

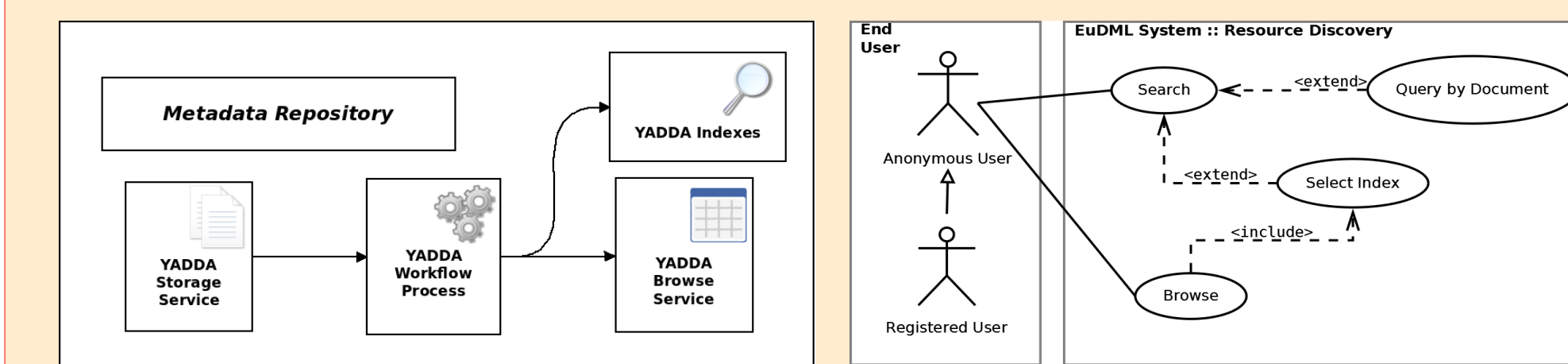
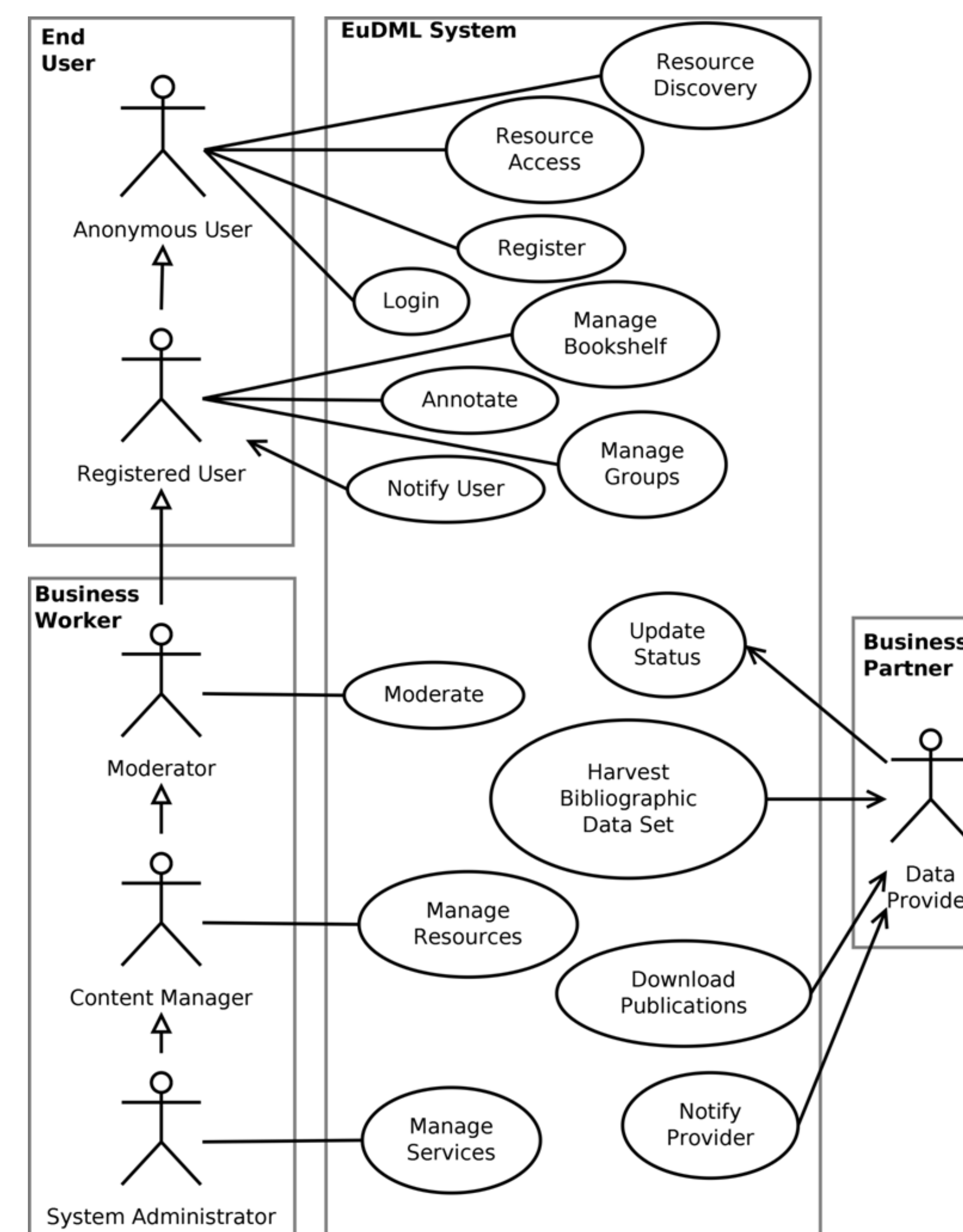
