HAL
open science

Toward a digital mathematics library?

Thierry Bouche

To cite this version:

Thierry Bouche. Toward a digital mathematics library?. Jon M. Borwein, Eugénio M. Rocha, and José F. Rodrigues. Communicating mathematics in the digital era, AK Peters, pp.47-73, 2008, 10.1201/b10587 . hal-00347682

HAL Id: hal-00347682

https://hal.science/hal-00347682v1

Submitted on 16 Dec 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TOWARDS A DIGITAL MATHEMATICS LIBRARY?

A FRENCH PEDESTRIAN OVERVIEW

THIERRY BOUCHE

ABSTRACT. After an overview of the worldwide mathematical journals ecosystem, we summarise some of the hopes and fears raised by the digital environment. We then review the underlying principles and main features of some of the projects launched by Cellule MathDoc that aim at settling robust foundations and giving wider access to academic mathematical research journals. In turn, they could prefigure some building blocks of the digital mathematics library, which is still to come into being.

INTRODUCTION

We are living a period of transition as for the diffusion of knowledge and the results of research. Printed paper form has been the vector, and to a certain extent the engine, of culture and scholarship during the past 500 years. This paradigm today is largely challenged by new means of production, diffusion, and conservation of the texts. While it is impossible to predict the future of scientific publishing now, one must be aware that search for documentation is made from now on mainly on the Internet, and that the access to reference works in a click is an immense asset for their diffusion in the whole world. Surely, the predictions about the death of the Gutenberg galaxy were a little hasty, as there is no better device to disseminate consolidated knowledge today than printed and bound books. Nevertheless, one has to admit that contemporary research is done primarily through electronic means (email, exchange of preprints over the networks, online bibliographic searches, access to the articles themselves through electronic publishing services). In order for the fundamental heritage that our forefathers left us not to be simply lost for generations to come because of those ongoing changes, it was necessary to undertake conversion from the existing paper literature into a sufficiently rich digital format to allow whatever use might seem appropriate to our descendants. Moreover, it is essential to think globally about this move, that is: let us try to avoid too specific designs in our applications so that they can eventually *merge* together within current born digital environment and with future developments. We consider that it is our generation's responsibility to initiate this task because it will probably be the last one which will have been familiar at the same time with the old medium and the new ones.

Preliminary version of an article to appear in *Communicating mathematics in the digital era*, book written by the participants of same title's conference, held in Aveiro, Portugal, August 15-18, 2006. Second revision with updated info: February 2008.

1. A GLANCE AT THE MATHEMATICS LIBRARY

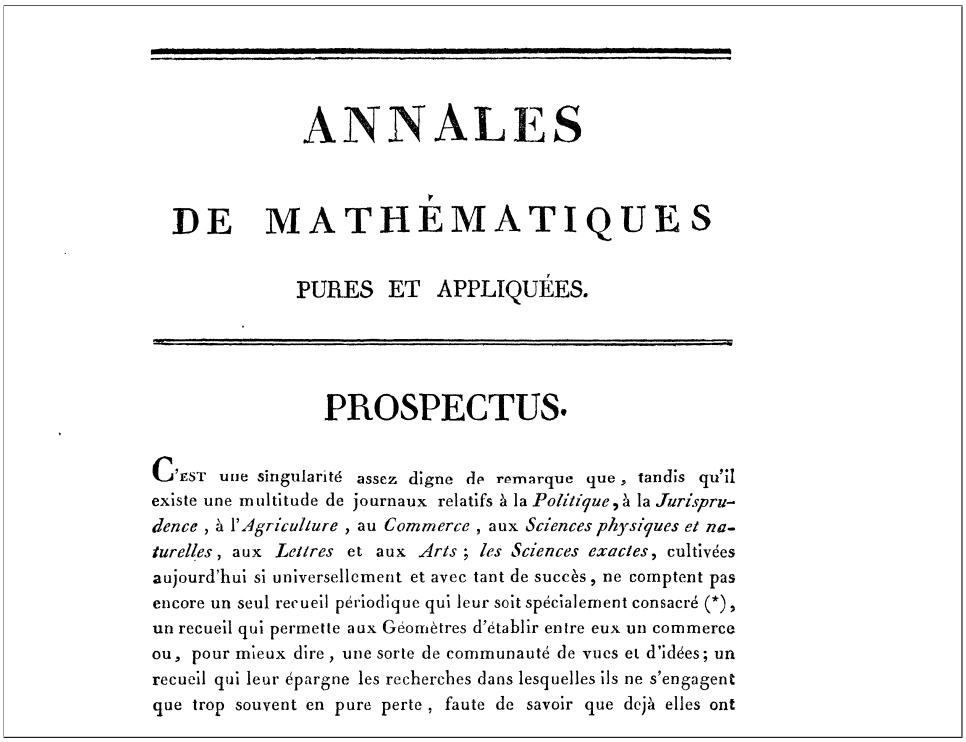
"But the number of the periodical repositories of mathematical literature has become so great, that papers consigned to them, although preserved, as we may hope, for all time, are in imminent danger of passing out of sight within a few years after their first appearance. They are preserved from destruction, but not from oblivion; they share the fate of manuscripts hidden in the archives of some great library from which it is in itself a work of research to disinter them."

