David Ruddy

The Evolving Digital Mathematics Network

In: Petr Sojka (ed.): Towards a Digital Mathematics Library. Grand Bend, Ontario, Canada, July 8-9th, 2009. Masaryk University Press, Brno, 2009. pp. 3--16.

Persistent URL: `http://dml.cz/dmlcz/702554`

The Evolving Digital Mathematics Network

David Ruddy

Project Euclid, Cornell University Library
107 D Olin Library, Ithaca, New York, 14853, USA
Email: `dwr4@cornell.edu`

Abstract. The grand vision of a Digital Mathematics Library (DML), coordinated by a group of institutions that establish polices and practices regarding digitization, management, access, and preservation, has not come to pass. The project encountered two related problems: it was overly ambitious, and the approach to realizing it confused local and community responsibilities. While the vision called for a network of distributed, interoperable repositories, we approached and planned the project as if we were building a single, unified library. After a discussion of this, a brief status report on Project Euclid is given. This is followed by a description of activities that local repositories and the mathematics community can engage in to encourage the development of network services.

Key words: digital mathematics library, network development, repositories, interoperability, Project Euclid

1 Introduction

The history of efforts to create a Digital Mathematics Library (DML) has been usefully summarized by Thierry Bouche [1,2]. In 2001 and 2002, motivated by the importance and use of past mathematical literature for current mathematical work and the attraction of a growing body of web-accessible literature, a group of mathematicians, librarians, and publishers articulated a vision for a unified online collection:

> *In light of mathematicians' reliance on their discipline's rich published heritage and the key role of mathematics in enabling other scientific disciplines, the Digital Mathematics Library strives to make the entirety of past mathematics scholarship available online, at reasonable cost, in the form of an authoritative and enduring digital collection, developed and curated by a network of institutions.* (Digital Mathematics Library [3], a one-year, NSF-supported planning project coordinated by Cornell University Library. See also [4,5,6].)

Today, the one thing about the DML on which everyone agrees is that this vision has not been realized. There is no single cause for this. Certainly the fact that the project was unfunded did not help. Further, what the DML proposed

Petr Sojka (editor): DML 2009, Towards a Digital Mathematics Library, pp. 3–16.
ISBN 978-80-210-4781-5

was overly ambitious, as many at the time pointed out. It called for solutions to problems that were only partially understood and that were far larger than this particular effort. It also called for a level of institutional cooperation that was operationally and politically beyond our capabilities, especially in light of the lack of funding. In essence, the technical and social infrastructure for a project of this scope was not in place. And yet another cause, we suggest here, was the overall approach taken to carrying out the project, revealed in the vision statement phrase “developed and curated” and in the use of “library” as an analogy for the effort’s result. While the notion that the DML could be implemented using a centrally organized, coordinated, and planned approach was perhaps merely wishful thinking, to the degree that it confused local and community responsibilities it was counterproductive.

That said, it is not at all clear that a different approach would have worked any better, given the scope of the project. A different approach may, however, have led to a longer-term perspective on the entire endeavor, more realistic expectations, and thus less frustration. In any event, coming to terms with this now will reveal some useful lessons and constructive ways to think about future work.

1.1 Libraries and Networks

The scope of the DML was large: the past mathematical literature in its “entirety,” which has been estimated at approximately 50 million pages [4]. This would indeed be an enormous collection of specialized material and it is understandable how the analogy to a library came about. At the time, there was a lot of talk about “digital libraries.” The analogy, however, went beyond the notion of a body of literature collected together, even virtually. The early discussions and planning documents of the DML reveal an approach strikingly familiar to anyone involved in library project management [3,4,5]. One develops a collection policy, selection criteria, standards for production and processing, and operational workflows. One builds centralized registries, indices, and access systems. Above all, one plans, coordinates, forms committees, and attends numerous meetings.

Such an approach may well be necessary within an organization that requires long-term stability and uniformity of practice to carry out its mission of ensuring access to materials far into the future. In other words, within a library or archives. Yet the DML was a large, loosely organized and diverse group of institutions, societies, and publishers with differing cultures and competing interests. Organizing the DML effort as if it were an internal library project may have seemed sensible at the time, but it produced limited results. Working groups were formed to address various components of the project: content selection, technical standards, metadata, intellectual property rights, archiving, and business modeling. But as the discussions deepened, many of the groups stalled. The issues were huge. Some of the topics were only beginning to be understood and engaged by the wider community (e.g., metadata, archiving). Some were intractably complex, involving issues well beyond the domain of mathematics (rights, economic models). This is where the approach bumped up against the ambitious nature of the project, and progress declined.

