

# A Library-Organized Virtual Science and Technology Reference Collection

Gerard McKiernan  
*Iowa State University*

---

Gerard McKiernan, "A Library-Organized Virtual Science and Technology Reference Collection." *Proceedings of the IATUL Conferences*. Paper 11.

<http://docs.lib.purdue.edu/iatul/1996/papers/11>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact [epubs@purdue.edu](mailto:epubs@purdue.edu) for additional information.

# CyberRef(sm): A 'Universally-Available' Technology Reference Collection

*McKiernan, Gerard*  
*Iowa State University*

---

## *“Save the Time of the User”*

### **I. Information Retrieval**

While there have been noteworthy efforts to facilitate the identification of relevant Internet and World Wide Web (WWW) data and information sources, users continue to be confronted with the cumbersome task of effectively selecting, evaluating, and using resources that best meet their information needs. In Fall 1995, CyberStacks(sm) <http://www.public.iastate.edu/~CYBERSTACKS/>

was formally established in response to limitations of early efforts to organize access to Web and other Net resources. [1] CyberStacks(sm) offers an alternative to inherent and perceived inadequacies of original Internet directories and search services by providing users with efficient access to significant sources in clearly defined and appropriate categories.

Although most Net search tools typically retrieve a substantial volume of candidate resources, too often the resulting sets provide the user with sources that are poorly described, unorganized, or of questionable value. Significant resources that are known to exist are often not included within a particular Internet directory or indexed by a particular search engine, or if included, not differentiated from less significant resources that cover a similar or related topic. In many cases, to identify sources that might be appropriate, users far too often are burdened with the task of reviewing numerous resources from a seemingly endless and disorganized list. At sites that have selected the more significant resources from those generally accessible, these typically are arranged alphabetically, or at best, organized under broad categories. While value has been added in these cases, users are still required to examine extensive lists to identify those resources of greatest potential importance. Although some sites do enhance access to selected resources by rudimentary organization, few offer appropriate description or annotation, thereby placing an additional burden on users to identify those resources which might satisfy an information need. As the conceptual relationship to other resources is also rarely indicated, users must also review a complete set to identify any and all *related* resources.

A number of the general concerns relating to efficient access to Internet resources and their effective organization have been of specific interest to those involved with Information Retrieval (IR) research for more than a generation. Such issues as browsing, recall and precision, relevance feedback, routing and filtering, information extraction, and effective retrieval are equally as germane to the optimal functioning of local or remote database systems as they are to productive searching of the Internet. The work of researchers such as those at the Center for Intelligent Information

Retrieval at the University of Massachusetts, Amherst, is particularly noteworthy for recognizing the similarity of problems faced in each environment. A recent paper by Croft offers a concise and clear summary of the common challenges and desires of those seeking more effective retrieval of information. [2]

## **II. Library of Congress Classification Scheme**

Our perception of the deficiencies and inadequacies of the initial efforts and attempts to provide more efficient and more effective use of World Wide Web (WWW) and other Internet resources led us to consider the use of a standard library classification scheme as a framework for enhancing retrieval of Net resources. The idea of creating an organized collection of selected Web resources using an established classification system originated in a series of listserv and newsgroup postings that began in summer 1994 [3]. Such discussions were directly influenced by an extensive weeding project of the Reference Collection that was underway at Iowa State University and by the recognition that the Internet offered an opportunity to provide Reference Desk staff, as well as library clientele, with direct access to significant resources that could augment or replace titles in our local Reference stacks.

The intensive review of the science and technology reference titles during this project reinforced the importance of the Library of Congress classification scheme as an effective system for indicating subject coverage, content, and relationships among published works. While we had identified some Web sites during an early review period that had adopted this scheme as an organizational framework, none had applied the system at the level of specificity nor provided the breadth of description that we believed could enhance effective use of selected resources. Although such sites had adapted the scheme, specific relationships to other resources were typically not indicated. In creating CyberStacks(sm), we believed that a comprehensive application of such an established classification system would offer the value-added structure, organization, and context that could offer more effective identification and retrieval of relevant and related Web and Net resources.

## **III. Virtual Library**

During the early implementation of CyberStacks(sm) it became obvious that to realize an Internet service as envisioned, it was fundamental that the nature and scope of resources to be managed within it be defined and that the criteria used for their selection be delineated. After reviewing a number of existing efforts, we decided to establish CyberStacks(sm) as a collection of significant Reference resources in science and technology. [4] As we wished to create an electronic counterpart to our physical Reference Collection, we decided initially to seek and include Net resources that were the equivalent, or an analog, to a print, or other electronic reference work, for our defined Internet collection. Therefore, nearly all resources selected for inclusion within CyberStacks(sm) are full-text, hypertext or hypermedia sources of a scholarly or research nature.