Henry John Stephen Smith, 1882.

1.1. **Looking back.** Mathematics are special among sciences in that there is no experiment: the main activities of mathematicians are reading, thinking, discussing, and writing. However, their writings are meant to be a piece in a collective effort to build a logically coherent elaboration from axiomatic foundations to rather sophisticated statements. While any human output is tightly related to previous works in the same field, typically in a dialectic fashion, mathematics might be the only domain where citation is almost never a tool for contradiction. This is why mathematicians usually see mathematical literature as a whole, and would not feel comfortable if some large parts of it got lost, or hidden behind strong barriers. Of course, permanent and immediate full access to all mathematical literature is not required because nice people summarise topics in books, but random access to the actual source is highly appreciated, because it is often illuminating. Reading the invention of a concept in the author's own words yields a much better understanding of the context and intentions behind the discovery. Moreover, as has been already stressed, see e.g. G. Michler [12], it might happen that the proof of an important theorem be so much untractable that it never ends up in a textbook, so that resorting to the original article is sometimes mandatory.

Another important aspect of mathematical literature is that mathematics are the common idiom of science, so that virtually any scientist's work relies on some part of the mathematical corpus, but in an *asynchronous* fashion: while most of current science and technology use well-established mathematics from the 20th century, specific schools of the past are rediscovered and are an impulse for new trends in some disciplines—it can also be said that many physicists use mathematical theories that have not yet been sorted out by mathematicians themselves...

In the sequel we will focus on cumulative and creative mathematical literature where current advances have been appearing for three centuries: scholarly journals. France has a long tradition in producing and publishing mathematics: while the first scholarly journals appeared roughly at the same time in Great Britain and France at the end of the 17th century, the first journal solely devoted to mathematics appeared in southern France in the early 19th century (*Annales de mathématiques pures et appliquées*, published in Nîmes by Joseph Gergonne). It was soon followed up by Crelle's and Liouville's journals, both of them had previously published in Gergonne's. Since that time, many scientific journals were published by academic institutions such as academies, big institutions such as Polytechnique in Paris, hence multidisciplinary; but some of them, published by smaller groups of people (scientific department, learned societies which began to formalise during the 19th century...) were more specialised. It is interesting to notice that the first journal was launched by Joseph Gergonne,



ANNALES

DE MATHÉMATIQUES

PURES ET APPLIQUÉES.

PROSPECTUS.

C'EST une singularité assez digne de remarque que, tandis qu'il existe une multitude de journaux relatifs à la *Politique*, à la *Jurisprudence*, à l'*Agriculture*, au *Commerce*, aux *Sciences physiques et naturelles*, aux *Lettres* et aux *Arts*; *les Sciences exactes*, cultivées aujourd'hui si universellement et avec tant de succès, ne comptent pas encore un seul recueil périodique qui leur soit spécialement consacré (*), un recueil qui permette aux Géomètres d'établir entre eux un commerce ou, pour mieux dire, une sorte de communauté de vues et d'idées; un recueil qui leur épargne les recherches dans lesquelles ils ne s'engagent que trop souvent en pure perte, faute de savoir que déjà elles ont

FIGURE 1. Gergonne's *Annales de mathématiques pures et appliquées* (1810).

