

# A Middleware Registry for the Discovery of Collections and Services

Ann Apps

MIMAS, The University of Manchester, M13 9PL, UK

ann.apps@manchester.ac.uk

**Abstract.** The JISC Information Environment Service Registry (IESR) publicises collections of resources applicable to UK academia, including many relevant to Social Scientists, along with details of how to access them, in a machine-readable format, aiming to assist applications such as portals to serve their users' interests. This paper describes the content of IESR, the metadata scheme used for its capture and the services providing its dissemination. Some possible uses of IESR are suggested, and its position within the information environment, both UK and global, is discussed.

## Introduction

The JISC Information Environment Service Registry (IESR, 2005) contains information about collections of resources that the JISC (Joint Information Systems Committee) makes available to researchers, learners and teachers within UK Higher and Further Education. In addition, IESR contains details of how to access technical services, both those that make the collections available, and other services that play a significant role in the information environment. The collections cover all disciplines, but include many relevant to Social Scientists.

The aim of IESR is to assist other applications, such as portals, virtual learning applications or research services, in the discovery and subsequent use of materials that are relevant to their users' interests. It is a middleware, shared service, initially a single central registry within the JISC Information Environment (JISC, 2003a), but potentially within a much wider landscape. It is primarily intended for machine-to-machine access, but ultimately should benefit end-users by facilitating more awareness of, and easier access to, relevant resources, for example through the single point of search that a portal provides.

## IESR Content

Resources are described in IESR as collections, the intention being to provide to another application discovery of collections of datasets covering a topic matching a researcher's interests. The more specific discovery of an individual dataset would be part of the functionality of a Social Science portal with which the user interacts directly. At present collections described in IESR include: those hosted by the JISC Data Centres, such as UK Census Datasets, ESDS International Macro data, and the Zetoc citation database at MIMAS, and Digimap and UKBORDERS at EDINA; Survey Datasets provided by the UK Data Archive; and the Resource Discovery Network subject hubs. Future content is likely to include electronic journals, library catalogues, e-learning resources and institutional repositories. It is probable that the IESR remit will be widened to include further content

relevant to UK practitioners, for example resources within the Common Information Environment (Miller, 2004), which includes the museums, archives and libraries communities.

IESR also contains descriptions of the machine-level services that provide access to a collection, known as 'informational' services, and all agents associated with the collections or the services. A further set of described services are broker services, called 'transactional' services. These do not provide access to explicit collections but play a significant role within the information environment, an example being an OpenURL resolver. The word 'service' is rather overloaded. Within IESR context the term 'service' refers to a technical, single access point to a resource collection, rather than to a 'JISC service', which roughly corresponds to an IESR collection.

Data is supplied by the resource providers, thus assuring its accuracy, with a further quality check by the IESR content manager. The benefit to a content provider of registering their resource with IESR is that its consequent publicity is anticipated to bring additional use.

## IESR Data Description

Resources are described within IESR by metadata according to a scheme that is specific to IESR, but open standards based. The IESR data model comprises three types of entity: a collection; a technical, machine service that is either 'informational', providing a single access point to a collection, or 'transactional', acting as a stand-alone broker; and an agent that owns a collection or administers a service. A collection may have many services, but must have at least one registered in IESR. A collection and a service may have one or many associated agents. An agent may be both a collection owner and a service administrator, and may be associated with multiple collections and services.

A unique, global identifier (URI) is assigned to each entity on registration with IESR, using the PURL-based Object Identifier (POI) scheme (Powell, Young & Hickey, 2004), unless an existing URI is available. As well as being available to identify IESR entities within the JISC Information Environment and beyond, these identifiers are used to link between entities within published descriptions.

The metadata model for each entity is consistent with the Dublin Core Metadata Initiative's (DCMI, 2005) data model (Powell et al., 2005), being a 'flat' set of repeatable properties each with a single value, rather than a hierarchical XML-style model. For machine processing the metadata is published within an XML encoding. The semantics and occurrence requirements of the properties are documented as a Dublin Core Application Profile following guidelines endorsed by the European information standards organisation (CEN/ISSS, 2003). IESR metadata description is described briefly below. An earlier paper describes data capture by the IESR implementation (Apps, 2004a).

