Chapter II

Literature Review
LITERATURE REVIEW

The literature and the Subject Gateway evaluation suggest that Subject Gateways are important elements in today’s information landscape. In contrast with Internet directories resource descriptions in Subject Gateways are of high quality and the services are of use to researchers, academics and other professionals not only in their specific field, but also could be of use to the wider community. However, the literature and evaluation also suggest that sustainability is one of the future problems that Gateways will face. User evaluations can help to identify the needs and views of a Gateway’s user base. It could be concluded, therefore, that more evaluations are needed for further research into Subject Gateways before they can be developed until they are secure in today’s information landscape.

2.1. Growth and Development of Subject Gateways

Subject Gateways offer a means of bringing control to the Internet and involve creating some kind of catalogue record for sites included. This literature review will therefore concentrate on Subject Gateways, their emergence and role in the electronic library environment in the United Kingdom, United States, Australia, aspects concerning cataloguing Internet resources, and the sustainability of Subject Gateways. The literature on Gateways and cataloguing the Internet is vast, even though they only really began to be widely considered during the 1990’s. The rapid growth of the Internet and wide use of the Internet, without any comprehensive indexing or classification system, led to a vast network of information that became increasingly difficult to navigate or to search.
Dempsey outlines the historical perspective to Gateways, which 'emerged in response to the challenge of 'resource discovery' in a rapidly developing Internet environment in the early and mid 1990's.' The term 'Subject Gateway' 'was popularized in the United Kingdom Electronic Libraries Programme (eLib), and it has been given currency by initiatives which have been influenced by the eLib Gateways(Dempsey,2000). The eLib programme was an initiative of the Joint Information Systems Committee (JISC) of the UK Higher Education Funding Councils, following the publication of the Follett Report on Libraries.(Law and Dempsey,2000). Kerr and Macleod emphasise the influence that the Follett report had on libraries, as it stated that 'the exploitation of IT is essential to create the effective library service of the future.' (Kerr and MacLeod, 1997).

The Follett report also states of the Internet that 'Its long term potential will be for information. Information which is provided quickly and reliably will eventually have a dramatic effect on the decision making functions of society.'(Follet, 1994). The Internet, then, has had and will continue to have a wide-ranging effect on culture, society, education and research. Ward states that 'The growth of the Internet has engendered an information-addicted community eager for quick access to current electronic data.' (Ward 2001).

2.2. Obstacles of Internet

The problem with relying on the Internet and Search Engines such to provide data is, as Pitschmann states, that: Search results can be so overwhelming that the user cannot be expected to evaluate them in a reasonable length of time. Moreover, organization within the Web presents most scholarly users with a technological labyrinth. Results may be far
removed from, or totally unrelated to, the desired findings. The artificial intelligence technologies employed by the major Web discovery tools are insufficient to retrieve and adequately evaluate scholarly content. (Pitschmann, 2001).

Searching the Web on any topic will retrieve the entire information pertinent to one’s query, but there will be no qualitative evaluation or filtering of the content. Only when sites have been reviewed, evaluated, selected, and catalogued will users be spared the ambiguities resulting from the randomness and ‘quantity without quality’ of Web search results. (Pitschmann, 2001).

Kerr and MacLeod agree, as they argue that ‘as the indexes of large search engines are created by indiscriminate trawling of the Internet, there is no emphasis on quality of resources retrieved.’ (Kerr and MacLeod, 1997). The explosion of information on the Internet, then, and the need to tame the potential chaos, was one of the reasons why gateways became so popular and prevalent; as Dempsey points out, there are now ‘portals everywhere.’ (Dempsey, 2000).

2.3. **Advantages of Subject Gateways**

Sladen and Spence have highlighted the advantages of using Subject Gateways: While offering a single point of access to Internet-based resources in a given field, selective subject gateways have one key feature which distinguishes them from more commercial enterprises – they are characterised by a quality control methodology based on skilled human input from the relevant academic discipline. (Sladen and Spence, 2000). Many authors have stated that library and information professionals are the ideal candidates to ‘tame’ the chaos of the Web. Ward argues that
'The library and its trusted information specialists are in a unique position to guide people to research quality information.' (Ward, 2001).

Bater argues that: We have let the Web techniques and their masters dumb down one of the biggest and most comprehensive information resources of them all, applying their Web crawlers with minimalist indexing strategies without even a nod of recognition for close-on 130 years of information science. (Bater, 2000). This contention between information professionals and computer professionals seems to present a constant problem in information environments, yet the literature does not seem to suggest that the solution may lie in better computer instruction in library schools, thus enabling newly qualified library professionals to contribute more knowledgably and meaningfully to the situation. Pitschmann states that ‘by selecting a subset of resources that meet predetermined criteria and by facilitating access to them, librarians impose a quality structure on those resources.’ (Pitschmann, 2001).