In some sense, there was a mismatch between the DML vision and our approach to implementing that vision. For what was really desired was a way for mathematicians to discover and access digitized mathematical literature across distributed individual systems, effortlessly navigating from one work to others—in essence, a network within which individual systems could seamlessly interoperate [4,7]. The vision called for a network, and yet we set out to build a library. We thought we needed to set standards and establish common practices, and to ensure that all participants conformed to a particular set of well-defined requirements. But this approach was doomed, for decisions regarding digitization methods, file formats, metadata, internal naming conventions, business models... these are all local responsibilities, and attempts to impose requirements from outside will, and did, fail.[1]

Given the ambitious nature of the DML vision and the large unresolved problems it faced, it is arguable whether this approach, or the adoption of the library analogy per se, actually obstructed or delayed progress much.[2] And in some respects, progress never really stalled on the DML. The work that was actually feasible, building individual repositories (libraries) of digitized mathematical literature, has proceeded steadily. But if we now accept that "library" was only an analogy for what we imagined (and perhaps not the best or most useful analogy), and that what we really want to see take shape is a large network of interoperating repositories, does this help us move forward?

To answer this, we can ask another question: how do we build a network of online mathematical literature? In some sense, we do not "build" it, but we adopt technologies and engage in activities that encourage and allow it. Two broad areas of work seem most appropriate today. One is the continued development, expansion, and maintenance of local repositories of mathematical content in ways that will support the growth of networks. This requires viewing one's local repository as a potential node in a network and asking how it can best interoperate with other nodes. The second effort is to work collaboratively as a community to design and adopt cross-repository tools, standards, and agreements that act as gateways between nodes, permitting the creation of network enabled services. This effort, creating "plugs and sockets," contributes to the infrastructure development necessary to allow networks to form.[3]

Below, some specific examples are provided of potential work in both these areas. Local repository practices that support interoperable network services are described, and then some possible community efforts are listed. First, however,

---

[1] In 2003, as the initial NSF supported effort was disbanding, Allyn Jackson reported that "Those running existing retrodigitization projects want to continue their work as they see fit rather than follow rules set by a larger authority" [6].

[2] John Ewing expressed frustration in 2003 with the dead-ends the project was encountering and attributed some of them to the library analogy [8]. The causes, though, as he recognized, were deeper.

[3] Much of the thinking here about networks and infrastructure is influenced by [9].

since Project Euclid [10] is an example of a repository of mathematical literature, and since the author has some familiarity with it, it seems appropriate to provide a brief status report on that project. For the most part, only those topics about Project Euclid that have to do with DML work or the later discussion are mentioned here.

2 Project Euclid

2.1 Background

In 1999 and early 2000, Cornell University Library conducted discussions with faculty, publishers, and librarians that led to the broad outlines of Project Euclid. The goal of the project was to provide a not-for-profit electronic publishing alternative for small and independent publishers of pure and applied mathematics and statistics, and thereby to promote affordable scholarship. The impetus behind Euclid was not only the rising cost of serial literature but a related yet larger problem academia faced, that of losing ownership and control of its intellectual output [11,12].

The delivery platform, early organization, business models, and operational workflows of Project Euclid were initially worked out during 2000–2002, a development phase supported by The Andrew W. Mellon Foundation. During this phase, Cornell University Library worked closely with five partner publishers, especially Duke University Press, to refine technical and business operations and procedures. In early 2003 Euclid was officially launched, with the recent issues of 19 journals. Since then the project has grown steadily and today hosts 56 journal titles as well as monographs and conference proceedings.

2.2 Current Status

Although originally conceived to deliver current published content, in 2002 Euclid managed the digitization of back issues of the *Michigan Mathematical Journal* and the preparation of the content for online delivery. Since then the project has digitized over a half million pages for Euclid publishers and prepared another 200,000 pages from existing files. The earliest serial literature in Euclid currently dates from 1891. Holdings for 38 of the 56 journals delivered by Euclid today begin with volume 1, and Euclid has complete coverage of 35 of these titles.