In seeking to expedite access to Internet resources within an experimental prototype, we sought not to create a digital library *per se*, but a true 'virtual' library in which resources would metaphorically 'stand together' in virtual library stacks. Underlying the approach taken by CyberStacks(sm) is the use of a hypertext outline of the Library

of Congress classification schedules. [5] Through a series of hypertext links, CyberStacks(sm) allows users to browse through broad classifications (e.g. Q [Science], R [Medicine], S [Agriculture], T [Technology]), less broad subclasses (T [Technology (General)], TA [Engineering (General) and Civil Engineering], TC [Hydraulic Engineering], TD [Environmental Technology & Sanitary Engineering], etc.) to a specific classification range (T1 - T53 (Technology (General), T54 - T55.3 (Industrial Safety), T 55.4 - T 60.8 (Industrial Engineering), Patents & Trademarks (T 201-T 342), etc.) and associated description that we believe best characterize the coverage and content of the resource (e.g., Dictionaries and Encyclopedias). For each, its title, along with a brief summary, are provided, as are specific instructions on effectively navigating and using the resource. Subject coverage, scope, record structure, special features, source acknowledgment, as well as the mode of access, when appropriate, are included in a typical profile. When and where possible, the summary data and information are excerpted from the original source description. No attempt is made, however, to standardize the format of the data.

#### **IV. Reference Works**

A Reference publication may generally be defined as a work which arranges and treats its subject matter for consultation for definite items of information rather than to be read comprehensively. Such works may be divided into two main groups:

- 1) compilations that furnish information directly, and
- 2) compilations that refer to other sources containing information.

Encyclopedias, dictionaries, almanacs, handbooks, yearbooks, biographical sources, directories, atlases and gazetteers are primary examples of the first group; catalogs, bibliographies and indexes are common examples of the second. [6] While we believed it most appropriate to develop a collection of selected Internet Reference resources to serve the practical and potential needs of our departmental specialists and clientele, a focussed collection of Net resources with Reference value was also particularly suitable in view of the function that Reference works historically have played in helping users locate primary and secondary information. Over the years, to assist users in accessing an ever-increasing volume of information, librarians and others have developed or applied a variety of selection and organizational tools and techniques. Guides, handbooks, review articles, literature reviews, abridgments and rankings, as well as indexes, digests and abstracts, are among the conventional tools created to facilitate access to such information . [7] The establishment of a service that extends access to these and similar types of tools via the Web is particularly fitting in view of the difficulty commonly expressed by many in their attempts to easily locate information via the Net. Indeed, many librarians have sought to facilitate access and use of Internet resources by creating the electronic analog of conventional guides and reviews.

#### **V. A 'Universally-Available' Collection**

In a recent article, Ensor recounts her initial attempts to locate information on the Internet, and outlines her use of an Internet search engine to successfully identify significant resources on a research topic. She subsequently describes her evaluation of these resources, and her plans to "arrange these items hierarchically by subjects in a way ... [she] believe[s] will be most helpful to those using the pages." [8] Ensor's

account is representative of many such efforts that have been made to enhance access to Internet resources by careful selection and organization. However, while Ensor applauds these singular initiatives, she also questions whether such separate efforts are the best “way to ‘organize’ the future” [9]. In considering the role that librarians can play in organizing significant Internet resources, we should heed her concern. Librarians must look beyond individual and local needs, and begin to give full consideration to the ways in which our professional skills can be applied not just to the creation of a special collection of Net resources, but also to the larger issue of the development of *centralized, integrated and unified* national collections of Internet resources which can be built through cooperative collection development projects in a networked environment. Britten’s earlier call for the creation of a central Virtual Library Collection, “in place of dozens, yet separate collections,” and for collaborative collection development by subject specialists using the Internet for the selection of Web and other Net resources, provides additional support for more comprehensive efforts to manage Internet resources. [10]

