Digitization Projects in Spain

E. Macías-Virgós · R. de la Viesca

Abstract   We present the current status of the Spanish projects for the digitization of the mathematical literature, as well as some related topics.

1 Introduction

The situation of mathematical research in Spain—at least from the organizational point of view—has substantially evolved since the publication of a previous report by the first author in the Proceedings of the DML 2006 workshop held in Aveiro [19]. Moreover, the first stage of the digitization project DML-E ended with the implementation of the portal http://dmle.cindoc.csic.es in June 2007. Our team is now planning new projects and looking for a more stable funding.

In the first part of this paper we give a brief account of the organization of research in Mathematics in Spain and of the existing platforms related to digitization and Open Access. Secondly, we list the Mathematics journals currently published in Spain and we explain the achievements and technical aspects of the DML-E project. Finally several objectives for the future are outlined, including a new repository of doctoral theses and the expected collaboration with Latin America.

E. Macías-Virgós (✉)
Institute of Mathematics, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain
e-mail: quique.macias@usc.es

R. de la Viesca
Publications Department, CSIC, Joaquín Costa, 22, 28002 Madrid, Spain
e-mail: rviesca@cindoc.csic.es

2 Research in Spain

A general description of the mathematical research in Spain can be found in the cited Aveiro report [19]. The present section is a brief summary which underlines the most recent changes.

There are about 350 research groups in Mathematics, which receive on the average 50,000 euros every 3 years. The mean size of these groups is 6 people and each researcher is evaluated every 6 years. The Spanish scientific production in Mathematics has increased extraordinarily in the last 20 years, passing from a 0.3% of the total worldwide published papers to a current 5%. Accordingly, Spain, which entered the group II of the International Mathematical Union (IMU) in 1952, passed to group III in 1986, group IV in 2004 and is now planning to ask for promotion to the top group V.

MICINN There is a new Ministry of Science and Innovation (MICINN) [23] which has been created with the aim to integrate more efficiently the efforts of universities, research centers and private companies. Also, a new "National plan for R&D&I 2008–2011" has been conceived as an instrument for increasing the scientific and technological development. The main novelty of this plan is that it is not structured in thematic areas but around strategic objectives [33].

I-MATH A supplementary financial resource for researchers is the so-called "Consolider" programme, which is running in Mathematics under the name "Ingenio Mathematica I-MATH" [17]. This initiative is currently funding research programmes, workshops, seminars, summer courses and fellowships for post-doc and senior visiting positions. Moreover, the "International Centre for Mathematical Meetings" (CIEM) in Castro Urdiales, Cantabria [4] in the north Cantabric coast is now open to all those groups interested in the organization of meetings and conferences.

IEMATH The "Spanish Institute of Mathematics IEMATH", a network of research centers with four main nodes (Barcelona, Granada, Madrid and Santiago de Compostela), will be a very important structure in the future development of Mathematics in Spain. The Experts Panel in charge of the design of this network wrote a detailed report on the perspectives of mathematical research in Spain [16].

3 Digitization and Open Access

"Digitization" and "Open Access" are lively but rather unorganized activities in Spain. The web page of the Spanish National Library [1] lists around twenty digitization projects, mostly directed to the preservation of the cultural heritage; also the central and regional governments are routinely funding similar initiatives, unfortunately without a clear common guidance.

IEDCYT The "Institute for Documentary Studies on Science and Technology IEDCYT" (formerly CINDOC) [15] is the oldest organization in the field of documentary support for the scientific community. Its mission is to analyze, disseminate and promote the scientific information in all areas of knowledge. It was founded in 1953 as a bibliographic inquiry service and since 1976 maintains a catalogue of references for all papers published by Spanish scientific journals. It is located in Madrid and is an institute of CSIC, the Spanish Council for Scientific Research.

