

Building a “Library in a Box” with an out of the box solution: creating a prototype engineering library using XTF

Rutgers University has made this article freely available. Please share how this access benefits you.
Your story matters. [\[https://rucore.libraries.rutgers.edu/rutgers-lib/43114/story/\]](https://rucore.libraries.rutgers.edu/rutgers-lib/43114/story/)

This work is an **ACCEPTED MANUSCRIPT (AM)**

This is the author's manuscript for a work that has been accepted for publication. Changes resulting from the publishing process, such as copyediting, final layout, and pagination, may not be reflected in this document. The publisher takes permanent responsibility for the work. Content and layout follow publisher's submission requirements.

Citation for this version and the definitive version are shown below.

Citation to Publisher Palumbo, Laura Bolton. (2012). Building a “Library in a Box” with an out of the box solution: creating a prototype engineering library using XTF. *Library Hi Tech News* 29(6), 8-10.

Citation to this Version: Palumbo, Laura Bolton. (2012). Building a “Library in a Box” with an out of the box solution: creating a prototype engineering library using XTF. *Library Hi Tech News* 29(6), 8-10. Retrieved from [doi:10.7282/T3P26W92](https://doi.org/10.7282/T3P26W92).

Terms of Use: Copyright for scholarly resources published in RUcore is retained by the copyright holder. By virtue of its appearance in this open access medium, you are free to use this resource, with proper attribution, in educational and other non-commercial settings. Other uses, such as reproduction or republication, may require the permission of the copyright holder.

Article begins on next page

Building a "Library in a Box" with an out of the box solution: creating a prototype engineering library using XTF

Building a "Library in a Box" with an out of the box solution: creating a prototype engineering library using XTF

Introduction

Rutgers University Libraries, in collaboration with the University of Liberia, and with grants from the Engineering Information Foundation (EIF), the United States Agency for International Development (USAID), and with contributions from the International Society of Electrical Engineers (IEEE), have created The EAKO System- Engineering Access to Knowledge Offline, a prototype "Library in a Box" for engineering information. The product runs on the open source platform XTF, created by California Digital Library, and is based on TEEAL, The Essential Electronic Agricultural Library, developed by Cornell University. The EAKO System, like TEEAL, will work entirely offline.

Internet access in Liberia, as it is in many developing countries, is extremely limited, and if available, excessively slow. Liberia's fourteen year civil war, which began in 1989, decimated the country's infrastructure and basic services. The University of Liberia is the largest university in the country with 23,800 current students (Dennis, 2012), yet their library lacks quality information resources and up-to-date technology. EAKO will provide a digital collection of open source engineering articles, books, and courses, which will be stored on a high capacity hard drive, accessed offline via a browser installed on a computer in the University library. As the computer resources in the library grow, the program will be migrated to a local area network. The program can also be used online when broadband Internet becomes available. Future development of the project would allow access via mobile devices.

Background

During the initial development of the product, the existing physical and informational resources of the library at the University of Liberia were evaluated, along with the current engineering curriculum. It was determined that the library's resources were extremely limited to non-existent. The engineering curriculum consists of programs in civil engineering, electrical engineering, and mining engineering, with plans to include a major in mechanical engineering as well. In order to provide as much useful information as possible, it was determined that the EAKO System prototype would provide resources in these areas only, with the addition of books and instructional materials in foundation courses in science and mathematics. Materials to be included in EAKO would be open source scientific journal articles and books, as well as donated archived articles in electrical engineering from IEEE.

Technical constraints

The EAKO System set out to emulate the success of TEEAL (<http://teeal.org/>), and hoped to utilize the same platform, Greenstone Digital Library software. However, early discussions with the librarians at Cornell University about the technical modifications which were necessary to produce TEEAL, as well as the extremely limited budget for the EAKO grant, persuaded team members that a simpler open source solution should be found. This alternative open source program would need to be able to search and manage a collection of articles and larger documents in a variety of formats, but primarily PDF's, on a high capacity hard drive. The operating system would be Windows-based, due to the almost exclusive use of PC's in Liberia. More than twenty programs were investigated, and as a result we initially hoped to use Fedora Commons. It was soon determined that like Greenstone, Fedora would require some fairly intensive technical modifications, which were beyond our budget to procure. I was very excited when I then discovered XTF, the eXtensible Text Framework by California Digital Libraries (<http://xtf.cdlib.org>). In addition to the six California Digital Library implementations using XTF, and the 18 highly customized implementations, there are currently 28 lightly customized uses in academic universities and archives in the US, UK, and Sweden (California Digital Library, 2012). XTF not only met our project requirements, but also provided easy to follow instructions for simple technical modifications. Documentation displayed on the website is extensive and well-written, and the program is actively maintained and supported by both the developers and a user's group.

