



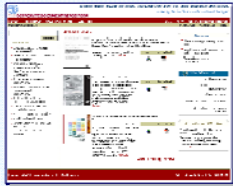
The Food and Agriculture Organization (FAO) Open Repository

Claudia.Nicolai, Imma.Subirats, Stephen.Katz@fao.org
IAALD XIIIth World Congress
26-29 April 2010

Background analysis

The FAO Open Archive (FAO OA) is based on the merging of the CDR-EIMS and the FAODOC.

The Corporate Document Repository (CDR) is a corporate output interface for FAO full text electronic publications. It contains around 40,000 records; the Electronic Information Management System (EIMS) is the workflow management tool which supports CDR.



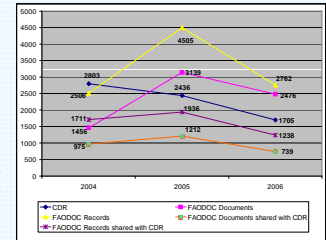
CDR/EIMS

FAODOC



The online catalogue for documents produced by FAO (FAODOC), which contains bibliographic metadata of electronic and printed documents produced by FAO since 1945; it contains around 125,000 records.

EIMS-CDR and FAODOC do not apply the same document selection criteria; EIMS-CDR and FAODOC have different cataloguing policies. In the CDR each record corresponds to one document, in FAODOC, cataloguing is done for documents and their analytics (e.g. a book and its chapters).



The overlapping at content level brings some inconsistencies that affect the proper dissemination of the FAO publications. In addition, FAO duplicates efforts in cataloguing and maintaining technically different systems.

A platform for FAO OA

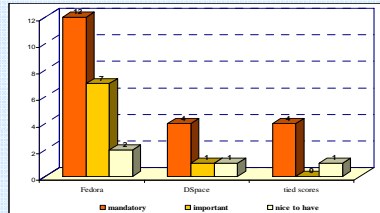
An evaluation is carried out to identify an Open Source Repository System to:

- Support FAO document workflows, for electronic publishing as well as cataloguing;
- Better disseminate FAO documentation;
- Place FAO in a community of users who can provide support and feedback; and
- Provide a reliable and efficient platform.



DSpace and Fedora Commons are evaluated against 32 criteria grouped into 10 categories:

- Community,
- Security,
- Functionality,
- Integration,
- Modularity,
- Metadata,
- Statistics and Reports,
- Preservation and Outputs



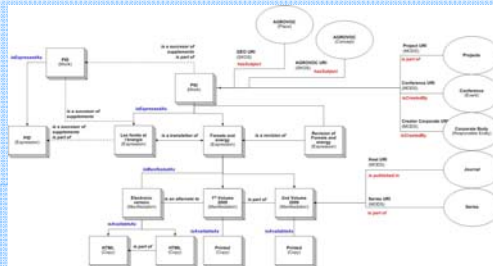
DSpace and Fedora Commons are evaluated against 32 selected criteria, 20 of which are considered mandatory. Fedora Commons ranks higher on 21 out of the 32 criteria and ranks higher on 12 out of the 20 mandatory criteria. The two systems tie on 5 of the remaining criteria. As evidenced in the Fig. above Fedora ranks higher on most of the evaluated criteria.

Fedora Commons provides:

- the necessary **flexibility** to implement all requirements without any modification to the core of Fedora,
- the possibility of implementing FAO OA functionalities **incrementally**,
- the ability to **manage** and **preserve** different type of resources (text, images, maps, audio, video) within the same repository framework,
- ease of **integration** that can allow future incorporation of other functionalities in the FAO OA.

Content model & workflow

The priorities in defining FAO OA content model were to clearly define relationships between the digital objects e.g. language versions, different editions and monographs/analytics relationships, and to adopt international standards



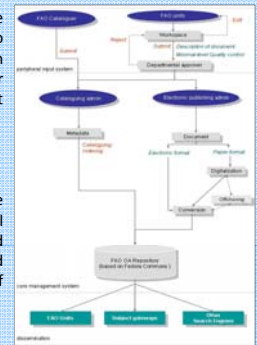
to improve interoperability and better share and disseminate FAO information. The international standard **Functional Requirements for Bibliographic Records (FRBR)** was chosen for the use in the FAO OA content object model.

FRBR is a conceptual entity-relationship model developed by the International Federation of Library Associations and Institutions (IFLA) that relates user tasks of retrieval and access in online library catalogues and bibliographic databases from a user's perspective.

The workflow of the FAO OA has to integrate the two main activities that so far have been carried out separately:

- electronic publishing
- cataloguing.

They both feed the FAO OA repository, full text and enriched metadata are managed without duplication of efforts.



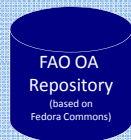
Implementing & releasing the FAO OA

Release of the new Electronic Publishing Module by the end 2010

E-publishing module

Release of the new Cataloguing module by mid-2010

Cataloguing module



Dissemination Interfaces

New FAO OA search engine based on GSearch (full text + metadata search) by mid-2010

The search engine will have two main interfaces:

- A **simple search** for regular users with a text box and the button search.
- An **intermediate search** with specific and selected metadata fields only.
- A **combined advanced search** using all metadata and combinations of them together with the full-text for specialized users such as librarians, cataloguers and information management specialists.

Both modules will be web-based, use Java as programming language on the server side, and interface with Fedora-Commons 3.3 through its public and documented APIs to access its built-in XML repository. FAO plans to distribute them with a suitable open source licence.

Installation of Fedora 3.3 and final import of merged records from CDR/EIMS and FAODOC by mid-2010