In fact, CSIC, which is the main research organization in Spain outside the universities, is now organizing a digital repository [6].

e-revistas e-revistas [12] is an Open Access platform—the first portal totally digital—that includes most of the open acces scientific electronic journals published in Spain and Latin America, covering 300 journals and up to 75,000 articles. It is under the authorship of the Publications Department of CSIC. In the near future, the Spanish National Library will collaborate in the update and management of e-revistas.

RECOLECTA RECOLECTA [27] is a portal harvesting OAI-PMH Spanish repositories. It has been organized by REBIUN, the national consortium of university libraries [26], with the aim of providing tools for the development of Open Access contents and to ease the visibility of research. It follows the guidelines of the DRIVER project [11], which has established a network of Open Access repositories in order to deliver any form of scientific output, including scientific and technical reports, working papers, preprints, articles and original research data.

DIALNET DIALNET [7] is a portal of bibliographic information supported by the University of La Rioja. It aims at the diffusion of the scientific production and the integration of different bibliographical resources for researchers and librarians. With 350,000 users and covering more than 5,000 journals and 1,800,000 references its success has been enormous.

DIALNET provides a powerful search engine and offers the index of every issue for all journals; sometimes the full text is available. A peculiarity is that each scientific event (Congress, Meeting, Workshop) has a specific page, including the relevant information and the proceedings; and each author has his/her own page too, offering such data as scientific field, affiliation and publications.

The portal is now planning to become more autonomous through a more sustainable model less dependent on a single university.

Revicien and RECYT Other scientific portals in Spain are Revicien [29] which is an on-line platform for the promotion and diffusion of more than fifty Spanish scientific journals (not only in Mathematics) and RECYT [28], which depends on "Fundación Española para la Ciencia y la Tecnología FECYT" [14]. They do not give access to the full text of the papers, but provide just links and information about the journals.

The contents of Revicien have been explained in [19]. On the other hand, RECYT is an acronym for "Spanish Journals in Science and Technology". The most innovative aspect of RECYT is that it offers support to the editors of journals—provided they fulfill some quality requirements, which include the existence of an Editorial Committee, a regular periodicity, the selection of papers by peer review and a minimum international impact of the journal. In exchange, RECYT offers advice about such topics as: instructions for authors and referees, protocols and check lists, composition of committees, how to increase impact and so on.

4 Spanish Journals

The initial design of the DML-E project included a number of journals, which was further increased due to the incorporation of older series of current journals and as soon as editors were aware of the success of the digitization effort. In Table 1 we have listed the main research journals in our area. Most of them are included in the DML-E digitization project.

The large majority of the Spanish scientific journals are published by university departments or learned societies, but there is an increasing presence of commercial publishers. Six mathematics journals are now included in the *Journal Citation Reports* [18]: *Archives of Computational Methods in Engineering*, *Revista Matemática Iberoamericana*, *TEST*, *Publicacions Matemàtiques*, *Collectanea Mathematica* and *Revista Matemática Complutense*.

5 DML-E

In the "World Mathematical Year 2000" only a few isolated Spanish universities were digitizing some of their collections in Mathematics. The situation changed dramatically after the participation of Spain in the preparation of the DML-EU proposals done under the auspices of the European Mathematical Society. Indeed, although European funding was never attained (until now) it became clear that it would not cover digitization costs, but only coordination tasks, definition of standards or organizational criteria for a long-term archiving. National funds were then needed for the physical process of digitization, metadata capture, web page access and links to databases.

