CRITERIA FOR EVALUATING INFORMATION ON THE INTERNET: PERSONAL PERSPECTIVE

by

Murali N *

ABSTRACT

Internet is a valuable information resource, but it also contains lot of information, which is incomplete, out-of-date, or inaccurate. Author considers the importance of standards and contents of sites on the World Wide Web, specifically in Indian context, a field where scope, coverage and validity of information is vital for both academic and research community. An attempt has been made to identify scope and policy for evaluating Internet resources. The criteria, which are to be taken into consideration while evaluating the Internet resources, including authority, accessibility, arrangement, currency, response time, stability, target audience, accuracy, coverage, completeness and style are discussed. An effort has been made to evaluate and catalog few Newsletters and Databases, which are freely available on the Web. The evaluation of the quality of information available on the Internet could require a large amount of human effort.

* Scientific Officer, INFLIBNET, Ahmedabad – 380 009, India

0.   Introduction: cataloguing and the Internet Resources

Modern descriptive cataloguing theory and practice has developed over the past 150 years as a means of organising information for retrieval in libraries. Library catalogues typically consist of a collection of bibliographic records that describe published materials. But at present, electronic information is also a part of our library collection and needs to be catalogued. There is a great deal of valuable information available through the Internet and these resources need to be organised for accessibility. Existing library techniques and procedures can be used along with metadata standards for cataloguing electronic resources, which are freely available on the Web. Some have cataloguing home pages or Web pages as a first step in providing access to electronic resources. Some have cataloguing e-journals and OPAC's. An effort has been made to evaluate and catalog few Newsletters and Databases, which are freely available on the Web. Internet resources to be cataloged should be selected as carefully as any other material that is added to a collection.

The present paper is an attempt to define criteria and procedures to select qualitative resources on the Web.

1 Scope Policy

1.1 Objectives and Audience

Objective: The main objective is to give access to a collection of high quality resources that are freely available on the Internet in the public domain.

Audience: Who is the intended audience for the identified source? The primary user group to be serverd is the Indian Academic and Research community. The service aims to meet the information needs of researchers, students and teachers in higher education.

1.2 Information Coverage

Subject Matter: Only information relevant to students, researchers, librarians and other user communities is considered. It aims to cover all Indian academic disciplines based on the collection profiles. This includes:
- ?? Websites published by Indian Organizations
- ?? Websites available from Indian Servers
- ?? Resources written by Indian Authors
- ?? Resources written in Indian Languages
- ?? Websites and documents about the Indian culture and society in remote
- ?? servers

Acceptable Sources: Information from academic, government, commercial, trade and industry, non-profit and private sources are all acceptable provided they fall under the subject matter criteria.

Levels of Difficulty: The content of the resources should be at a level suitable for higher research and education. Information should be scholarly rather than popular.
Types of Resources: Resources should be freely available on the Internet. Resources intended for an individual should not be selected. Resources that consist exclusively of links to other resources will not be selected. Resources that give a very comprehensive overview of a subject area and are well maintained are to be selected. In this case, the added value is the completeness of the coverage.

Level of Selection: Only separate "information entities" will be selected
- ?? The Website of an organization
- ?? A document (article(s), book(s), OPAC(s), discussion fora, software's, e-journals, newsletters etc.,)
- ?? A secondary information resource (Bibliography, information gateway)

What an "information entity" is not

Advertisements: Resources that consist for a considerable part of publicity materials, more advertisements information will not be selected.

1.3 Geographical Coverage

Restrictions: There are no geographical restraints. First preference is to cover the Indian resources and later other resources will be covered outside India.

Language: There are no restraints in terms of Indian Languages.

1.4   Access

Technology: There are no restraints on selection of resources that use advanced WWW technology.

Registration: Resources, which require registration (without costs) will be selected.

1.5   Resource Descriptors

Metadata: The following Dublin Core elements are partially used to describe the resource.

