Vittorio Coti Zelati
An Update on bdim the Italian Digital Mathematical Library

In: Petr Sojka and Thierry Bouche (eds.): Towards a Digital Mathematics Library. Bertinoro, Italy, July 20-21st, 2011. Masaryk University Press, Brno, Czech Republic, 2011. pp. 15--18.

Persistent URL: http://dml.cz/dmlcz/702598

An Update on bdim
the Italian Digital Mathematical Library

Vittorio Coti Zelati

Dipartimento di Matematica e Applicazioni
Università degli Studi di Napoli "Federico II"
via Cintia, M.S. Angelo, 801 26 Napoli, Italy

Abstract. bdim has been in operation since the fall of 2010 and has been slowly growing in the last year. I will report here: 1) on what is new in bdim with respect to the first presentation in DML 2010; 2) on some of the technical aspects of our implementation; 3) on the projects for the near future; 4) on some of the issues related to possible integration of bdim with EuDML.

Keywords: bdim, digital libraries, metadata, mathematics

1 The State of bdim

The *Biblioteca Digitale Italiana di Matematica*, bdim, is a joint project of SIMAI (Società Italiana di Matematica Applicata e Industriale) and UMI (Unione Matematica Italiana) and has been in operation since the fall of 2010. The goal is to offer to the many Italian mathematics journals a common repository.

Up to now it contains part of the journal *Bollettino dell'Unione Matematica Italiana*, more precisely all the third series (years 1946–1967) and part of eight series (years 1998–2004).

The collection can be accessed at the web site `http://www.bdim.eu`.

Since the presentation of the project at DML 2010 (see [5]), bdim has got a stable url, and started operation. The collection has grown, in particular we have added part of the eight series of BUMI. We have at the moment 1,670 articles in the collection.

We have a new search engine, Lucene based, through which a full text search has been implemented. We have added MathJax to display math, and we will soon offer MathML as an option.

The collection has also been integrated in the mini-DML project.

2 The bdim Project: Some Technical Aspects

We started our project with the aim to make it compatible both with the international effort to produce a Digital Mathematics Library and with the Italian project of the Biblioteca Digitale Italiana (see [1]). To this end we have followed

– the international guidelines contained in *Some Best Practices for Retrodigitization*, see [4]

Petr Sojka, Thierry Bouche (editors): DML 2011, Towards a Digital Mathematics Library, pp. 15–18.
 ISBN 978-80-210-5542-1

- indications from NUMDAM see [3], which has been very helpful in starting and developing our project;
- the Italian guidelines for the Biblioteca Digitale Italiana, see [6] (in Italian), with their required set of metadata (in MAG format).

For these purpose we have decided to describe the metadata related to objects in our repository (where for us objects are not only articles but also issues, volumes and journals) with XML files in a particular format which we have developed, see the schema file `http://www.bdim.eu/schema/bdim.xsd`.

Examples of XML files describing an article and the corresponding issue can be found at `http://www.bdim.eu:8080/fedora/get/bdim:BUMI_2004_8_7B_1_23_0/FILE_BDIM`, and `http://www.bdim.eu:8080/fedora/get/bdim:BUMI_2004_8_7B_1/FILE_BDIM`

We have collected in these files all the information that we have gathered, even that which is not meant to be distributed, like the references (in BibTex format) which we have downloaded from MathSciNet while looking for matches for the bibliographies.

For articles, the XML file also gives information (size, format, dimension, md5, location) on the files associated to the given article (at least the PDF files, the DjVu file and the XML file containing the OCR of the article). For issues, volumes and journals the XML file also gives additional information about their internal structure (the articles that make up an issue, etc.). This is required by the standards of the Biblioteca Digitale Italiana.

We store and disseminate our collection using a FedoraCommons repository (see [2]). FedoraCommons uses a "compound digital object" design which aggregates one or more content items into the same digital object. Following this approach we created a Fedora object for each article, issue, volume and journal. Each Fedora object aggregates different contents (datastreams in the FedoraCommons language): the XML file describing the object, the relations involving the object itself and, for articles, their pdf and djvu files.