a mathematician who had failed to be an academician, and lived and worked away from the capital: he probably felt more urgently the need of a live record of the ongoing research, rather than relying on peer-to-peer correspondence or a proximity network of colleagues. Looking backwards at the mathematical press since the 18th century, it is also interesting to observe the deep intrication of commercial publishers, learned societies, and public institutions in the making of those journals. However, while most editorial boards were acting under the auspices of some trusted moral authority (not always explicitly stated: for instance, the *Bulletin des sciences mathématiques et astronomiques* was launched under the auspices of the École pratique des hautes études, by contract between the ministry and the commercial publisher Gauthier-Villars), mathematicians would not generally supervise themselves the making and marketing of the volumes.

One can wonder about the main benefits these journals have brought to the scholarly community, whether we should list standardisation in exposition and style, up to notation, organisation of the flow of discoveries, or a long-lasting record of the present activity, afterwards resting on library shelves to be easily checked whenever needed. Nevertheless, an important feature of the first journals was to provide bibliographic records of the ongoing publications (books, but translations or abstracts of contemporaneous 'competitors' as well). This was then passed along to specific projects like the *Jahrbuch*.

1.2. **Current trends.** Today, the general entropy of the system has gone far further, with hundreds of mathematics journals published worldwide, from those

published without an economic model by a couple of colleagues, through the many publications of math departments and small societies, up to the big science, technics and medecine (STM) publishers with hundreds of titles in their catalogues. Where scope, audience, or reputation vary wildly, with no obvious correlation between any of these and price, availability, production quality, etc.

This is a fragile ecosystem which is currently undergoing a deep crisis. The most visible facts are concentration, bundling, price raises on one hand, and free access advocacy on the other one, which is inspired by the open source movement in the software industry, which is at first sight a surprising paradigm for designing new ways to produce and disseminate intellectual achievements. Over the past decades, it is amazing to observe how often academic publishing has been outsourced when it was done in-house, and otherwise internalised. The reasons for outsourcing typically are that the scholars feel at some point of time that commercial vendors can better achieve the technical tasks such as proof-reading, printing, but also managing subscriptions or branding and selling the journals: they are more reliable and recover their higher costs and profits by better marketing. The reasons for internalisation are symmetrically based on disappointments with regard to the quality-price ratio of services by external providers compared to what scholars think they can easily do themselves. This is a circular movement that won't reach an equilibrium state before long. As the then Elsevier Science chairman (now Springer Science+Business Media CEO) Derk Haank puts it: "It is ironic that the whole world is talking about outsourcing and the academic community would in-source a tedious job like publishing." [8]

One possible answer to this remark is that librarians are not satisfied with many directions that some publishers opted for recently, thus leading them to undertake themselves some of the tasks traditionally associated to publishers. For instance, publishers are more concerned with their immediate future—the rate of profit they'll present to their shareholders at the end of the fiscal year—, than with taking care of their unsold items, which are usually sold at paper's value after a short period of time. It is generally considered that it is libraries' job to store carefully the publisher's products and index them so that researchers can access them when required. The early years of electronic publishing have placed a certain amount of burden on the librarians because the typical offer was a site license which allowed a limited time access to electronic resources hosted on servers operated by the publishers. This placed the library in a situation of full dependance upon the publisher, as the first mission of libraries—storing intellectual output in order to provide local access to it—was now resized to a minimal one: managing subscriptions and controlling access to third-party services.

While these issues don't look very specific to mathematics, they have a special meaning to the mathematical community because we are probably the scientific community which relies most on its past publications, and we might be the one in the STM area which has kept the largest diversity, where the mainstream publishers don't yet entirely shape the landscape (see some convincing discussion by Pierre Bérard [1]).

It is natural that new ways of producing and disseminating mathematical research should be inquired as new tools and methods appear, but we must beware that what drives choices of the present is mostly vanity and greed. For instance, when the paper industry found cheaper ways of producing something that looked like premium white paper made out of tree trunks rather than softer fibers, publishers adopted quickly what was soon to be known as acid paper, and would start distroying itself later. The use of the Linotype, and then electronic typewriters, in the production of complex mathematical articles was never for the sake of mathematical expression's clarity or better typography, but of cutting costs (for a nice set of samples, see Donald E. Knuth [10]).

