# EuDML Assessment and Evaluation - Final Report

Romeo Anghelache, Jiri Rákosnik, Aleksander Nowiński, Wojciechowski Krzysztof, Thierry Bouche, Michael Jost, Olaf Teschke, Petr Sojka, Vlastislav Dohnal

**Project Acronym:**      **EuDML**

**Grant Agreement number:**  **250503**

**Project Title:**      **The European Digital Mathematics Library**

# D11.4: EuDML Assessment and Evaluation — Final Report

**Revision: 3.0 as of 15th May 2013**

**Authors:**

| | |
|---|---|
| Romeo Anghelache | FIZ Karlsruhe, FIZ |

**Contributors:**

| | |
|---|---|
| Jiří Rákosník | IMAS Prague, IMAS |
| Aleksander Nowiński, Krzysztof Wojciechowski | ICM Warsaw, ICM |
| Thierry Bouche | UJF Grenoble, UJF/CMD |
| Michael Jost, Olaf Teschke | FIZ Karlsruhe, FIZ |
| Petr Sojka, Vlastislav Dohnal | Masaryk University, MU |

# Revision History

| Revision | Date | Author | Organisation | Description |
|---|---|---|---|---|
| 1 | 18 Jan 2013 | Romeo Anghelache | FIZ | survey answers integrated and main text written |
| 1.11 | 23 Jan 2013 | Petr Sojka | MU | typo edits |
| 1.2 | 24 Jan 2013 | Romeo Anghelache | FIZ | added info on schema/best practices |
| 1.3 | 30 Jan 2013 | Jiří Rákosník | IMAS | internal review |
| 1.4 | 30 Jan 2013 | Thierry Bouche | UJF/CMD | internal review |
| 1.5 | 30 Jan 2013 | Petr Sojka | MU | polishing and consistency editing |
| 1.6,1.7 | 31 Jan 2013 | Romeo Anghelache | FIZ | fixed conditions of proposed parameters |
| 1.8 | 3 Apr 2013 | Romeo Anghelache | FIZ | added section 2 describing evaluation processes and methodology, inserted tables with evaluation criteria to make them explicit in Appendix E, described users and content-providers 3.3.3 |
| 1.9 | 4 Apr 2013 | Michael Jost | FIZ | internal review |
| 1.91 | 4 Apr 2013 | Aleksander Nowiński | ICM | input on sources |
| 1.92 | 4 Apr 2013 | Krzys Wojciechowski | ICM | input on sources |
| 2.0 | 5 Apr 2013 | Romeo Anghelache | FIZ | applied changes as suggested |
| 2.1 | 5 Apr 2013 | Thierry Bouche | UJF/CMD | internal review |
| 2.2 | 8 Apr 2013 | Romeo Anghelache | FIZ | made explicit that the missing documentation is the **optional, public** documentation, and is not part of the initial criteria of evaluation |
| 2.21 | 10 Apr 2013 | Petr Sojka | MU | tidying tables |
| 2.22 | 12 Apr 2013 | Romeo Anghelache | FIZ | incorporating changes from internal review |
| 3.0 | 14 May 2013 | Romeo Anghelache | FIZ | moved surveyed entities in new appendices (C, D); edited the definition of 'functional' |

**Statement of originality:**

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

## Executive Summary

This evaluation report's findings are that the project was delivered on schedule, meeting almost all the defined evaluation criteria, while almost all its performance parameters are found to be at or above their expected values.

The content-providers and user feedback on the EuDML Release (version 1.4) show that the system is currently stable, functional and useful.

Most of the suggested improvements made by users while answering the surveys were translated and tracked as bugs through the Mantis system (the EuDML bug-tracking system). Many of them have been attended to and the result is the current Release (version 2).

Few work areas are left to be refined, mostly documenting some services or processes.

This report ends with three recommendations:

1. Two new global performance (public effectiveness) parameters should be defined and monitored:
   - the ratio between the items freely available to the public and the total number of searchable items in EuDML (currently >87%);
   - the ratio between full-text indexed items and the total number of searchable items in EuDML (currently 59%).

   Thus, the overall aim of EuDML would be, between adding new collections, to achieve a value of 1 for both of these parameters.
2. Emphasis should be put on documenting clearly and concisely EuDML's functionalities for content-providers, Scientific Advisory Board and different categories of users (e.g. a Help/FAQ section should be created).
3. More robust steps should be taken by the partners to ensure the future growth and sustainability of EuDML.

# Contents

# 1   Introduction

This report shows how EuDML satisfies the requirements described previously in Deliverable D11.2 EuDML— assessment and evaluation plan (the Evaluation Framework) [12] while preserving as much as possible the structure of the D11.3 EuDML— assessment and evaluation plan — First Report [13].

In an attempt at completeness, this evaluation is performed from different perspectives: organizational, functional, content, user. On the technical side, EuDML is using state of the art software designed for searching and indexing, metadata extraction, validation and exportation, most of which is open-source and evolving continuously.

This document is structured so that a clear picture about the project's status emerges, that means its core is enumerating critical results without enumerating each evaluation requirement which was satisfied. However, for easy referencing, the full list of requirements and indicators defined in D11.2 [12] is listed in Appendix E along with the information source on how and if they were satisfied.


# 2   Description of the evaluation process and methodology

The evaluation process consisted in monitoring the evolution of the project and its parts by assessing its status against the requirements previously defined in the evaluation plan (D11.2 EuDML— assessment and evaluation plan), while the evaluation methodology consisted in mapping the assessments' results into a restricted space of qualifiers as follows:

- **functional and documented** (technically functional and publicly documented, on schedule),
- **functional** (technically functional but incomplete due to public documentation pending),
- **delayed** (technically functional, but implemented behind the schedule),
- **acceptable** (partially functional, on schedule) and
- **unsatisfactory** (not functional, behind the schedule).

The only difference here between **functional** and **functional and documented** resides in the public documentation/description. By "public documentation" we mean henceforth the availability of short documents, or on-line help, describing comprehensively what does EuDML provide and what its users should expect from it. It should be noted that the public documentation is optional, beyond the Description of Work, however, the EuDML Initiative association, which will be maintaining and developing the EuDML in the post-project period, will subsequently supplement this documentation. The technical documentation describing the EuDML system is provided in the project deliverables. **Functional** here means that almost all of the evaluation criteria defined for that facet were met, the criteria which were not met, if any, were not critical to the functionality of the EuDML as a system and/or as a digital mathematics library.

In the previous evaluation deliverable (D11.3 [13]) we used the terms 'good' and 'excellent' for the present terms 'functional' and 'functional and documented', respectively.

Assessments were made based on continuous direct interactions with the EuDML system, its designers, its content-providers and its users, as well as providing continuous

feedback to the project. This hands-on approach ensured the effectiveness of the evaluation in keeping the project focused on its performance coordinates. The presentation of this document is a direct result of the evaluation's evolution: we started with the entire list of criteria to be met, and dropped them off one by one whenever the project has satisfied them.

The typical monitoring activities consisted in:
- periodical complete downloads of the EuDML metadata followed by validating the data against the EuDML schema, checking its quality (verifying the links to the provided items), communicating the results to the project developers and using the EuDML's bug-tracking system,
- curating a particular collection (ELibM) and updating its EuDML data to verify directly if the mechanism of performing a collection-update is functional and is well supported by the EuDML system's administrators,
- surveying the content-providers on the quality of their interactions with the EuDML system, as well as the problems they encountered and the benefits they obtained,
- setting up groups of users and surveying them periodically on their EuDML experience,
- presenting the results of the intermediary evaluations in a summarized form at the project's periodical technical meetings inviting discussions on the next steps to be followed.