There are several local, national and international projects that can serve as models for a broader cooperative effort to organize Net resources. The Engineering Electronic Library, Sweden (EELS) [<http://www.ub2.lu.se/eel/>], a cooperative project of The Swedish University Technology Libraries, offers an excellent example of the use of cooperative collection development within defined subject areas in a wide range of engineering disciplines among universities. The Edinburgh Engineering Virtual Library (EEVL) project [<http://eevl.icbl.hw.ac.uk>], a more recent endeavor led by staff at the Heriot-Watt University Library, Scotland, provides a general model for broader collaboration among universities, government and a professional society. The NISS Information Gateway [<http://www.niss.ac.uk>] is an outstanding example of a cooperative Internet cataloging project in Great Britain that can well serve as a model for cooperative cataloging as well as collection development. Within the United States, the OCLC Internet Cataloging Project [<http://www.oclc.org/oclc/man/catproj/catcall.htm>] is an excellent example of the application of an established cooperative cataloging model for managing access to Net resources.

As recognized by Britten, the Internet offers the opportunity for a network ‘copy’ to serve the needs of many libraries simultaneously. [11] Indeed, during the process of identifying resources for a virtual collection, it gradually became evident that it was not necessary to limit our selection of candidate resources only to those that met the needs of our local specialists or clientele. As we did not physically acquire such works nor did such works compete for shelf space with publications of greater relevance to our local programs, we began to recognize the potential of creating a centralized, *universally-available* Reference collection that could serve the needs of Reference specialists and clientele at other institutions and organizations as well. [12] With appropriate HTML coding, any and all libraries could link to the CyberStacks(sm) collection either as a primary source of significant Reference resources in science and technology, or as a supplement to local or other efforts.

## **VI. CyberRef(sm)**

Although we had not surveyed sites to determine the potential number of resources that might be appropriate for inclusion within an experimental prototype, we believed

that these would not exceed several dozen unique sources. However, after an extensive review of obvious and potential sites undertaken during the Winter of 95/96, more than *five hundred* (500) relevant resources that met the general and specific criteria for potential inclusion, were identified. These not only included a significant number of resources in Science (Q), but in Medicine (R), Agriculture (S) and Technology (T), as well as related areas (e.g., Military Science (U)). Within Technology (T), these included resources in the full range of conventional Reference works, including biographies, classification outlines, data sheets, dictionaries and glossaries, directories, encyclopedias, handbooks and tables, as well bibliographies, catalogs, and a variety of databases and indexes and abstracts.

While a range of conventional and innovative forms and formats of Reference resources have been made available over the Net, subject coverage is not yet as comprehensive nor complete as print publications. However, a number of notable examples for many of the more common categories of Reference works were identified in the review period.

### **Abstracts and Indexes**

While more and more commercial abstract and indexing services are providing access to their respective databases over the Web for a direct or an auxiliary fee, a number of professional societies offer free access to their Net indexes and abstracts. *The SPE Publication Index* [[http://pumpjack.tamu.edu:80/cgi-bin/speindex?Source=spe\\_publications](http://pumpjack.tamu.edu:80/cgi-bin/speindex?Source=spe_publications)], a forms-based index to the publications of the Society of Petroleum Engineers, is accessible at no cost. It not only offers conventional search options by author, title or year of publication, but range searching by date and combined field searching as well (e.g, author and journal title).

### **Acronyms**

Sources that provide access to acronyms in a variety of technological and non-technological areas were among the most common Net Reference works identified during the review period. These took the form of gateways, glossaries, expanders, handbook indexes and reference lists. A notable example of this type of Reference work that covers various fields of technology is the *Acronym Gateway* [<http://haleakala.jpl.nasa.gov/cgi-bin/acrom.sh>]. This searchable acronym database was originally compiled by Charles Voge of the Jet Propulsion Laboratory California Institute of Technology, and focusses on the subjects of data processing, telecommunications, remote sensing, and space related topics. Although there is greater opportunity and interest in collaboration via the Net, often it is quite common for a specialist, as in this case, to spearhead the development of a new Net resource or service.

### **Bibliographies**

One of the more widespread initiatives undertaken by many organizations in an effort to create a presence on the Net is the conversion of existing paper references aids to Web documents through simple HTML markup. Selected or comprehensive bibliographies on timely topics not unlike the *Dredging and Dredge Disposal Bibliography* [<http://www.lib.noaa.gov/edocs/dredge.html>] are typical.

## Catalogs

Many professional organizations have adopted the Web as a medium by which they can promote their services and products. It is not uncommon for such sites to include not only a profile of their activities, but also a database or directory of their publications. The *ANSI Catalog* [<http://www.ansi.org/catalog.html>] is a noteworthy example of such a directory that serves also as an index to over 11,000 approved American National Standards.