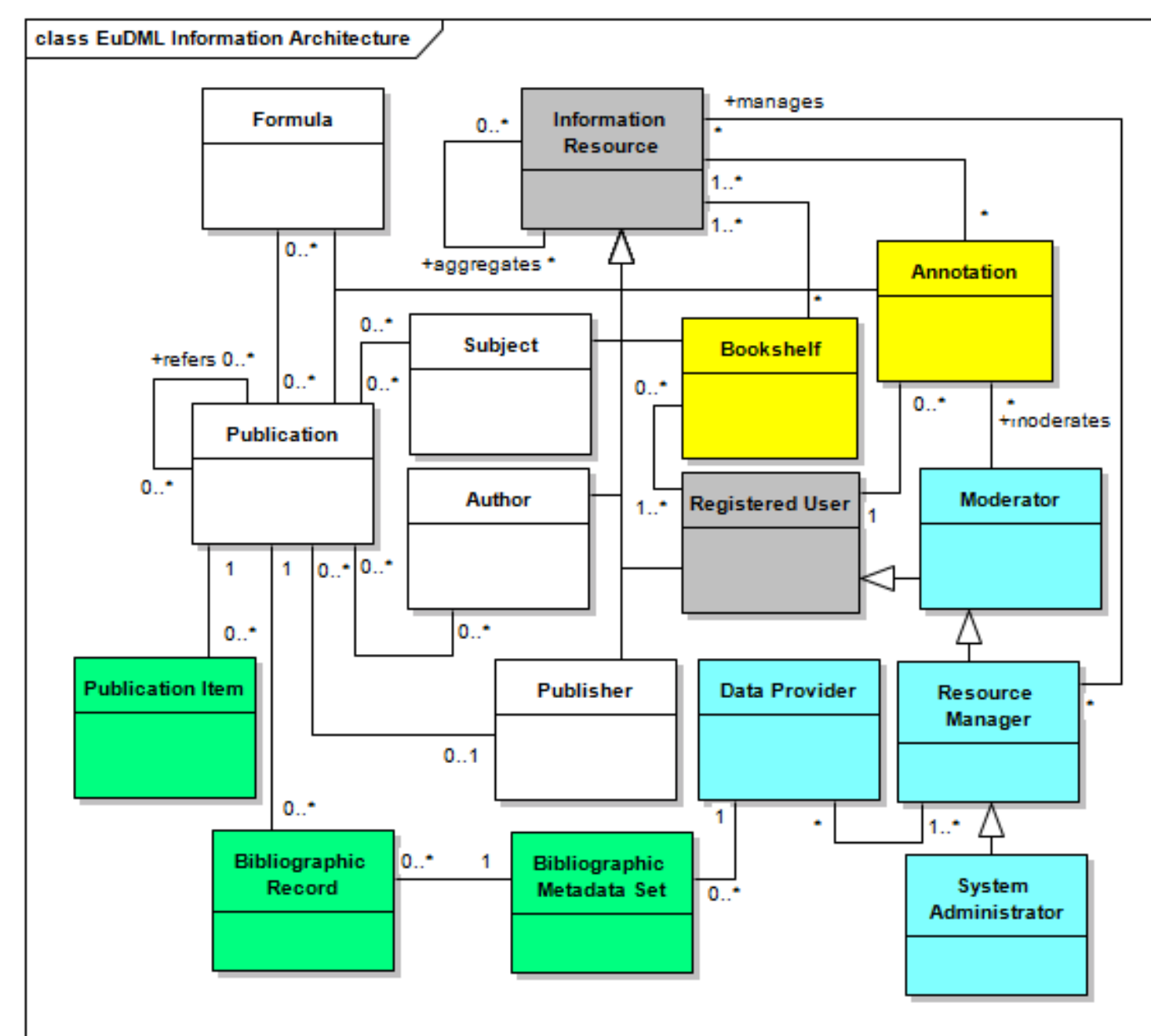
http://www.eudml.eu