## 3    Evaluation results by facets

We will mention below if the qualifiers changed from the last evaluation. Even if a qualifier did not change, it does not mean that progress has not been made (e.g. some features were not scheduled to be implemented at the last evaluation, and now they are implemented and on schedule).

### 3.1    The EuDML as a project

The last external review of the EuDML project issued several recommendations in September 2011. The EuDML partners answered most of them successfully since, the only issue left is:
- although some important progress has been made since our last evaluation, a basic but comprehensive documentation for the various categories of EuDML users is still missing

This evaluation facet's overall score is **functional** (unchanged from the previous evaluation reported in D11.3 [13]): the EuDML project is on schedule and provides a functional and partially documented Digital Mathematics Library prototype.

### 3.2    The EuDML as a generic digital library

For this evaluation facet, we follow the criteria defined and listed in the DL.org project's Digital Library Reference Model [1], which are grouped by theme (content-oriented 2, user-oriented 4, architecture-oriented 10, functionality-oriented 6, policy-oriented criteria 9).

We will preserve the overall structure, but, to be concise, and because EuDML satisfies almost all the criteria, we will only enumerate here those ones where work still needs to be done. The optional part of the DL model is omitted here in its entirety and replaced with the specific criteria for a *mathematical* Digital Library enumerated in [12] and addressed below, in section 3.3.

Evaluation themes:

- Content
  The Policies regulating every information object are documented in EuDML metadata schema specification (v2.0 – final) and the associated best practices [18]. The complete list of assessments is shown in Table 2.
  Overall score for this theme: **functional and documented** (changed from **functional**).

- User-oriented
  The user roles and profiles are functional but not properly documented. The complete list of assessments is shown in Table 4.
  Overall score for this theme: **functional** (unchanged).

- Functionality
  The functions to manage actors and information objects are not yet documented, in particular how these functions can be organized in workflows, however, the functions enabling collaboration among DL actors/users are implemented and functional in the current EuDML Release as well as those dealing with information objects. In particular EuDML offers now a set of REST services (at `https://project.eudml.org/rest-services`) which are well documented, also including usage examples.
  EuDML offers also the standard OAI-PMH interface (at `https://project.eudml.org/oai-pmh-server`). The complete list of assessments is shown in Table 6.
  Overall score for this theme: **functional** (changed from **acceptable**).

- Policy
  The EuDML policies [2] follow the best practice recommendations endorsed by the International Mathematical Union and are described in [3]. The complete list of assessments is shown in Table 9.
  The overall score for this theme: **functional and documented**.

- Architecture
  The Digital Library Management System [1] (the set of facilities to set up and maintain the DL) is documented in deliverable D4.3 – EuDML global system functional specification and design – Revision. The complete list of assessments is shown in Table 10.
  The overall score for this theme: **functional**.

The overall score for this facet is **functional**: EuDML qualifies well as a generic digital library according to the *mandatory* and *recommended* criteria defined in the Digital Library Reference Model [1].

### 3.3 The EuDML as a digital mathematics library

For this facet we follow that part of the evaluation framework described in D11.2 EuDML— assessment and evaluation plan [12], which is *specific* to EuDML. Out of that checklist, we will only enumerate here the criteria which are not yet completely satisfied. (The complete lists are available in Appendix E).

#### 3.3.1 Content

Of the 9 content-related criteria (Table 3) specified in D11.2, only one is not met:

- No European publishers have been persuaded yet to cooperate with the library, but discussions are ongoing.

  All the performance indicators mentioned in D11.2 are met:

- **Diversity of partnership**: EuDML 12 content-providers, at expectation (expected: 12),
- **Critical Mass**: almost 230,000 digital items are currently present in EuDML, above expectation (160,000)
- **Diversity of content**: there are 317 integrated collections available in EuDML to date, above expectation (200).

  The project partners built a statistics page (`https://eudml.org/statistics`) which is updated on the fly at EuDML.

  Overall score for this facet: **functional**.

#### 3.3.2 System

The digital library system is the integration of various tools and processes, enabling and helping its users to locate and discover the information they need.

Of the 13 criteria (Table 11) specified in D11.2, all are met to date. However, the issue remaining to be fixed is listed below:

- No public and comprehensive documentation exists yet for the system as a whole (information workflow description, description of EuDML services relevant for the typical users).

  Progress has been made in integrating various components, the result is the current EuDML Release (version 2).

  Overall score for this facet: **functional** (changed from **acceptable**).

#### 3.3.3 Users

The objective of this evaluation facet is to understand and follow the expectations users have from the EuDML service. There are two complementary categories of users: the typical users, persons looking to locate and discover mathematical resources, and the content-providers, who are looking to enrich/curate/increase visibility of their own collections by interacting with EuDML.

Surveys have been submitted to potential users and the results are available below (4). The surveys themselves are listed in Appendix A for typical users and in Appendix B for content-providers.

The 'typical users' groups which answered the surveys below were formed with mathematicians, mathematics editors and librarians from Sofia (Bulgaria), Prague (Czech

Republic), Berlin (Germany) and Grenoble (France); there were 14 'typical users' in total, their names are among the list of contributors on the title page of this documents. They are enumerated in Appendix C.

The content-providers answering the surveys were IMI-BAS/BulDML (Bulgaria), SIMAI-UMI/BDIM (Italy), CMD/NUMDAM+CEDRAM (France), EDP Sciences/EDPS (France), SUB Goe/GDZ (Germany), IU/HDML (Greece) and FIZ/ELibM (Germany). They are enumerated in Appendix D.

Because of the relatively small number of users, weighting would be meaningless; the next best thing we could do is to de-duplicate their answers, categorize them into (neutral) notes, expectations, recommendations (as two degrees of urgency for improvement suggestions) and praises, and post the expectations and recommendations as issues in the EuDML's bug-tracking system to be addressed by the project developers. We also tried to preserve the authenticity of the users' message, whichever it was.

The complete list of user-related criteria is available in Table 5.

- **User's involvement**: 148 registered users, under expectation (expected 1,000 registered users contributing annotations in the 3-rd year)
- **User generated content** (**annotations and external links**): 187, under expectation (expected 10,000)

These two particular performance parameters have not been met because the user registration and annotation system was fully functional much later than originally estimated in the workplan (M30 instead of M18).

Overall score for this facet: **delayed** (changed from **acceptable**).

### 3.3.4 Usability (**user-system interaction**)

This facet focuses on the quality of user-system interaction. The general aim is to make it easy, efficient and enjoyable using the EuDML system.

Of the 10 criteria specified in D11.2 (Table 12), all 10 are met to date.

The usability study recommendations in [9] and [11] were being followed (see tables 13, 14 and 15) in the process of the EuDML system implementation.

Overall score for this facet: **functional** (changed from **acceptable**).

### 3.3.5 Usefulness (**user-content relation**)

This evaluation facet concerns the user-content relationship, i.e. relevance of content to the user.

Of the 18 criteria specified in D11.2 (tables 7 and 8), 15 are met to date, the remaining issues are listed here:

- User profiles are not yet derived in order to rank the query results depending on the user's mathematical background.
- No social network for each work in EuDML exists yet.
- There is no simple navigation and browsing through networks of related and interconnected documents and linking of elements (e.g. names of people, theorems or concepts) in these items to and from external resources such as encyclopaedia entries (including Wikipedia), historical information and cultural references is possible.

