



HAL
open science

Digital libraries: Comparison of 10 software

Mathieu Andro, Emmanuelle Asselin, Marc Maisonneuve

► **To cite this version:**

Mathieu Andro, Emmanuelle Asselin, Marc Maisonneuve. Digital libraries: Comparison of 10 software. Library Collections, Acquisitions, and Technical Services, 2012, 36 (3-4), pp.79-83. 10.1016/j.lcats.2012.05.002 . hal-00746713

HAL Id: hal-00746713

<https://hal.science/hal-00746713v1>

Submitted on 29 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INTRODUCTION

In France, most digitized books in libraries, cannot be found online, except from the French National Library ones.

The main reason for this, is that Gallica the French National Library digital library, does not yet allow, other libraries to upload their electronic documents on this platform.

These electronic documents will only appear in Gallica if libraries are able to build their own digital library and their own OAI repository.

However, most of them don't have either the resources, the money nor the expertise to develop such tools, and the minority who are indeed able to build digital libraries, usually build bad ones, with old specifications and bad Web Page Ranking.

That's why most books French libraries digitized are "sleeping" on CD-ROMs, DVDs or on external hard drives.

There might be another reason to explain this situation, most French Librarians do not know which software or platforms to use to build their own digital libraries, even though a lot of them are available for free and user-friendly.

Indeed, there are very few surveys on software to develop digital libraries either in French or in English.

That's why, we have joined effort (Tosca consulting and digital project manager of an Academic Library) to conduct such a survey.

We sent a questionnaire of 160 questions to 10 software companies:

Software name	Editor (country)	License	Respondents
Mnesys	Naoned Systèmes, SARL (France)	Editor software	Alexis Moisdon, director of Naoned Systèmes
DigiTool	Ex Libris (International)	Editor software	Frédéric Lefèvre, director in France for Ex Libris
Yoolib	Amanager (France)	Editor software	Foudyl Zaouia, Director of Amanager
ContentDM	OCLC (International)	Editor software	Christian Négrel, director in France for OCLC
Invenio	Invenio-software.org (CERN) (Switzerland)	Open source	Flavio Costa, Project Manager at Invenio software.org
Greenstone	Department of Computer Science at the University of Waikato (New Zealand)	Open source	John Rose (former employee of the division of the Information Society of UNESCO).
Omeka	Fondation Roy Rosenzweig Center for History and New Media, Department of History and Art History of the George Mason University (USA)	Open source	Bernadette Vincent, in charge of electronic resources of the Bibliothèque universitaire des langues et civilisations and Julien Scot, in charge of digital library of the documentation service of Rennes 2 University.
EPrints	School of Electronics and Computer Science at the University of Southampton (UK)	Open source	Sebastien Francois, software developer for EPrints
ORI-OAI	National Consortium ORI-OAI. University of Valenciennes et du	Open source	Yohan Colmant, technical coordinator of the ORI-OAI project at the University of

	Hainaut Cambrésis (France)		Valenciennes and Nolwen Clement-Huet, responsible for document information system at the University of Poitiers
DSpace	DuraSpace, a U.S. nonprofit society (USA)	Open source	Isabelle Le Bescond, head of digital library service and theses and François Lefebvre, a computer engineer in the documentation service of the University of Lille 1

This selection of 10 solutions is fairly representative of what can be used to develop a digital library. Open source software (Invenio, Greenstone, Omeka, EPrints, ORI-OAI, DSpace) and proprietary software (Mnesys, DigiTool, Yoolib, CONTENTdm), software for old documents (Mnesys, DigiTool, Yoolib, CONTENTdm, Greenstone, Omeka) and software originally intended to open archives (Invenio, EPrints, ORI-OAI, DSpace). A lot of Countries are represented (United States, UK, France, Switzerland, New Zealand etc.).

The answers and the analysis have been published in a book written in French by: Mathieu Andro, Emmanuelle Asselin, Marc Maisonneuve, Bibliothèques numériques: logiciels et plateformes (Paris: ADBS, 2012).

Nevertheless, we decided to publish an English synthesis (not an extract) of the study in this Journal, for a wider readership. Among the 160 questions(x 10 software = 1600 responses), we have decided to publish here a selection of 43 questions (x 10 software = 430 responses) representative of approximately a quarter of all responses received and published in our book.

These selected responses were then classified into six original tables which have not been published in our book but only in this article.

1 - DOCUMENT MANAGEMENT

Software name	The software manages collections of documents	The software manages document assembly unit for reconstructing a document constructed	The software supports export to formats suitable for permanent archiving	The software supports the identification of each document by a permanent URL	The software manages the metadata structure to bring the constituent files	The software manages the metadata structure to reconstruct the digital document
Mnesys	Yes	No	No	Yes	No	Yes
DigiTool v. 3	Yes	Yes	Yes	Yes	No	No
YooLib	No	No	No	No	Yes	Yes
CONTENTdm v. 5.4	No	No	No	Yes	No	No
Invenio v. 1.0.0-rc0	Yes	No	Yes	Yes	Yes	Yes
Greenstone v. 3.05	No	In progress	Yes	Yes	Yes	Yes
Omeka v. 1.4.1	Yes	Yes	Yes	Yes	Yes	
EPrints v. 3	No	No	Yes	Yes	Yes	Yes
ORI-OAI	Yes	No	No	Yes	No	No
DSpace v. 1.7.2	Yes	Yes	No	Yes	No	No

