UNIVERSITY OF
BATH

Citation for published version:
Russell, R 1997, UKOLN MODELS 4: evaluation of cross-domain resource discovery. in P Miller & D Greenstein (eds), Discovering online resources across the humanities: a practical implementation of the Dublin Core. UK Office for Library and Information Networking, on behalf of the Arts and Humanities Data Service, Bath, pp. 18-21.

Publication date:
1997

Link to publication

University of Bath

UKOLN MODELS 4: Evaluation of Cross-Domain Resource Discovery

Rosemary Russell[1]

UK Office for Library and Information Networking, University of Bath, Bath BA2 7AY, United Kingdom

2.1 The Models 4 workshop: integrating access to resources across domains

This first workshop in the AHDS/UKOLN series involved nearly 50 participants representing a range of curatorial professions (e.g. libraries, museums, archives, and data archives), systems experts, and scholarly users in a discussion of whether and how to integrate access to information resources across domains. The topic reflected the hypothesis that people want to search for and locate relevant information resources, irrespective of their format and of whether they are held in libraries, museums, data archives, or any other organisational structure. The workshop set out to test the hypothesis, and if found to be true, to explore how resources and systems might be structured in order to realise it in information service environments.

The workshop was the fourth in a series of MODELS workshops (MOving to Distributed Environments for Library Services) conducted by UKOLN and sponsored by the Electronic Libraries Programme (more information about MODELS may be found at <URL: http://www.ukoln.ac.uk/models/>). MODELS was ideally situated to launch the AHDS's and UKOLN's work on metadata for resource discovery. As an ongoing concern, it aims to develop a framework for managing distributed library and information services by facilitating informed and focused discussion of salient strategic and practical issues, and by articulating the technical service models which emerge from those discussions. Even prior to this workshop, MODELS had established a track record for involving key stakeholders in discussions which ultimately initiated work of national significance.

---

[1] First published in: *Discovering Online Resources Across the Humanities: A Practical Implementation of the Dublin Core*, ed. Paul Miller and Daniel Greenstein (Bath: UK Office for Library Networking; Arts and Humanities Data Service, 1997), pp. 18-21. It formed the first section of Chapter 2: Reports from the Front: Domain-Specific Perspectives on Cross-Domain Discovery. The section headings follow those in the published version. Online version available at: http://www.ukoln.ac.uk/metadata/publications/ahds-ukoln/

2.2 Significant findings

2.2.1 Cross-domain resource discovery

It was agreed at an early stage in the workshop that cross-domain searching is highly desired by both scholarly and curatorial communities, making technical and service models worth pursuing. That is, the group agreed about the desirability of being able to search across a range of potentially complementary online information resources in a manner which, from the users point of view, obscured any differences that existed in the underlying resources' hardware, software, record structures, and query languages. What constituted cross-domain discovery, however, was seen to differ. For some it involved searching across the holdings of several institutions operating within a single curatorial tradition. In the archival community, for example, finding aids differ so considerably in both their structure and content that cross-domain discovery could be seen as searching across the holdings of two or more archives. From other perspectives, cross-domain discovery could entail searching across the different databases maintained by a single institution for different aspects of its collection - for example, by a library about its book, manuscript, and print collections. For still others, cross-domain resource discovery might entail searching for information across a range of library catalogues, museum databases, archival finding aids, data archive catalogues, and subject-based catalogues of World Wide Web resources. The different definitions of cross-domain discovery encouraged the group to focus upon...

2.2.2 Defining domains

Flexibility in approach was deemed desirable in defining the domains which surrounded discrete collections of information resources. Thus, domains could reflect the curatorial traditions which shaped the way in which those resources were managed (libraries, museums, archives, etc.), the academic disciplines within which they were principally created or used (archaeology, film studies, geology), or the regional settings where they were stored (north Wales, south-east). Crucially, by facilitating cross-domain discovery the group felt it possible to break down the institutional, disciplinary, geographical, and other barriers which may impede access to and use of information.