## Collection Description

IESR collection metadata is based on the 'de facto' standard RSLP Collection Description Schema (Powell, 2000), with some simplification to describe electronic resources. It is also consistent with the DCMI Collection Description Application Profile (DCMI, 2004), IESR metadata definition having informed the DCMI working group and vice versa.

The collection metadata includes properties to enable resource discovery such as a collection's title, description and subject terms. Restrictions on property values are defined by vocabulary encoding schemes that are specified in the application profile. To provide quality results from subject searching, terms are limited to a small set of controlled vocabularies, but including those widely used within UK academia, such as HASSET for Social Science. To further enhance coherent, consistent searching over all collections, IESR requires at least one Dewey term to furnish a single, common, 'backbone' vocabulary.

A further, IESR-specific, metadata property, 'usesControlledList', captures the controlled vocabularies used by a collection to describe its items. The list of such schemes is an IESR-defined list, extensible on request, the growth of which reflects the multitude of different domain-specific vocabularies. This information may be useful to a terminology service that maps between vocabularies. It would inform a portal that, having discovered a collection of possible interest to a user, wants to provide an item level search over the collection using an appropriate vocabulary.

Figure 1 shows an example collection description, the UK Census Datasets, in XML format, as published by IESR. The actual IESR record has been abbreviated for inclusion in this paper, cutting short the abstract and omitting many of the subject terms.

```

<iesr:Collection iesr:id="coll-1084445724-12653">
  <dc:title xml:lang="en-gb">UK Census Datasets</dc:title>
  <dcterms:alternative xml:lang="en-gb">
    UK Census of Population Datasets and Related Datasets held by MIMAS
  </dcterms:alternative>
  <dc:identifier xsi:type="dcterms:URI">
    http://purl.org/poi/iesr.ac.uk/1084445724-12653</dc:identifier>
  <dcterms:abstract xml:lang="en-gb">
    Census results provide a unique and invaluable source of detailed current and historical
    information...</dcterms:abstract>
  <dc:type xsi:type="dcterms:DCMIType">Collection</dc:type>
  <dc:type xsi:type="rslpcd:CLDT">Collection.Dataset.Virtual</dc:type>
  <dc:language xsi:type="dcterms:RFC3066">en-gb</dc:language>
  <dc:rights xml:lang="en-gb">(c) Census output remains Crown copyright</dc:rights>
  <dcterms:accessRights>
    Available to UK FE and HE. Conditionally free. Registration required.
  </dcterms:accessRights>
  <iesr:useRights xml:lang="en-gb">
    For Terms and Conditions please see http://census.data-archive.ac.uk/licence.asp
  </iesr:useRights>
  <iesr:hasService xsi:type="dcterms:URI">
    http://purl.org/poi/iesr.ac.uk/1084445405-10420</iesr:hasService>
  <iesr:logo xsi:type="dcterms:URI">http://census.ac.uk/cdu/images/CDU_splash.jpg</iesr:logo>
  <dc:subject xsi:type="dcterms:DDC">300</dc:subject>
  <dc:subject xsi:type="dcterms:DDC">900</dc:subject>
  <dc:subject xsi:type="dcterms:LCSH">Census</dc:subject>
  <dc:subject xsi:type="dcterms:LCSH">Demography</dc:subject>
  <dcterms:spatial xsi:type="dcterms:ISO3166">gb</dcterms:spatial>
  <dcterms:spatial xsi:type="iesr:UNESCO">uk</dcterms:spatial>
  <dcterms:temporal xsi:type="dcterms:W3CDTF">1971</dcterms:temporal>
  <rslpcd:owner xsi:type="dcterms:URI">
    http://purl.org/poi/iesr.ac.uk/1084445246-9103</rslpcd:owner>
  <dcterms:isReferencedBy xsi:type="dcterms:URI">
    http://census.ac.uk/cdu/datasets/</dcterms:isReferencedBy>
</iesr:Collection>

```

Figure 1. An Example IESR Collection Description

## Service Description

Service descriptions inform another application how to access or search a discovered collection. Services are described in IESR using a bespoke scheme to provide a simple set of metadata for resource discovery. When designing the metadata scheme it was apparent that there was a requirement to capture more information about a service than simply RSLP's 'locator', which defines the access point to a collection. Links between the IESR entities, service type, and authentication information are needed, as well as title and description for a transactional service. In fact, in data modelling terms, an IESR service is actually a conflation, made for pragmatic reasons, of the 'location' of a collection and a 'service' provided at that location. Several standard service description schemes were considered, including Web Services Description Language (WSDL) (Christensen et al., 2001) and ZeeRex (ZeeRex, 2004), which is used to describe digital library service interfaces. But IESR, needing a metadata schema to cover a wider range of service types, preferred a simpler scheme to record basic information about a service for resource discovery and data entry, and to include some IESR-specific details. Additionally a scheme consistent with the Dublin Core 'flat' data model, rather than a hierarchical XML model, was sought.

