Chapter IV

Subject Gateway Features
SUBJECT GATEWAY FEATURES

Specialised subject gateways have become an essential tool for locating and accessing digital information resources, with the added value of organisation and previous evaluation catering for the needs of the varying communities using these. Within the framework of a research project on the subject, a software tool has been developed that enables subject gateways to be developed and managed.

The tools for retrieving and accessing information available on the Internet to answer users' needs take three main forms, namely search engines, subject gateways and vertical portals. Among these, the subject gateways are gaining in popularity and importance as a source of digital information that has been chosen and assessed and which is being provided with added value services, as described by Tramullas (2006).

Need for Gateways

The gateways emerged from this engagement between an emerging service model and a policy aspiration to create a network use environment which effectively supported research and learning. They responded to the need to provide managed collections of resources, supported by effective resource description and subject access, available for browse and search access. In effect, they aimed to give shape and definition to an information space in a particular subject area. They aim at saving time for their users, to connect them to resources which supported their learning, teaching and research interests, and to make sure that information about useful resources was effectively disclosed.

- anyone can publish anything on the Internet, the Internet reflects the real world

58
it contains the good, the bad and the ugly! It is left up to us to decide which is which!

- difficult to find QUALITY relevant information from hundreds of millions of web pages
- no real organisation and many poor quality resources

Definition

"Subject Gateways - are subject-based resource discovery guides that provide links to information resources (documents, collections, sites or services), predominantly accessible via the Internet." The service is based on resource description. Browsing access to the resources via a subject structure is an important feature.

A subject gateway may be defined as: "A Web-based mechanism for accessing a collection of high quality, evaluated resources identified to support research in a particular subject discipline where the resources are evaluated and described by information specialists in the field, such as science librarians."

Gateways: "A gateway is a set of interrelated tools that enable users to identify and locate materials relevant to their study and research. The tools are presented in a common structure or framework that facilitates the rapid identification of and navigation to the material. A gateway may include the catalog, databases to other online resources such as e-journals, tools that link different databases together, real-time assistance, and help or user guidance components"(Casswell, 2004, p.72).

A gateway can be defined in general terms as a link between different elements. Within computer or telecommunication it is used to
denote an object linking two different networks or different communication protocols or standards. It is often used to denote a converter or translator between different formats. Such translation is one task that serves as gateways between different networks or standards have to do. In some cases the term is used in a broader sense where even a standard or a whole network can be seen as a gateway between different computers or applications.

A Gateway is translating between different standards and accordingly and enabling the interconnection of networks based on different standards. This means that we see gateways as complementary to standards.

In the traditional information environment human intermediaries, such as publishers and librarians, filter and process information so that users can search catalogues and indexes of organised knowledge as opposed to raw data and disparate information. Subject gateways work on the same principle - they employ subject experts and information professionals to select, classify and catalogue Internet resources to aid search and retrieval for the users. Users are offered access to a database of Internet resource descriptions which they can search by keyword or browse by subject area. They can do this with the knowledge that they are looking at a quality controlled collection of resources.

**Common gateway features**

1. The host institutions of each gateway have a mandate to be information providers, but they are not required to be information creators.
2. Each gateway has a distinctive logo and a reflective name, thereby positioning themselves for a significant Web presence. Distinctive branding also bears fruit as a common recognition technique.

3. Each gateway has selected and utilized a standard metadata schema for describing the resources incorporated into the gateways. The metadata is applied by librarians or educators with experienced knowledge of the disciplines.

4. The resources in each gateway have all been selected according to pre-determined criteria, published at each gateway site as part of a content coverage policy. A gatekeeper function, often a mix of computer and human intervention, ensures adherence to the selection criteria. This function lends itself to a high level quality rating for the site as a whole, but in conjunction with the commitment by each gateway's partners to a valid, valuable service. While the gateways have targeted academic communities (due to the nature of their source funding and development), the currently free access means that they may be used by any member of the public with an interest in the discipline.