2.2.3 A search model and its implication for metadata

With regard to cross-domain discovery itself, a reiterative or staged approach was identified. At the first stage a user requires rudimentary information about relevant information resources in order simply to be made aware of their existence. To sustain this stage, generic metadata such as the Dublin Core is required. At a second stage, the user having found a potentially interesting information resource, might need richer metadata to determine whether to acquire, browse, or analyse it. At a third stage, the same user having acquired to a resource might need still further descriptive information in order to use the resource effectively. At both of these later stages, metadata in more specialist formats (e.g. MARC records for books, ISAD(G) records

for archival material, TEI headers for electronic texts) would likely be required. It was here that the group articulated a search model which enabled the user, in a single search environment, to 'drill down' or move progressively through a hierarchy of increasingly rich and specialist metadata as they moved through a continuum from resource discovery to resource evaluation, access, and use.

2.3 Areas for further investigation

2.3.1 Metadata for resource discovery

Domain-specific approaches to resource description, though fundamentally different, need not impede cross-domain resource discovery, particularly if discovery is seen as part of the continuum described above. Some consensus is required across domains, however, about the minimum level of metadata that needs to be associated with an information resource if it is to be located meaningfully by those who might wish to gain access. As a first step towards that consensus, domain specialists should, in light of the documentation standards and best practice that are current within their domains, formally identify their resource discovery requirements and express them using a common formalism for comparative purposes. Within this process the Dublin Core could provide the common formalism but also a starting point for discussions about domain-specific resource discovery requirements. This recommendation was particularly formative for the six more specialist workshops convened by the AHDS and UKOLN.

2.3.2 Collections description

The issue of collections description first arose at the third MODELS workshop (Dempsey and Russell 1996) and focused on mechanisms for providing users but also resource discovery tools some forward knowledge about the contents of a particular collection catalogue. In the context of cross-domain resource discovery, collections description was seen as a higher level of resource discovery metadata; that is as metadata to help users select from a range of online catalogues, finding aids, Web-based gateways, etc. those worth including in a particular search. Discussion touched on the possibilities for using Centroids (Knight and Hamilton 1997). More importantly it called for further investigation into communities' collections description practices and into the collection description requirements of users involved in real cross-domain discovery. Whereas the former investigation requires traditional research (an eLib supporting study of collections description is currently being co-ordinated by UKOLN and is due for completion in autumn 1997), the latter requires the development of cross-domain discovery services such as the one being built by the AHDS and reported in Chapter 4 of this volume.

2.3.3 The Z39.50 protocol for search and retrieve

The Z39.50 protocol seemed potentially capable of permitting users to implement the reiterative cross-domain search and retrieve model outlined above. Although a relatively unexplored area, some investigative work was at the time of the MODELS

workshop either underway (for example, by a group of UK-based archivists) or intended (for example by the AHDS). The Z39.50 Digital Collections Profile appears particularly promising (Library of Congress, 1996). It provides a generic or over-arching framework capable of accommodating more specialist profiles each of which is designed to navigate heterogeneous collections databases developed by particular communities of, for example, museums, libraries, curators of geospatial information.

2.3.4 Controlled vocabularies in a cross-domain environment

Data description standards ensure that information resources are described with reference to a common range of attributes. Controlled vocabularies ensure a degree of consistency in the use of attribute values, and thus a degree of consistency in search and retrieval. In a cross-domain discovery scenario, the user is likely to encounter different, possibly competing controlled vocabularies which reflect the underlying domains' different resource discovery and description requirements. Resolving these conflicts is essential if users are to be assisted in meaningfully searching a wide range of information resources. Mapping between controlled vocabularies may be one option although here too, it was felt that experience of user behaviour in testbed cross-domain discovery environments was an essential first step in addressing this issue.

2.4 Conclusion

The MODELS 4 workshop brought together a wide range of representatives from relevant communities, and initiated discussion on the hitherto little explored area of cross-domain resource discovery. It confirmed the desirability of this approach to resource discovery and debated possible strategies and implementations. It was particularly successful in identifying the metadata requirements and how these might build upon work already conducted under the Dublin Core initiative. As such it laid an important foundation for the more specialist AHDS/ UKOLN workshops which followed.

References

Dempsey, L. and R. Russell, 1996. "National Resource Discovery Workshop: Organising Access to Scholarly Printed Materials". UK Office for Library and Information Networking <URL: http://ukoln.ac.uk/models/models3.html>

Knight, J. and M. Hamilton, 1997. "Dublin Core Qualifiers". ROADS <URL: http://www.roads.lut.ac.uk/Metadata/DC-SubElements.html>

Library of Congress, 1996. "Z39.50 Profile for Access to Digital Collections". Draft for review. Library of Congress <URL: http://lcweb. loc.gov/z3950/agency/profiles/collections.html>