Every IESR service description has a single technical access method, for example Z39.50, SOAP, or Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), and a location URL. A simple 'web page' is another possible service type, although this is not a machine-to-machine interface, many collections within the JISC Information Environment being accessible only in this way. An 'interface' property, pointing to a separate document in an appropriate format, is included for service types for which more details are needed to make a connection. For Z39.50 the 'interface' is a ZeeRex specification detailing the search attributes and result formats supported, whereas for SOAP it is a WSDL definition.

Figure 2 shows an example IESR service description, a Z39.50 service that provides access to SOSIG, the Social Science Gateway of the Resource Discovery Network.

```
<iesr:Service iesr:id="svc-1084526746-17952">
  <dc:title xml:lang="en">SOSIG Internet Resource Catalogue Z39.50</dc:title>
  <dc:identifier xsi:type="dcterms:URI">
    http://purl.org/poi/iesr.ac.uk/1084526746-17952</dc:identifier>
  <rsrpcd:locator xsi:type="dcterms:URI">z3950s://sosig.ac.uk:2104/sosig</rsrpcd:locator>
  <iesr:interface xsi:type="dcterms:URI">
    http://iesr.ac.uk/registry/intf/zeerex/1084526746-17952-zx.xml</iesr:interface>
  <dc:type xsi:type="iesr:AccMthdList">z3950</dc:type>
  <dc:type xsi:type="dcterms:DCMIType">Service</dc:type>
  <dcterms:accessRights xsi:type="iesr:AuthList">None</dcterms:accessRights>
  <rsrpcd:seeAlso xsi:type="dcterms:URI">
    http://www.sosig.ac.uk/about_us/connect.html</rsrpcd:seeAlso>
  <rsrpcd:administrator xsi:type="dcterms:URI">
    http://purl.org/poi/iesr.ac.uk/1084526500-15966</rsrpcd:administrator>
  <iesr:serves xsi:type="dcterms:URI">
    http://purl.org/poi/iesr.ac.uk/1084527323-22034</iesr:serves>
</iesr:Service>
```

Figure 2. An Example IESR Service Description

## Agent Description

Agents are collection owners and service administrators. They are described using an IESR scheme to capture their details. Figure 3 shows as an example the IESR agent description for MIMAS. For inclusion in this paper this record has been abbreviated omitting many of the

'owns' and 'administers' links, the ones shown here being for the UK Census Datasets collection of Figure 1.

```
<iesr:Agent iesr:id="agt-1084445246-9103">
  <dc:identifier xsi:type="dcterms:URI">
    http://purl.org/poi/iesr.ac.uk/1084445246-9103</dc:identifier>
  <dc:title>MIMAS</dc:title>
  <dc:description>Service administrator</dc:description>
  <iesr:phone>+441612756109</iesr:phone>
  <iesr:email>info@mimas.ac.uk</iesr:email>
  <dc:relation xsi:type="dcterms:URI">http://www.mimas.ac.uk/</dc:relation>
  <iesr:logo xsi:type="dcterms:URI">
    http://www.mimas.ac.uk/images/mimas/mimas_d2.gif</iesr:logo>
  <iesr:owns xsi:type="dcterms:URI">
    http://purl.org/poi/iesr.ac.uk/1084445724-12653</iesr:owns>
  <iesr:administers xsi:type="dcterms:URI">
    http://purl.org/poi/iesr.ac.uk/1084445397-10330</iesr:administers>
  <iesr:administers xsi:type="dcterms:URI">
    http://purl.org/poi/iesr.ac.uk/1084445405-10420</iesr:administers>
</iesr:Agent>
```