Element: Title	
Definition:	A name given to the resource.
Element: Creator	
Definition:	An entity primarily responsible for making the content of the resource.
Element: Subject or Keywords	
Definition:	The topic of the content of the resource.
Element: Description	
Definition:	An account of the content of the resource.
Element: Publisher	
Definition:	An entity responsible for making the resource available
Element: Contributor	
Definition:	An entity responsible for making contributions to the content of the resource.
Element: Date	
Definition:	A date associated with an event in the life cycle of the resource.
Element: Type	
Definition:	The nature or genre of the content of the resource.
Element: Format	
Definition:	The physical or digital manifestation of the resource.
Element: Identifier	
Definition:	An unambiguous reference to the resource within a given context.
Element: Source	
Definition:	A Reference to a resource from which the present resource is derived.
Element: Language	
Definition:	A language of the intellectual content of the resource.
Element: Relation	
Definition:	A reference to a related resource.
Element: Coverage	
Definition:	The extent or scope of the content of the resource.

Element: Rights	
Definition:	Information about rights held in and over the resource [12].

2 Selection Criteria

2.1 Evaluating Internet resources

Many of the methods used to evaluate print sources, such as journal articles and books, can also apply to the evaluation of resources on the Internet. In addition, there are some evaluation criteria that are unique to Internet resources. Before publishing, most books and journal articles undergo a peer review process but anyone can publish anything on the Internet. So information often does not go through traditional quality 'filters' such as publishers, editors or peer reviewers. Resources may not be what they appear to be or what they say they are, as on the Internet, there is no guarantee that the resource is accurate or honest. For this reason, it is imperative to critically evaluate all information before use.

Resources should be evaluated based on three types of criteria:

(a) The Information Content
(b) The form in which the information is presented and
(c) The process that support/ guarantee Internet Access

2.2 Quality Criteria

The information content is a primary consideration when evaluating Internet resources for cataloguing. It aims to point to primary information and not just lists of links. Information should be valid, accurate and current. Content criteria concern the evaluation of the information that the resource contains.

Validity: The validity of a resource depends on how well researched, well substantiated and trustworthy the content of the resource is.

Accuracy: It is closely related to validity. The accuracy of the resource depends on how accurate the information is.

Authority: The authority of the resource depends on the expertise, reputation and status of the author/producer. On the Internet, the origin of a resource is not always clearly stated.

Uniqueness: Uniqueness depends on the quantity of primary information that is contained in the resource and is not available from other sources. On the Internet, a lot of resources contain relatively little primary information, but consist mainly of links to other, external resources, available from other servers, or that reproduce information available elsewhere. A resource that provides information that is available nowhere else (online), is in most important.

Substantive ness: The substantive ness of the resources is the quantity of information that is actually provided on the site. On the Web, where information is hyper linked, it is not always clear where one resource ends and another begins. Many resources contain very little real information.

Comprehensiveness: Comprehensiveness means that the information is provided in a fairly complete form. Ex: On the Web information is often provided that is not yet completed ("under construction"), or that is only partly made available online and refers to printed or off line versions for the complete resource.

Coverage: Depth and extent of the information.

2.3   Form Criteria

Form criteria concern the presentation and the structure of the information and the interface.

Navigation: It aims to point to resources that are user-friendly. They should be easy to navigate, provide adequate user-support and make appropriate use of technology.

Provision of user support: User support includes various facilities to help end users answer their questions and solve any problems they may have. Users may need support concerning the content of the resource or the mode of access to the technology used.

Use of recognized standards and technologies: The use of technology and standards is preferred that gives users, access and allow them to make use of all components of a site.

Internet standards and technologies are continuously under development and often influence the level of access for users who have various hardware and software at their disposal.

2.4   Process Criteria

Process criteria refer to the processes and systems, which exist to support the information resource. The system that is situated between the creation of the information by the author and the access to the information by end users is determined by many parameters that may influence the value of a resource for an end user. Internet information is volatile and likely to change over time. Resources may at any moment be changed, moved or deleted. Old resources or older versions will not always be archived.