EuDML

The EUROPEAN DIGITAL MATHEMATICS LIBRARY

YADDA Services...

- **Yadda Storage** manages the collection of records, where each record is a collection of XML metadata files and content files.
- **Yadda Indexes** is based on Apache Solr, and is used by the **Search Service** and **Browse Service**. The data in all these services is kept updated by the **Yadda Workflow Process** framework.
- The **Search Service** uses and extends many of the Solr advanced capabilities. Its main features are:
 - fully automatic Solr instance configuration and initialization
 - special Solr plug-in for mathematical formulae indexing and searching
 - use of advanced Solr's search functions, such as: Faceted Search, Filtering, Dynamic Fields
 - Solr's features with improved or customized processing logic, such as: Sorting, Highlighting, Wildcard Search
 - new features, designed to improve search in such cases as: authors name search, full text search
- The **User Interface** is a lightweight web application enabling searching and browsing. It is oriented on mathematical content, displaying MathML data and allowing to search for mathematical formulae. Searching supports faceted search, while browsing is possible by authors, journals, etc. Users also can create personal bookshelves and report possible errors and mistakes in the published data (through the Annotation Component). The **User Interface** seamlessly integrates with Twitter, Facebook, OpenID infrastructure and Google Scholar...

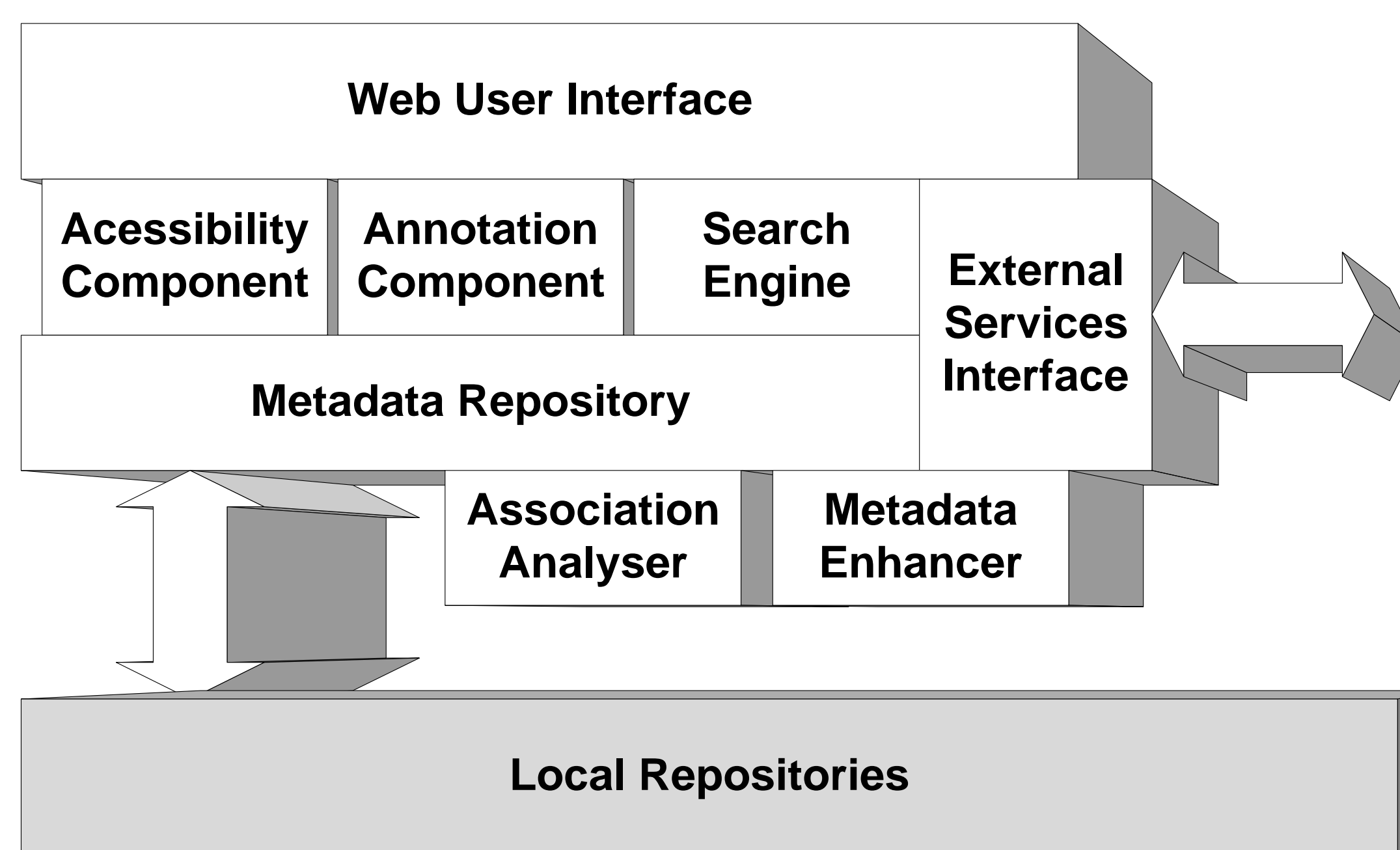


Annotations