Figure 3. An Example Agent Description

## Administrative Metadata

Every IESR entity has an associated set of administrative metadata, which includes provenance data and requirements. IESR descriptions are freely available for non-commercial use under a Creative Commons licence, provided attribution of provenance and the same licence are maintained. As an example, Figure 4 shows the administrative metadata for the UK Census Datasets collection of Figure 1.

```
<iesr:admata iesr:about="coll-1084445724-12653">
  <dc:creator>MIMAS</dc:creator>
  <dc:creator>http://purl.org/poi/iesr.ac.uk/1084445246-9103</dc:creator>
  <dc:publisher>JISC Information Environment Service Registry (IESR)</dc:publisher>
  <dc:publisher>http://iesr.ac.uk</dc:publisher>
  <dcterms:modified xsi:type="dcterms:W3CDTF">2004-05-13</dcterms:modified>
  <dc:rights xsi:type="dcterms:URI">
    http://creativecommons.org/licenses/by-nc-sa/2.0/uk/</dc:rights>
  <dc:rights>
    This IESR administrative metadata must always be retained with its associated entity
    description.</dc:rights>
</iesr:admata>
```

Figure 4. An Example Administrative Metadata Record

## IESR Services

IESR provides access to its metadata via several services, based on standard interfaces, and further services are planned. Although of secondary purpose, it has a web interface that allows searching and browsing of the content of the registry. Thus discovery of IESR descriptions is available to humans, maybe data suppliers checking their records, builders of portals looking for relevant material to include, or anyone for general resource discovery.

The first two machine-to-machine services provided by IESR, Z39.50 and OAI-PMH, reflect its development being within a digital library environment. A Web Services SOAP service is planned, and the viability of a UDDI view on IESR data is under investigation. Eventually RSS services will be available providing both informative news and new data alerts. Accesses

to all IESR services are logged, providing the possibility of future evaluation studies of preferred uses of the different options.

## Z39.50 Service

The Z39.50 service, implementing the Z39.50 standard machine protocol for information retrieval, supports searching using a range of standard bibliographic attributes and provides results as either simple text or XML, for which there is a choice of simple Dublin Core or full IESR XML records. Simple Dublin Core results are provided for conformance with the digital library 'Bath Profile' but it is unclear whether they will be useful, being very 'dumbed down' from the rich IESR descriptions.

An IESR XML record provided by the Z39.50 interface is composite, returning details of a collection and all its associated services and agent entities, including the administrative metadata for all these entities. The advantage of a composite result to an application using IESR for collection discovery is that it will obtain details of the collection and all its service interfaces in one operation. It could then decide, and perform, the best access option for an end user and their environment, with no need for subsequent retrievals from IESR.

## OAI-PMH Service

The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) (Lagoze et al., 2004) interface allows harvesting of records from IESR and subsequent maintenance of the harvested data. Supporting most of the 'verbs' of the protocol, it is possible to retrieve single records, lists of records and lists of record identifiers according to modification date criteria, as well as documentation about the service. Again records are made available in XML, both simple Dublin Core versions and full IESR XML descriptions, but in this case as single entities, with the administrative metadata in the 'about' field of the OAI-PMH record. Partly to alleviate the lack of specificity in the simple Dublin Core records, the supply of which is a requirement of the OAI-PMH protocol, an additional 'relation' property is included that is a by-reference pointer to the full IESR XML description of the entity. This is implemented using a standard way of passing information, the NISO standard Z39.88-2004, 'The OpenURL Framework for Context-Sensitive Services'. This is a novel use of the OpenURL Framework, which originated within the scholarly information environment for article reference linking (Apps & MacIntyre, 2005).

## Web Services Interface

Development of a SOAP interface is planned based on the World Wide Web Consortium's standard Web Services server-to-server protocol. Currently many Web Services application interfaces use proprietary XML schemas for request and response message data. A possible standard interface, which IESR will support to promote interoperability, is Search/Retrieve Web Service (SRW, 2004). SRW, developed by the Z39.50 community, is a standard Web Services protocol within the digital library community, which has recently been registered by NISO. In addition to a WSDL interface definition, an SRW service's functionality is specified using ZeeRex. Searching is by CQL (Common Query Language), the required, official query language of SRW. In practice, the Dublin Core context set for CQL is widely supported to enable interoperable cross searching. SRW supports the definition of the fields within the result set, the implementation of an interoperable basic Dublin Core result being encouraged, as well as providing fields to split a result set into suitable sized sections for reasonable transmission. The design of the IESR SOAP interface will also be informed by the NISO

Vendor Initiative for Enabling Web Services (VIEWS). Similar to other IESR services it will provide XML results optionally as simple Dublin Core or full IESR descriptions.