5. The coverage policies have made electronic resources, both 'born digital' and digitised, the highest priority for inclusion in the gateways; but they are all extensible to include books as well as bytes, and under-utilised or unknown resources such as databases, and descriptions of people.

6. The resources are supported by similar architectures, based on distributed creation and maintenance of their metadata, with a centralised facility for access. The metadata, if embedded in resources prior to their selection, may be enhanced or merely augmented but not ignored. The access facility, a navigation interface which provides for the metadata to be queried transparently
or explicitly, is provided in at least one of three forms - keyword searching, index browsing, or structured pathways.

7. The navigation points are usually supported by the use of at least one thesaurus, to ensure reliability of resource description and subsequent search results. AVEL, for example, uses the Library of Congress Subject Headings (LCSH). While a thesaurus enables subject terms to be selected for semantic consistency, the use of "street" terms for the creation of a dynamic thesaurus has also been considered.

8. **Gateway as one of internet services:** There are a number of services which bring subject expertise to bear on the selection, classification and description of resources. These services, subject gateways, select, classify and describe quality resources in a specified subject area. They effectively fill the role of information broker for information seekers in that subject and the people selecting and describing resources are usually subject-specialists, for example librarians. These gateways can be relied on to identify useful quality online resources, and to be an important resource for anyone working in a field in which there is a significant mass of online source material.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Name of the Subject Gateways</th>
<th>Year of Establishment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>SOSIG</td>
<td>1994</td>
</tr>
<tr>
<td>2.</td>
<td>OMNI</td>
<td>1994</td>
</tr>
<tr>
<td>3.</td>
<td>AGRIGATE</td>
<td>1998</td>
</tr>
<tr>
<td>4.</td>
<td>METACHEM</td>
<td>1998</td>
</tr>
<tr>
<td>5.</td>
<td>EEVL</td>
<td>1999</td>
</tr>
</tbody>
</table>
How to Develop Gateways

Preliminary planning

The First and foremost for developing a gateway must mean that we have some aims and objectives. This might be to establish a service for a specific national user community, or perhaps it is to set up a gateway for a University Library. Each different gateway will have a different set of aims and objectives. Some of the key deliverables that helped to put the broad aims and objectives are SMART: Specific, Measurable, Achievable, Relevant, Time-based. Gateways differ widely in aims and scope, but a better understanding of the costs and benefits of gateways would enable improvements in future decision making.

Staff and skills required

Depending on the exact technology used, there is going to be relatively large up front cost in terms of time and unique skills, in the setting up of a gateway. The information management issues will require research and documentation. It is likely that the people involved with this side of the setting up, will continue to play a part in the project, most usually in the building of the resources database and the day to day running of the project. There will also be a large up front cost in terms of the technical implementation of the infrastructure software that the gateway will operate on. How large this cost will be depends on whether or not an existing set of gateway technology is being used or a new system is being developed. Either option will require people with the appropriate technical skills.

The key staff needed for the running of a gateway are subject specialists who will be involved in the expansion and development of the resources catalogue. The exact number of these will depend on the scope
of the gateway. If the gateway aims to catalogue all resources in a given field within a short period, then a larger number of cataloguers will be required. The more subject specialist and resource cataloguers there are, then the faster the number of resources in the gateway can grow.

System requirements

Subject gateways operate in a Web environment. This means that they must be available all the time. End-users expect reasonable response times while they browse the gateway and fast and predictable performance when they search the database. Subject gateway cataloguers expect reasonable response times as they add resource descriptions to the database. Subject gateway managers want to be able to deliver all this at a reasonable cost - both in terms of the initial cost of establishing the gateway and in terms of ongoing hardware and software support costs. We can achieve this through the use of appropriation: network connectivity, hardware configuration (memory, CPU speed, disk space), operating system software, subject gateway database and associated software, Web server software.

Co-operation between Gateways

The Internet offers great potential for co-operation between one gateway service and another since it allows geographically distributed databases and people to communicate with one another and to work together to build integrated services. Co-operation between gateways is increasingly being seen as a strategy for:

- enhancing Internet resource discovery for end-users
- improving the efficiency and sustainability of gateway services
There are a number of different models for collaborative work, and, as gateways are still a relatively new type of information service, there is still much scope for exploring the potential of co-operation. Those running gateways should consider the benefits of, and opportunities for, co-operation with other gateways.