Table 1 Main Spanish journals in mathematics

Journal	Publisher	Year[a]
Appl. Gen. Topol.	Univ. Politécnica Valencia	
Arch. Comput. Methods Eng.	Int. Center Num. Methods Eng., Springer	
Bol. Soc. Estad. Investig. Oper.[a]	Soc. Estad. Investig. Oper.	2000–
Butl. Soc. Catalana Mat.	Soc. Catalana Mat.	
Collect. Math.[a]	Univ. Barcelona	1970–
Disertaciones del SMF[a]	Univ. Nacional Educ. Distancia	1980–
Extr. Math.[a]	Univ. Extremadura	1980–
Gac. R. Soc. Mat. Esp.	R. Soc. Mat. Española	
Matemàtiques	Univ. València	
Mathware Soft Comput.[a]	Eur. Soc. Fuzzy Logic Tech.	1990–
Pub. Mat., Barc.[a]	Univ. Autònoma Barcelona	1980–
Qual. Theory Dyn. Syst.[a]	Birkhäuser	1999–
RACSAM[a]	R. Acad. Cienc. Exactas Fís. Nat.	2000–
Rev. Mat. Complut.[a]	Univ. Complutense Madrid	1990–
Rev. Mat. Iberoam.[a]	R. Soc. Mat. Española	1980–
SORT[a]	Inst. Estad. Catalunya	2000–
TEST	Soc. Estad. Investig. Oper., Springer	
Top	Soc. Estad. Investig. Oper., Springer	

[a] Covered by DML-E

The Spanish project DML-E started in 2004 as a initiative of the Royal Spanish Mathematical Society (RSME) and the Catalan Society of Mathematics (SCM). It was initially funded by the Ministry of Education and Science by means of a so-called "Special Action 2005-07". The main partners of the project were IEDCYT [15] and the Spanish IMU Committee (Comité Español de Matemáticas, CEMAT) [3], where the main Spanish mathematical societies (RSME, SCM, SEMA, SEIO, FESPM, SEHCYT, SEIEM) are represented.

The members of the Electronic Information and Communication subcommittee of CEMAT took in charge the mathematical aspects of the project. For the period 2007-2010 they are Antonio Alonso (URJC, Madrid), Ernesto Aranda (UCLM, Ciudad Real), Manuel González Villa (UCM, Madrid), Rafael de la Llave (Austin, Texas), Josep Masdemont (UPC, Barcelona) and Enrique Macías Virgós (chair, USC, Santiago de Compostela). Former members were Jaume Amorós (UPC, Barcelona) and Juan Luis Varona (ULR, Logroño).

From IEDCYT the participants in the project were Rosa de la Viesca (project leader), Elena Fernández and Pablo de Castro. They brought their expertise as documentalists and informaticians. The physical scanning of the journals was contracted with an external enterprise.

The collaboration of researchers, librarians, journal editors, learned societies led finally to the DML-E portal at http://dmle.cindoc.csic.es/ [8]. On 13 June 2007 the portal was launched, with an official presentation where several authorities from the Ministry of Education and Science and the mathematical societies were present. Later, the portal was presented at several places, including the University Complutense of Madrid (2 Feb 2007), the congress ICHFM07 at Zaragoza (13 July 2007) and the meetings of DOCUMAT at Madrid (26 Oct 2007) and Logroño (17 Oct 2008). Several newspapers and radio programmes showed their interest in the project and published news and interviews.

Currently we are receiving support from "Ingenio Mathematica I-MATH" and from CEMAT with funds provided by the Ministry of Science and Innovation, but we expect to obtain a more stable funding.

6 Technical Aspects

The DML-E portal [8] manages 16 journals, 4,000 papers and 70,000 pages. All issues since 1980 have been digitized, while for some journals this period extends back to 1940.

Access DML-E offers access to the full text of the papers, although most journals apply a so-called embargo period or moving wall, which reserves to subscribers the last few years (usually five).

Scanning For scanning, the project followed the standards recommended by IMU [5]. Each page is kept in a TIFF file (bitonal, lossless compression, 600 dpi) named as `Journal_year_vol_num_art_page.tiff`. A quality control was made with JHOVE (JSTOR/Harvard Object Validation Environment).