1.3. **Looking forward.** Plenty of literature has already been devoted to the future of our academic publication system: we're not going to add yet another prediction to the list. While scholars were arguing, the industry was moving fast. The near future is already largely dictated by the big players: very large technological platforms have been set up (Extenza, Highwire press, Ingenta, Metapress, Scitation, Crossref...) and market leaders have asembled gigantic bundles (ScienceDirect, SpringerLink, JSTOR, AIP, etc.) where one can observe that no access is given for free. Ironically, it is often not even easy to know what content is actually available to subscribers: sometimes even browsing the collections is not allowed. In some occurences, messages like "references secured to subscribers" are displayed for papers having no references at all, or for papers that are not yet accessible at the website. But maybe the most evil, although becoming routine, practice is to secure access to texts that are in the public domain.

The only device for interoperability among STM publishers, for the sake of linking, is Crossref, which provides a general infrastructure based on a central unique registry allowing to link to some other publisher's website at article level without having to dig into its catalogue. In fact, the full catalogue is exported and exploited by commercial third parties like generic Internet search engines, instances of them dedicated to the academic literature, or highly integrated tools designed for the rich scholar like Scopus.

Once again, the idiosyncrasies of mathematical literature are left along the wayside of these important manœuvres. Probably the most influential proposal for a more responsible shaping of the digital near future of the mathematics library has been the recommendations by the CEIC of the IMU [13], and this is one of the bases on which we settled when designing the French projects to be discussed below—mainly: build on structured open formats, keep all navigational items free, and provide eventual free access to the articles themselves. For a publisher, obeying these recommandations is kind of a bet: offering much to anyone while charging only for a small subset of the features whose development you invested in might raise sympathy, but will it drive revenue?

We're getting to a world where we have a standard implicit commercial practice spread accross big publishers, which is to secure everything to subscribers, but to tune the subscription fees so that, in fact, anyone *can* be a subscriber, and to acquire such large journal portfolios that research centers can't bypass them, implying that anyone *must* be a subscriber. On the other hand, if the standard practice of gentle publishers—maybe those that have a stronger connection to the scholars themselves, that publish few journals, in a restricted

disciplinary area—is that recommended by the CEIC, one might wonder how long they will get enough income to survive. Having the researchers archiving themselves their preprints is a honest provision for alternative access to greedy publishers. But, if a publisher takes care of providing a high quality interface, which is not a light investment, including digital backfiles which are freely accessible after some reasonable moving wall, then the only electronic pieces of information that remain to be sold (or exchanged) are the full text of recent articles, which is something you often can live without, especially when the preprint is widely available, or when you can get an 'electronic reprint' from the author himself. The main value of the refereed journals system as it has been in existence since few centuries is the complex process of validation of the sheets sent to the press and the duration of their content in well curated libraries shelves, so that the texts are safe and well ordered. The traditional paper publishing paradigm assumed that publishers produce carefully the volumes that are bought by libraries in order to be stored. With the vanishing of paper editions, we cannot ask the publishers to take on the tasks traditionally carried out by the librarians, without charging these new services. We may hope though that some digital libraries will devise protocols to transfer newly generated content in their virtual shelves, letting the publishers concentrate on their skill: choosing content and presenting it in a desirable form.