## Classification Outlines

With the migration of many standard Reference works to the Web, many sites are not only enhancing access by creating a local searchable database of such works, but also converting associated print user aids as well, providing users with an up-to-date guide that can enhance the effectiveness of searching a Reference work in a new format. The HTML markup of the *U.S. Patent Classification* [<http://patents.cnidr.org/pto/classes.html>] is a notable case in point. Users are provided not only with a description of each class, but also information about related classes and subclasses.

## Data Sheets

Data sheets present essential data in a structured format that can be conveniently reviewed and/or printed on-demand. In sources such as the *Material Safety Data Sheet Index (MDAS)* [<http://www.enviro-net.com/technical/msds/>], users can readily locate summary information and need not be concerned with conducting an extensive search of Web sites for required data.

## Databases

The development and application of Web-based search engines for datasets that historically have been available primarily in a print format, has enhanced their usefulness significantly. Like data sheets, databases can provide ready access to data that would otherwise require users to search extensively in a variety of sources. Databases as such the *Hazardous Chemical Database* [<http://odin.chemistry.uakron.edu/erd/>] not only provide users with direct access to data by means of a keyword search, but also offer multiple keyword, wildcard and truncation searching, in addition to basic Boolean searching. Many Web databases, such as the *U.S. Patent Database* [<http://patents.cnidr.org:4242/access/access.html>] offer not only simple keyword and Boolean searching, but field searching and command-line searching as well.

## Dictionaries and Glossaries

Among the more common resources linked to a subject-oriented Web site are dictionaries. The *SEMATECH Official Dictionary* [<http://www.sematech.org/public/dictionary/home.html>] is an excellent example of a Web resource not only defining a profession's or an industry's standard terminology, but one of providing users with specific concepts as well as words and phrases that serve as candidates for a more refined Net search.

## Directories

Internet sites that are an electronic equivalent of a print directory, or serve this function, are also among the more common resources available on the Net. While many sites are only analogs to print publications, many have expanded the directory function to include not only a conventional profile of selected entries, but also a hotlink to the homepage of the listed source. The *Electrical Engineering Programs* site [[http://www.ee.umn.edu/schools/ee\\_programs.html](http://www.ee.umn.edu/schools/ee_programs.html)] provides not only an organized listing of academic programs in this field, but also a direct link to associated homepages. A more comprehensive example of this service is the *TechExpo Directory of Hi-Tech Organizations in the Engineering and Medical/Life Sciences* [[http://www.techexpo.com/tech\\_soc.html](http://www.techexpo.com/tech_soc.html)] which offers an extensive alphabetical listing of societies and other organizations involved in various aspects of engineering, and the medical and life sciences.

## Handbooks

Recognition of the benefit of providing users direct access to an electronic form of an established Reference work has led many organizations to convert their standard handbook to a Web document. The *Airline Handbook* [<http://www.air-transport.org/handbk/INTRO.HTM>] is a good example of a standard Reference book that has been published as an Internet document. Societies, trade groups and other organizations have quickly come to realize that the Net can provide wide access to their publications and permits easy maintenance and correction. For Reference resources where current information may be particularly important to users, such as in the case of the *Pesticide Poisoning Handbook* [<http://hammock.ifas.ufl.edu/txt/fairs/pp/19729.html>], the Web offers an opportunity to update a source quickly or link to others that are considered appropriate.

## VII. WWW Participatory Database Development

We believe that user participation in the development of a service such as CyberRef(sm) is not only preferred but essential. Through a variety of links within CyberStacks(sm) users have the option of nominating resources for inclusion within any one of its Reference collections, or can request that a specific type of publication be made available as a Web document that is accessible from within it. [13] Thus, as within other proactive collection development policies and practices, users themselves are involved with the enhancement of a collection they use and consult. For individuals interested in collaborating more formally, four 'virtual' advisory boards have been established, [14] including one devoted to Selection and Collection Development.

## VIII. Conclusion

In an effort to augment a local print collection of Reference sources, a demonstration prototype Internet database Web resources was established on the homepage server at Iowa State University. In the process of selecting and organizing resources for inclusion in this local virtual collection, it quickly became apparent that the Net offered the opportunity to create a 'universally-available' collection that could also serve the interests and needs of other libraries and organizations.