- No comprehensive documentation is publicly available yet for the new services relevant to content-providers, such as metadata capture, enhancement and merging, document and reference matching and cross-repository document linking. This issue is partially solved by the EuDML REST services, which were proven to be well documented and robust.

Performance indicators:

- **Enhancement progress**:
  There are currently 229,425 searchable items in the EuDML out of which, 135,990 (59%) are full-text indexed (of which 55,233 containing MathML); the number of freely accessible items is strictly greater than 200,424 (at least 87%) and the maximum limit of a moving wall for any item present in EuDML is 5 years; above expectation (expected: over 50,000 indexable full-text items with MathML).
- **Internal networking**: more than 1,000,000 links between database objects, considering an average of 5 similarity links per item; above expectation (expected: 500,000);
- **External networking**: more than 1,100,000 external links (expected: 500,000). Overall score for this facet: **functional** (changed from **acceptable**).

### 3.3.6 Performance (system-content relation)

Based on the user experience (described below), the overall score for this perspective: **functional** (changed from **acceptable**).

## 4  Evaluation by typical users

The following is a list of observations ("note:" and "fix:") and recommendations ("rec:") from mathematicians who experimentally interacted with the EuDML (Release 1.4) while answering the user-survey (Appendix A), as well as their expectations ("exp:") and praises ("praise:"). Many of their observations were already taken into account and the fixes are provided in the current EuDML Release (version 2). The observations left un-addressed were put in the Mantis bug-tracking system and will be attended to in the next EuDML Releases.

### 4.1  Homepage, facilities provided

**note:**

- the discontinuous black and orange vertical line on the left of the page is distracting and annoying, some less annoying visual aid would be better

**fix:**

- descriptions are sometimes too brief

**exp:**

- some help is missing on what a registered user can do
- "Recent Notes" should be moderated
- more adaptable to various screens/devices

**praise:**

- homepage is well designed
- the math formulas in titles and abstracts seem to be correctly displayed and handled
- "it's very good and handy to have the second column listing the document types, journals, authors and years of the found items and in that second column, things are even listed by the number of hits, and that number of hits is explicit"
- "available title, authors with affiliations, the precise references in the journal, abstract – excellent!"
- document citations in various formats (with good mathematic formulas)
- in the second column, there are the subjects with the classification number and the describing text, with links
- links to other databases

## 4.2 Search

**note:**

- complicated tree (AND, OR, NOT would be clearer)
- "takes a while to grasp that 'match all' means boolean AND, but OK afterwards"

**fix:**

- for German umlauts, ä→ae, ü→ue and ö→oe mappings are missing, thus search results are affected
- the "advanced search" is difficult to understand
- textual help should be improved (should work also for including years, languages etc., not only titles.)

**exp:**

- the downloadable search results file should contain EuDML in its name
- some journals of EuDML are not really mathematical journals; the query answers involving them should be given at the end of the list or not at all
- a clear/reset button would be useful

**praise:**

- the organisation in rules+sub-clauses which can be added/removed is clear and efficient.
- useful to have journals, authors and years listed along with the results
- overall satisfied with the functionality

## 4.3 Formula search

**note:**

- disappointing
- could do without
- not very efficient (quality of scans?)
- the user has to learn formulating requests properly
- a clear/reset button would be useful

**fix:**

- provide hints and examples on how to use
- provide info on using web browsers with MathML support

**praise:**

- works as expected
- good direction of development

Rated with an average score of 5 (individual scores are ranging from 2 to 8, so the relevance of this score is questionable).

## 4.4 Browse by Subject

**exp:**

- some filters, other than the title's first letter, would be useful
- journals could also be classified according to topics covered, publisher/country/ language, level (target readers), dates (living titles, very old ones. . . ).
- the mathematical classification could be used in combination with other criteria but not alone.

**rec:**

- add numbers to lower levels in MSC (e.g. 31-XX gives a list, containing 31-03 (which is a tree leaf) but also 31Axx (which is a tree node containing more subcategories)
- new navigation systems like finding papers which refer to some paper or mathematician which is most often cited by another one.
- use topic maps

## 4.5 Browse by Journals

**note:**

- "clear and easy if one knows what one's looking for"
- "OK, but don't see its usefulness"
- basic, but works (Portugaliae Mathematica has one entry in the list while the full-text are scattered over two repositories)

**fix:**

- "it is simply a list with incomplete information not good enough to be displayed"

**exp:**

- there is no journal information indicating successor journals and other relations
- journal could also be classified according to topics covered, publisher/country/ language, level (target readers), dates (living titles, very old ones. . . ).
- more information about ways to full text (e.g. for the period. . . see. . . )

**rec:**

- a kind of classification would be useful
- search capability in the local context of a currently selected journal.

**praise:**

- the list of journals is impressive
- "perfect"

## 4.6 The article-description page

**note:**
- strongly dependent on available sources and thus far away from a possible unification
- links in the section "In Other databases" can be improved
- link to Google Scholar is a search after the author name which can return confusing results

**fix:**
- publisher should not be mixed with pagination
- incorrect syntax in bibtex field for multiple authors
- info on moving wall is missing

**exp:**
- target links to a new tab in the browser, otherwise it's difficult to return to the initial page
- should be clearer that links to full-text point to external websites
- if full-text is embedded in/served directly by EuDML, there should be a second link

**praise:**
- looks quite comprehensive
- clear
- MathML looks fine

## 4.7 Similarity

**note:**
- can be improved
- difficult area

**fix:**
- facility is not documented: what does similarity mean? a few simple examples would make it clear, the algorithm description is not required

**exp:**
- to be offered a choice on what kind of similarity, if several available

**praise:**
- interesting and promising

## 4.8 Language issues

**note:**
- is it important to have the full article translated (score 0–10)? 3
- is it important to have the abstract translated (score 0–10)? 7
- is it important to search keywords translated in all the EuDML languages (score 0–10)? average of 6 (but large spread: from 1 to 10; a larger pool of users might be needed to settle the answer).
- "any translation should be treated as additional information only".

## 4.9 User environment and social network

**note:**
- a step in the right direction
- not clear how the user feedback would be useful
- scientists don't comment papers, most of the comments would be controversial
- monitoring the feedback would be a tedious and time-consuming task

**rec:**
- some guidance/documentation for those unfamiliar with this facility would be helpful
- "read status" property should be added as a feature to the personal list: it would help for personal reading planning
- preferences saved in the profile for a formula display form (TEXor MathML)

**praise:**
- capability of creating and managing private lists

## 4.10 General experience with EuDML

The scoring scale is 0 to 10 and refers to Release 1.4. At the time of writing this report, Release 2 fixed many of the issues reported by the users, so these scores should be understood as prefixed with "at least":

- general functionality: 7
- task accomplishment speed: 7
- information provided: 7
- ease of use: 7
- documentation: 5

The following user observations do not belong to a specific EuDML functionality category. They either refer to general EuDML features or to ways in which particular collections have curated their EuDML shared content:

**note:**
- would use EuDML for full-text
- "the aim is always the full text. The other facilities are only a way to reach it."
- cannot use EuDML as a single tool for bibliographic research (the database does not cover the whole mathematics literature)
- EuDML is more useful for mathematicians and historians of mathematics, or to mathematical societies and libraries with mathematical focus
- not all journals are there, would not use EuDML for a first search on some specific topic.
- would not use EuDML for citations.
- would not use EuDML if looking for a comprehensive hit list
- would not use EuDML for metadata, there are other, more complete databases for that
- would use EuDML if it would offer browsing by topic maps and/or a useful and/or user-configurable similarity search, or other equivalent browsing enhancements

- difficult to decide – should one search somewhere else if not in EuDML? (is it that full text does not exist, source is unreliable, illegal etc.?)
- links to other databases: very useful; link to the journal:useful; links to Google scholar: not useful
- usage tasks in decreasing importance order: full-text, abstract, bibliography, formulas, citations, related links, bookmarks and lists

**rec:**
- classification of journals by public and/or topics and/or language
- specific sample tours for specific users.
- a brief description of actual limitations (European source, non/collaborating institutions...)
- broader coverage, important journals are not yet included
- more documentation on search/similarity would be useful
- more info and links to authors' pages, profiles, or journals', publishers' home-pages, etc.