The holdings of Project Euclid, as of June 1, 2009, are as follows:

- 1,253,299 total pages
- 855,738 pages open access (68% of total)
- 96,821 total journal articles
- 67,925 journal articles open access (70% of total)
- 56 journals (all but 5 are active)[4]

---

[4] "Active" means that current content is regularly submitted by publishers. The journals are published worldwide: 33 in the US, 10 in Japan, 8 in Europe, 2 in the UK, 2 in Egypt, and 1 in Iran.

- 5 monographic series (99 books, 1,123 chapters)
- 1 conference series (23 volumes, 651 proceedings articles)
- 32 publishers regularly submitting content [13]

2.3 Additional Content in 2009

Thus far in 2009 four new journals have joined Euclid, and we expect to add another four to five before the end of the year. We will add the entire back run of two of the recent titles (*Illinois Journal of Mathematics* and *Journal of Mathematics of Kyoto University*), and the articles in these approximately 65,000 pages will push the total Euclid journal article count above 100K. In the fall of 2009, we will fold into Euclid the *Cornell Historical Mathematics Monographs*, a collection of approximately 570 public domain books currently available through a separate Cornell University Library system [14]. This popular collection will have enhanced chapter-level access (and chapter-level metadata) and contribute another 191,000 pages of mathematical content to Euclid. The monographs will also include links to Amazon, where individual print copies can be purchased at reasonable prices. By the end of 2009, we anticipate the total number of pages accessible through Euclid to surpass 1.5 million.

2.4 Sustainability

While initial capital investment for Project Euclid came from the Mellon Foundation, it was recognized from the beginning that the project would need to be self-supporting. Although the Library wished to provide as much open content as possible, there was an acknowledgment that the project, in order to survive, would need to permit some access restrictions based on subscription, and that both Euclid and publishers would sell access to content under various arrangements. A flexible business model was developed to accommodate the needs of a range of publishers with various business goals and requirements. By 2005 Euclid was covering its costs.

In mid-2008, Cornell University Library and Duke University Press entered into a formal agreement to co-operate Project Euclid [15]. This partnership, between organizations with shared values and mission, has expanded the project's capabilities in critical areas of acquisitions, marketing, order fulfillment, and customer relations. Management responsibilities have been divided along lines of expertise: Cornell operates the technical infrastructure, while Duke supports Euclid's business operations. The partnership will allow Euclid to offer improved services to publishers, subscribing institutions, and users, and thereby better support long-term stability.

2.5 Summary

Project Euclid has demonstrated a successful model for a not-for-profit repository of formally published mathematical literature. It has attracted publishers worldwide who share the project's values of a non-commercial, academic-based effort to keep scholarship affordable. Euclid and its publishers have developed an economic model that keeps prices reasonable, sustains growth, and at the same time provides access to a significant amount of literature without any restrictions. Today, over two-thirds of the content in Euclid is openly accessible.

3 Repository Practices

In what ways can a single repository of mathematical literature support the development of network services? Below is a description of repository practices that would allow other repositories, or independent service providers, to develop network services. Although presented here as a set of recommendations, these practices are described at a general level. How, specifically, any of these recommendations might be implemented needs to be worked out at a local level, where different circumstances will require different solutions. They are, rightfully, the responsibility of local repositories. Further, in an open, loosely coupled network of the type advocated here, there is no need for explicit and formal relationships among participants. Thus the recommendations here should not be viewed as a set of potential requirements. A successful network will create its own incentives for participation.

Digitization. A primary responsibility of most repositories will be to acquire more digital content and to do so in a way that builds as much value into the resulting image files as possible. Value here means 1) that it is unlikely that the materials will ever need to be re-digitized, and 2) that there is a high likelihood that the resulting files will support future processes (re-OCR'ing with new tools, for example; or the creation of new derivative files). Those planning or involved in digitization projects should be aware that digitization and image manipulation techniques will, and should, continue to evolve. The latest research and recommendations need to be consulted.

As an aside, two comments related to digitization are offered here: a caveat and a suggestion.

600 dpi imaging. 600 dpi bitonal TIFF format has been something of a *de facto* scanning standard for black and white journal content. It should be remembered, however, that this recommendation emerged from digital scanning benchmarking guidelines which stressed that all scanning decisions needed to begin with close analysis of the source materials and their meaningful attributes, and specifically with the identification of the smallest significant character in the materials [16,17]. While 600 dpi bitonal has proven adequate for most serial literature, it is possible that higher resolutions or greater bit depth may be required in certain cases. This is especially true for mathematics, where very small font sizes can occur within math expressions. This same methodology applies to all image processing. One should identify examples of the smallest meaningful characters in the content being digitized and then use these to measure and calibrate the effects of image post-processing. Despeckling in particular can be dangerous, changing i to ι in small sub- or superscripts. Deskewing and text block centering, if applied indiscriminately across dissimilar page sizes or source materials, can also create inferior images.