Through an extensive search of relevant Net sites, more than five hundred (500) resources with 'reference value' were identified and selected for inclusion in a Title Index. These included a number of resources in a variety of fields of technology and related disciplines which were either electronic equivalents of print resources or enhancements that provide sophisticated display and search options heretofore not available for all Reference works. Resources in a variety of traditional Reference formats were identified during this review and included abstracts and indexes, acronyms, bibliographies, catalogs, classification outlines, data sheets, dictionaries, directories and handbooks, as well as a variety of conventional and unconventional database forms.

The increasing availability of Net substitutes for print publications, as well a clear trend toward the conversion of many others to versatile database formats, provide a sufficient resource base to pursue the further development of an organized Technology Reference collection that could concurrently expedite and facilitate access to these resources and meet local as well as non-local Reference needs. CyberStacks(sm) provides users with a framework for building such a collection and specialists within various fields of technological librarianship are particularly encouraged to contribute to the development of one of its many Reference collections.

## References

1. MCKIERNAN, Gerry. (1995). CyberStacks(sm): a 'library-organized' virtual science and technology Reference collection. D-Lib Magazine [Online]. Available HTTP: [http:// www.dlib.org/ dlib/december95/ briefings/12cyber.html](http://www.dlib.org/dlib/december95/briefings/12cyber.html) [1997, June 14]
2. CROFT, W. Bruce. (1995). What do people want from Information Retrieval. D-Lib Magazine [Online]. Available HTTP: [http://www.dlib.org/ dlib/november95/ briefings/11croft.html](http://www.dlib.org/dlib/november95/briefings/11croft.html) [1997, June 14]
3. MCKIERNAN, Gerry. 1996a. Build it and they will come: a case study of the creation, development and refinement of an organized database of Internet resources in science and technology. Paper prepared for presentation at The Digital Revolution: Assessing the Impact on Business, Education and Social Structures, the 1996 Mid-Year Meeting of the American Society for Information Science, May 18-22, 1996, San Diego, California. Medford, NJ., Information Today, 1996, pp. 284-297.
4. MCKIERNAN, Gerry. 1996b. Casting the Net: the development of a resource collection for an Internet database. In: Duda, Andrea L., ed. Untangling the Web: proceedings of the conference sponsored by the Librarians Association of the University of California, Santa Barbara and Friends of the UCSB Library, University Center, University of California, Santa Barbara, April 26, 1996. [Online]. Available at HTTP: <http://www.library.ucsb.edu/untangle/mckiernan.html> [1997, June 14]
5. LIBRARY OF CONGRESS. OFFICE OF SUBJECT CATALOGING POLICY. LC Classification Outline. Washington, D.C., Library of Congress, 1990. 47p.

6. BOPP, Richard E. and SMITH, Linda C., eds., Reference and information services: an introduction. Englewood, Colorado, Libraries Unlimited, 1995. xxiv + 626p.
7. HOPKINS, Richard L. Countering information overload. Reference librarian, 49/50 1995: pp.305-333.
8. ENSOR, Pat. Organizing the Web: a contradiction in terms? Technicalities, 15(9) 1995: pp. 6.
9. ENSOR, Pat. Organizing the Web: a contradiction in terms? Technicalities, 15(9) 1995: pp. 7.
10. BRITTEN, William A. Building and organizing Internet collections. Library acquisitions: practice and theory, 19(2) 1995: pp. 247.
11. BRITTEN, William A. Building and organizing Internet collections. Library acquisitions: practice and theory, 19(2) 1995: pp. 249.
12. MCKIERNAN, Gerry. 1996b. Casting the Net: the development of a resource collection for an Internet database. In: Duda, Andrea L., ed. Untangling the Web: proceedings of the conference sponsored by the Librarians Association of the University of California, Santa Barbara and Friends of the UCSB Library, University Center, University of California, Santa Barbara, April 26, 1996. [Online]. Available at HTTP: <http://www.library.ucsb.edu/untangle/mckiernan.html> [1997, June 15]
13. MCKIERNAN, Gerry. 1996c. The New/Old World Wide Web Order: the application of 'neo-conventional' functionality to facilitate access and use of a WWW database of science and technology Internet resources. Journal of Internet cataloging: The international quarterly of digital organization, classification & access, 1(1) 1997: pp. 47-55.
14. MCKIERNAN, Gerry. (1995). CyberStacks(sm): a 'library-organized' virtual science and technology Reference collection. D-Lib Magazine [Online]. Available HTTP: <http://www.dlib.org/dlib/december95/briefings/12cyber.html> [1997, June 15]