**praise:**
- EuDML "...is now quite nice, compared to the version I previously tested."
- the possibility to look for a formula is interesting and not always present in other databases.
- "EuDML covers all necessary tools for working mathematicians"
- "plenty of relevant things to be found in one place"
- "delivers good stuff especially for graduate students and researchers"
- references to other EuDML documents very useful
- references to other databases very useful
- navigating inside full text (following references to a theorem and getting back) is very useful
- bibliographic links useful

## 5    Evaluation by content-providers

Here follows a list of notes, as well as expectations and recommendations, made by the content-providers (CP) while answering to the content-providers survey (Appendix B):

**note:**
- (re)harvesting happens often
- issues with managing multiple ISSN numbers for articles published in printed editions and published in digital editions under different ISSNs. ISSN doesn't tell which media they refer to (printed or digital).
- ingestion workflow is stable and has improved
- the JATS NISO 1.0 schema [15] could be used at its full power
- "is there a way to make mathematics that is not in the public domain generally available?"

**fix:**
- there are links of types other than 'cite' or 'cited by', EuDML should use these types too (some of them very relevant scientifically, e.g. errata).

- parts of some collections are not yet ingested in EuDML
- decide if the enhancements produced by EuDML should be public domain or not
- OAI-PMH relies on outdated Dublin Core scheme (oai_dc); the CP would face many obstacles in case of changing metadata standard and migration.
- improve the partnership and exchange between CPs and EuDML
- create a page describing the EuDML providers and their contributions to the EuDML effort
- apparent schema limitation: there is no possibility of giving volume titles (this is possible only for issues). In mixed-citation, both authors and editors are coded in the same way (inside a <string-name>), with no possibility of marking the name (only the surname)
- EuDML should create and use an urn:eudml:author identifier

**exp:**

- the enhancements produced by EuDML should not be public domain
- participation in EuDML should produce links from the reference databases (Zentralblatt Math and MathSciNet) to the CP items
- plans to continue working with EuDML after the end of the project, if funding/resources available
- increase volume and quality of content
- open access to journals - use and share, setting up community groups.
- keep in touch with similar CPs, worldwide, in order to be aware of emerging and best practices around.
- a larger audience gets access to CP content

**rec:**

- improve communication with all CPs
- set up a more generally available documentation
- would be helpful to have some qualifiers to distinguish, filter and search digitized and digital born content.
- involve digital publishers and their software in EuDML
- use alternatives to OAI-PMH and get rid of some limitations of Dublin Core dependency and OAI-PMH protocol.
- EuDML could provide compliance directly in official software releases of (DSpace, Eprint, Invenio) (Similar to how OpenAIRE integration has been done and officially approved in the latest versions of DSpace, Invenio, etc.)
- modify and update OAI-PMH or provide alternative ways for non DC metadata aggregation. (e.g. XOAI Core Library [14])
- build a second tier EuDML network from "local" CPs routed through the current CPs (e.g. [16])
- ensure that handling data stored from publishers is a trusted process
- a gentler introduction in the EuDML schema would be helpful for CPs
- the ingestion workflow was opaque and CPs would like to know what happened and how it is done

**praise:**

- large OCR-ed files produced by EuDML are (re)used at some CP sites
- can use links to EuDML objects in the CP's items references
- being part of EuDML definitely improves the CP's visibility
- the processes of collecting and serving the CP items are currently considered good
- making digitalized publications easily available in one location
- most of the content is full-text searchable, some CPs are completely full-text searchable in EuDML
- the CPs providing books consider the change from the previous EuDML release a big improvement
- EuDML has become a very usable search system; structure is clear and clean

# 6     Concluding recommendations

To simplify the tracking of the evolution of EuDML as a service, two new global performance parameters should be used:

1. the ratio between the publicly readable items and the total number of searchable items in EuDML
2. the ratio between the full-text indexed items and the total number of searchable items in EuDML

These performance parameters are able to track the global effectiveness of EuDML as a full-text service to the public (if both of them are close to 1, then EuDML is as effective as possible).

Emphasis should be put on clearly and concisely documenting EuDML's functionalities for (potential) content providers, Scientific Advisory Board, and different categories of users (mathematicians, researchers, redactors, editors, reviewers, laymen) such that all the functionalities available in EuDML are described in appropriate terms and accessible either in a "Help"/"FAQ" section of the EuDML site, or context-based, or, ideally, in both forms.

More robust steps should be taken by the partners to ensure the growth and sustainability of the EuDML service after the end of this project.

# References

[1] Deliverable D3.2b of Project DL.org (Coordination Action on Digital Library Interoperability, Best Practices & Modelling Foundations), ICT Programme "Cultural Heritage and Technology Enhanced Learning", Project Number: 231551, L. Candela, A. Nardi, G. Athanasopoulos, D. Castelli, K. El Raheb, P. Innocenti, Y. Ioannidis, A. Katifori, A. Nika, G. Vullo, and S. Ross. The Digital Library Reference Model, 2011; http://bscw.research-infrastructures.eu/pub/bscw.cgi/d222816/D3.2b %20Digital%20Library%20Reference%20Model.pdf

[2] The EuDML policies description, 2012; http://project.eudml.org/policies

[3] Recommendations on electronic information communication, 2006. http://www.mathunion.org/fileadmin/IMU/Report/CEIC-Recommendations-2006.pdf

[4] D3.1 – Annexes - Complementary description of metadata schemas, Thierry Bouche, Hugo Manguinhas, 2010; https://wiki.eudml.eu/eudml-w/images/D3.1_annex.pdf

[5] D3.6 – EuDML metadata schema - final version, Claude Goutorbe, Nicolas Houillon, Thierry Bouche, Jean-Paul Jorda, 2013; https://wiki.eudml.eu/eudml-w/images/D3.6.pdf

[6] D4.3 - EuDML global system functional specification and design, Aleksander Nowiński, Wojtek Sylwestrzak, Krzysztof Wojciechowski, José Borbinha, 2012; https://wiki.eudml.eu/eudml-w/images/D4_3_v1_0.pdf

[7] D4.4 - The EuDML Information Life Cycle Process, JosÃľ Borbinha, Aleksander Nowiński, Gilberto Pedrosa, Krzysztof Wojciechowski, JosÃľ Delgado, 2012; https://wiki.eudml.eu/eudml-w/images/D4.4-v1.0.pdf

