)
14. Ed. Beckett, D, *RDF/XML syntax specification (revised). W3C recommendation 10 February 2004*, 2004, Cambridge, MA, W3C.
    http://www.w3.org/TR/rdf-syntax-grammar/ (Accessed 16 April 2009)
15. Berners-Lee, T, *Primer: Getting into RDF & Semantic Web using N3*, 2005, MA, W3C.
    http://www.w3.org/2000/10/swap/Primer.html (Accessed 16 April 2009)
16. Ed. Beckett, D and Berners-Lee, T, *Turtle - Terse RDF Triple Language*, 2008, MA, W3C.
    http://www.w3.org/TeamSubmission/turtle/ (Accessed 16 April 2009)
17. Ed. Bray, T, Paoli, J, Sperberg-McQueen, C M, Maler, E and Yergeau F, *Extensible Markup Language (XML) 1.0 (Fifth edition)*, 2008, Cambridge, MA, W3C.
    http://www.w3.org/TR/2008/REC-xml-20081126/ (Accessed 16 April 2009)
18. W3C Semantic Web Activity, ref. 11.
19. Dublin Core Metadata Initiative:
    http://dublincore.org/ (Accessed 16 April 2009)
20. Nilsson, M, Powell, A, Johnston, P and Naeve, A, *Expressing Dublin Core metadata using the Resource Description Framework (RDF)*, 2008, Seoul, Dublin Core Metadata Initiative.
    http://www.dublincore.org/documents/dc-rdf/ (Accessed 16 April 2009)
21. Creative Commons:
    http://creativecommons.org/ (Accessed 16 April 2009)
22. Authorities & Vocabularies (Library of Congress):
    http://id.loc.gov/authorities/ (Accessed 30 April 2009)
23. Morla:
    http://www.morlardf.net/ (Accessed 16 April 2009)
24. Altova Semantic Works 2009:
    http://www.altova.com/products/semanticworks/semantic_web_rdf_owl_editor.html (Accessed 16 April 2009)
25. W3C RDF Validation Service:
    http://www.w3.org/RDF/Validator/ (Accessed 16 April 2009)
26. The FOAF project:
    http://www.foaf-project.org/ (Accessed 16 April 2009)

27. Graves, M, Constabaris, A and Brickley, D, FOAF: Connecting people on the Semantic Web, *Cataloging & Classification Quarterly*, 2007, 43 (3/4), 191–202.

28. Brickley, D and Miller, L, *FOAF Vocabulary Specification 0.91: Namespace Document 2 November 2007 – OpenID Edition*, 2007.
http://xmlns.com/foaf/spec/ (Accessed 16 April 2009)

29. Ding, L, Zhou, L, Finin, T and Joshi, A, How the Semantic Web is being used: an analysis of FOAF documents. In: *Proceedings of the 38th Annual Hawaii International Conferences on System Sciences (HICSS'05), 3–6 January, Hawaii, USA*, 2005, IEEE Computer Society, USA.
http://ebiquity.umbc.edu/_file_directory_/papers/120.pdf (Accessed 16 April 2009)

30. Golbeck, J and Rothstein, M, Linking social networks on the web with FOAF: A Semantic Web case study. In: *Proceedings of the Twenty-third AAAI Conference on Artificial Intelligence National Conference on Artificial Intelligence, 13–17 July, Chicago, USA*, 2008, California, AAAI Press, 1138–1143.
http://www.cs.umd.edu/~golbeck/downloads/foaf.pdf (Accessed 16 April 2009)

31. Malmsten, M, Making a library catalogue part of the Semantic Web. In: *Proceedings of the International Conference on Dublin Core and Metadata Applications: Metadata for semantic and social applications, 22–26 September 2008, Berlin, Germany*, Ed. Greenberg, J and Wolfgang, K, 2008, Go_ttingen, Go_ttingen Univ.-Verl.
http://dcpapers.dublincore.org/ojs/pubs/article/view/927/923 (Accessed 16 April 2009)

32. Kruk, S R, Woroniecki, T, Gzella, A and Dπbrowski, M, JeromeDL – a semantic digital library. In: *Proceedings of the Semantic Web Challenge Workshop, 6th International Semantic Web Conference, 11–15 November 2007, Busan, South Korea*, 2007.
http://library.deri.ie/resource/81n2JzC5 (Accessed 16 April 2009)

33. JeromeDL:
http://www.jeromedl.org/ (Accessed 16 April 2009)

34. Kruk, S R, Zimmerman, K and Sapkota, B, Semantically enhanced search services in digital libraries. In: *Proceedings of the International Conference on Internet and Web Applications and Services (ICIW 2006), 23–25 February, Guadeloupe, French Caribbean*, 2006.
http://library.deri.ie/resource/iwqOhveX (Accessed 16 April 2009)