Mnesys	Yes	Reservation requests, research (SOAP)	RDF	Yes	In progress	No	No	Yes, RDF
DigiTool v. 3	Yes	20 available web services (SOAP)	Other	Yes	Yes	No	No	Yes, export, online journal
YooLib	Yes	Several services and REST API	RDF	No	In progress	No	No	No
CONTENTdm v. 5.4	Yes	WorldCat API		Yes	Yes	No	No	No
Invenio v. 1.0.0-rc0	No	REST API		Yes	Yes	Yes	No	Yes
Greenstone v. 3.05	Yes	All services (SOAP and WSDL)		In progress	No	No	No	Yes, social networks
Omeka v. 1.4.1	Yes	A REST API is provided	RDF	Yes	Yes	Yes	Yes	
EPrints v. 3	Yes	No	RDF	Yes	Yes	No	No	Yes, webservice
ORI-OAI	No	Research, data import, control of entry procedures, access to vocabularies, harvesting (SOAP and WSDL)		Yes	No	Yes	Yes	Yes
DSpace v. 1.7.2	Yes	Harvesting (SOAP)	RDF and other	Yes	Yes	No	No	Yes, JSON

5 - USERS MANAGEMENT

Software name	The software can perform access control based on the IP address	The software offers a service user self-registration	The software allows to distinguish between user groups and assign for each specific rights	The software allows to choose those services freely accessible or only if identified	The software manages the access rights to the digital document	The software manages to use the digital document	The software distinguishes the rights granted to the following four types of users: administrator, metadata producer, producer of digital documents, simple user	The software references the users	Third-party tools for the analysis of access to web
Mnesys	Yes	Yes	No	In progress	No	No	No	Yes	Google Analytics
DigiTool v. 3	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	
YooLib	No	No	Yes	Yes	No	No	Yes	No	Google Analytics
CONTENTdm v. 5.4	Yes	Yes	Yes	Yes		No	Yes	Yes	COUNTER statistics harvested using a SUSHI server.
Invenio v. 1.0.0-rc0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Google Analytics, Piwik, AWStats

Greenstone v. 3.05	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Google Analytics or other
Omeka v. 1.4.1	Yes	No		No	Yes	Yes	Yes	No	Google Analytics, Piwik
EPrints v. 3	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Google Analytics
ORI-OAI	No	No	Yes	No	No	No	No	No	Piwik or Google Analytics
DSpace v. 1.7.2	No	Yes	Yes	Yes	Yes	No	No	Yes	Urchin, par exemple. Urchin, for example.

6 - WEB 2.0

Software name	Collaborative indexing	Users can annotate digital documents	Users can comment digital documents	Users can select documents to build their own library
Mnesys	Yes	Yes	Yes	Yes
DigiTool v. 3	Yes	Yes	Yes	Yes
YooLib	No	No	No	No
CONTENTdm v. 5.4	No	No	No	No
Invenio v. 1.0.0-rc0	No	No	No	Yes
Greenstone v. 3.05	Yes	Yes	No	Yes
Omeka v. 1.4.1	Yes	Yes	Yes	Yes
EPrints v. 3	Yes	No	No	Yes
ORI-OAI	No	In progress	In progress	Yes
DSpace v. 1.7.2	Yes	Yes	In progress	Yes

CONCLUSION

The 10 solutions we surveyed were all of good quality.

The choice of software will depend mainly on the type of documents you will want to upload (contemporary or old documents), on the political criteria (open source or proprietary software) and particularly in one of the criteria found in the 160 questions of our survey.

For smaller libraries, it is very often interesting to join a collective and shared platform like Hathi Trust, Internet Archive, or e-corpus, a French platform. The costs are shared (HathiTrust) or free (Internet Archive and e-corpus), and the tools often offer a higher degree of specifications because they are developed with more important financial and human resources.

Their inconvenient being that they will slightly lack in autonomy and freedom of action, and you may sometimes have the impression that its identity is part of a huge global market.

However, solutions such as white label products allow them to retain their identity, their graphics, their domain names and their own statistics.

It is mainly in terms of visibility on the web that this choice may be interesting, as a digital library of several thousand books will generally have a much lower PageRank than a digital library of millions of digitized books, the number of links pointing to the domain name of the biggest digital library being potentially much more important.

Unfortunately, webpage ranking has rarely been sought in digitization projects and the consultation statistics of the multitude of small digital libraries often prove to be disappointing.

Books they digitized often appear beyond the first page of results in search engines, Google type search engines, those types of search engines being the main ones used by our virtual users.

The existence of these collective platforms probably also explain why the market for software for digital libraries is still quite limited as many libraries favour pooling.

However, in France, for example, much of which has been digitized (outside of the National Library) has not yet been broadcasted online and is still "sleeping" on CD-ROMs, on DVDs or on external hard drives which lives unfortunately are not eternal...

So, there can still be a lot of opportunities for digital libraries software.

NOTES AND REFERENCES

1. Andro, M., Asselin, E., Maisonneuve, M. (2012). Bibliothèques numériques: logiciels et plateformes. Paris: ADBS.
2. Cervone, H. F. (2006). Some considerations when selecting digital library software. *OCLC Systems & Services*, 22, 107–110.
3. DeRidder, J. L. (2007). Choosing software for a digital library. *Library hi tech news*, 24, 19–21.
4. Goh, D. H. L., Chua, A., Khoo, D. A., Khoo, E. B. H., Mak, E. B. T., Ng, N. W. M. (2006). A checklist for evaluating open source digital library software. *Online Information Review* 30, 360–379.
5. Maisonneuve, M. (2012). Logiciels pour bibliothèques, vers la généralisation des services en ligne. *Livres Hebdo*, 897, 20-25.
6. Maisonneuve, M. (2010). L'offre de logiciels métier pour bibliothèques en 2010. *Documentaliste Sciences de l'information*, 47, 16 -17