[8] D05.4 – The EuDML Metadata Registry and Repository - Final, Gilberto Pedrosa, José Borbinha, João Edmundo, António Higgs , 2013; https://wiki.eudml.eu/eudml-w/images/D5_4_v1.0.pdf

[9] D6.1 – Usability study, Tim Kitchen, 2010; https://wiki.eudml.eu/eudml-w/images/D6.1.pdf

[10] D7.4 – Toolset for image and text processing and metadata enhancement - Final release, Petr Sojka, Krzyś Wojciechowski, Nicolas Houillon, 2013; https://wiki.eudml.eu/eudml-w/images/D7.4-v1.0.pdf

[11] D11.1 – Report on evaluation of existing services and content providers, Radoslav Pavlov, Peter Pelikan, Michael Jost (ed.), 2011; https://wiki.eudml.eu/eudml-w/images/D11.1.pdf

[12] D11.2 – EuDML assessment and evaluation plan, Michael Jost, Radoslav Pavlov, 2011; https://wiki.eudml.eu/eudml-w/images/D11.2.pdf

[13] D11.3 – EuDML assessment and evaluation plan – First report, Romeo Anghelache, Michael Jost, 2012; https://wiki.eudml.eu/eudml-w/images/D11.3.pdf

[14] OAI 2.0 Server, 2012; https://wiki.duraspace.org/display/DSDOC3x/OAI+2.0+Server

[15] JATS: Journal Article Tag Suite, NISO 2012; http://jats.niso.org/

[16] Report on the Current State of the French DMLs, Thierry Bouche, 2009; https://eudml.org/doc/220779

[17] Mathematical Subject Classification 2010, editorial staffs of Mathematical Reviews (MR) and Zentralblatt für Mathematik (Zbl) in consultation with the mathematical community, 2010; http://msc2010.org/msc2010final2-Aug10.pdf

[18] EuDML metadata schema specification (v2.0 – final) and the associated best practices, EuDML project, 2013; https://project.eudml.org/eudml-metadata-schema-specification-v20-final

# Appendices

## A    Typical users survey

The EuDML entry page is http://eudml.org

**EuDML as a tool:**

Please offer short comments (dis/like, missing, suggestions/expectations) on your experience with the EuDML as a tool for mathematical resources discovery:

- facilities provided to you as a *registered user*
- homepage
- search
- browse by subject
- journal
- article-description page
- links
- "find similar documents"

**Satisfaction:**

What would be your satisfaction factors of choice (score 0–10)? How does EuDML fare at them (score 0–10)?

- general functionality
- task accomplishment speed
- information provided
- ease of use
- documentation

**Information handling strategies:**

Accomplishing tasks using EuDML:

- What "feature" bothers you the most? What is there which should NOT be changed?
- for what general kind of tasks would you use EuDML (looking for full-text, bibliography, citations, abstracts, formulas)?
- for what general kind of info-retrieval tasks would you NOT use EuDML?
- what kind of EuDML-provided links would you consider useful, very useful, not useful?

**Recommendations:**

What recommendations would you have for EuDML as a public service for different kinds of users (researchers, redactors, students, layman)?

**Multi-linguality:**

Language issues:

- is it important to have the full article translated (score 0–10)?
- is it important to have the abstract translated (score 0–10)?
- is it important to search keywords translated in all the EuDML languages (score 0–10)?

**Formula display:**

Your comments on EuDML's mathematical formulas display: suggestions/preferences?

**Similarity:**

> How would you rate (on a 0–10 scale) the formula search facility provided by EuDML? If you have any special comments/advice/expectations on this issue, please write them here.

**Missing features:**

> Is there anything missing from EuDML that you would badly need in your specific activities?

## B    Content-providers survey

**The EuDML schema:**

- any further recommendations on tuning the EuDML schema towards an effortless content export to the EuDML collection?
- is there anything important missing or not yet addressed in the current schema?

**Ingestion workflow:**

- how would you qualify the evolution of the EuDML ingestion workflow of your content since the last survey (improved|same|degraded)?
- any recommendations on improving the EuDML ingestion workflow on the content provided by you?
  - technically
  - organizationally

**Benefits:**

- are there currently any benefits for you from participating as a content-provider in EuDML?
- what benefits would you expect in the future from continuing your participation as a content-provider in EuDML?
- do you have plans for collaborating as a content-provider with the EuDML, after the end of the EuDML *project*?

**The content collection:**

- do you plan to enlarge the content offered to EuDML?
- any plans to increase the amount of full-text in your collection to be indexed by the EuDML?
- any suggestions for EuDML on the ways to further enlarge its content collection?
- any suggestions about how the content collections should be processed or served by EuDML?

**The end result:**

- what would be, in your opinion, the main drawbacks/obstacles in your interaction with EuDML?
- any suggestions for improving the current situation?

**Missing features:**

> Is there anything else we missed or you have a strong opinion about, relative to EuDML?

## C  List of surveyed users

- Jiří Veselý, Assoc. Professor, mathematician, Charles University Prague
- Jarmila Štruncová, M.A., librarian, Head of the Library, Institute of Mathematics AS CR Prague
- Jiří Rákosník, Dr., mathematician, Deputy Director, Institute of Mathematics AS CR Prague
- Georgi Simeonov, BSc, Informatics and Mathematics Software Developer, Institute of Mathematics and Informatics at Bulgarian Academy of Sciences (IMI-BAS)
- Radoslav Pavlov, Professor of Computer Science, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences (IMI-BAS)
- Peter Stanchev, Professor of Computer Science, IMI-BAS, Kettering University
- Jean-Luc Archimbaud, senior engineer at Cellule Mathdoc (CNRS/CMD)
- Brigitte Bidegaray-Fesquet, math researcher at CNRS (Laboratoire Jean-Kuntzmann, Grenoble)
- Yves Laurent, senior math researcher at CNRS (Institut Fourier, Grenoble) and former director of Cellule Mathdoc (CNRS/CMD)
- Julien Puydt, Ph.D., math teacher (EPA, Montbonnot)
- Klaus Kiermeier, Dipl.-Math., Section Editor Mathematics, FIZ-Karlsruhe
- Lucia Santamaria Lara, Ph.D. (Physics), Author Database Manager and Section Editor Mathematics and Physics, FIZ-Karlsruhe
- Helena Mihaljevic-Brandt, Ph. D. (Mathematics), Deputy Head of Department Mathematics and Computer Science and Section editor Mathematics and Computer Science, FIZ-Karlsruhe
- Aleksandar Perovic, Dipl.-Math., Manager of the Electronic Library of Mathematics and Section Editor Mathematics, FIZ-Karlsruhe

## D  List of surveyed content-providers

- BDIM (Biblioteca Digitale Italiana di Matematica), Vittorio Coti Zelati, www.bdim.eu, SIMAI (Società Italiana per la Matematica Applicata e Industriale) & UMI (Unione Matematica Italiana)
- BulDML, Georgi Simeonov, sci-gems.math.bas.bg, Institute of Mathematics and Informatics at Bulgarian Academy of Sciences (IMI-BAS)
- EDPS, Marie-Louise Chaix, www.esaim-cocv.org, EDP Sciences
- NUMDAM, Thierry Bouche (scientific advisor) and Hélène Falavard (project manager), www.numdam.org, Cellule Mathdoc (CNRS/UJF)
- CEDRAM, Thierry Bouche (scientific advisor) and Claude Goutorbe (project manager), www.cedram.org, Cellule Mathdoc (CNRS/UJF)
- GDZ-Mathematica, Rolf B. Röper, Göttingen State and University Library, Mathematica, gdz.sub.uni-goettingen.de/mathematica