Search Facilities There is the possibility of searching into the metadata by journal, title (original, English, Spanish), author or keywords (English, Spanish). Metadata also include the complete bibliographical reference, affiliation of the authors, the UNESCO classification and in most cases a link to the databases *Mathematical Reviews* and *Zentralblatt MATH*.

It must be emphasized that CINDOC had already a very good database of articles published in Spain, so the project took advantage of the work previously done. In particular, for this reason, titles and keywords were translated into Spanish—although most papers (80%) had been written in English—and UNESCO codes were included.

Keywords In some cases translation of keywords supposed an amazing difficulty. Terms like "Betweenness", "Very ampleness" or "Cojection" required consultation to experts in the field and even to the Academy in order to have a correct Spanish version. Sometimes, this resulted in very interesting discussions: for example, the term "Catalecticant", introduced by J. J. Sylvester in the 19th century, comes from "iambicus trimeter catalecticus", which in prosody refers to a line of verse lacking one syllable in the last foot [25].

IPR Issues Accordingly to the Spanish Law, intellectual property rights are preserved 75 years after the death of all the co-authors of a paper. Copyright of the articles included in the DML-E database belong to the publisher. For each journal, the Spanish Mathematics Committee (CEMAT) asked for permissions and agreed with the editors that the rights of the digitized version, when digitization is done with public funds, will pass to the public domain—always respecting the embargo period. The collection is understood to give a service to the research community on a *bona fide* basis and its commercial use is not allowed. In any case, if an author feels that his/her rights have not been respected he/she is invited to contact the administrators of the portal (this has never happened until now).

Old Series The portal also covers the old series of some journals. They are listed in Table 2. Typically, these issues correspond to years when the Spanish mathematical activity was not so internationalized or did not require professional editorial standards as today. Many of the papers of that age are written in Spanish.

7 Digitization of Doctoral Theses

In 2007 a new project was born from the collaboration of the Royal Spanish Mathematical Society (RSME) [31], with DOCUMAT and DIALNET.

DOCUMAT [10] is a network formed by the majority of the Mathematics libraries of the Spanish universities and research centers. Its activities started in 1989 and DOCUMAT has been a pioneer in the implementation of cooperative efforts such as elaborate collective catalogues, inter-library loan or consortia. Its members meet every year and also organize seminars and workshops.

DIALNET is the portal that we presented in page 245. Although earlier committed to Humanities, they were later seeking for new partners, thus contacting mathematicians through DOCUMAT and the Bibliographical Committee

Table 2 Old series of current Spanish journals, covered by DML-E

Journal	Years	Publisher
Gac. Mat.	1949–1988	R. Soc. Mat. Española
Historia de la Matemática	1980–1998	R. Acad. Cienc. Exactas Fís. Nat.
Qüestiió	1980–2002	Inst. Estad. Catalunya
Rev. Mat. Hisp.-Am.	1919–1982	R. Soc. Mat. Española, Consejo Sup. Investig. Científ.
Rev. Mat. Univ. Complutense Madr.	1988–1997	Univ. Complutense Madrid
Rev. R. Acad. Cienc. Exactas Fís. Nat.	1980–2000	R. Acad. Cienc. Exactas Fís. Nat.
Stochastica	1975–1992	Univ. Barcelona, Univ. Politéctica Catalunya
Trab. Estad.	1986–1991	Soc. Estad. Investig. Oper.
Trab. Estad. Invest. Oper.	1980–1985	Soc. Estad. Investig. Oper.
Trab. Invest. Oper.	1986–1992	Soc. Estad. Investig. Oper.

of RSME. The people in charge of this project are Marta Magriñá, Joaquín León, José M. Pérez Izquierdo and David Martín, among many others.

The common project issued from this collaboration is called "Digital Access to Doctoral Thesis and Scientific Documentation in Mathematics". It is funded by the Ministry of Science and Innovation, the "I-MATH" programme and several scientific organizations and is officially supported by the Conference of Deans and Directors of Mathematics [2]. A beta version of this repository of Ph. D. theses is available since October 2008 at http://documat.unirioja.es. It satisfies all the standards for Open Access and interoperability (WCAG 1.0, XHTML, OAI-MPH) and is multilingual. Currently it gives access to the full text of 500 doctoral dissertations.