35. FOAFRealm:
http://www.foafrealm.org/ (Accessed 16 April 2009)

36. Kruk, S R, et al, ref. 34.

37. Kruk, S R, Decker, S, Gzella, A and Grzonkowski, S, Social Semantic Collaborative Filtering, *Journal of Web Semantics*, 2008 (pre-print).
http://library.deri.ie/resource/790k6n8a (Accessed 16 April 2009)

38. BibSonomy:
http://www.bibsonomy.org/ (Accessed 16 April 2009)

39. FOAF Explorer:
http://xml.mfd-consult.dk/foaf/explorer/ (Accessed 16 April 2009)

40. Semantically Interlinked Online Communities (SIOC) Browser:
http://sparql.captsolo.net/browser/browser.py (Accessed 16 April 2009)

41. FOAF ('friend-of-a friend') Database:
http://swordfish.rdfweb.org/rweb/who (Accessed 16 April 2009)

42. Mozilla Firefox:
http://www.mozilla-europe.org/en/firefox/ (Accessed 16 April 2009)

43. Semantic Radar for Firefox:
http://sioc-project.org/firefox

44. Ed. Miles, A and Bechhofer, S, ref. 4.

45. W3C Semantic Web Deployment Working Group:
http://www.w3.org/2006/07/SWD/ (Accessed 16 April 2009)

46. Miles, A and Pérez-Agüera, J R, SKOS: Simple Knowledge Organisation for the Web, *Cataloging & Classification Quarterly*, 2007, 43(3/4), 69–83.
47. Ed. Dean, M and Schreiber, G, *OWL Web Ontology Language Reference*, 2004, Cambridge, MA, W3C. http://www.w3.org/TR/owl-ref/ (Accessed 16 April 2009)
48. Shadbolt, N, et al, ref. 6.
49. Ed. Miles, A and Bechhofer, S, ref. 4.
50. UNESCO Thesaurus: http://databases.unesco.org/thesaurus/ (Accessed 16 April 2009)
51. Ed. Sauerman, L and Cyganiak, R, *Cool URIs for the Semantic Web - W3C Working Draft, 17 December 2007*, 2007, Cambridge, MA, W3C. http://www.w3.org/TR/2007/WD-cooluris-20071217/ (Accessed 16 April 2009)
52. Panzer, M, Cool URIs for the DDC: Towards Web-scale accessibility of a large classification system. In: *Proceedings of the International Conference on Dublin Core and Metadata Applications: Metadata for semantic and social applications, 22-26 September 2008, Berlin, Germany*, Ed. Greenberg, J and Wolfgang, K, 2008, Göttingen, Göttingen Univ.-Verl. http://dcpapers.dublincore.org/ojs/pubs/article/viewFile/932/928 (Accessed 16 April 2009)
53. Summers, E, Isaac, A, Redding, C and Krech, D, LCSH, SKOS and linked data. In: *Proceedings of the International Conference on Dublin Core and Metadata Applications: Metadata for semantic and social applications, 22–26 September 2008, Berlin, Germany*, Ed. Greenberg, J and Wolfgang, K, 2008, Göttingen, Göttingen Univ.-Verl. http://dcpapers.dublincore.org/ojs/pubs/article/view/916/912 (Accessed 16 April 2009)
54. Authorities & Vocabularies, ref. 22.
55. Berners-Lee, T, linked data, 2007, Cambridge, MA, W3C. http://www.w3.org/DesignIssues/LinkedData.html (Accessed 16 April 2009)
56. Linked Data - Connect Distributed Data across the Web: http://linkeddata.org/ (Accessed 16 April 2009)
57. Berners-Lee, T, ref. 55.
58. Ed. Miles, A and Bechhofer, S, ref. 4.
59. Efthimiadis, E N, Query expansion, *Annual Review of Information Systems and Technology (ARIST)*, 31, 121–187. http://faculty.washington.edu/efthimis/pubs/Pubs/qe-arist/QE-arist.html (Accessed 16 April 2009)
60. Authorities & Vocabularies, ref. 22.
61. Explicator project: http://explicator.dcs.gla.ac.uk/ (Accessed 16 April 2009)
62. Gray, A J G, Gray, N and Ounis, I, Searching and exploring controlled vocabularies. In: *Proceedings of Exploiting Semantic Annotations in Information Retrieval (ESAIR 2009), 09 February, Barcelona, Spain*, 2009, New York, ACM, 1–5. http://www.zaragozas.info/esair/final_6.pdf (Accessed 16 April 2009)
63. Gray, A J G, Gray, N and Ounis, I, Finding data resources in a virtual observatory using SKOS vocabularies. In: *Proceedings of 25th British national conference on databases: sharing data, information and knowledge, 07–10 July, Cardiff, UK*, 2008, Berlin, Springer-Verlag, 189–192.
64. Isaac, A, Schlobach, S, Matthezing, H and Zinn, C, Integrated access to cultural heritage resources through representation and alignment of controlled vocabularies, *Library Review*, 2008, 57(3), 187–199.
65. Semantic Technologies for Archaeological Resources project (STAR): http://hypermedia.research.glam.ac.uk/kos/STAR/ (Accessed 16 April 2009)
66. Tudhope, D and Binding, C, Experiences with Knowledge Organization System services from the STAR Project, *Signum*, 2008, 5. http://pro.tsv.fi/STKS/signum/200805/3.pdf (Accessed 16 April 2009)
67. Macgregor, G, McCulloch, E and Nicholson, D, Terminology server for improved resource discovery: analysis of functions and model. In: *Proceedings of the 2nd International Conference on Metadata and Semantics Research, 11–12*