Information Integrity: Information integrity refers to the stability of the content of the resource over time. This depends on the author or the information provider.

Site Integrity: Site integrity means guaranteed access over time. This is normally the work of the site manager or the Webmaster. Sites may be moved or removed by those that are responsible for making the resource available over the Internet.

System Integrity: System integrity has to do with the permanent accessibility of the site over time - usually the work of the systems administrator. If a server is not permanently accessible, this may influence the value of the resource, independent of the intrinsic value of the information content.

3   Organizing Resources
Resources are planned to be organized in one or more ways, including:
   ?? By subject (e.g., sports can be broken up into soccer, cricket, etc.)
   ?? By format (e.g., mailing lists, Web sites, etc.).

?? By audience (e.g., academic vs. general users).
?? By chronology (for a history guide).
?? By geography (for a guide to a region).
?? By authors (for a literature guide).

4   Practical Implementation

Cataloguing of Internet Resources, defined as the creation of surrogate records, are be used to facilitate the identification, location, access, and use of resources (Levy, 1995). These descriptions are usually created in accordance with certain standards (cataloguing rules and metadata formats).

Apparently, it is a relatively easy task to evaluate Web resources available on the Web. Here there is a list of Newsletters and Databases, which are freely available on the Web. Define all these databases and newsletters using minimum metadata elements. All these Electronic databases turn up as very handy tools for searching vast bibliographic data within shortest time. There are very good number of such databases available free on the Internet today, which can be accessed free of cost. Most of them mentioned here are bibliographical databases and some give links to full text articles too.

> **Databases**

Database: MEDLINE

Description:
It is a database on international literature, produced by the National Library of Medicine - NLM, which gathers bibliographic references and summaries of more than 4000 biomedical journals published in the United States and in other 70 countries. There are approximately 11 million records on biomedical literature since 1966 which cover medicine, nursing, odontology, veterinary and preclinical sciences. Updating is done on a monthly basis. MEDLINE is divided into sections by date.
Coverage: 1966 to till date
URL: http://www.bireme.br/bvs/I/ibd.htm   (WWWISIS)
    http://www.ncbi.nlm.nih.gov/PubMed/   (PubMed)

Database: LILACS - Latin American and Caribbean Health Sciences

Description:
Latin American and Caribbean Literature on the Health Sciences is a BIREME System's cooperative database which covers literature related to the health sciences and has been published in countries of the Region since 1982. It contains articles from about 670 of the most well-known journals in the medicine field, reaching approximately more than 150,000 records and other documents, such as: theses, chapters of theses, books, chapters of books, congress and conference proceedings, technical and scientific reports and governmental publications.
Coverage: Not Available
URL: http://www.bireme.br/bvs/I/ibd.htm

Database: PubSCIENCE

Description:
PubSCIENCE is a natural evolution of OSTI tools and services dating from the late 1940s. Until recently, the two major end tools for announcing OSTI's collection of journal information have been Nuclear Science Abstracts (NSA) and the Energy Science and Technology Database (EDB). Both these databases contain metadata -- offering citations and abstracts and stating the availability of documents for users who sought to obtain full-text copies. NSA is an historical record of nuclear research beginning with the Manhattan Project and following throughout the life of the Atomic Energy Commission. It contains international nuclear science and technology references from the early 1940s through June 1976. With the creation of the Energy Research and Development Administration and then the Department of Energy (DOE) in the 1970s, the scope of the agency was broadened. EDB then supplanted NSA to provide a comprehensive source of worldwide energy-related information, both nuclear and non-nuclear. It covered 1974 to the present and included information energy sources, use and conservation, environmental effects, waste processing and disposal, regulatory considerations, as well as basic scientific studies. The scope of NSA and EDB mirrored the scope of the R&D program of the parent agency. Together these databases offered more than 5 million records in energy science and technology. OSTI's newest tool, Energy Citations Database (ECD) contains bibliographic records for energy and energy-related scientific and technical information from the DOE and its predecessor agencies from 1948 through the present.