The **Annotation Component** allows users to add additional metadata to items referenced or stored within EuDML, such as comments, reading lists or corrections. Annotations allow users to interact with items by exposing the annotation support to user via features on the EuDML website and user interface. The annotations component is available as a SOAP service interface for creating and managing annotations and collections. It follows a Linked Data (RDF) approach based on the OpenAnnotation specification.

Metadata Enhancement

The purpose of **Metadata Enhancement Toolset** is to provide tools that will allow algorithmic enhancements of the heterogeneous article metadata acquired from content providers. Tools gradually enhance collected [meta]data by conversions (e.g. TeX to MathML), OCR (e.g. getting text from page images), extraction (e.g. Text and formulae mining from PDF), analysis (e.g. Bibliography citations from plain text) or refinement (e.g. PDF optimization).



Association Analyser

The **Association Analyser** aims to discover rich semantic associations between documents. Mathematical works always build on previous results, and therefore can be considered a network of literature with rich links between individual documents. Sometimes such links are explicitly provided by metadata or can be inferred by subject classification. However, such metadata is often missing or incomplete. Associations between papers are discovered via:

- Citation Indexing: Bibliographies are used to create a network of documents within the collection by automatic parsing and linking of citations.
- Similarity Clustering: Unsupervised clustering techniques are used to cluster and group documents in the collection according to semantic similarity. Such techniques can also be used to compare papers or provide a quantitative analysis of content.