Machine readability. It should also be noted that early scanning specification development, while attuned to “full informational capture,” was essentially geared toward human readability. In the future, highly accurate machine readability will be critical for analysis of large text collections. As optical character recognition is put to more sophisticated tasks, such as processing mathematical notation [18], those involved in such research should inform the community as to whether and how we should adjust scanning specifications and methods.

Identifiers. A repository should create, use, maintain, and in some way publicize persistent identifiers for all objects contained within it. In most cases today, objects in mathematical literature repositories are documents, but in the future they may also be sections or features within documents. Since the navigation of networks is largely facilitated by machines, whether an identifier is “opaque” (containing no semantic information) or “meaningful” is not very important. Opaque identifiers are generally easier to create and manage. In either case, one should avoid using characters that require URL encoding, such as URI reserved characters and characters outside the ASCII set. These characters create complications in the rendering, transfer, and use of identifiers on the WWW. (CrossRef has recently disallowed them in the formation of DOIs [19].) For this reason, one should avoid using SICI identifiers.

A repository should provide a reliable method for resolving identifiers on the WWW (see below). It should publicize this method, so that other systems can systematically build links to the repository.

HTTP Accessible Record Pages. It is extremely useful for every document in a repository to have a record page that can be referenced and openly accessed via HTTP, even if the full text of any particular document may have access restrictions. Because it frustrates users to be directed to a page that cannot be accessed, system builders will tend not to link to a repository that does not provide such a record page to every user. These records should unambiguously identify the document. Ideally, complete metadata is displayed, along with any access restrictions. Such a record page should be the target for the resolution of a document identifier, described above. The use and usefulness of these pages will likely increase in time. It is now possible to embed in these pages structured, machine parseable metadata in formats that various tools or services can utilize.

Exposed Metadata. Those who wish to provide network services require machine-aided methods of discovering what resources (documents) are in a repository. One common way to meet this need is for repositories to allow metadata about their resources to be machine harvesting via a protocol such as OAI-PMH [20]. There are also smaller scale ways of exposing metadata. One can embed metadata within a web page in ways that allow it to be used by applications such as Zotero, a personal library tool [21], or utilized by various social bookmarking services. One could also generate and expose more complex, machine-readable OAI-ORE resource maps, which describe web resources as aggregations of data (author, title, date, identifiers, references, available formats, etc.). These maps are expressed in an RDF serialization format and made available to client software via a process of content negotiation [22]. Such initiatives as OAI-ORE are potential components in the larger semantic web effort [23].

The functional requirement to support the export of any particular metadata format will impose requirements on internal metadata management. As the number of tools and services that can make use of well-structured metadata is rapidly increasing, new formats will continually emerge, creating unanticipated requirements. With respect to metadata, the best preventative approach is to keep internal metadata as finely grained as possible. For example, distinguishing individual authors names, rather than lumping them together in a single author field, is good practice. Taking this further, one may want to distinguish different forms or parts of a single name, such as surname and given name. Such an approach, applied to all metadata elements, provides insurance that one can successfully map to new metadata formats as they emerge.

Repository APIs. A repository that can be queried in useful ways will be an important network node. Others will tend to hook into it in innovative ways, primarily because they can. Simple, open APIs that allow others to perform lookups, retrieve metadata, or other functions, are extremely useful. They promote an open, loosely coupled network that will support innovative interoperable services with minimal coordinated planning.

Archiving and Preservation. Preservation of digital content was only beginning to be understood when the DML was first proposed, and it still involves many uncertainties. Most discussions of the topic today distinguish bitstream from content preservation. The former is sometimes called archiving. Practices that support the preservation of bitstreams are becoming better understood (reliable storage provision and administration, backup, transfer to new media, checksum auditing, file inspection, etc.). What is clear is that bitstream preservation goes well beyond mere storage provision and involves active and effective management practices. Content preservation aims to preserve the functionality of the original object and is a more complex issue. Potential strategies might involve file migration or emulation techniques. While successful content preservation is less well-understood, it also will certainly depend upon effective management of digital content.