Program Features and Customization

XTF's standard searching and indexing capabilities are quite powerful, yet it is also adaptable in that it is easily modified by users with limited programming experience. Some of the easily performed customizations of XTF which are important to its use in The EAKO System are the ability to add custom metadata and create our own hierarchical facets. Another key feature is its ability to handle large documents in a variety of formats, including PDF's, Word, HTML, text, and XML. Boolean search capabilities, as well as wildcard/truncations/exact phrase search allowances are "out of the box" features. Browsing and advanced search by facets are also standard. Misspelled words are recognized, and spelling suggestions are provided. Full details of these and all XTF features are available on their website, xtf.cdlib.org.

Knowing that most of the potential users at the University of Liberia will have little experience in searching using the Internet, we attempted to simplify the search interfaces as much as possible, while retaining XTF's powerful searching capabilities. We discarded proximity searching, freeform searching (similar to autocomplete), as well as alphabetic browsing for authors and titles. The default basic search is by keyword in the full text, with a single search box. (Figure 1) The advanced search filters by field, allowing input of author, title, subject, date range by years, and engineering category. This can be used in conjunction with the keyword search (see Figure 2 below).

Building a "Library in a Box" with an out of the box solution: creating a prototype engineering library using XTF

Figure 1. EAKO Basic Search Interface