Sreenivasulu, in his article which reads almost like a job description for a digital librarian, states: The navigation of the future would tend to integrate with the human-assisted information retrieval from the networked universe and would support rapid information navigation and precision retrieval. The digital librarian is an expert in navigation, browsing and filtering digital reference services and electronic information services. The digital librarian acts as an intermediary in the task of massive Digitization of information, its storage, dissemination, managing the archive, and making available digitized networked information to the end users. Digital librarians and computers depend on each other for processing and dissemination of digital information and both are interrelated. (Sreenivasulu, 2000). Perhaps the digital librarian should also be an expert
in computer science. Ward states that 'The success and marketability of the twenty-first century library hinges upon its response to the Internet.' (Ward, 2001). The way that some librarians seem to be responding to it is by creating or collaborating in the creation of Subject Based Gateways.

2.4. **Subject Gateways and Resource Descriptions**

There are many definitions of the term 'Subject Gateway' in the literature; Koch states that: A Subject Gateway is an Internet service with a primary focus on distributed Internet resources which support systematic resource discovery, the service is based on resource description. Browsing access to the resources via Subject structure is an important feature. (Koch, 2000). Wiseman states that 'Subject Gateways provide lists of quality-tested resources in specific disciplines, and often a variety of value-added services relevant to the specific disciplines.' (Wiseman, 1998). Dempsey, Gardner and Day suggest that Gateways have some or all of the following characteristics: Services based on resource descriptions; high level of manual creation/intervention often by information and/or Subject specialists; search and browse access; collection development policy, supported by selection and quality criteria; collection management policy, supported by maintenance and updating Procedures. (Dempsey, Gardner and Day, 1997).

The Desire information Gateways Handbook states that: One of the key roles of Internet Subject Gateways is the creation of descriptive metadata about networked resources which can be used as a basis for searching and browsing the Gateway. These descriptions can also help Gateway users to identify whether the resources are really what they need, potentially saving them a considerable amount of time browsing through
the information available elsewhere on the Internet. Resource description, or cataloguing, has been a library practice for a very long time.

Taylor states that ‘The Anglo-American Cataloging Rules, Second Edition including all its updates and revisions, has served successfully in providing guidance in cataloging materials of all kinds for two decades.’(Taylor, 1999). But as Hsieh-Yee points out, 'The nature and variety of resources on the Internet underscore the need for information organization in the digital age the question is whether cataloguing is the answer.'(Hsieh-Yee, 2000).

Medeiros is in no doubt that bibliographic cataloguing is not the answer: ‘the point is clear: traditional MARC/AACR based catalogs and cataloging cannot handle the Web.’(Me deiros, 1999). Dempsey suggests that: What is required is a description which is simple to create yet full enough for effective retrieval and relevance judgment. This implies a description which falls between the terseness of the crawlers and the fullness of a research library catalogue. (Dempsey, 2000).

Ayres argues that traditional cataloguing is too complicated, especially for the user: Users are often baffled or more often frustrated in attempting to use a catalogue and it is not surprising if they sometimes wonder if the complexity is really necessary or even worse whether it is based on some sort of witchcraft. (Ayres, 1999).

Medieros comments on the slow change process that he sees as dogging the library profession: Changes rarely take a quick route in libraries tradition and complacency hook the profession, often preventing
its practitioners from seeing better alternatives or complements to the way things have always been done. (Medeiros, 1999).

Taylor makes a valid point when he states that 'it is not as if the 3 x 5 card is still governing the amount of space available.' (Taylor, 1999). Commentators as little as four years ago were warning against the perils of being swept away by the Internet: We should not allow ourselves to be seduced by the current popularity and 'trendiness' of electronic resources and the Internet into either totally rejecting or existing, traditional means of providing access and organization of information where those means are appropriate for the type of information, or jumping on to technological bandwagons and disregarding the fact that resources on the Internet are purely other potential resources. The Internet per se does not make data into knowledge, and does not make electronic information more or less desirable than any other media in which it is held users may, indeed have come to regard the Internet as their preferred research tool. Information at the touch of a button is what students of today have come to expect (Strutt, 1997).

Medeiros counteracts Strutt’s comments with the now prevailing view of the information seeking environment: While the library community forges new ground in an attempt to catalog portions of the Web, we are at the same time diminishing the prominence of print resources and the online catalog. More and more, users want, expect and pursue full text (Medeiros, 1999).

Some librarians consider Internet materials too much unstable to be effectively cataloged others wondered if increasingly powerful search and
browsing tools, automated indexing tools and 'intelligent agents' would obviate the need for catalogs.(Vinh-The Lam, 2000).

There have been changes in traditional cataloguing that have tried to take account of the changing cataloguing environment for 'computer files'.(Desire Information Gateways Handbook). Other strategies have been devised in an attempt to deal with this problem, such as the Dublin Core Metadata Initiative which 'provides card catalog-like definitions for defining the properties of objects for Web-based resource discovery systems and which serves as a 'simple alternative for library catalog records for Web resources.' (Weibel & Koch, 2000). There is also Nancy Olson's 'Cataloging Internet Resources: a manual and practical guide' which was published in response to OCLC's Internet Cataloguing Project begun in 1991.(Parris Sibley, 1998).