**Strategic advantages of co-operation**

The development of a myriad of information gateways on the Web is, ironically, making it increasingly difficult for users to search the Internet effectively. Many gateways are claiming to offer a 'one-stop shop' for finding information and this may work for certain users; however, other users will benefit from searching more than one gateway. With lots of independent and uncoordinated gateways, this can involve making a series of searches in a number of services, all of which have different interfaces and ways of working. Collaboration can help gateways to offer integrated services for end-users. The advantages of this for users (depending on the co-operative model used) may include:

- Access to far broader collections than any single gateway could offer, including high quality Internet resources on many subjects, from many countries, written in many languages
- access to a large number of metadata records via a single user-friendly interface
- the ability to locate new gateways that they may not have heard about
- the possibility of searching a selection of gateways simultaneously as opposed to one by one
- Improving the efficiency and sustainability of gateway services

As more organisations invest in building gateway services, more opportunity for collaborative work arises. Collaboration can help organisations to develop their gateways more efficiently and
effectively. It can also help them to sustain the gateways in the longer term. The advantages of co-operation for organisations may include being able to:

- use established technologies, methods and practices - and avoid starting from scratch
- divide responsibilities for creating or sharing metadata records - and avoid duplication of effort
- combine effort for technical development - and avoid repetition of work and errors
- create joint publicity, training and promotion
- share staff effort (management/technical/administrative/cataloguing) - to make organisational efficiencies
- create shared strategies for long-term sustainability

All of these factors have the potential to improve the service that an organisation can offer to its target users.

Disadvantages of co-operation

There can be political or funding issues that rule out co-operation; indeed in some cases gateways will see competition as a natural alternative to collaboration! Disadvantages of gateway co-operation may include:

Extra expense

To make some models for co-operation work, some extra effort will be required to set up the necessary systems. For example, to make gateways interoperable some work needs to be done on making different classification schemes, metadata formats and collection development policies compatible. In the longer term, savings may be made from having co-operative strategies but the initial setup may be too expensive to consider.
**Intellectual property rights**

There is an issue surrounding ownership of metadata records which may stand in the way of co-operation. Gateways may have invested considerable resources into creating records and be unwilling to share them or give them away for free. The issue of intellectual property rights on the Internet is still a new one with some unresolved issues, and gateways would need to investigate these before entering co-operative agreements.

**Agreeing on aims and objectives**

Gateways may have incompatible aims and objectives. Having developed with particular audiences in mind, they may have reservations about the value of co-operation for their users which need to be resolved. There may also be issues for sponsors of gateways who have vested interests which need to be considered.

**Models for Co-operation**

In the library world, co-operative agreements that support information search and retrieval are commonplace. For example, national libraries each take responsibility for collecting materials published in their country and then offer user’s access to these collections via inter-library loans. Another example is the sharing of cataloguing effort, where groups of libraries work together to create union catalogues and where the catalogue records are shared and re-used by many libraries, regardless of which library actually created the record.

This co-operation enables libraries to:

- offer users access to far broader collections than could be offered by any single library
- offer users a more comprehensive catalogue than could be created by a single library
• achieve efficiencies in cataloguing and collection development without reducing the level of service to users

Such co-operation translates well into the Internet environment and the development of information gateways. Collaboration is particularly pertinent to organisations with a remit for providing access to scientific, cultural and educational resources on a large scale.

A number of different models for co-operation between gateways exist:
• Co-operative agreements for metadata records
• Gateways can create co-operative agreements regarding metadata records:

Co-operative agreements for creating metadata records
Gateways can share the effort required to create metadata records by dividing responsibilities. For example, a group of gateways can agree that each should spend time creating records for different parts of the Internet, each focusing its efforts on records for resources in a particular subject, language or from a particular country.