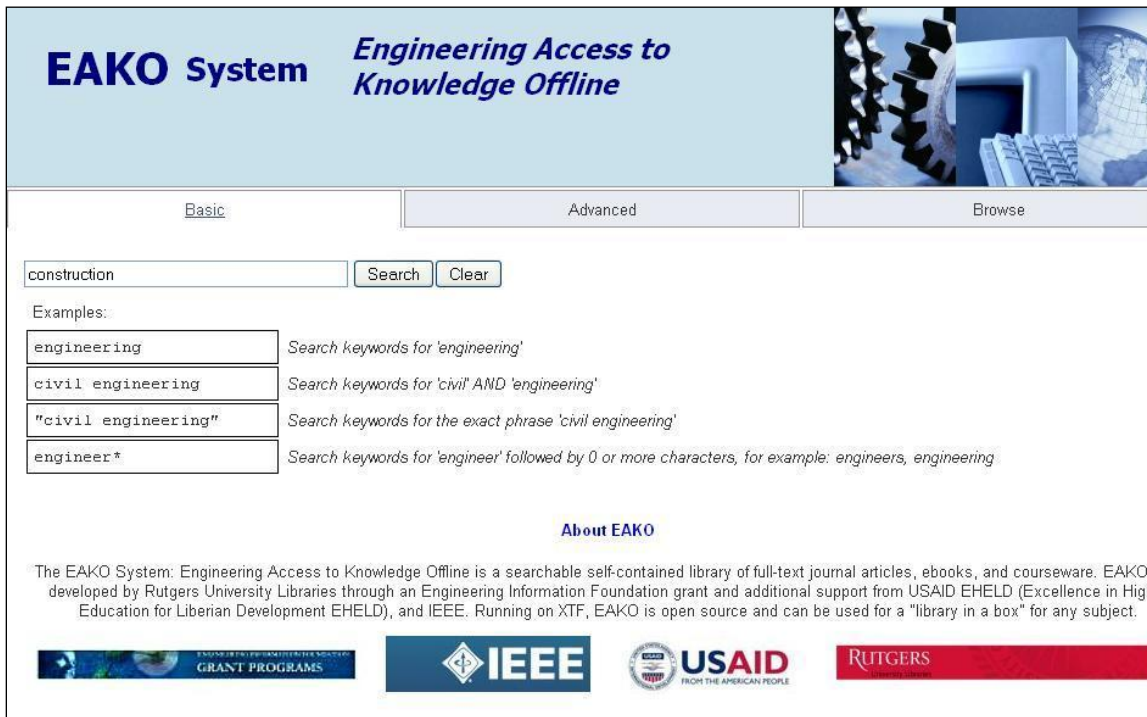
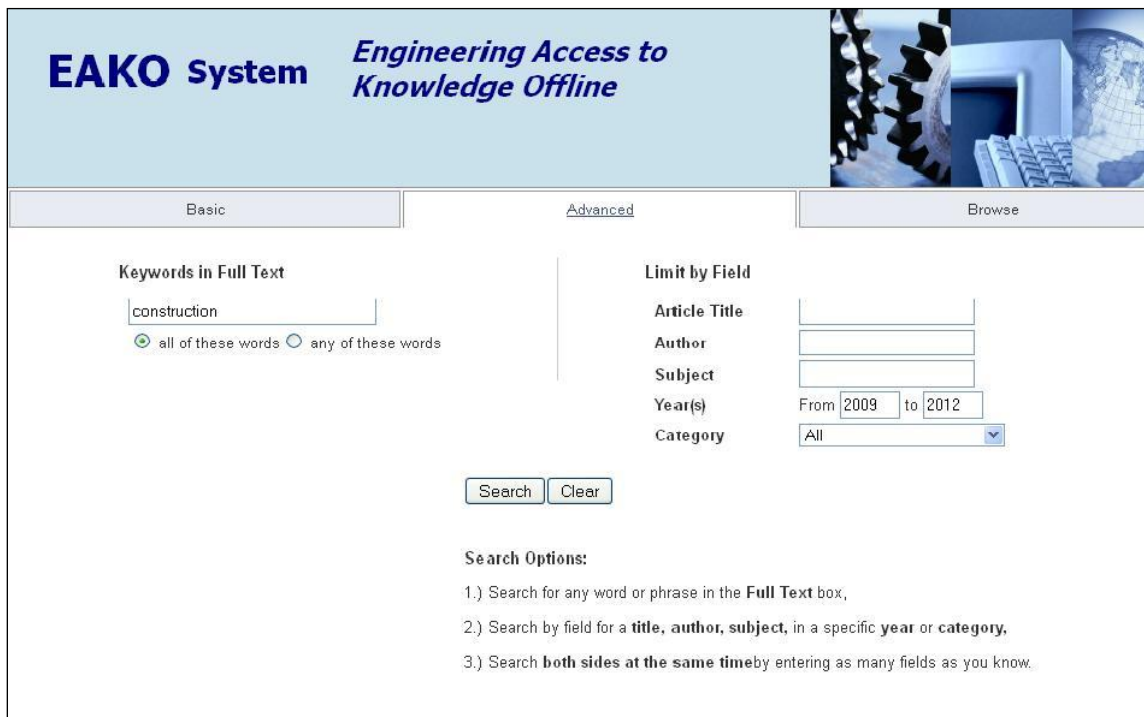


Figure 2. EAKO Advanced Search Interface



Building a "Library in a Box" with an out of the box solution: creating a prototype engineering library using XTF

Creating new and hierarchical facets was easily accomplished thanks to detailed instructions provided on the XTF website. New facets were created for books, courses, and categories, as well as hierarchical facets for journal names, volumes, and issues. Browsing in EAKO is performed by alphabetic listing by journal title, book titles by category, and course titles by category.

Input requirements

For each article, book or other object to be indexed into the EAKO System collection, a file must be created in XML, using Dublin Core metadata. A template form for creating the input for the surrogate records is provided by XTF. Any XML editor can be used to create these files, but we chose to use Oxygen, an inexpensive program recommended by XTF developers (www.oxygenxml.com). Styling is done automatically by XTF. The document and XML file are linked in a separate folder, and then indexed by the program. One drawback to this program is the amount of time necessary to create individual data files for each document to be indexed. We found that each document required approximately 10 to 15 minutes to complete the process of downloading an article, writing the XML, and saving and naming the files. This time intensive process would require funding for approximately 1250 hours of data entry for every 5000 documents, or sufficient number of staff to volunteer this much time. We are currently looking into minor technical modifications which will streamline this process.

