

MITRA: A Metadata Aware Web Search Engine for Digital Libraries

Jorge Machado, José Borbinha

Biblioteca Nacional Digital
http://bnd.bn.pt



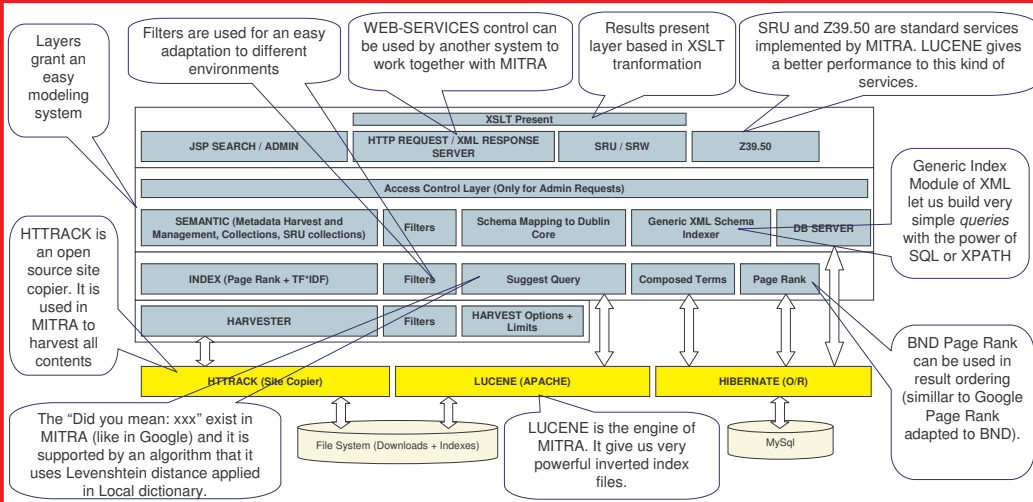
MITRA is an open-source advanced search engine developed by the BND, the National Digital Library initiative. It supports several services, namely the generic search service for the National Library of Portugal, BUSCA (http://busca.bn.pt).

MITRA can harvest and index on-line contents in multiple formats, as also structural and descriptive information in XML. If each site to index provides structural and descriptive information coded in XML, MITRA can be easily configured to index it accordingly, and use that information to enrich the searching functions or the presentation of the results.

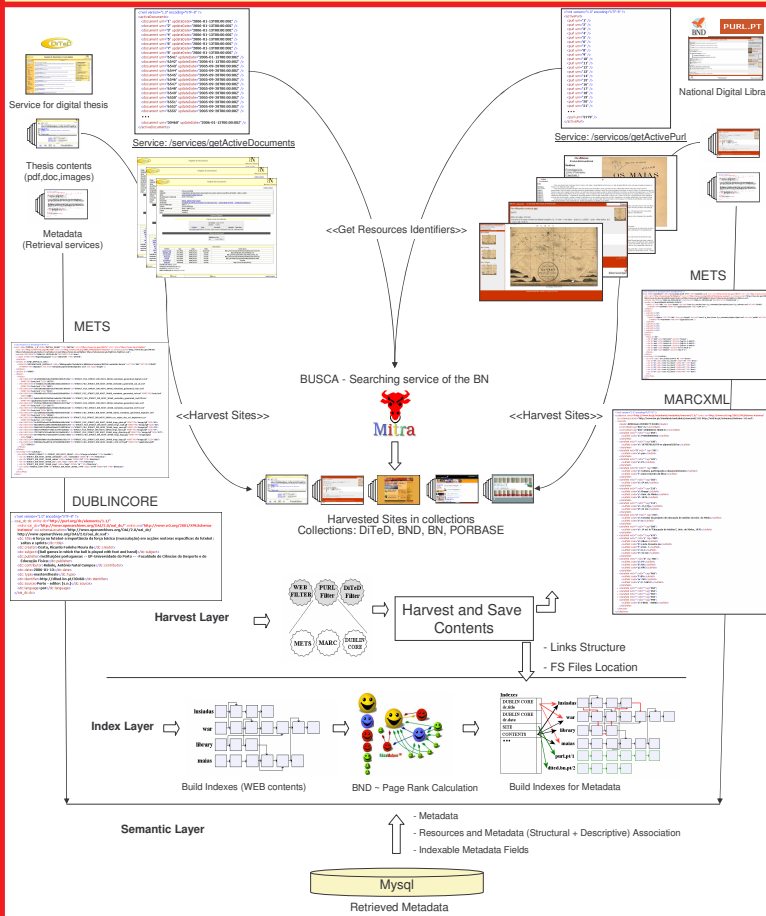
In the case of the collection of the BND (one of the collections available for searching in the service BUSCA), all the resources (both digitized and born-digital) have descriptive records in UNIMARC or Dublin Core, coded in XML. Also, all the digitized resources have a structural description in METS, which is an XML schema too, that makes it possible to describe the images according their location in the work (page number, chapter, part, etc.). In these cases, when the images are processed by OCR, the resulting text for each page can be also referenced by the METS structure, for convenient content indexing (even if the OCR has errors, it is useful for this purpose). All these schemas of descriptive and structural descriptions can be parameterized in MITRA, so it can take it all in consideration for the searching functions and presentation of the results (especially for ranking and clustering).

MITRA is based on the LUCENE basic search engine, but it uses also MySQL for internal databases. Besides the human search interface, the MITRA package provides also service for Z39.50, SRW/SRU, and OAI-PMH (from which all the descriptive metadata is available in Dublin Core).

Finally, MITRA has also a fully functional management interface available as web-services, making it possible to be controlled by external services. This is very convenient for applications using MITRA as their own indexing and public search service, like for example in DEPTAL, another open-source solution of the BND for institutional repositories.



MITRA na BN (http://busca.bn.pt)...



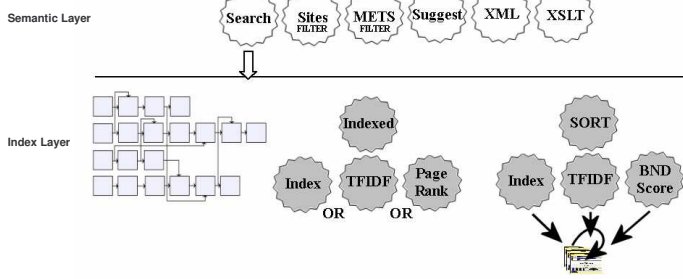
MITRA em ação...

BUSCA National Digital Library

This section shows the BUSCA search interface and XSLT transformations:

- Search Interface:** Multiple screenshots showing search results, filters, and detailed record views.
- XSLT:** Examples of XSLT transformations used for result presentation, including a 'GetRecordsByRequest Form'.

Search Process



Other Systems Interoperation using WEB-SERVICES

