AC 2009-1061: THE NATIONAL SCIENCE DIGITAL LIBRARY AS A PLATFORM FOR AN ENGINEERING EDUCATION DATABASE.

Nestor Osorio, Northern Illinois University

Nestor L. Osorio is professor and subject specialist for science and engineering at Northern Illinois University, DeKalb, IL, e-mail: nosorio@niu.edu.

Andrew Otieno, Northern Illinois University

Andrew W. Otieno is associate professor at the Department of Technology, Northern Illinois University, DeKalb, IL, e-mail: otieno@ceet.niu.edu.

© American Society for Engineering Education, 2009

The National Science Digital Library as a Platform for an Engineering Education Database

Abstract

Studies about the literature of engineering education have clearly demonstrated that a substantial gap exists between the yearly production of articles in this field and the current and historical coverage included by major standard engineering databases. It is estimated[1] that half of the literature produced in English, including those from the USA and Canada, do not have a permanent archival record. This paper proposes the creation of a bibliographic database for engineering and technology education.

The National Science Digital Library (NSDL) would be an appropriate platform for such an initiative. NSDL has built, or provides access to, extensive collections of documents and objects in all areas of science and engineering. Well known NSDL projects, such as Annals of Research on Engineering Educations (AREE) or National Engineering Education Delivery System (NEEDS), whether fully or partially funded, are integrated within the NSDL Engineering Pathway (EP). An engineering education database would be a valuable supplement to the robust collections and services provided by EP. The proposed database would be built utilizing the already existing and highly-recognized technical infrastructure of the NSDL. This will be a collaborative project where professional organizations such as the Engineering Library Division (ELD) of the American Society for Engineering Education (ASEE), the Science and Technology Section (STS) of the Association of College and Research Libraries (ACRL), and the Engineering Division (ED) of the Special Libraries Association (SLA) can play an important role.

Introduction

The literature of engineering education is complex and extensive. In a recent publication, Powell[2] describes the components of engineering education that include, for example, courses and programs; assessment and evaluation; learning resources and practices; advising; research opportunities; retention; teaching methods; research methods; hiring; promotion; tenure; strategies and tools used in the classroom; assessment and evaluation to improve specific courses; models for engineering programs; and many others. Other topics not mentioned by Powell are: the teaching of fundamental engineering concepts at the K-12 level, international education collaboration, and lifelong learning. Jesiek, et al.[3] however, point out that there have been improvements and "an impressive expansion of engineering education has been underway since at least the early 2000s. This domain now boasts an infrastructure comprised of funding and granting agencies, publication outlets, conference venues, and academic units."

The engineering education literature is produced in a variety of formats. De Petro[4] describes a selection of the formats and indicates that articles in journals and articles in conference proceedings constitute the largest percentage. Other formats include books, book chapters, theses, reports, manuscripts and others. De Petro also states that engineering education is often

underutilized; libraries do not have consistent collection development programs to acquire and maintain subscriptions of journals and conference proceedings, and that the bibliographic records of the many contributions to the field are not found in the standard databases related to engineering or education. It also is estimated that half of the literature produced in English, including those from the USA and Canada, do not have a permanent archival record[1]. These problems, which De Petro labeled a 'bibliographic deficiency,' have often been a topic of discussion in some professional groups of science/engineering librarians. The De Petro article is an accurate confirmation of the finding demonstrated by Osorio[5,6] at two conferences of the American Society for Engineering Education (ASEE).

Studies in analyzing the literature of engineering education[7, 8,9] demonstrate that the field is robust, with an increased emphasis on the utilization of research methodologies, and is international in coverage. It is an active scholarly field with participation from several professional societies that produce a good number of dedicated journals and conference proceedings. The literature of engineering education also receives contributions from publications from almost all fields of engineering. In summary, its annual output is significant.

Responding to the 'bibliographic deficiency,' we propose the development of a repository. This repository would be a specialized and comprehensive database for all published documents related to engineering and technology education. The term 'published documents' includes articles in journals, conference papers, book chapters, and other publications that have gone through a peer review process and have been published by recognized professional publications.

The proposed project for improving access to the literature of engineering and technology education has four specific goals: 1. To identify the body of literature from 1980 to the present that has been subject to a peer review process. This will include articles in journals, articles in conference proceedings, books, book chapters, reports, dissertations, and other documents. 2. To construct a thesaurus for technology and engineering education, which will be used in the metadata of the database. 3. To construct a database repository in an open-source environment such as EPrint, Fedora, or DSpace that will comply with the existing sophisticated technical requirements of cyberspace, including the future allocation for high density data collection, images, graphs and other objects. 4. To communicate the availability of this dedicated engineering education database to the target audience.