# E     List of evaluation requirements

Table 1: Project evaluation criteria/Performance indicators

| Nr. | Indicator name | Goal | Result | Source |
|---|---|---|---|---|
| 1 | Timeline and budget compliance | to achieve | achieved | DoW |
| 2 | Conformance to the work plan | to achieve | achieved | see criteria below |
| 3 | Number of participating content repositories (partners or associated partners) | 12 | 12 | OAI metadata |
| 4 | Number of integrated collections (e.g. journal runs, book series, conference proceedings, Ph.D. theses) | 200 | 317 | OAI metadata (total # of journals + # of "other serials" + # of "book collections") |
| 5 | Number of integrated resources (digital items reused from the network) | 160,000 | 225,809 | OAI metadata |
| 6 | Number of supported resources (items with enhanced MathML full text and metadata) | 90,000 | 135,990 | OAI metadata |
| 7 | Links between database objects | 500,000 | 1,000,000 | internal metadata |
| 8 | External links (generated for third parties using dedicated component, or detected through referrer website) | 500,000 | 1,100,000 | internal metadata |
| 9 | User's involvement | 1,000 | 148 | live statistics |
| 10 | User generated content (annotations and external links) | 10,000 | 187 | partners' activities |

Table 2: Digital library content-oriented criteria [12]

| Nr. | Criterion | Result | Source |
|---|---|---|---|
| 1 | The Digital Library must manage a set of Information Objects and the set cannot be empty. | satisfied | OAI metadata |
| 2 | Every Information Object must have a unique identifier (Resource Identifier). | satisfied | OAI metadata |
| 3 | Every Information Object must have at least one element of Metadata associated with it. | satisfied | OAI metadata |
| 4 | Every Information Object must belong to at least one Collection. | satisfied | OAI metadata |
| 5 | Every Collection must have a unique identifier (Resource Identifier). | satisfied | OAI metadata |
| 6 | Every Collection must have at least one element of Metadata associated with it. | satisfied | OAI metadata |
| 7 | Every Information Object should conform to an explicit and known format (Resource Format). | satisfied | EuDML schema |
| 8 | Every Metadata should conform to an explicit and known format (Resource Format). | satisfied | EuDML schema |
| 9 | Every Annotation should conform to an explicit and known format (Resource Format) | satisfied | class eu.eudml.service. annotation.Annotation |
| 10 | Every Collection should have a well-defined intension, i.e., the set of criteria characterising Collection membership, and should have a well-defined extension, i.e., the set of Information Objects belonging to the collection. | satisfied | OAI metadata |
| 11 | Every Information Object should be regulated by Policies. | satisfied | EuDML policies |

Table 3: EuDML-specific content-oriented criteria [12]

| Nr. | Criterion | Result | Source |
|---|---|---|---|
| 1 | Content consists of published texts holding mathematical knowledge that has been validated through a scientific editorial process (criterium to be understood in a broad sense) | satisfied | EuDML policies |
| 2 | Content typically open access within 5 years of its publication date. | satisfied | EuDML policies |
| 3 | Less than 5% of the whole content being subject to any form of restricted access. | satisfied | EuDML policies |
| 4 | The network of documents is constructed by merging and augmenting the information available about each document from each collection, and matching documents and references across the entire combined library. | satisfied | internal metadata |
| 5 | Additional repositories (in Europe) have been stimulated to join. | satisfied | BDIM, math-net.ru |
| 6 | European publishers have been persuaded to cooperate with the library, licensing it for open access with a reasonable moving wall licensing policy. | not satisfied | |
| 7 | Collections contain texts that would be qualified as belonging to physics, economics and social sciences, etc. | satisfied | EuDML system |
| 8 | Existing thesauri of mathematical keywords are used | satisfied | system module 'eudml-term-translator' |
| 9 | Translation lists (Mathematical Subject Classification, UNESCO thesaurus, full text analysis) are used so that a query in one language can return documents in any other. | satisfied | system module 'eudml-term-translator' |

Table 4: Generic Digital Library User-oriented criteria

| Nr. | Criterion | Result | Source |
|---|---|---|---|
| 1 | The Digital Library must serve a clearly identified set of Actors and this can not be an empty set. | satisfied | D4.3 [6] |
| 2 | Every Actor must have a unique Resource Identifier. | satisfied | D4.3 [6] |
| 3 | Every Actor must be described by at least one Actor Profile. | satisfied | D4.4 [7] |
| 4 | Every Actor must act with at least one Role. | satisfied | D4.3 [6] |
| 5 | the set of managed Roles must include the DL Manager Role. | satisfied | D4.4 [7] |
| 6 | The set of managed Roles must include the DL Software Developer Role. | satisfied | D4.4 [7] |
| 7 | The set of managed Roles must include the End-user Role. | satisfied | D4.4 [7] |
| 8 | Every Actor should perform Actions that apply Functions and concern Resources. | satisfied | D4.4 [7] |
| 9 | Every Actor that interacts with a digital library should be able to perform certain Actions that involve the application of Functions and specific Resources. | satisfied | D4.4 [7] |

## Table 5: EuDML-specific User-oriented criteria

| Nr. | Criterion | Result | Source |
|---|---|---|---|
| 1 | User surveys will be carried out in order to receive user feedback and better understand user needs. | satisfied | project evaluation deliverables |
| 2 | Replicated, synchronised management of user annotations to items in the content repositories are in place. | satisfied | EuDML site |
| 3 | Annotations can be comments, discussion threads, tutorials, reviews, reading lists, or other user contributed elements that can be attached to individual items in the collection. | satisfied | EuDML site |
| 4 | The user interface supports viewing of such content and the search engine supports searching within it. | partially satisfied | EuDML site |
| 5 | Annotations can themselves contain references to items and are analysed by the association analyser toolset. | partially satisfied | EuDML site |

## Table 6: Generic Digital Library Functionality-oriented criteria

| Nr. | Criterion | Result | Source |
|---|---|---|---|
| 1 | The purpose of the DL is to offer functions, i.e., a particular processing task that can be realised on a Resource or a Resource Set as the result of an Action of a particular Actor. | satisfied | D4.3 [6] |
| 2 | Every Function must have a unique identifier (Resource Identifier). | satisfied | D4.3 [6] |
| 3 | Every Function must be performed by Actors. | satisfied | EuDML site |
| 4 | Every Actor must be provided with Functions to Access Resources. | satisfied | EuDML site |
| 5 | Every Actor must be provided with Functions to Discover Resources. | satisfied | EuDML site |
| 6 | Every DL System Administrator must be provided with Functions to Manage & Configure DLS. | satisfied | D4.3 [6] |
| 7 | Every Function should be able to interact with other Functions. | satisfied | D4.4 [7] |
| 8 | Functions to Acquire Resources should be provided. | satisfied | EuDML site |
| 9 | Functions to Browse the Resources should be provided. | satisfied | EuDML site |
| 10 | Functions to Search the Resources should be provided. | satisfied | EuDML site |
| 11 | Functions to Visualize the Actor's requested Resources should be provided. | satisfied | EuDML site |
| 12 | Functions to Manage Information Object(s) should be provided. | satisfied | EuDML site |
| 13 | Functions to Manage Actor(s) should be provided. | satisfied | sysadmin |
| 14 | Functions to Manage DL specific domains in a large scale should be provided. | satisfied | EuDML REST services |