## UDDI

The viability of a UDDI (Universal Description Discovery and Integration) standard service based on the IESR data is currently under investigation. UDDI is a standard protocol to build a registry of, and thus to publicise, businesses and the services they offer, described according to an XML grammar. Although UDDI is not exclusively for the registration of Web Services that is its general use, dynamic access to a UDDI registry being via a SOAP interface. The investigation is experimenting with mapping IESR data into a UDDI format with the intention of populating a prototype registry. At present a UDDI version of IESR does not appear to be a stakeholder requirement, there being little evidence of UDDI use within the JISC or digital library communities, it being more prevalent in the eBusiness sector. However UDDI registries are available and used by applications within the eScience community.

## Using IESR

IESR envisages portals being significant users. JISC defines a portal as (JISC, 2003b):

“A network service that brings together content from diverse distributed resources using technologies such as cross searching, harvesting, and alerting, and collates these into an amalgamated form for presentation to the user.”

To support use by portals, IESR affords discovery of resource collections. It supplies ‘up to date’ descriptions of collections and how to access them, with the caveat that IESR relies on its data suppliers to maintain currency. Resources are described as collections rather than at item level. Thus IESR provides discovery of, for example, a collection of datasets that may cover a topic in which a user is interested, rather than the more specific discovery of an individual dataset. However a Social Science portal with the capability of detailed dataset discovery could then provide more detailed searching over items from a collection it has discovered through IESR.

## Portal Z39.50 Use Scenario

Possibly a Z39.50 enabled Social Science portal would search IESR for collections of probable interest to an end user, maybe by a specific topic provided by that user. From the descriptions returned by IESR it would select the collections that have a Z39.50 interface. Using the Z39.50 service details the portal would perform a cross search, or ‘metasearch’, over those collections on behalf of, and displaying the results to, the end user. Because in this scenario the portal uses IESR dynamically, the portal builder does not have to be aware of all available collections. Further, using a Z39.50 metasearch, the portal does not need to maintain its own registry of services. New collections would appear in the search as they were registered in IESR. The advantage to an end user is the potential discovery of resources of interest of which they were unaware.

## Harvesting by Portals

Alternatively a portal may wish to harvest, via the OAI-PMH interface, IESR descriptions to cache and use locally for its own functionality. Possibly this would be to maintain its own service registry, or for ingest into a ‘knowledge base’ after transformation into an appropriate format, or for conversion into a local configuration file. If new and updated IESR records are

harvested on a frequent basis, this becomes a similar scenario to the Z39.50 portal use described above. IESR would be used effectively dynamically with no need for a portal builder to specifically include a particular resource.

### Example Application Using Harvesting

An illustration of a possible harvesting and reuse use of IESR descriptions is provided by a planned development of the MIMAS Metadatabase (Apps, MacIntyre & Morris, 2002). This is a searchable catalogue of resources that are provided by MIMAS, which has Web, Z39.50 and OAI-PMH interfaces. Support staff are obviously reluctant to maintain descriptions of MIMAS resources in both IESR and the Metadatabase. A resolution of this problem will be to re-implement the Metadatabase as a copy of IESR including only MIMAS supported resources, and with a MIMAS-specific Web interface, but maintaining the current access addresses for the Metadatabase interfaces. The Metadatabase will harvest changed IESR records frequently and use any retrieved MIMAS records to update its own database.

### RSS Aggregator

Another type of application that may make use of IESR is an RSS (RDF Site Summary) aggregator, there being several RSS services already described in IESR. An RSS aggregator could discover RSS services in IESR relevant to Social Science, and bundle these into an amalgamated news feed. Similar to a portal's use of IESR, if the RSS aggregator were to use IESR dynamically the application builder would not need to know about RSS service availability, and new RSS feeds would be included as they were registered in IESR. At present an RSS aggregator would use either the Z39.50 or the OAI-PMH IESR service to discover RSS feeds, but in the future IESR may have an RSS interface available. A possible RSS service built from IESR data may be an aggregated news feed about new developments in Social Science, similar to the JISC news service (Davey, MacLeod & Moffat, 2004). Another possibility would be a Social Science journals table of contents syndication, similar to such a service in science publishing (Hammond, Hannay & Lund, 2004).