It should be noted that since 1976, there is an official database, called TESEO [32] which systematically receives information on all Ph.D. dissertations defended in Spain. Each doctoral thesis is unambiguously identified by a code, while the data (author, title, area of knowledge, jury) come directly from the universities.

DIALNET provides links not only to theses, but to about 70,000 bibliographical items related to the mathematical production of Spanish authors (journals, books, preprints and theses) and has collected metadata directly from different universities. Apart from the repository, the activities of this project include training meetings for librarians, workshops and other initiatives—as the translation to Spanish of all the terms in the Mathematics Subject Classification MSC 2000, which is almost finished.

8 The Future

We are now planning the continuation of the digitization project in three directions: enlarging the number of journals and years, attracting journals from Latin America and improving some technical aspects.

OCR One important issue is OCR (Optical Character Recognition) that will allow searches in the full text and not only in the metadata. The experience of other projects as NUMDAM [24] or DML-CZ [9] will serve us as a guide.

MSC 2000 With the cooperation of librarians and mathematicians from different universities, DOCUMAT-DIALNET will begin to classify all doctoral theses in Mathematics by their MSC (Mathematics Subject Classification) code. One very useful aspect of this project is that every library can access on-line the database in order to incorporate

Table 3 Some mathematical journals in Latin America

Argentina	*Revista de la Unión Matemática Argentina*
Brazil	*Bulletin of the Brazilian Mathematical Society*
Chili	*Notas de la Sociedad de Matemática de Chile*
Colombia	*Revista Colombiana de Matemáticas*
Cuba	*Boletín de la Sociedad Cubana de Matemática y Computación*
Mexico	*Aportaciones Matemáticas, Sociedad Matemática Mexicana*
Peru	*Pro-Mathematica*
Venezuela	*Boletín de la Asociación Matemática Venezolana*

data—for example, the text of a thesis, information about an author or a link to *Zentralblatt MATH* or *Mathematical Reviews*.

Additions Several journals—e.g. *Applied General Topology* (Univ. Politécnica de Valencia), *Gaceta de la RSME* or *Matemàtiques* (Univ. de Valencia) remained outside the first phase of the project, for different reasons.

Two other journals with an historical interest for the Spanish mathematical community are: *Publicaciones del Seminario Matemático García Galdeano* and *Arquímedes*.

Finally, two journals (*TEST* and *Top*), formerly edited by SEIO, signed a contract with a commercial publisher (Springer) and are not yet in the project, although our firm expectations are to get their permissions to accommodate them.

We should like to include several journals from Latin America, like those indicated in Table 3. We hope that the existing "Network of Latin American Mathematical Organizations ROLMa" [30] will play a fundamental role in this process.

Digitization of old books is another possibility. Spanish Mathematics at the turn of 19th century produced some original books of interest from the point of view of research—not only historical—as well as several very good translations to Spanish of famous foreign mathematical works.

9 Final Remarks

From the beginning of the project DML-E we have been regularly reporting about it in different papers (see for example [19–22]). But these reports do not reflect well the real difficulties that such a project must afford: too many bureaucratic forms to fill, the lack of support from the European community, the political changes which make useless many contacts and negotiations. Despite the fact that the situation has really improved during the last years, several problems remain unsolved and undermine a long term solution. To achieve a sustainable model for a portal which will satisfy the expectations of both researchers and journals seems to us the most difficult one. Initiatives such as the "preparatory meeting" organized by the European Science Foundation and the European Mathematical Society in Santiago de Compostela, Spain, in March 2009 [13] have given us the opportunity to discuss specific proposals to solve those problems and ease the way towards a European Virtual Library in Mathematics.