Centralized preservation services are beginning to develop and this trend will probably increase (for example, Portico [24]). But such services will not exempt repositories from developing strong local content management practices. For now, this appears to be the foundation upon which preservation strategies will be implemented. Libraries have a key and familiar role to play here. Their mission of providing long-term access has always required an attentive focus on the long-term preservation of materials. Academic research libraries are at the forefront of efforts to understand what it means to be a "trusted digital repository" [25]. While it is not likely that a repository can easily or quickly "solve" the preservation problem, by keeping informed of the latest research and by developing and enacting policies, methods, and technologies in support of active data management, it is at least positioning itself to implement successful preservation strategies and technologies as they emerge.

Sustainability. Since a network of mathematical literature will connect data and content from individual repositories, the persistence and reliability of these repositories is crucial for a comprehensive and dependable network. While larger entities (communities, societies, nations) share certain responsibilities for sustaining persistent access to digital information, much of the responsibility for developing a sustainable organization, and the economic model to support it, will likely fall to individual repositories and their sponsors.

One of the greatest challenges for early digitization efforts has been to shift from a project focus (typically short-term), to a more programmatic approach, one in which repository activities are well-integrated into the internal operations of the hosting organization or institution. This goes beyond the mere identification of funding. Especially in support of sustainable preservation activities, numerous other obstacles exist, including confusion over responsibilities, lack of incentives, and complacency [26]. In other words, sustainability is a complex financial, legal, organizational, and social challenge. Those responsible for a repository need to assess the sustainability of their organizations in an honest and systematic manner, and to plan for the future.

One practical strategy may be to consider consolidation of repositories, or at least collaboration in certain areas of repository support. Since all of the repository activities described above require resources to carry out and maintain, and since operational economies of scale will surely develop, consolidation could offer cost-effective solutions to the problem of sustainability. While inter-institutional relationships present obvious political challenges, mutually beneficial partnerships are possible between organizations that share similar values.

4 Community Work

What follows is a preliminary and admittedly underdeveloped list of work areas where some level of community support will be necessary for us to see progress. These areas involve issues beyond the scope of any single repository and will likely require collaborative effort. While a number of these topics involve technical solutions, most of them are not primarily technical problems.

Digitization Support. A substantial amount of mathematical literature is available online. It has been estimated that perhaps as much as 65% of the entire corpus has already been digitized, almost all of it in the last ten years [27]. While the precise number may be debated, it is clear that progress is being made. Nationally focused efforts seem particularly energetic and successful. (See the reports on national projects in [28] and [29].) Digitization today is more of a financial problem than a technical one, and efforts such as these need to be encouraged and supported. Community effort, at various levels, may be required to find sufficient funding to complete the digitization of formally published mathematical literature. Of course, completion of this work will raise questions of potential further collections of value, such as dissertations and theses, unpublished conference proceedings and other gray literature, and archival materials. The latter will present a new set of problems as well as rich rewards.

Metadata Exchange Format. We need a better way of exchanging rich metadata about digitized mathematical literature. The standard method of open exchange in academia is OAI-PMH, and the default metadata format is simple, unqualified Dublin Core. Simple DC has at least two problems. Perhaps because it is so simple, or because the lack of element qualification is such an annoying restriction, simple DC is often inconsistently applied. The other, related, problem is that simple DC is relatively impoverished as a set of metadata. Even if it adheres to best practice guidelines, there is only so much one can do with it.

Recognizing this, representatives form NUMDAM, University of Göttingen, and Project Euclid have worked sporadically over the last few years on a set of recommendations for simple DC to be used for mathematical literature [30]. Work on this has proceeded slowly. A logical next step is to develop an extension of this format that allows qualification (and thus greater specification of data elements), and which would support higher levels of functionality.

Name Authority. It would be extremely useful to have a way to refer to published mathematicians unambiguously. This would allow systems to collocate works by author, generate authoritative publication lists, and no doubt offer other valuable services. While individual organizations can and have developed such authority files, it would be a tremendous savings in cost and effort to pool resources in support of a common, openly available name authority database. The most useful system would establish unique URIs for every published author, as well as a method for discovering, matching, and disambiguating author names.