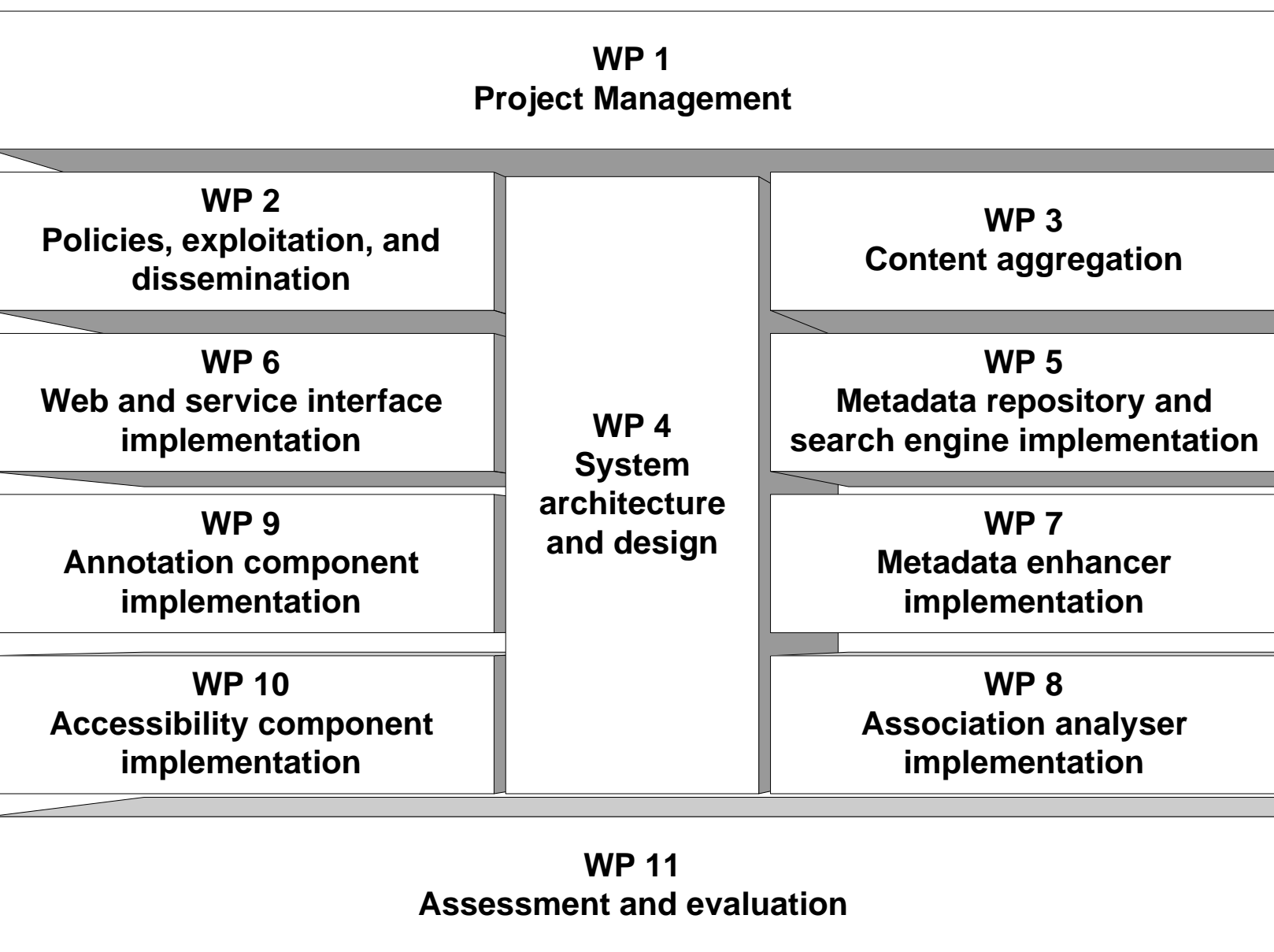
Metadata and Content Aggregation

REPOX (<http://repol.ist.utl.pt>) is the technology supporting the **data aggregator service**, a part of the Metadata Repository. REPOX manages the harvesting processes of the original metadata and full-text provided by the Data Providers, plus the data transformation and normalization. REPOX is the technology also used by the TEL – The European Library and Europeana services!!!