### Description Reuse

One of the simplest uses of IESR metadata by another application is to reuse the description of a collection, having ascertained it using one of the IESR service interfaces. If this information reuse were implemented dynamically, there would be some certainty that this description, when displayed by the consuming application, would always be current. This scenario implies that a definitive description for a particular collection is created by the resource provider, this single description then being shared by multiple registries and applications. At present IESR holds descriptions for resource collections funded by JISC. This method of using IESR has been demonstrated by the Subject Portals Project (Clark, 2001), although their use was not dynamic, by populating their descriptions of resource collections with data taken from IESR. Similarly the builders of the ESRC Information Centre investigated using IESR to discover definitive descriptions of Social Science resources.

### Using the Web Interface

The IESR Web interface is available for human resource discovery, even though the main purpose of IESR is to provide a middleware service to other applications. Resource providers are beginning to link to their collection descriptions within the IESR Web interface, effectively providing a two-way quality assurance mark.

The simplest use a portal could make of IESR would be to give its users a simple link to the IESR search interface to enable discovery of resources available in the Information Environment. However the form of information displayed may be confusing to end users, with its inclusion of service connection details in addition to the collection description. It would seem preferable for the portal to retrieve XML descriptions using one of the IESR machine interfaces and transform these into a Web display appropriate to its particular users.

## Web Services Scenarios

Use scenarios by portals, similar to those based on Z39.50, could be suggested using Web Services, especially if the IESR content were extended to describe more Web Services access points to collections. One could envisage 'mix and match' scenarios where a portal discovers other types of services using IESR's SOAP service, or where a portal uses OAI-PMH to harvest details of Web Services.

To date IESR has registered very few SOAP and no SRW services, which may reflect a paucity of Web Services implementations within the JISC community, but this situation will probably change in the near future with the introduction of a wider range of resources. The opening of IESR to a wider community that includes UK eGovernment sponsored resources, where Web Services is the expected machine interface, should increase their registration. The JISC e-Learning Framework indicates that the eLearning community is moving towards a service-oriented architecture, as are the Virtual Research Environments domain, implying that such resources will have Web Services interfaces.

A Web Services interface will also enable the use of IESR within an eScience context. Discovery of, with subsequent access to, a relevant social science collection that provides a Web Services service, could be incorporated into a composite service workflow, similar to those performed within the Bioinformatics domain using the Taverna workbench facilitated by the myGrid service-oriented architecture (Oinn et al., 2004). Another scenario may include discovery of bibliographic resources, such as articles within Zetoc (Apps, 2004b), to be used in a literature investigation workflow, IESR providing the connection details for the bibliographic collection's SOAP service. Provision of an IESR UDDI service, or a means of communicating between IESR and an eScience UDDI registry, may further enable the use of IESR within the eScience arena.

## Distributed Registries

The initial remit for IESR was to provide a single registry of collections and services within the JISC Information Environment. It is one of a set of 'shared services' in a middleware layer of the information environment providing an interface, consisting of machine readable information, between portals that provide a display to end users and content providers that provide access to the data. However, questions about data ownership, data maintenance workflow, scalability, quality assurance and the possibility of extension to an international landscape, all point towards the development of distributed registries. A parallel development by the Grid Engineering Task Force of a network of UDDI-based service registries at the eScience centres, albeit focused on transactional services rather than collections, also suggests a shift towards a distributed registry model.

Provision of IESR's OAI-PMH interface opens up the possibility of sharing data between service registries, both within the JISC Information Environment and beyond, assuming they use the same, or a derivable, metadata schema. This leads to the idea of a distributed grid of

service registries, each node being responsible for descriptions of collections and services under its own remit, but contributing to a shared, searchable, virtual service registry. However this is still a vision, which is leading to discussions, but as yet there are no concrete plans for development within the UK. Questions need to be resolved about how data will be shared and how an 'up to date' version of the total virtual registry could be searched. Thus development of a single registry will continue for some time, with IESR serving descriptions of data according to its own metadata schema.