NSDL and the Engineering Pathway

The National Science Digital Library (NSDL) is an NSF program. Until October of 2008, its central administrative body was the Core Integration Team. In addition, it had the Policy Committee, with elected representatives of the projects, and it also had five standing committees: Community Services, Content, Educational Impact, Sustainability, and Technology. A great deal was accomplished under this organizational structure from its beginning in 2000. Since then, the NSDL has funded more than 200 projects for several main tracks: core integration system, collections, services, targeted research, and pathways. In addition, other funding agencies that have provided significant co-funding of these projects, particularly during the beginning years, are the Directorate for Geosciences (GEO) and the Directorate for Mathematical and Physical Sciences (MPS)[10].

NSDL resources are intended to serve a diverse clientele; these resources are produced with the participation of multiple organizations and they are presented using different formats. The engineering collection is presently within the division of Engineering, Computing and Technology; a browsing function can be used to explore its components. The Engineering Pathway (Engineering Education Resources) has been available for the last two years to better facilitate access to materials related to engineering. This is one of the thirteen pathways now in operation. Their purpose is to organize, present and make access available to resources in a specific area. There are pathways for biology, chemistry, mathematics, and physics, for example.

The Engineering Pathway <http://nsdl.org/about/?pager=pathways&subpager=EP > contain resources for higher education educators and students as well as K-12 level materials. Some of their features are: Browse by subject/disciplinary content area; Advanced Search over keyword, title, author, discipline, grade, etc; Personalized workspace; Diversity Resources; Highlighted Resources; and EP Weblog. Obviously, one additional feature could be the Technology and Engineering Education Database.

Open Access and Repositories

Open Access has become a tool of the trade in today's generation of information technology. Swan[11] has suggested that with the information technologies available in the 21st century, if we would invent 'scholarly communication', it would not be in its 'traditional' form; it would be something greater, influenced by the capability of researchers preparing their work and communicating electronically. Suber[12] explains that "Open-access (OA) literature is digital, online, free of charge, and free of most copyright and licensing restrictions. What makes it possible is the Internet and the consent of the author or copyright-holder". Suber also indicates that "OA is entirely compatible with peer review." As for its cost, Suber clearly states that is not free of cost; the key is creating new business models to make professional literature less costly. The fundamental principle of OA is to provide free access and produce publications at a lower cost. In his book, Willinsky[13] offers an extensive discussion of the problems encountered in the system of scholarly communication and their potential solutions. Open access has tremendous opportunities in the field of engineering education, which is a relatively new and expanding area.

It is well known that an increasing number of academic and research institutions are building repositories of electronic documents and objects. Institutional Repositories (IRs) have been created to display the creativity and productivity of faculty and students. IRs are found in many different frameworks; some are institutionally-oriented, while others are subject-oriented. The number of IRs is continuously growing. OpenDOAR – the Directory of Open Access Repositories < http://www.opendoar.org/ > includes 1,200 entries. Subject-focused IRs are also abundant. For instance, the well-known arXiv at Cornell University < http://arxiv.org/ > which covers documents on physics, mathematics, non-linear science, computer science, quantitative biology, and statistics, contains a half-million openly-accessible articles. The project proposed in this article would have an open access bibliographic database with a partial full-text component made of archived documents. Therefore, a repository of publications on engineering education contained as part of the Engineering Pathway of the open access National Science Digital Library would be a valuable resource. This database will complement the rich collections and services already present in the Engineering Pathway.

Proposed Database

The Technology and Engineering Education Database (TEED) will be a repository database developed using open-source software such as the Flexible Extensible Digital Object and Repository Architecture (Fedora), EPrint, or DSpace, with approximately 100,000 records – partial and full-text - covering the period from 1980 to the present. The initial efforts would be concentrated in creating a retrospective bibliographic database with abstracts covering the period from 1980 to1995. This bibliographic database would then be extended to the present. Full-text uploads to the database would be done as the documents become available. The logistics involved in uploading documents not only involve finding them, but also the process of getting the proper copyright permission for making them available. The final goal is that when the construction of this database is completed, users will be authorized to upload their proposed documents using a proper registration and log-in system.

The major tasks for the development of this project are: 1. Identification of documents, 2. Procurement of documents, 3. Thesaurus development, 4. Database architecture development, and 5. Development of the public interface development.

1. Identification of documents

The identification of publications related to engineering education represents an essential step. The publications to be included in the database are: A. Conference proceedings, B. ASEE sectional conferences, C. Journals, and D. Books and other publications. De Petro[4] and Osorio[5,6] have already identified core lists of conference proceedings, a core list of dedicated engineering education journals, and other journals (most of then related to engineering fields) that publish on educational topics. As it has been indicated, standard engineering and educational databases such as Compendex, INSPEC, and ERIC have shown inconsistency in indexing articles on engineering education.