## Table 7: EuDML-specific functionality-oriented criteria (part 1/2)

| Nr. | Criterion | Result | Source |
|---|---|---|---|
| 1 | Users can navigate by browsing the collections, following links to related items (same author, text citing or cited by the given one, similar subject, similar mathematical content, etc.). | satisfied | EuDML site |
| 2 | Users are guided by tips or additional keywords left by other users, and leave their own annotations as well. | satisfied | EuDML site |
| 3 | User profiles are derived in order to rank the query results depending on the user's mathematical background. | not satisfied | EuDML site |
| 4 | Linked resources, such as Zentralblatt by partner FIZ can be used to explore further with other methods, bringing the user back when a reference to an interesting EuDML item is finally found. | satisfied | EuDML site, Zentralblatt-MATH site |
| 5 | All item pages provide a link to the associated full text, 95% of which point to a partner's repository serving the file under open access. The remaining 5% will be hosted at their publisher's platform, possibly charging for access. | satisfied | EuDML site and OAI metadata |
| 6 | A number of features will help locate related items, refine queries, and support quickly retrieving the most relevant documents. | satisfied | EuDML site |
| 7 | Content providers enjoy new services such as metadata capture, augmentation and merging, document and reference matching and cross-repository document linking. | satisfied | content providers surveys |
| 8 | EuDML will make historical and comparative analysis, or even serendipitous discovery, of the development and achievements of European mathematics much easier. | satisfied | EuDML site |
| 9 | It is an essential feature of EuDML to address the difficult question of providing access to articles based on their subject, or scientific meaning, rather than on their language. | partly satisfied | EuDML site |

Table 8: EuDML-specific functionality-oriented criteria (part 2/2)

| Nr. | Criterion | Result | Source |
|---|---|---|---|
| 10 | Interlinking is used as a powerful access tool to the mathematical resources regardless of their language, but rather according to their subject and/or their scientific importance. | satisfied | EuDML site |
| 11 | A link to a documents review in one of the reviewing databases (Jahrbuch, Zentralblatt MATH, Math. Reviews) is given. | satisfied | content-providers and metadata |
| 12 | The interlinking infrastructure deployed in the project allows the exploitation of links to and from other related resources, such as citations from reviews and from subsequent works. | satisfied | EuDML site |
| 13 | Webs of citations provide a language-neutral way to find resources across multiple languages. | satisfied | EuDML site |
| 14 | A social network designed for each work in EuDML makes it possible to offer a powerful scenario for resource discovery by serendipity. | not satisfied | EuDML site |
| 15 | Mathematical content is encoded in a semi-structured format. | satisfied | EuDML site and metadata |
| 16 | Mathematical knowledge management techniques will be solicited to assess its novel technologies such as mathematical OCR, XML/MathML full-text generation from (La)TeX source files or PDF, formula representation and searching, and mathematical similarity metrics. | satisfied | different parts of the system including 'eudml-processing' module and searching module |
| 17 | Tools are in place to identify various types of referential and semantic connections between different items in the EuDML Metadata Repository, and also between such items and external resources. | satisfied | EuDML site |
| 18 | Simple navigation and browsing through networks of related and interconnected documents and linking of elements (e.g. names of people, theorems or concepts) in these items to and from external resources such as encyclopaedia entries (including Wikipedia), historical information and cultural references is possible. | partially satisfied | EuDML site |

Table 9: Generic Digital Library Policy-oriented criteria

| Nr. | Criterion | Result | Source |
|---|---|---|---|
| 1 | The Digital Library must be regulated by a clearly defined set of Policies and this can not be an empty set. | satisfied | sysadmin |
| 2 | Access Policies must regulate the use of the Digital Library by Actors. | satisfied | sysadmin |
| 3 | Every Policy must be addressed at least to an Actor. | satisfied | sysadmin |
| 4 | Every Policy must have clearly defined scope(s) and characteristics (Policy Quality Parameter). | satisfied | sysadmin |
| 5 | Every Policy should be expressed by an Information Object. | satisfied | sysadmin |
| 6 | Every Policy should have a unique identifier (Resource Identifier). | satisfied | sysadmin/java |
| 7 | Every Policy should have a known format (Resource Format) | satisfied | sysadmin/java |

Table 10: Generic Digital Library Architecture-oriented criteria

| Nr. | Criterion | Result | Source |
|---|---|---|---|
| 1 | The Digital Library System underlying the 'digital library' must have a well-defined Software Architecture. | satisfied | D4.3 [6] |
| 2 | The Digital Library System underlying the 'digital library' must have a well-defined System Architecture. | satisfied | D4.3 [6] |
| 3 | Every Architectural Component must have a unique identifier (Resource Identifier, identifiedBy) | satisfied | D4.4 [7] |
| 4 | The Software Architecture must consist of at least one well identified Software Architecture Component. | satisfied | D4.3 [6] |
| 5 | The System Architecture must consist of at least one Hosting Node and one Running Component. | satisfied | D4.3 [6] |
| 6 | The 'digital library' service is deployed and operated by means of a Digital Library Management System. | satisfied | D4.4 [7] |
| 7 | Every Software Component should be regulated by a License. | satisfied | sysadmin |
| 8 | The Software Architecture should be composed of more than one identifiable Software Architecture Components. | satisfied | D4.3 [6] |
| 9 | The System Architecture should be composed of more than one identifiable System Architecture Components. | satisfied | D4.3 [6] |
| 10 | Every Architectural Component should conform to a Framework Specification. | satisfied | D4.3 [6] |

Table 11: EuDML-specific architecture-oriented criteria

| Nr. | Criterion | Result | Source |
|---|---|---|---|
| 1 | The services are accessible to humans through a web portal with many innovative discovery features, and to other machine services through a set of common digital libraries' interoperability protocols (namely Z39.50, SRU and OAI-PMH with an optional dedicated schema). | satisfied | EuDML site |
| 2 | EuDML tools are able to search, browse and exploit a distributed network of resources as if it were a single well-managed library. | satisfied | EuDML site |
| 3 | Metadata schemas from the content providers have been identified. | satisfied | D3.1 [4] |
| 4 | A common EuDML metadata schema has been defined, and present a low barrier of entry for potential EuDML partners in the future. | satisfied | D3.6 [5], EuDML schema (JATS NISO 1.0) |
| 5 | A framework to federate the metadata in a central metadata repository is in place. | satisfied | system |
| 6 | A metadata registry makes it possible to easily integrate new data providers with new metadata schemas, reusing the aggregated metadata in any new required schema. | satisfied | D5.4 [8] |
| 7 | Discovered metadata are validated and merged with those already registered for the target items. | satisfied | system module 'eudml-processing' |
| 8 | Tools are in place to identify, from the Metadata Repository, items that may benefit from metadata enhancement and automatically apply such enhancement processes to them. | satisfied | system module 'eudml-processing' |
| 9 | As much of the full texts as possible has been converted to structured XML with MathML representation of formulas and English metadata. | satisfied | system module 'eudml-processing' |
| 10 | A set of dedicated tools has been packaged in order to generate metadata (structured textual OCR, mathematical OCR, keyword extraction, subject classification, bibliographic linking and citation, etc.). | satisfied | system module 'eudml-processing' |
| 11 | Common authority and interlinking structures have been developed. | satisfied | system module 'eudml-processing' |
| 12 | Tools and workflows to extract textual (coded in a proper XML schema) and mathematical (MathML) metadata (i.e. titles, keywords, authors, references etc.) from items in the content repositories, namely various types of mathematical documents, including scanned images, TeX/LaTeX sources, PDF documents, etc., are in place. | satisfied | D7.4 [10], system module 'eudml-processing' |
| 13 | Important requirements for interoperability were considered. | satisfied | EuDML REST services |