Partners

Our partners are universities, public institutions and leading technology providers, dedicated to producing quality scientific information. Participants include mathematicians, librarians, digital library specialists, publishers, professional information service and document engineering specialists, and computer scientists.

- Instituto Superior Técnico: Computer Science Department (Lisbon, Portugal)
- Université Joseph Fourier (Grenoble, France)
- University of Birmingham (United Kingdom)
- FIZ: Fachinformationszentrum (Karlsruhe, Germany)
- Masarykova univerzita (Brno, Czech Republic)
- University of Warsaw, ICM (Poland)
- Consejo superior de investigaciones científicas (Madrid, Spain)
- Édition Diffusion Presse Sciences (Paris, France)
- Universidade de Santiago de Compostela: Instituto de Matemáticas (Spain)
- Institute of Mathematics and Informatics, BAS (Sofia, Bulgaria)
- Matematický Ústav Av Cr VVJ, (Prague, Czech Republic)
- Ionian University (Corfu, Greece)
- Made Media UK (Birmingham, United Kingdom)
- Centre national de la recherche scientifique (France)



Data and Content Providers

The EuDML project aims to design and build a collaborative digital library service that will collate the mathematical content brought by its partners and make it accessible from a single platform, tightly integrated with relevant infrastructures such as the Zentralblatt MATH. As such, it is the first attempt toward a large-scale implementation of a Digital Mathematics Library (DML), and is expected to pave the way towards a truly inclusive and global DML. Integration of the collections will be an iterative process. It will start by just registering basic metadata for each contributed item (a book, a journal article, a contribution to a collective work such as conference proceedings, a Ph.D. Thesis, etc.), together with a unique identifier...

Provider	Collection	Website	OA/PSM server
CMO	CEDRAM	http://www.cedram.org/	http://math-thur.gf-grenoble.fr/repo/OAHandler
NUMDAM	NUMDAM	http://www.numdam.org/	http://math-thur.gf-grenoble.fr/repo/OAHandler
CSIC	DM-E	http://dmla.csic.es/	http://math-thur.gf-grenoble.fr/repo/OAHandler
EDPS	Math-J	http://publitions.edpsciences.org/	http://www.edpsciences.org/
FIZ	EBM	http://www.emis.de/ebm/	
ICM	DML-PL	http://www.icm.edu.pl/	http://www.icm.edu.pl/
IMBAS	BuDM	http://www.geni.math.bas.bg/	http://www.geni.math.bas.bg/9000/doi/
IJ	HM	http://www.hilmi.org/	http://www.hilmi.org/9000/doi/
SIROBIO	Math-M	http://www.math.mcgill.ca/sirobio/	http://www.math.mcgill.ca/sirobio/
RuDM	RuDM	http://www.ru.nl/	http://www.ru.nl/
BNF/CMD	JMPA	http://www.math.cnrs.fr/JMPA/	http://math-thur.gf-grenoble.fr/repo/OAHandler
BNP/ST	PM	http://www.pmf.uni-lj.si/	http://www.pmf.uni-lj.si/

... Next steps will use the outcome of a number of work packages to endow each contributed item with a sufficient number of fields ensuring effective exploitation of the whole aggregated content (seamless navigation, interlinking, language-neutral and math-aware retrieval, accessibility, etc.).

Accessibility

The purpose of this line of work is to identify or develop accessibility tools that can be used to provide access to mathematical literature to visually impaired users as well as print impaired users (i.e., people with specific learning disabilities like dyslexia).

There is currently a plethora of tools available that aim to provide accessibility support for visually or print impaired users. We have surveyed the more important ones and concluded that the support for advanced mathematics in accessibility tools is rather limited. It can also be stated that many tools work less well with mathematical expressions than promised. Generic screen readers, document readers and voice based browsers have very limited ability to work with math. However, dedicated modules could change this. There are a number of dedicated tools for mathematics that work fairly reliably with MathML.