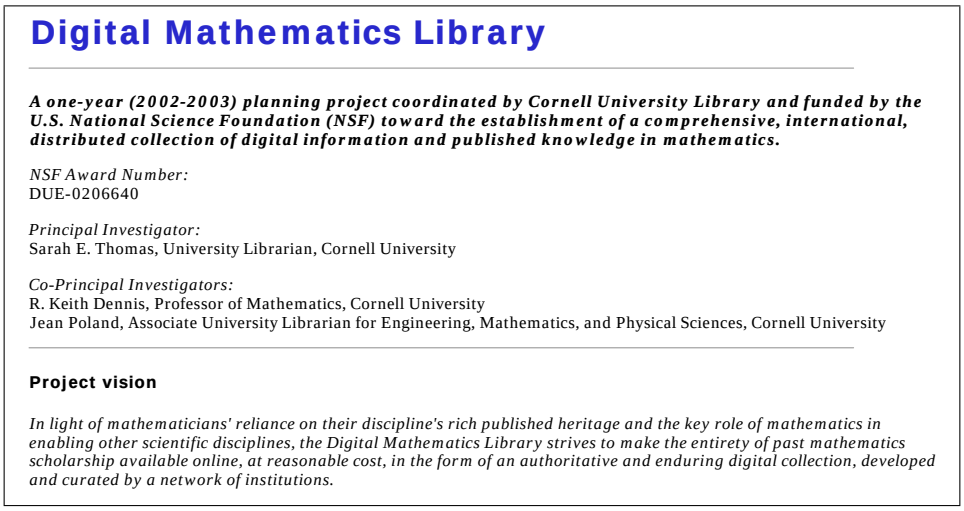
2. THE DIGITAL MATHEMATICS LIBRARY

"With all the money that has been spent on prototypes that aren't working anymore we could have built a simpler but more sustainable digital environment."
Martin Grötschel [7]

At the turn of the century, there has been a lot of activity which we can now look back as "wish for a DML", where DML means *Digital mathematics library*, which in turn had as many meanings as interested parties.

2.1. **Beginnings.** While some of the protagonists focused primarily on retro-digitisation, others felt from the start that what was needed was a comprehensive thus heterogeneous, searchable, interrelated distributed network of all the existing mathematical literature [11, 5]. Since its inception, however, the DML concept has provoked many discussions, but no significative outcome. To some degree, the initial direction was indicated by John Ewing's 'white paper' [6] as an endeavour to digitise all the past mathematical literature, which required some sort of worldwide cooperation, competition, and organization. So far, some competition, and some cooperation between 'competitors', have occured, but the lack of a real organization, which might be related to the lack of large scale funding, has stalled the project.

Notice that the amount of valuable mathematics already available digitally is considerable, but all those individual bits do not glue into anything resembling a library. Moreover, as so much of the new additions are driven by commercial apetite, many basic principles are left out. For instance, the Springer 'Online Journals Archive' bundle has been asembled according to principles that make sense commercially, but not scientifically, at least for mathematics: the concerned journals are those in Springer group that are alive today, mostly written in English. They are scanned at 300 dpi and OCRed using an English-only OCR



Digital Mathematics Library

A one-year (2002-2003) planning project coordinated by Cornell University Library and funded by the U.S. National Science Foundation (NSF) toward the establishment of a comprehensive, international, distributed collection of digital information and published knowledge in mathematics.

NSF Award Number:
DUE-0206640

Principal Investigator:
Sarah E. Thomas, University Librarian, Cornell University

Co-Principal Investigators:
R. Keith Dennis, Professor of Mathematics, Cornell University
Jean Poland, Associate University Librarian for Engineering, Mathematics, and Physical Sciences, Cornell University

Project vision

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.

FIGURE 2. Home page of the DML project at Cornell.

(which means for instance that the French letter *é* can be recognised as a 6, yielding such poetic statements as"D'aprbs la d6finition de la d6riv6e g6om4trique"). We might be concerned by the fact that only the latest series of the Italian journal *Annali di matematica pura ed applicata* was recognised to meet those selection principles, thus we will have to look elsewhere for the seminal works of Beltrami, Betti, Cremona, etc. The actual state of disorganisation is pretty well depicted by Eugénio M. Rocha and José F. Rodrigues in [15], see their chapter in this volume for an update.

To date, the DML has been

(1) An informal club of enthusiasts that met at San Diego, California, in January 2002;
(2) An NSF planning grant hosted at Cornell university library (2002-2004);
(3) An IMU project coordinated by the CEIC through its WDML subcommittee which was formed by IMU's president during summer 2003, but never started working;
(4) A wdml.org website with some sketchy information provided by the CEIC since 2003.

The second item was the most productive, see [14]. Six working groups were formed, addressing the following issues: economic model, archiving, metadata, content, rights, and technical standards. They are here listed in rough order of success, or reverse order of difficulty. Although the simple task of writing down the consensus on technical standards was achieved at that time, this is the only area where the CEIC has proposed a new text, while the questions known to be inhibiting the whole process have stayed untouched.