## Related Initiatives

### Metadata Standards

When designing the metadata to capture descriptions of collections, services and agents within IESR, an objective was to use existing open standard properties and schemas wherever possible. The collection metadata description is consistent with the DCMI Collection Description Application Profile, which is also largely utilised for collection description within the NISO Metasearch Initiative (NISO, 2005). IESR uses a few RSLP and local properties where a standard alternative is not available. Members of the IESR project team are active in the DCMI Collection Description Working Group, and one, who is the Chair of that group, also has significant input to the NISO Metasearch Initiative collection description group. Thus IESR is providing a practical application to prove the designs of these developing standard collection description designs.

The service metadata description is IESR specific, for reasons given above in the 'Service Description' section, but most of the individual properties are Dublin Core. There is a divergence here from the NISO Metasearch Initiative, which is using ZeeRex to describe services. Their remit is narrower than that of IESR, covering only standard interfaces, such as Z39.50 and SRW, for cross searching of scholarly information resources provided by libraries. Thus they have no requirement to describe simple Web page services, and they do not describe transactional, broker services. Although having its own, simpler and 'flatter', scheme for service description, IESR does also make use of ZeeRex to describe connection details for services that are implemented using digital library standards.

### Related Projects

A significant, separate development, within the USA, is OCKHAM (Morgan, Frumkin & Fox, 2004). This is a project to build a distributed registry of digital library services, initially National Science Digital Library content. OCKHAM have focused on developing software to be deployed at each of the distributed 'hubs', data being circulated using OAI-PMH in a 'DNS-like' model to ensure its currency and entirety at every hub, where local searching of the registry is performed. OCKHAM is using the IESR metadata schema for data description and possibilities of collaboration between the two projects are being explored. There are clearly opportunities at this experimental stage to share data descriptions because of the common data model.

## IESR Future

The IESR development project is entering its third funding phase, lasting until April 2006. The various services referred to in previous sections will be implemented during that time, as well as improvements to data supply. The content of IESR will include more resources from

the JISC Information Environment such as institutional digital repositories and library catalogues and services. It will also be extended to a wider landscape covering resources from other UK Common Information Environment partners, which includes the Museums, Libraries and Archives Council, the British Library, and the National Health Service.

Demonstrating viable use of IESR is still in its early stages. There has been some initial interest in testing IESR use by JISC projects investigating portal development, and there is also interest from some library systems vendors. It is hoped that further opportunities for the exploitation of IESR will develop during this phase of the project.

It became apparent when developing the service metadata description and describing the initial content of IESR that many collections within the JISC Information Environment do not yet have interoperable machine readable interfaces, but rather all searching and access is by human users directly through a Web interface. Thus it seems that the IESR vision of providing a dynamically accessible middleware registry is a bit ahead of the field. It is expected that further machine interfaces to collections will become available as part of the move towards a service-oriented architecture, in particular within the eLearning community.

It will be important to continue maintenance of the IESR metadata schema to remain consistent with developing standard initiatives and to respond to requirements to capture further resource properties. For example the current implementation of Shibboleth for authentication within the JISC Information Environment may imply a support requirement within the IESR metadata schema. Similarly new interfaces into IESR content, and new ways of communicating with other services, should be developed when appropriate to remain inline with any technical developments of the information environment. But most significantly, IESR's continuation on a longer-term basis is anticipated to ensure persistence of its content.

## Acknowledgments

The author wishes to acknowledge the assistance of colleagues, Pete Johnston and Andy Powell of UKOLN at the University of Bath, in the development of the IESR metadata, and for the contribution of ideas about using IESR and future expected developments of distributed service registries. Contribution towards IESR development by other project colleagues is also acknowledged including Amanda Hill, the project manager, and Leigh Morris, of MIMAS at The University of Manchester, and Rachel Heery of UKOLN. The IESR development project is funded by the Joint Information Systems Committee (JISC) of the UK Higher and Further Councils as part of its 'Shared Services' programme, with partners at MIMAS, UKOLN and the Cheshire Development team at the University of Liverpool. The registry is developed and hosted by MIMAS.

## References

- Apps, A. and MacIntyre, R. (2005): 'Emerging Uses for the OpenURL Framework', in *ELPUB2005: Proceedings of the 9<sup>th</sup> ICCO International Conference on Electronic Publishing, 8-10 June 2005, Leuven, Belgium*, (accepted for publication).
- Apps, A. (2004a): 'A Registry of Collections and their Services: from Metadata to Implementation', in *DC-2004: Proceedings of the International Conference on Dublin Core and Metadata Applications, October 11-14, 2004, Shanghai, China*, Shanghai Scientific & Technical Literature Publishing House, pp. 67-73.
- Apps, A. (2004b): 'zetoc SOAP: a Web Services Interface for a Digital Library Resource', *Lecture Notes in Computer Science*, issue 3232, September 2004, pp. 198-208.