Co-operative agreements for using metadata records
Metadata records can be shared and re-used, and are not confined to the service which created them or to being used in only one service. Agreements on intellectual property rights would need to be established, and work is being done in this area, but the potential exists for gateways to create agreements that enable them to offer users access to records that have been created through a distributed network of gateways. Building integrated services
Co-operation can lead to the development of integrated gateway services, which offer users access to a number of gateways via a single interface. This interface might offer different levels of functionality:

**Guiding users to other gateways/mirrors of gateways**

The simplest form of co-operation is for gateways to point to other gateways that might support the user group. This may involve offering a set of hyperlinks to other related gateways, or offering mirrors of related gateways where access could be improved by keeping a local copy of the service. Although each of the gateways would have to be searched serially, the user would be alerted to other gateway services which they might not have otherwise found.

**Fully integrating distributed gateways into a single service**

In some cases it may be easier for users if they can access many gateways simultaneously. A fully integrated service offers users the chance to select a number of gateways and then to cross-search or cross-browse all the gateways in one go. A single interface offers users a single point of access to distributed gateway services. In some cases it will not be necessary to disclose to users the fact that they are searching distributed databases.

Libraries and other organisations still have a lot of work to do on the political and organisational issues involved in co-operative work. However, a number of gateway projects are now able to demonstrate some of the ways in which issues of technical and data interoperability can be solved.
Table: 5 Co-operation between gateways

<table>
<thead>
<tr>
<th>Example</th>
<th>Gateway 1</th>
<th>Gateway 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>An example of a gateway pointing to the front pages of other gateways</td>
<td>EEVL and Pinakes</td>
<td>EEVL (The Edinburgh Engineering Virtual Library) offers users a page of links to other high quality information gateways. This is simply a page that has hyperlinks to the front pages of other gateways; however, it may help users to find gateways which they did not know about.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://www.eevl.ac.uk/">http://www.eevl.ac.uk/</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://www.hw.ac.uk/libWWW/irn/pinakes/pinakes.html">http://www.hw.ac.uk/libWWW/irn/pinakes/pinakes.html</a></td>
</tr>
<tr>
<td>An example of gateways mirroring one another's services</td>
<td>SOSIG/Scout Report</td>
<td>The UK's SOSIG (Social Science Information Gateway) and the USA's Scout Report for the Social Sciences have a reciprocal agreement to mirror one another's services, to improve access for users on both sides of the Atlantic.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://scout.cs.wisc.edu/addserv/mirror/sosig">http://scout.cs.wisc.edu/addserv/mirror/sosig</a></td>
</tr>
<tr>
<td>Examples of cross-searching two gateways simultaneously</td>
<td>SOSIG and Biz/ed</td>
<td>In the UK, two gateways (SOSIG and Biz/ed) are offering users a service where two separate databases are simultaneously cross-searched via a single interface. Users are unaware that they are in fact searching two gateways, as the results are fully integrated.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Go to SOSIG: and search for industrial psychology.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>You will retrieve records from both the SOSIG and the Biz/ed databases - displayed in a single list. Both gateways use the ROADS software which enables cross-searching</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EELS and EEVL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>This is an example of two gateways based in different countries being cross-searched. Both are engineering gateways - EELS is based in Sweden and EEVL in Scotland. This is a demonstration service, but illustrates the potential for cross-searching two gateways, regardless of the fact that they are geographically separated.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><a href="http://roads.ukoln.ac.uk/eels-eevl/">http://roads.ukoln.ac.uk/eels-eevl/</a></td>
</tr>
</tbody>
</table>
An EXAMPLE of gateway standards and software that support co-operative work | CrossROADS and Interoperability | The ROADS software has been developed specifically to support the development of gateways and to ensure that those gateways are interoperable. A demonstration of how distributed gateways can be cross-searched is available from the ROADS Web site:  
http://www.ukoln.ac.uk/metadata/roads/crossroads/  
http://www.ariadne.ac.uk/issue14/metadata/  
A paper discussing interoperability issues with metadata |

