

Chapter VI

Comparing Open Source Digital Library Software

George Pyrounakis

University of Athens, Greece

Mara Nikolaidou

Harokopio University of Athens, Greece

ABSTRACT

In the last years, a great number of digital library and digital repository systems have been developed by individual organizations, mostly universities, and given to the public as open-source software. The advantage of having many choices becomes a great headache when selecting a digital library (DL) system for a specific organization. To make the decision easier, five well-known and extensively used systems that are publicly available using an open source license are compared, namely DSpace, Fedora, Greenstone, Keystone, and EPrints. Each of them have been thoroughly studied based on basic characteristics and system features emphasizing multiple and heterogeneous digital collection support. Results are summarized in a score table. Cases for which each of these systems is considered as the most suitable are proposed.

INTRODUCTION

In the last years, a great number of digital library and digital repository systems have been developed by individual organizations, mostly universities, and given to the public as open-source software. The advantage of having many choices

becomes a great headache when selecting a digital library (DL) system for a specific organization. To make the decision easier, we compared five such systems that are publicly available using an open source license, are compliant with open archives initiative protocol for metadata harvesting (OAI-PMH) (Lagoze & Sompel, 2001), and already have

a number of installations worldwide. Using these basic restrictions, we selected for comparison the following five broadly used DL systems:

- DSpace (DSpace Federation), developed by the MIT libraries and Hewlett-Packard Labs (BSD open source license)
- Fedora (Fedora Project), jointly developed by Cornell University and the University of Virginia Library (educational community license)
- Greenstone (Greenstone digital library software), produced by the University of Waikato (GNU general public license)
- Keystone (Keystone DLS), developed by Index Data (GNU general public license)
- EPrints (EPrints for digital repositories), developed by the University of Southampton

Each of these systems has been thoroughly studied based on basic characteristics and system features described in the following sections. The latest versions of those systems were examined. When writing this chapter, the versions provided were: DSpace 1.4, Fedora 2.2, Greenstone 3, Keystone 1.5, and EPrints 3. The DL systems are compared based on stated characteristics and the level of support on each of them. In the following section, the characteristics needed by a modern DL system are discussed. In the third section the five DL systems are compared based on each of the DL characteristics and the results are summarized in a score table. Finally, in the fourth section, the results of this comparison are commented on and cases for which each of these systems is suitable are proposed.

DL SYSTEMS CHARACTERISTICS

The basic characteristics and features that are expected from modern integrated DL software are:

1. **Object model:** The internal structure of the digital object (Kahn & Wilensky, 2005) (entity that integrates metadata and digital content) in the DL system. Existence of unique identifiers for the digital object and every part of it is also important to ensure preservation and easy access.
2. **Collections and relations support:** Collection description metadata, definition of subcollections, and templates that describe the format of the digital objects or the presentation of the collection. It is the definition of relations between objects of the same or different types.
3. **Metadata and digital content storage:** The storage capabilities are stated, along with the preservation issues. It is important for the DL system to ensure standards as well as user-defined metadata sets and multiple formats of digital content.
4. **Search and browse:** The mechanisms used for indexing and searching the metadata. It is important for the DL system to support indexing not only for a restricted metadata set, but also for selected metadata fields.
5. **Object management:** Methods and user interfaces provided from the DL system to manipulate (i.e., insert, update, and delete) metadata and digital content.
6. **User interfaces:** Provided user interfaces for end-user access on the DL, its collections, and the digital objects.
7. **Access control:** Support for users and groups and authentication and authorization methods. It provides a level of restriction for access and update (e.g., DL, collection, digital object, and content).
8. **Multiple languages support:** Multiple languages should be supported in the user interface, in the metadata fields, and in the digital content. The character encoding is of great importance in order for the DL systems to be fully multilingual.

8 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/comparing-open-source-digital-library/19869

Related Content

A Presentation-Preserved Compositional Approach for Integrating Heterogeneous Systems: Using E-Learning as an Example

Fang-Chuan Ou Yang (2011). *International Journal of Digital Library Systems* (pp. 67-85).

www.irma-international.org/article/presentation-preserved-compositional-approach-integrating/59889

The Issue of 'Orphan' Works in Digital Libraries, Especially as Treated by the Directive 2012/28

Maria-Daphne Papadopoulou (2012). *International Journal of Digital Library Systems* (pp. 26-53).

www.irma-international.org/article/the-issue-of-orphan-works-in-digital-libraries-especially-as-treated-by-the-directive-201228/99593

QiVMDL - Towards a Socially Constructed Virtual Museum and Digital Library for the Preservation of Cultural Heritage: A Case of the Chinese "Qipao"

Yin-Leng Theng, Yanling Luo and Gladys Theng Sau-Mei (2010). *International Journal of Digital Library Systems* (pp. 43-60).

www.irma-international.org/article/qivmdl-towards-socially-constructed-virtual/48202

A Unified Algorithm for Identification of Various Tabular Structures from Document Images

Sekhar Mandal, Amit K. Das, Partha Bhowmick and Bhabatosh Chanda (2011). *International Journal of Digital Library Systems* (pp. 27-54).

www.irma-international.org/article/unified-algorithm-identification-various-tabular/54186

Core Topics in Digital Library Education

Seungwon Yang, Barbara M. Wildemuth, Jeffrey P. Pomerantz and Sanghee Oh (2009). *Handbook of Research on Digital Libraries: Design, Development, and Impact* (pp. 493-505).

www.irma-international.org/chapter/core-topics-digital-library-education/19914