October 2007, Ionian Academy, Corfu, Greece, Ed. Miguel-Angel, S and Lytras, M D, 2007, Corfu, Ionian Academy.
http://www.mtsr.ionio.gr/proceedings/mcculloch.pdf (Accessed 16 April 2009)

68. High-Level Thesaurus (HILT) project:
http://hilt.cdlr.strath.ac.uk/ (Accessed 16 April 2009)
69. Svennson, L G, National libraries and the Semantic Web: requirements and applications. In: *Proceedings of the International Conference on Semantic Web and Digital Libraries (ICSD-2007), 21–23 February 2007, Bangalore, India*, Ed. Prasad, A R D and Madalli, D P, 2007, Bangalore, DRTC, 101–108.
70. Authorities & Vocabularies, ref. 22.
71. *STW Thesaurus for Economics*:
http://zbw.eu/stw/versions/8.04/about.en.html (Accessed 16 April 2009)
72. AGROVOC (SKOS):
http://www.fao.org/aims/ag_download.htm (Accessed 16 April 2009)
73. GEneral Multilingual Environmental Thesaurus (GEMET):
http://www.eionet.europa.eu/gemet/webservices?langcode=en (Accessed 16 April 2009)
74. Ed. Adida, B, Birbeck, M, McCarron, S and Pemberton, S, *RDFa in XHTML: syntax and processing: A collection of attributes and processing rules for extending XHTML to support RDF. W3C recommendation 14 October 2008*, 2008, Cambridge, MA, W3C.
http://www.w3.org/TR/rdfa-syntax/ (Accessed 16 April 2009)
75. Ed. Axelsson, J, Birbek, M, Dubinko, M, Epperson, B, Ishikawa, M, McCarron, S, Navarro, A and Perberton, S, *XHTML 2.0*, 2006, Cambridge, MA, W3C.
http://www.w3.org/TR/xhtml2/ (Accessed 16 April 2009)
76. RDFaWiki:
http://rdfa.info/wiki/RDFa_Wiki (Accessed 16 April 2009)
77. Ed. Adida, B and Birbeck, M, *RDFa primer: Bridging the human and data webs*, 2008, Cambridge, MA, W3C.
http://www.w3.org/TR/xhtml-rdfa-primer/ (Accessed 16 April 2009)
78. RDFa Distiller:
http://www.w3.org/2007/08/pyRdfa/ (Accessed 16 April 2009)
79. Neubert, J, Bringing the 'Thesaurus for Economics' on to the Web of linked data. In: *Proceedings of linked data on the Web 2009 (LDOW2009), 20 April 2009, Madrid, Spain*, Ed. Bizer, C, Heath, T, Berners-Lee, T and Idehen, K, 2009, CEUR-WS.
http://events.linkeddata.org/ldow2009/papers/ldow2009_paper7.pdf (Accessed 16 April 2009)
80. *STW Thesaurus for Economics*, ref. 71.
81. German National Library of Economics (ZBW):
http://www.zbw-kiel.de/index-e.html (Accessed 16 April 2009)

■ **George Macgregor**
Information Strategy Group
Information Management & Systems
Liverpool Business School
Liverpool John Moores University, John Foster Building
98 Mount Pleasant, Liverpool L3 5UZ, UK.
E-mail: g.r.macgregor@ljmu.ac.uk

Biographical note

George Macgregor is a lecturer in Information Management and a member of the Information Strategy Group based in the Information Management & Systems section of Liverpool Business School, Liverpool John Moores University. Prior to this George was a Research Fellow based in the Centre for Digital Library Research at the University of Strathclyde. His research interests lie in the areas of distributed digital libraries, metadata, networked Knowledge Organization Systems (KOS), and the use of Semantic Web technologies within digital libraries and repositories.