Network Use of the MSC. Similarly, the ability to use MSC codes in network provided services would be greatly enhanced if a URI were assigned to every code (uniquely associated also to a MSC version: 2010, 2000, 1991, etc). This is relatively straightforward. The more complex work would be in providing a method to collocate like subjects across the different versions of the MSC. Ideally, one could send an MSC code and a date of publication to a service and receive back a current equivalent code or an appropriate set of related codes.

Math Encoding. Agreement on and adoption of a system-independent, nonproprietary format to encode mathematical expressions would allow numerous advantages in document production and analysis, as well as promote tool building. This obviously involves complex technical and business issues, and there would be costs to making a commitment to one solution or another. But there are also costs to not making a decision. More mathematical scholarship is published every day. If in twenty years all the mathematical content published in those years was accessible to machine analysis, surely no one can dispute the enormous range of possibilities this would allow.

DML Research. Encouraging advanced research into DML related topics, such as the studies collected in [28] and [29], is an important community role that can be carried out in various ways and should be richly rewarding in the future. The very scale of data available to query and manipulate will increasingly alter the research questions asked. Similarly, network related research will lead to progressively innovative ways of exposing relationships between and among works of mathematical scholarship. These efforts will evolve in sophistication over time as our ability to process natural language improves, and as reliable methods of identifying and processing mathematical expressions develop.

Publisher Relations. Commercial publishers control significant portions of the published corpus of mathematical literature. While these publishers have business reasons for restricting access to the full-text of this scholarship, the mathematics community should encourage them to open up access to metadata about that content. These publishers already provide detailed metadata to Google and the large suppliers of e-resource access and management services, such as Serials Solutions and Swets. They do this no doubt because they believe it increases visibility, demand, convenience to customers, and sales. As network services mature, the mathematics community should work to demonstrate similarly positive benefits that increased openness can bring. Likewise, at some point, it is probable that we will consider full-text, in some non-presentational format, as a type of metadata, and develop mechanisms and agreements to share it. Such data will be valuable to discovery systems of all kinds. Commercial publishers will need to be encouraged to participate in these activities by demonstrating to them the business advantages of doing so.

Sustainability. While it is not precisely clear what role the larger community should or needs to play in sustaining a network of mathematical literature, that role will likely occur where issues of sustainability and collaboration intersect. As repositories mature and we learn more precisely what organizational structures, management practices, and technical systems are successful at supporting network interoperability and ensuring long-term preservation and access, we will learn what economies of scale are possible. Where benefits are significant, then organizational collaboration, consolidation, or merger would produce cost savings to the community as a whole. In these cases, larger entities within the community, such as societies or national libraries, could serve a role in encouraging appropriate alliances and partnerships.

5 Conclusion

Networks evolve when sufficient infrastructure is in place to support the formation of cross-system services. The growth of digitized mathematical literature available online has reached a point where network services can begin to develop. There are numerous unanswered questions here, many having to do with ownership and control. The approach suggested here, however, argues against the need for centralized control of such a network. Flexibility is limited when ownership, management, and control is concentrated, and information networks today appear to thrive in flexible environments where requirements for participation in the network are low and the boundaries between participating services can be easily reconfigured. We are creating the infrastructure for a network, not building a library.

Why did we call the DML a “library”? For one, it was a familiar concept and seeking the familiar in new innovations is a human trait. The earliest printers made their books resemble manuscripts. Today, much e-only scholarship is made to appear like the printed page. The ubiquity of PDF on the web, in fact, is a testament to our procedural and psychological investment in the printed page, which has intimations of fixity and is referenceable in ways to which we are accustomed, even when not literally printed. In a similar fashion, when we conceived of a large collection of mathematical literature, we thought of it as a library. But when the goals of the DML are achieved, it will not have much in common with a library, at least as we know them today. It will be a fluid network of resources, richly interconnected in various and fluctuating ways. It will be dynamic and flexible. It will not merely connect an article to each work it cites, but will identify and graph relationships based on all sorts of criteria. As a result, it will allow unforeseen and deeply revealing observations on the evolution of mathematical thought.