- Apps, A., MacIntyre, R. and Morris, L. (2002): 'Exposing Cross-Domain Resources for Researchers and Learners', in *DC-2002: Proceedings of the International Conference on Dublin Core and Metadata for e-Communities, 13-17 October, 2002, Florence, Italy*, Firenze University Press, pp. 71-80.
- CEN/ISSS. (2003): 'CWA 14855, Dublin Core Application Profile Guidelines', <ftp://ftp.cenorm.be/PUBLIC/CWAs/e-Europe/MMI-DC/cwa14855-00-2003-Nov.pdf>.
- Christensen, E., Curbera, F., Meredith, G. and Weerawarana, S. (2001): 'Web Services Description Language (WSDL) 1.1', <http://www.w3.org/TR/wsdl>.
- Clark, J. (2001): 'Subject portals', *Ariadne*, issue 29, October 2001.
- Davey, P., MacLeod, R. and Moffat, M. (2004): 'Improving Communications with JISC through News Aggregation', *Ariadne*, issue 41, October 2004.
- DCMI. (2005): 'The Dublin Core Metadata Initiative', <http://www.dublincore.org/>.
- DCMI. (2004): 'DCMI Collection Description Application Profile', via <http://www.dublincore.org/groups/collections/>.
- Hammond, T., Hannay, T. and Lund, B. (2004): 'The Role of RSS in Science Publishing', *D-Lib Magazine*, vol. 10, no. 12, December 2004.
- IESR. (2005): 'A machine readable registry of electronic resources', <http://www.iesr.ac.uk/>.
- JISC. (2003a): 'Investing in the Future: Developing an Online Information Environment', [http://www.jisc.ac.uk/index.cfm?name=ie\\_home](http://www.jisc.ac.uk/index.cfm?name=ie_home).
- JISC. (2003b): 'Portals: Frequently Asked Questions', [http://www.jisc.ac.uk/index.cfm?name=ie\\_portalsfaq](http://www.jisc.ac.uk/index.cfm?name=ie_portalsfaq).
- Lagoze, C., Van de Sompel, H., Nelson, M. and Warner, S. (2004): 'The Open Archives Initiative Protocol for Metadata Harvesting', <http://www.openarchives.org/OAI/openarchivesprotocol.html>.
- Miller, P. (2004): 'Towards the Digital Aquifer: introducing the Common Information Environment', *Ariadne*, issue 39, April 2004.
- Morgan, E.L., Frumkin, J. and Fox, E.A. (2004): 'The OCKHAM Initiative – Building Component-Based Digital Library Services and Collections', *D-Lib Magazine*, vol. 10, no 11, November 2004, (In Brief).
- NISO. (2005): 'NISO Metasearch Initiative', [http://www.lib.ncsu.edu/niso-mi/index.php/Main\\_Page](http://www.lib.ncsu.edu/niso-mi/index.php/Main_Page).
- Oinn, T., Addis, M., Ferris, J., Marvin, D., Senger, M., Greenwood, M., Carver, T., Glover, K., Pocock, M.R., Wipat, A. and Li, P. (2004): 'Taverna: A tool for the composition and enactment of bioinformatics workflows', *Bioinformatics Journal*, vol. 20, no. 17, 2004, pp. 3045-3054.
- Powell, A., Nilsson, M., Naeve, A. and Johnston, P. (2005): 'DCMI Abstract Model', <http://www.dublincore.org/documents/abstract-model/>.
- Powell, A., Young, J. and Hickey, T. (2004): 'The PURL-based Object Identifier', <http://www.ukoln.ac.uk/distributed-systems/poi/>.
- Powell, A. (2000): 'RSLP Collection Description Schema', <http://www.ukoln.ac.uk/metadata/rsdp/schema/>.
- SRW. (2004): 'Search/Retrieve Web Service', <http://www.loc.gov/z3950/agency/zing/srw/>.
- ZeeRex. (2004): 'ZeeRex: The Explainable "Explain" Service', <http://explain.z3950.org/>.