References

1. BNE. http://www.bne.es/BDH/coleccionesdigi.htm
2. CDM. http://www.usc.es/mate/cdm/
3. CEMAT. http://www.ce-mat.org
4. CIEM. http://www.ciem.unican.es/
5. Committee on Electronic Information and Communication CEIC. Some Best Practices for Retrodigitization. Endorsed on 20 August 2006 at Santiago de Compostela by the General Assembly of the International Mathematical Union IMU. http://www.ceic.math.ca/Publications/retro_bestpractices.pdf

 6. CSIC Digital Repository. http://digital.csic.es
 7. DIALNET. http://dialnet.unirioja.es/
 8. DML-E. http://dmle.cindoc.csic.es/
 9. DML-CZ. http://dml.cz/
10. DOCUMAT. http://wzar.unizar.es/documat/
11. DRIVER. http://www.driver-support.eu/
12. e-revistas. http://www.erevistas.csic.es
13. ESF preparatory meeting "Towards a European Virtual Library in Mathematics", Santiago de Compostela, Spain, 13–14 March 2009. http://xtsunxet.usc.es/macias/ESFMATHS
14. FECYT. http://www.fecyt.es
15. IEDCYT. http://www.cindoc.csic.es/eng/
16. IEMATH. http://www.iemath.es/IEMathVersionResumidaEnglishDefNew.pdf
17. I-MATH. http://www.i-math.org/
18. ISI Thomson Reuters JCR. http://www.thomsonreuters.com/products_services/scientific/Journal_Citation_Reports
19. Macías-Virgós, E.: The DML-E Digitization Project and Related Topics. The Spanish initiative. In: Borwein, J.M., Rocha, E.M., Rodrigues, J.F. (eds.) Communicating Mathematics in the Digital Era. Selected Papers Based on the Presentations at the Meeting CMDE 2006, Aveiro, Portugal, 15–18 August 2006, pp. 87–96. A. K. Peters (2008)
20. Macías Virgós, E.: El mundo matemático digital: el proyecto WDML (World Digital Mathematics Library). Arbor: Ciencia, pensamiento y cultura, vol. CLXXXIII, No 725 mayo-junio (Volumen dedicado a: Fronteras matemáticas del siglo XXI), pp. 433–444 (2007). http://arbor.revistas.csic.es/index.php/arbor/article/view/116
21. Macias-Virgós, E.: Some digitization initiatives in Spain. In: Becker, H., et al. (eds.) New Developments in Electronic Publishing of Mathematics. ECM 4 Satellite Conference on Electronic Publishing at KTH, Stockholm, Sweden, 25–27 June 2004. FIZ Karlsruhe, pp. 137–142 (2005)
22. Macías, E.: La publicación electrónica de la investigación matemática. In: Actas de la primera jornada DOCUMAT: La difusión de la documentación científica matemática. Madrid 2007. http://dialnet.unirioja.es/servlet/libro?codigo=274003
23. MICINN. http://web.micinn.es/contenido.asp?menu1=&menu2=&menu3=&dir=01_Portada&idioma=en
24. NUMDAM. http://www.numdam.org/
25. Parshall, K.H.: James Joseph Sylvester: Life and Work in Letters, Letter from J. J. Sylvester to T. A. Hirst, 19 December 1862. Oxford University Press, pp. 111–113 (1998)
26. REBIUN. http://www.rebiun.org/
27. RECOLECTA. http://www.recolecta.net/buscador/
28. RECYT. http://recyt.fecyt.es/
29. REVICIEN. http://www.revicien.net/
30. ROLMa. http://www.colmatelat.ehu.es
31. RSME. http://www.rsme.es
32. TESEO. https://www.micinn.es/teseo/
33. The National R&D&I Plan. http://web.micinn.es/contenido.asp?menu1=&menu2=&menu3=&dir=03_Plan_IDI&idioma=en