Table 12: EuDML-specific usability criteria

| Nr. | Criterion | Result | Source |
|---|---|---|---|
| 1 | EuDML is accessed via a web interface for human users, and a web service interface for tools and systems. | satisfied | EuDML site |
| 2 | Human interfaces all share the same graphical style and look and feel, promoting an "EuDML brand" which is multilingual and provides accessibility options to visually impaired and dyslexic users. | satisfied | EuDML site |
| 3 | A set of tools improves, to the extent that can be reached within this project's duration and resources, the accessibility to the corpus for visually impaired users. Born digital content has been converted to MathML and Daisy and made usable to Braille readers and text-to-speech engines. | partially satisfied | EuDML site |
| 4 | Explicit support is provided, using the latest technologies, for visually impaired or dyslexic users as well as automatic language translation support. | satisfied | EuDML site |
| 5 | Web 2.0 features are used. | satisfied | EuDML site |
| 6 | Insufficient metadata are augmented to a minimal level of quality among all integrated collections. | satisfied | D4.4 [7] |
| 7 | Mathematical knowledge management techniques are applied to overcome language barriers and connect various items related by their subject. | satisfied | EuDML site |
| 8 | Machine interfaces are based in common standards, or in effective web-services, following the Representational State Transfer (REST) paradigm when relevant (or SOAP when recommended), and outputting common standards-based representations including XML (agreed schema), RSS and JSON. | satisfied | REST services |
| 9 | Functional interfaces and widgets have been developed making it possible to include a "EuDML Search Box" in other local systems and portals. | satisfied | EuDML site |
| 10 | A widget configuration facility has been developed, making it easy for users to create tailored search interfaces for their own websites. Other functional interfaces have been designed and implemented for services related to interoperability. | satisfied | EuDML site |

Table 13: D6.1 "Usability study" recommendations (part1/3)

| Nr. | Recommendation | Result | Source |
|---|---|---|---|
| 1 | the site should have clearly tailored 'landing pages' for each journal, volume, paper and author, designed to catch searches on search engines and other indexing sites. | partially followed | EuDML site |
| 2 | our metadata and link structures are attractive and interoperable with Google and Google Scholar. | followed | `https://eudml.org/robots.txt`, `https://eudml.org/sitemap.xml` |
| 3 | In these circumstances we should integrate with or link to other platforms. | followed | REST services |
| 4 | We should consider citations to be a critical navigation feature. | followed | EuDML site and metadata |
| 5 | We should research Mendeley as a possible platform for delivering some of the more social aspects of EuDML that we are considering under 'Annotations'. | followed | EuDML site |
| 6 | We should consider adding obvious links to Authors' home pages. | not followed | |
| 7 | The ability to quickly and easily get to a page of results is perhaps more important than the ability to refine a search in precise detail. | followed | EuDML site |
| 8 | Provide an easy-to-use listing of all journals on the site, and a dedicated page for each journal, with a list of volumes and an easy search to find any paper from that journal. Link back to this page from any paper for which this journal is the source. | partially followed | EuDML site |
| 9 | Journal 'landing pages' should also appear in general search results when they provide a close match for a search. | not followed | EuDML site |
| 10 | Consider alpha-listing journal titles beginning with the phrase 'journal of' by the third word. | not followed | EuDML site |

Table 14: D6.1 "Usability study" recommendations (part2/3)

| Nr. | Recommendation | Result | Source |
|---|---|---|---|
| 11 | Call the journal list link 'Journals' rather than 'Browse' or similar. | followed | EuDML site |
| 12 | Have a separate page for each volume (listing papers) and a top-level listing page ordering from most recent to oldest. | not followed | EuDML site |
| 13 | Consider adopting the boolean search style, but hiding chained boolean search fields until the user indicates their wish to add an additional filter (see Figure 17). This offers the following benefits: The relationship between the field, the boolean operator and the next field are made explicit by the interface. The transition between 'simple' search and 'advanced' search is simple and subtle, rather than jarring: You just add extra filters. | followed | EuDML site |
| 14 | Include a NOT operator. | followed | EuDML site |
| 15 | Support the following operators within the search, this could be achieved with synonyms and compilation to SQL:<br>• "exact phrase"<br>• AND & +<br>• OR \|<br>• Go* matches Gordon<br>• published>1975 & ( Author:*Frey \| Author:*) | followed | EuDML site |
| 16 | Alternatively: Consider adopting Google's approach. Certain operators are supported, but the search is mostly over a full-text index, and the 'advanced search' field provides a clear indication of how to perform boolean text searches. | followed | EuDML site |
| 17 | Include a 'sort by' control on the search, and on the results page. | followed | EuDML site |
| 18 | Enabling searching with LaTeX is more important to mathematicians than enabling searching with MathML. | followed | EuDML site |
| 19 | Link the title of the paper, either directly to the paper, but more likely to an intermediary page which features the abstract, the download, references and citations, and various available formats. This will keep the results page cleaner and more Google-like and provide a natural, indexable 'home' for each paper. | followed | EuDML site |
| 20 | Each result from the list page should feature a 30 word excerpt from the abstract, with keyword-search terms highlighted in bold. This provides reassurance that results are relevant and allows the user to scan the results quickly to identify relevant papers. If keywords are identified in the abstract, use the 30 words surrounding those keywords, if not use the first 30 words. | followed | EuDML site |

Table 15: D6.1 "Usability study" recommendations (part3/3)

| Nr. | Recommendation | Result | Source |
|---|---|---|---|
| 21 | PDFs with BiBTex for citations should be clearly promoted as the key option. Other formats should be promoted less. | followed | EuDML site |
| 22 | Clicking on authors within search results should lead to a dedicated page on that author where available, with a list of papers by them, or otherwise a search for that author name. Where possible that page should include further information about the author and a link to their home page. | partially followed | EuDML site |
| 23 | Consider making journals a faceted navigation filter on sets of results. | followed | EuDML site |
| 24 | Consider integrating with CiteSeer for citations. | not followed | EuDML site |
| 25 | Where using standard classifications include a link to the standards body. Link the written description of the classification as well as the classification code. | followed | EuDML site |
| 26 | Careful consideration needs to be given to keywords. In order to establish trust in keywords a relatively complex user trust system may be required. We may need to question whether we will get a quantity of responses based upon expected user volumes that will make this effective. Otherwise we may consider automatic keyword generation algorithms. | followed | EuDML site |
| 27 | If we can find effective ways to associate keywords with papers we should consider using them as a faceted navigation filter within results pages, i.e. display high-frequency keywords within the overall result set which can be switched on as optional search filters (see delicious.com). | not followed | EuDML site |
| 28 | Regardless of whether we implement a save for later feature or not, result pages should always include a full GET query string in the URI so that pages of results can be bookmarked and emailed. | followed | EuDML site |
| 29 | Ideally authors should be normalised, modeled and have synonyms (including foreign accents etc.), making Author Searching less hit and miss. | partly followed | EuDML site |
| 30 | This would also enable us to include a mini-profile page. This can list all of the author's papers on the website and link to author's home page where known. | partly followed | EuDML site |