Several standard databases and a variety of bibliographies would be used to identify books, book chapters, reports, dissertations, and other publications. These types of materials are also not consistently indexed by engineering or educational databases. It is estimated that the total number of documents to be identified - from 1980 to the present - is near 100,000[1].

2. Procurement of documents

The procurement of documents will be a major effort for the completion of this project. Since most of the documents pre-1995 are in paper format, their procurement will require the expertise of document delivery services and the active participation of the director and co-directors, to contact several engineering, educational, and library-related professional organizations. The ability to find copies of hard-to-find publications would be an important task in this project. While the bibliographic data would be obtained from the original documents - a straight-forward operation - securing permission for archiving the documents can be a more time-consuming task. The members of the External Advisory Group would also be extremely valuable in this phase. The idea is to create a network of volunteers from several professional societies that would help in the finding of documents. This community involvement would also help in the dissemination process of the project.

3. Thesaurus development

A Thesaurus of terms for engineering education is not available. The development of such a thesaurus would require the compilation and examination of vocabulary terms being used in recognized publications such as conference proceedings and journals of the ASEE and the European Society of Engineering Education (SEFI), and other major publications. In this project, the expertise of engineering librarians would be extremely valuable in the construction of a thesaurus for engineering and technology education.

4. Database architecture development

This process involves the installation of open source software platform for the repository and underlying architecture by project technical staff. In this proposed project, Fedora is expected to be the leading choice with Java SE Development Kit, MySQL, Tomcat to handle other functions in the database. Linux servers are proposed as the hardware of choice because of their cost effectiveness, stability and security. Staff will create the Fedora database that will store registries and searchable metadata. They will design and create specifications for the digital object model, including choosing appropriate persistent identifiers, object properties (including Dublin Core metadata and administrative metadata), content datastreams, and disseminators (pointers to external programs that provide views of the object). This digital object model is expressed in FOXML, a flexible way to map Fedora concepts to an XML format. The system will be built with an eye on possible future expansion, including functions for viewing multiple formats of items (HTML, PDF, etc.) and potential hierarchical relationships, such as associating a particular journal article with a journal title. Technical staff will also devise an appropriate data entry mechanism (in either Access or MySQL) to capture the full bibliographic records for this project. These are the items that will serve as the "datastreams" for the Fedora module. Staff will also examine the possibilities of utilizing OpenURL link resolvers to provide access to items that may have a full-text representation in a subscription database, or to link to MARC records for items in library catalogs.

5. Development of the public interface

The public interface will have the capability of searching several fields using Boolean operators (“AND”, “OR”, and “NOT”). Some of the searchable fields will be: subject term, title, abstract, author, conference, journal title, and book title. An "all fields" option would also be available. Limits to the search would be by: a range of years, type of document (article in journal, conference papers, books, book chapter, dissertation, report, etc.), and audience. A free keyword search capability would also be an option. Access by browsing the list of journals, conference proceedings and books will be also available.

The project would be divided into three phases:

Phase 1. Build the database architecture, create the public interface, develop a thesaurus, initial testing, beginning identification and procurement of documents.

Phase 2. Continuing the identification and procurement of documents, load metadata of document without full-text, load full-text documents and their metadata, testing and evaluation.

Phase 3. Open to the public, continuing loading metadata and full-text materials. At this time users will also be allowed to upload their documents using a proper registration and log-in system.

Administrative Issues

A project of this magnitude would need a flexible organizational structure that would allow for the best utilization of technical and human resources to accomplish its goals. We envision having an executive group that, following the structure of a typical grant funding project, will have a director, one or two co-directors, a technical coordinator and a software expert. This group will serve as the Internal Management Team. This group will be involved in monitoring the progress of the project, its evaluation, and making recommendations for possible technical problems.

The principal directors and the two co-principal directors will oversee all operations of the project and will be the contact persons with other organizations. Among their responsibilities are: contact publishers and professional organizations in order to secure the usage of bibliographic data; obtain the procurement of documents; identify publications such as conference proceedings and journals as target sources for the database; work with other staff in the creation of a thesaurus; participate in the dissemination and evaluation processes of the project; obtain required agreements to include the documents in the database; participate in testing the database; identify other potential users for an initial pilot study of the database; oversee the technical aspects of the project; be part of the Internal Consultation Team; and participate in the evaluation of the project. In addition, other staff will include a Project Technical Coordinator, a Database Technical Expert, a Metadata/Indexing Consultant, a Document Procurement Consultant; a Programmer, and Indexer/Data Entry staff.

The project will also have an External Advisory Group. This committee will be formed by members from the Engineering Library Division (ELD) of the ASEE, the Science and Technology Section (STS) of the ACRL, and the Engineering Division (ED) of the SLA. Also, it will have three teaching faculty members; one member from the Information Systems Division (ISD) of ASEE, and the other two from international organizations such as SEFI. This committee will advise on matters concerning the design and interface of the database, the procurement of documents (for example, older conference proceedings), the dissemination of the database to the targeted audience, and its evaluation.