The repository `http://www.bdim.eu:8080/fedora/` runs on a dedicated tomcat web server. One can get the different datastreams of an object from the repository as follows: `http://www.bdim.eu:8080/fedora/get/bdim:BUMI_2004_8_7B_1_23_0/BUMI_2004_8_7B_1_23_0.pdf`.

The FedoraCommons repository is quite flexible and gives us complete control on access to the single objects and datastreams.

Moreover FedoraCommons provides an OAI server, which has been implemented at the address `http://www.bdim.eu:8080/oaiprovider/?verb=Identify`, and which disseminates our objects in oai_dc and minidml formats.

There is also a search engine distributed with FedoraCommons, based on Lucene, which indexes the objects in our repository. Using this, we have implemented the advanced search at the address `http://www.bdim.eu/ricerca`, a highlighting mechanism for hits and different methods to sort them (by relevance or "score", author and year).

We have set up our web site in such a way that the user does not have to interact with the FedoraCommons repository directly. The web site

at http://www.bdim.eu, written in PHP, fetches the relevant data from the repository, transforms it via XSL an builds the web page.

3 The Projects for the Future

Our aim is to include in our collection all the Italian mathematical journals willing to do so. Several have already expressed their desire to join bdim:

- *Rendiconti di Trieste* (since 1969, 11,000 pages),
- *Le Matematiche, Catania* (since 1945 23,000 pages),
- *Rivista di Matematica, Parma* (23,000 pages),
- *Ricerche di Matematica, Napoli* (up to 2005),
- *Rendiconti Lincei* (up to 2004, 80,000 pages)
- *Rendiconti dell'Accademia delle Scienze Fisiche e Matematiche, Napoli* (since 1862, 40,000).

Unfortunately, due to lack of funds, we have not yet been able to enlarge our collection. But we have been applying for funding to several places and we hope to get some funds in the near future.

We have also started a project for digitization of collected works of Italian mathematicians. We are working on the collected works of Salvatore Pincherle (1853–1936), Italian mathematician, first president of the Unione Matematica Italiana. This project presents several interesting aspects, in particular related to copyright issues and to the fact that we will not be dealing with journal articles only. Also, part of the relevant material has already been digitized within other digitization projects, so we have to decide how to proceed in these cases.

4 Integration with EuDML

The next step we would like to take is to integrate our collection in EuDML. We have been looking at the standard for data exchange, and we believe that our metadata could easily be cast in the required format. We are very interested in the EuDML enhancer toolset and we would like to share ideas and methods concerning the tools we have developed in our project. In particular, our expertise in the area of MathML is very limited and we could certainly benefit from the international cooperation.

References

1. *Biblioteca Digitale Italiana*, http://www.bibliotecadigitaleitaliana.it/.
2. *FedoraCommons*, http://www.fedora-commons.org/.
3. NUMDAM, http://www.numdam.org/.
4. Committee on Electronic Information and Communication (CEIC) of the International Mathematical Union (IMU), *Some best practices for retrodigitization*, Endorsed on August 20, 2006 at Santiago de Compostela by the General Assembly of the International Mathematical Union. http://www.mathunion.org/CEIC/Publications/retro_bestpractices.pdf.

5. V. Coti Zelati, *bdim: the Italian Digital Mathematical Library*, Towards a Digital Mathematics Library. Paris, France, July 7–8, 2010 (Brno, Czech Republic), (P. Sojka, ed.), Masaryk University Press, 2010, pp. 79–81, http://dml.cz/dmlcz/702576.
6. Istituto Centrale per il Catalogo Unico, *Standard mag – versione 2.0.1*, http://www.iccu.sbn.it/opencms/opencms/it/main/standard/metadati/pagina_267.html.