Search Results

Results are displayed by relevance, with keywords highlighted. Results can also be sorted by title, author, journal, publication date with oldest first, or reverse date with newest first. Links are provided to the item record, as well as to other items by subject keywords. Because the item is stored in EAKO on the hard drive, there is no waiting for the document to download. There is also a nice feature which will find similar items for individual results. Another useful feature is the "bookbag", which was originally set up in XTF to e-mail search results. For our use, e-mailing has been disabled, but it is still a nice way to create a saved search. Facets are displayed in a sidebar on the left, and are clickable hyperlinks. Since facets may be customized, these can be anything that is needed. Figure 3 below is a result page from a keyword search for construction.

Building a "Library in a Box" with an out of the box solution: creating a prototype engineering library using XTF

Figure 3. EAKO Search Results

EAKO System *Engineering Access to Knowledge Offline*

Search: **construction** in the full text [X]
2009 - 2012 in year [X]

Results: 713 Items

Sorted by: relevance [v] Go!

Bookbag [C] Modify Search | New Search
Browse by Journal | Books | Course
Page: 1 2 3 4 5 ... Next

Category

- [mining](#) (1)
- [mechanical engineering](#) (143)
- [general engineering](#) (313)
- [electrical engineering](#) (1)
- [civil engineering](#) (305)

Journal-vol-Issue

- [World Journal of Science and Technology](#) (3)
- [Turkish Journal of Engineering and Environmental Sciences](#) (6)
- [Thermal Science](#) (18)
- [The Pacific Journal of Science and Technology](#) (53)
- [THERMAL SCIENCE](#) (1)
- [Polish Maritime Research](#) (35)
- [Machine Design](#) (97)
- [Journal of Urban and](#) (11)

1

Author: Christopher J. Willis [Add](#)

Title: [Measuring the Maturity of Guyana's Construction Industry Using the Construction Industry Macro Maturity Model \(CIM3\)](#) pdf

Journal: Journal of Construction in Developing Countries

Published: 2010

Subjects: [construction industry](#) | [macro maturity model CIM3](#) | [Guyana](#) | [construction performance](#)

Category: [civil engineering](#) | [general engineering](#)

Matches: ...issues and concerns4 HS13: **Construction** organisations adhere...
248 hits ...to a national **construction** H&S code of practice as a...
...process improvement for **construction** enterprises (SPICE):...

Similar Items: [Find](#)

2

Author: Agung Budiwibowo [Add](#)

Title: [Competitiveness of the Indonesian Construction Industry](#) pdf

Journal: Journal of Construction in Developing Countries

Published: 2009

Subjects: [construction industry](#) | [economic development](#) | [Indonesian construction cluster](#) |

Conclusion

XTF provided an easy to use and readily customizable open source solution for The EAKO System. Because it runs on Java and Apache Tomcat, the program can be installed on a standalone computer without access to the Internet. Its powerful out of the box searching and indexing features make accessing stored documents simple, which was an important goal for our purpose. While this is a user-friendly program, one drawback to its implementation is the amount of time required to enter data. We are currently investigating some minor technical modifications to reduce the amount of time required to create XML files for the documents being entered. The program can be adapted to use in a variety of applications, and we hope that future development will lead to access via mobile devices.

Acknowledgements:

The author wishes to thank Martin Kesselman and Connie Wu for their support and guidance.

References

California Digital Library. *XTF- eXtensible Text Framework*. Available <http://xtf.cdlib.org>

Dennis, E. (2012, May 10). President Dennis' Report University of Liberia 92nd Commencement Convocation May 9, 2012. *Daily Observer*. Retrieved July 6, 2012, from <http://liberianobserver.com/index.php/opinioneditorial/item/1285-president%E2%80%99s-dennis%E2%80%99-report-university-of-liberia-92nd-commencement-convocation-may-9-2012>

The Essential Electronic Agricultural Library. Available <http://teeal.org>