EXAMPLES of plans for integrated gateway services on a national scale | RDN - The Resource Discovery Network | In the UK, government funding is being used to create the Resource Discovery Network - a gateway service for the higher education and research sectors. RDN will offer a single interface to a number of national subject gateways. Each of the services has its own identity and interface, but the RDN will offer another level of service to users - the ability to search for resources across several hubs at the same time.  
http://www.rdn.ac.uk/  
Within this project, a network of Danish libraries aims to form a virtual system to make the libraries' collective information resources (digital and traditional) available to users everywhere in the country in a simple, transparent way.  
http://www.deflink.dk/english/defihtml |

EXAMPLES of plans for an integrated gateway service on an international scale | REYNARD | The REYNARD project proposal suggests that national libraries in Europe should each assume responsibility for creating metadata records that describe high-quality Internet resources created in their own country. An integrated broker service will then be set up to enable each of the gateways to be accessed from a single interface and to allow users to cross-search the gateways.  
http://www.renardus.org |

Source: http://www.desire.org/handbook
Maintenance of Gateways

Information gateways need to be maintained in two key areas:

- collection management
- server integrity and functionality

Without adequate maintenance in these two areas a gateway is vulnerable to undermining its core aims and objectives; being a quality-controlled portal to online information resources. The key strength of an information gateway is in the quality of its data and the reliability of its service. Without adequate maintenance both of these areas are susceptible to developing weaknesses and problems.

Studies of gateway users

It is important to connect user behaviour regarding Internet resource discovery with wider issues relating to the use of information in the learning and research processes. Understanding users' behaviour in relation to gateways will enable gateway managers to position themselves within the mesh of existing gateways and meet the needs of their target audiences. The undergraduate may want to identify 'key texts', perhaps the top ten resources in a gateway, whereas the subject expert may be investigating the 'borders' of their knowledge, perhaps looking at treatment of their research topic in other disciplines. Such considerations begin to suggest opportunities for gateways to develop targeted services, providing expert and naïve users with varied interfaces to available data.

Whether child or professor is faced with the compelling option of using the global services, such as Yahoo and Google, as a first step. The undifferentiated experience offered by such services can be compared with the specialist view offered by information gateways. Gateways offer the user an alternative to the generalist approach of the commercial global
search engines, but in order to optimise the gateway service we need to gain a better understanding of user's requirements for particular types of search during the research and learning process. It would be instructive to compare information seeking behaviour and success rates for a variety of uses of global search engine as compared to gateways. Likewise one could analyse the differences in users' search strategies within the context of the traditional library, hybrid library and subject gateway.

It may be helpful to liken the subject gateway approach to the traditional 'departmental library' as the user's first port of call, a place where the user feels comfortable in a known environment and is able to gain skills to navigate a limited area of information. It would be interesting to see how far we could draw parallels between the requirements of users of subject gateways and the users of 'subject based' libraries. The users of both services benefit from an understanding of the boundaries and content of the information space they are accessing. Managers face much the same issues of specialist subject collections whether they exist on the web or in a traditional service environment, in both there is a need to relate one particular subject area to wider cross disciplinary information, and to manage the inter-relation of services.

**Importance of Subject Gateways**

- Subject gateways cover an important and the fastest growing part of resources and services a research library mediates to. The quality work, selection, description, maintenance and mediation of this constantly changing material is very demanding.
- Subject gateways correspond to traditional library functions. Quality selection, description and cataloguing, classification and indexing, annotation and recommendation, searching and browsing in
catalogues are the functions a gateway carries out for digital resources on the net. "Traditional" equivalents are

- annotated bibliographies, literature guides, pathfinders and finding aids, reference materials and services etc.
- Subject gateways transcend the limitations of their predecessors by offering catalogues of Internet resources which are not part of locally stored collections and by integrating mediation, discovery and direct access and retrieval of documents.
- Many library services and educational efforts need to make use of these Internet resources.
- Gateways are important sources for distance education and e-learning. They already play a big role in infrastructure projects in the educational sector, e.g. the European schoolnets and common information systems of universities (JISC/DNER in the UK).

The importance of subject gateways for libraries is closely connected to the question of the future role of libraries.