Funding issues are critical in the development and maintenance of such a repository. Some repositories are built with the efforts of a dedicated group of volunteers while receiving server space from friendly organizations; such is the case with E-Lis[14]. Large repository projects are usually funded by not-for-profit foundations. It is also possible to initially receive funding from a government agency in the form of a grant and continuing the project with funds received from the private sector. An NSF NSDL initial grant for two or three years would be sufficient to start the project and develop it to the point that it could be continued with other resources.

Assessment is important to ensure the success of the project. An assessment mechanism would have to be introduced in each of project components. This assessment would be done by the Internal Advisory Group, the External Advisory Group, user surveys, and focus groups.

Conclusion

In their recent article, Jesiek et al.[3] have observed that "an impressive expansion of engineering education has been underway since at least the early 2000s. This domain now boasts an infrastructure comprised of funding and granting agencies, publication outlets, conference venues, and academic units." This expansion is occurring at a time when information technology can be utilized not only for the dissemination of new contributions to the field but to the preservation of accumulated experiences and knowledge as well. The proposed bibliographic database and repository for technology and engineering education would make the literature of the field permanently available to the community of researchers, faculty, students, leaders of society, and the general public. The platform of the NSDL and its Engineering Pathway would be the appropriate place for this proposed database.

Bibliography

1. Osorio, N. L., Otieno, A. W. (2008). The Technology and Engineering Education Database. Unpublished, 12 p.
2. Powell, J. H. (2006). Engineering Education, in: *Using the Engineering Literature*, Bonnie A. Osif, editor, New York: Routledge, p. 269-285.
3. Jesiek, B. K., Newswander, L. K., and Borrego, M. (2009). Engineering Education Research: Discipline, Community, or Field?, *Journal of Engineering Education*, Vol 98 no 1 p. 39-52.
4. De Petro, T. G. (2000). Engineering Education Periodicals and Proceedings: Increasing Awareness of and Access to the Literature, Proceedings of the 2002 ASEE Gulf-Southwest Annual Conference, Session IC3. [Available online] http://engr.louisiana.edu/asee/proceedings/IC3.pdf
5. Osorio, N. L. (1998). Engineering Education as Knowledge Discipline, American Society for Engineering Education, Annual Conference, Engineering Libraries Division, Poster Paper Session.
6. Osorio, N. L. (2005). What Every Engineer Should Know about Engineering Education, Proceedings of the 2005 Illinois/Indiana Sectional Conference, American Society for Engineering Education, D1-1. [Available online] http://eprints.rclis.org/archive/00008579/01/ASEE-IL-IN-2005-OSORIO.pdf
7. Whitin, K. and Sheppard, S. (2004). Taking Stock: An Analysis of the Publishing Record as Represented by the Journal of Engineering Education, *Journal of Engineering Education*, Vol 93 no 1 p. 5-12.
8. Wankat, P. C. (2004). Analysis of the First Ten Years of the Journal of Engineering Education, *Journal of Engineering Education*, Vol 93 no 1 p. 13-20.
9. Osorio, N. L.; Osorio, M. A. (2002). Engineering Education in Europe and the USA: An Analysis of Two Journals, *Science & Technology Libraries*, Vol. 23 no 1 p. 49-70.
10. Osorio, N. L. (2006). The National Science Digital Library (NSDL): An Examination of its Role as a Digital Engineering Education Library. Proceedings of the American Society for Engineering Education, Illinois-Indiana and North Central 2006 Conference. Technical Session B, Track 3, 10 p., in CD-ROM. Fort Wayne, IN: Indiana University – Purdue University Fort Wayne, 2006, [Available online] http://www.asee4ilin.org/Conference2006program/Papers/Osorio-P77.pdf
11. Swan, A. (2007). Open Access and the Progress of Science: The power to transform research communication may be at each scientist's fingertips, *American Scientist*, Vol. 95 no 3 p. 198-200. [Available online] http://amsciadmin.eresources.com/libraries/documents/2007327142636_307.pdf http://www.americanscientist.org/issues/pub/open-access-and-the-progress-of-science
12. Suber, P. (2009). Brief Introduction to Open Access, [Available online] http://www.earlham.edu/~peters/fos/brief.htm
13. Willinsky, J. (2006). *The Access Principle: The Case for Open Access to Research and Scholarship*. Cambridge, MA: MIT Press.
14. Morrison, H., Subirats C. I., Medeiros, N. and De Robbio, A. (2007). E-LIS: the Open Archive for Library and Information Science, *The Charleston Advisor*, vol. 9, n. 1 p. 56-61(6).