References

1. Bouche, T: Toward a Digital Mathematics Library? In: Borwein, J.M., Rocha, E.M., Rodrigues, J.F. (eds.): Communicating Mathematics in the Digital Era, pp. 47–73. A K Peters, Wellesley, Mass. (2008)
2. Bouche, T: Some Thoughts on the Near-Future Digital Mathematics Library. In: Sojka Petr (ed.): DML 2008 – Towards Digital Mathematics Library, pp. 3–15. Masaryk University, Brno (2008)
3. Digital Mathematics Library, Cornell University Library, `http://www.library.cornell.edu/dmlib`
4. Ewing, J.: Twenty Centuries of Mathematics: Digitizing and Disseminating the Past Mathematical Literature. Notices of the AMS 49, 771–777 (2002). Circulated as a white paper in 2001.
5. Digital Mathematics Library: A Proposal to the National Science Foundation Submitted by The Cornell University Library, (2001) `http://www.library.cornell.edu/dmlib/NSFnarr.html`
6. Jackson, A.: The Digital Mathematics Library. Notices of the AMS 50, 918–923 (2003).
7. Committee on Electronic Information and Communication (CEIC) of the International Mathematical Union: Digital Mathematics Library: A Vision for the Future (2005) `http://www.ceic.math.ca/Publications/dml_vision.pdf`
8. Ewing, J.: DML: Moving Forward. (2003) `http://www.ams.org/ewing/dml-moving-forward.pdf`
9. Edwards, P.N., Jackson, S.J., Bowker, G.C., Knobel, C.P.: Understanding Infrastructure: Dynamics, Tensions, and Design. Report of a Workshop on “History & Theory of Infrastructure: Lessons for New Scientific Cyberinfrastructures.” NSF Report (January 2007)
10. Project Euclid, `http://projecteuclid.org`
11. Koltay, Z., Hickerson, H.T.: Project Euclid and the Role of Research Libraries in Scholarly Publishing. Journal of Library Administration 35, no. 1 83–8 (2002) `DOI:10.1300/J111v35n01_06`
12. Thomas, S.E.: Publishing solutions for contemporary scholars: The library as innovator and partner. Publishing Research Quarterly 22, no. 2, 27–37 (2006) `DOI:10.1007/s12109-006-0013-5`
13. Project Euclid, Participating Publishers, `http://projecteuclid.org/Publishers`
14. Cornell Historical Mathematics Monographs, `http://resolver.library.cornell.edu/misc/5618008`
15. Ehling, T, Staib, E.: The Coefficient Partnership: Project Euclid, Cornell University Library, and Duke University Press. Against the Grain 20, no. 6 32–36 (December 2008–January 2009)
16. Kenney, A., Chapman, S.: Tutorial: Digital Resolution Requirements for Replacing Text-Based Material: Methods for Benchmarking Image Quality. Commission on Preservation and Access, Wash. D.C. (1995)
17. Chapman, S., Kenney, A.: Digital Conversion of Research Library Materials: A Case for Full Informational Capture. D-Lib Magazine (October 1996) `http://www.dlib.org/dlib/october96/cornell/10chapman.html`
18. Suzuki, M., Tamari, F., Fukuda, R., Uchida, S., Kanahori, T.: INFTY: an integrated OCR system for mathematical documents. In: Proceedings of the 2003 ACM Symposium on Document Engineering, pp. 95–104. ACM, New York (2003)
19. DOI Name Information and Guidelines, CrossRef (January 22, 2009), `http://crossref.org/02publishers/doi-guidelines.pdf`

20. Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), `http://www.openarchives.org/pmh`
21. Zotero, `http://www.zotero.org`
22. Open Archives Initiative Object Reuse and Exchange (OAI-ORE), `http://www.openarchives.org/ore`
23. W3C Semantic Web Activity, `http://www.w3.org/2001/sw`
24. Portico, `http://www.portico.org`
25. Research Libraries Group: Trusted Digital Repositories: Attributes and Responsibilities (2002) `http://www.oclc.org/programs/ourwork/past/trustedrep`
26. Sustaining the Digital Investment: Issues and Challenges of Economically Sustainable Digital Preservation. Interim Report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access (December 2008) `http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf`
27. Borwein, J.: IMU on the Web #25: Looking Back After a Decade (2008) `http://www.ceic.math.ca/News/IMUonWeb.shtml#CEIC25`
28. Borwein, J.M., Rocha, E.M., Rodrigues, J.F. (eds.): Communicating Mathematics in the Digital Era. A K Peters, Wellesley, Mass. (2008)
29. Sojka, P. (ed.): DML 2008 – Towards Digital Mathematics Library. Masaryk University, Brno (2008)
30. Digital Math Library Dublin Core (dml_dc): A Recommended Best Practice for Unqualified Dublin Core Metadata Records `http://projecteuclid.org/collection/euclid/documents/metadata/dml_dc.html`