The ROADS (Resource Organization and Discovery in Subject based services), which was funded by JISC via the eLib programme was: A collaborative project which contributed to the broad aims of the electronic libraries programme by providing a software tool-kit and a standard framework for the information gateways being developed under the Access to Network Resources (ANR) initiative. This project produced software which provided support for gateways and which many of the original eLib gateways used.

A comparative evaluation of Subject Gateways by Haynes in 1998 stated that 'it was argued that the prime role for subject gateways is to make critical assessments of resources.'(Haynes, et al., 1998). This implies, then, that it is not solely the catalogue record which is important, but the fact that resources are critically assessed, reviewed and 'bad' sites
filtered out; but an interesting flip-side to this argument is voiced by Sladen and Spence: If subject gateways only select good quality resources, to what extent do they take away the user's ability to decide for themselves whether to consult a resource? This has been one of the major challenges to the whole concept of subject gateways. (Sladen & Spence, 2000).

The constant exposure of readers to quality websites, with catalogue records which indicate author details and other items of information to suggest that the information comes from a reliable source, would alert the user when and if they come across unreliable sites which do not display this information. It is a matter of concern that this information is sometimes difficult to find, even on the most reputable sites. Another question to be addressed is if Internet resources can be included in a library's catalogue, why then create separate Subject Gateways for them?

Medeiros argues that 'Believing that all quality resources deserve a place in the online catalog is honorable, though no longer practical. We must think beyond the scope of the OPAC and apply a more holistic approach to information provision.' (Medeiros, 1999). By including them in a quality controlled Subject Gateway gives the user peace of mind, users can be assured that access to these sites will be stable, and that resource discovery will be tailored to the characteristics of the collection and its content, rather than to features of a specific search engine. (Pitschmann, 2001).

2.5. Subject Gateways in Future

One of the problems for the future of Subject Gateways is why collect free Web resources? The obvious answer is that current users need facilitated, value-added access to these resources to ensure that they will retrieve sites with high-quality content. The primary question for the future
is whether broad application of enhanced metadata standards and next generation search engines will allow end users to mine the Web themselves with greater precision than is currently possible and, in so doing, bypass the current need for facilitated access. (Pitschmann, 2001)

Mining the web using search engines will not, however, provide evaluated sites but users will still have to perform the evaluation procedure for each site themselves, until technology can facilitate the harvesting and cataloguing processes, manual practices will continue to be used.

Subject Gateways would appear to have a future problem which is that of sustainability. Developing and managing collections of free Web resources have wide-ranging, long-term implications for human resources, organizational issues, and fiscal matters that extend well beyond the circle of individuals responsible for selecting those resources.

Pitchmann argues that the creation and maintenance of Subject Gateways calls into question the work priorities of library staff: Staff members responsible for selecting and cataloguing analog materials have full-time jobs. Increasing their responsibilities to include developing collections of free Web resources calls into question pre-existing priorities which is a higher priority: processing new books that are not free, or cataloguing free Web sites? (Pitchmann, 2001). This argument does not consider the rapidly changing library environment, where electronic resources are replacing 'analogue' materials, but the point is valid: Subject Gateways do take up staff time and effort. As Heery states: 'Creating resource descriptions is one of the most time-consuming and costly tasks for gateways.' (Heery, 2000).
The 'business' side of Gateways is something that the literature suggests has been ignored by Gateway organizers: It is probably fair to point out that those developing information gateways have tended to spend more time considering technical and intellectual issues like interoperability and gateway selection criteria rather than to review the business and legal context in which gateways exist. The focus now needs to change; it is becoming clear that the development of sustainable gateway services will be dependent upon a clear understanding of the business and legal context in which information gateway services operate. (Day, 2000). These include

Collective activity through membership where Gateways indulge in collective activity through membership of some kind of organization, but would still need ‘to develop services that might attract paying members. ‘Shared public investment’ where ‘Gateways are funded as a service by some public body. ‘Publicly funded research and development’ which has been used by many Gateways. Public investment as part of the role of cultural, educational or scholarly institutions’, where Gateway funding is seen ‘as part of the role of developing cultural environment’, as a type of public investment.

Finally there is the ‘commercial model’ which could include ‘investment in resource description activity as ‘added value’ to a range of subscription-type services.’ Or ‘providing data supported by advertising or other services’ based on the value of the attention of visiting users’. Gateway services (or those who host them) own the IPR in the resource descriptions (metadata) that they have created and maintained. These are resources of significant value in their own right. The Intellectual Property Right in the resource description could again, then, be seen as a potential means of creating revenue. Choosing an appropriate business model is
important for gateways because it has a strong influence on how sustainable any given service is.

The literature discussed in this review illustrates that the information profession has come a long way in taming the web through the devising and implementing of quality-controlled Subject Gateways. It also indicates that there are still hurdles to be overcome if they are to become permanent, sustainable elements of the information landscape. To find out exactly what users think of them and how they use them is one more step that could help us in their future development.