Coverage: 1940's to till date
URL: http://pubsci.osti.gov/srchfrm.html

Newsletters

Subject: Library Science
Publisher: INFLIBNET
Title: INFLIBNET Newsletter

Description:
The objective of this newsletter is to create awareness about INFLIBNET and its activities. It also brings information on latest developments in Library and Information Science. Quarterly publication and published 4 times in a year.

Keywords: INFLIBNET Newsletter, Information and Library Network Centre.
Periodicity: Quarterly
Online Availability: 1998+
Format: html/printed
Country: India, Ahmedabad
Access Type: Online
URL: http://www.inflibnet.ac.in/newsletter/oldnewsletter.htm
Date: 2001-07-05 (YYYY-MM-DD)
Creator: Murali
ISSN: 0971-9849

Subject: Technology, Intellectual Property rights, Forecasting
Publisher: Technology Information, Forecasting and Assessment Council (TIFAC).
Title: TIFAC News (Newsletter of Technology Information, Forecasting and Assessment Council)

Description:
With a view to disseminate information about the current projects and programmes of TIFAC, a quarterly newsletter - 'TIFAC News' is being brought out on regular basis so as to forge strong forward and backward linkages between the academic, R&D labs and industrial sectors respectively. It also deals with the wide range of subjects on IPR (Intellectual Property rights), including user-friendly explanation of different types of IPR issues.

Keywords: TIFAC News, Technology Information, Forecasting
Periodicity: Monthly
Online Availability: February 1999+
Format: text/html
Country: India
Access Type: Online
URL: http://www.tifac.org.in/news/newlett.htm
Date: 2001-07-01 (YYYY-MM-DD)
Creator: Murali

5   Conclusion

Keep in mind how best to identify the quality of an Internet resource in this volatile, continually changing environment. Practical efforts have been made in this context to evaluate quality resources on the net and catalog them to serve primary user group of Indian Research and Academic community. The Web is easy to use, both for finding information and for publishing it electronically. Because so much information is available on the Internet, and because that information can appear to be fairly "anonymous", it is necessary to develop skills to evaluate and catalog them. When using a research or academic library, scholars, publishers and librarians have already evaluated the books, journals and other resources. When you are using the World Wide Web, none of this applies.

6   References

1. ALEXANDER (J) and TATE (M A). Teaching critical evaluation skills for World Wide Web resources. Computers in Libraries. 16, 10, 1996 (November/December). p 49-55.
2. ALEXANDER (J) and TATE (M A). How to evaluate and create information quality on the Web: Web wisdom, Mahwah. 1999. Lawrence Erlbaum Associates.
3. BRANDT (D S). Evaluating information on the Internet. Computers in Libraries. 16,5; 1996. p 44-46.
4. AGUILLO (I). A new generation of tools for search, recovery and quality evaluation of World Wide Web medical resources. Online Information Review. 24,2; 2000. p 138-143.
5. MERRILL (J). The Internet and classical civilization. Acquisitions Librarian. 23, 2000. p 97-115.
6. JACOBSON (T). Evaluating Internet Resources.

http://library.albany.edu/internet/evaluate.html

7. AUER (NICOLE). Library Research at Virginia Tech. Evaluating Internet Resources.
    http://www.lib.vt.edu/research/evaluate/evaluating.html
8. HARRIS (ROBERT). Vanguard University of Southern California.
    Evaluating Internet Research Sources.
    http://www.virtualsalt.com/evalu8it.htm
9. COOKE (ALISON). Medical Library, How to evaluate an Internet-based information
    Source.
    http://biome.ac.uk/guidelines/eval/howto.html
10. CORE (DUBLIN). Metadata Element Set.
    http://dublincore.org/documents/dces/
11. CHAPMAN (ANN) et al. Cataloguing practice and Internet subject-based information
    gateways.
    http://www.ariadne.ac.uk/issue18/metadata/