In parallel, the EMANI project was launched as an initiative where Springer would partner with libraries in order to archive its mathematical digital content, at a time when it was not obvious that the group would significatively change in size and launch its own digitisation program. Two European proposals were attempted under the auspices of the European mathematical society and its Electronic publication committee: although quite different from each other, both were unsuccessful.

At the time of writing these notes, there is no evidence that any DML-related international effort will ever succeed. The recent history shows that the stakeholders are eager to share their expertise and discuss the issues, but that no trusted entity has yet emerged in favour of which they would agree to loose some control over their collections.

2.2. **Retrodigitisation and archiving.** It is a strange fact of contemporary life that, while everyone faces problems reading electronic documents a couple years old, or even just finding one that had been stored on some local drive, but enjoys daily reading books decades old, we're convinced that paper will soon vanish and have to be replaced by some electronic device. In fact, we even believe that we're going to learn from our bad experience so that we will come up with very good tools for long-term preservation of electronic files. There is no evidence so far that supports this optimism, but we will have to wait some time to know!

However, it seems obvious to anyone that mathematical literature will be more useful once it has been converted into a suitable digital format, especially if we succeed in converting the legacy article stock into an active web of references. The first action should thus be to digitise the existing literature for which no satisfactory digital surrogate is already available. The next obvious actions are to store the newly created files in some virtual library system, to build multiple preservation and access services on top of that, and to set up a mechanism so that the ongoing production ends up in the same system.

As for long-term preservation, one can share the skepticism of Christian Rossi [17], namely that: the more meaning or structure you store in a file, the more likely it is that you loose the environment or documentation that allows you to exploit its content (notice, e.g., that the TEXBook is a TEX program, that the PDF format is described in a PDF file, or that a very well-structured XML file does not bear by itself any information about its own structure, its meaning, and its expected usage). So all we can hope is that very clever people will devise high-end systems that will enable us to store safely all relevant data and metadata together, in some format so simple that future generations won't have difficulties reverse-engineering it, and on some medium that won't be already unreadable when their curators will learn that it has become obsolete.

In fact, one lesson of successful digitisation projects might be that it is possible to build relatively easily, and at a relatively low cost, a full-featured electronic edition of a journal, starting from paper volumes. Due to this observation, together with the difficulty of reuse of sophisticated digital formats out of their context, we're inclined to think that the safest digital archive of a paper based publication (but maybe of a born digital one as well) is a flat directory of bitmapped pages in some straightforward uncompressed format.

2.3. **Born digital content and integration.** The questions regarding long-term archiving of electronic journals, secured away from publishers' fate, have yet to be answered: some concurrent proposals are on their ways, many of them being held by organisations which are not necessarily economically safe for a long life.

We are also concerned by advances in digital rights management (DRM) technologies together with laws enforcing full control to the rights owner over the usage restrictions at the time of publishing. This might lead to the massive dissemination of unarchivable data. For instance, an existing DRM technology

(protectedpdf, by Vitrium Systems) claims that it enables a publisher to produce apparently benign PDF files, which you can print a given number of times, whose use is monitored by the publisher, whose content can be modified at any time: updated or merely removed—all these features, thus the very possibility to access the published text, relying on a quite volatile state of the art: current software and networking technology, and quite some commercial partners alive, some of which are start-ups!

We could consider that the paper version is a better authority for long-term archiving, but we should realise that current paper printouts are paper surrogates of electronic systems, no longer the other way round. For instance, given the low diffusion of specialised mathematical journals, many publishers will provide a paper edition that is made using print-on-demand systems. But, while liquid ink chemically alters the paper that absorbs it and insures maximum longevity for the text (in some cases, the text has been observed to be the only lasting piece of some books printed on acid paper!), laser 'printing' consists in cooking some rigid polymer over the surface of the paper which will eventually reject it.

Digitisation might be the only way to secure an archive independant from the possibly bad choices of today, still for quite some time ahead...

2.4. **Concluding remarks.** Currently, the DML is a generous idea that drives some activity, but most actions are done in competition and independantly, with very short-term outcomes in mind. The best we can hope is that we will still have somewhere, in some usable format, the mathematically valid content that is output nowadays, and that we won't loose the one that is already at rest on our library shelves.

An effective digital mathematics library should address the three fundamental functions of a library: archiving, for access and for preservation, and acquisition. The first issue, electronic access, has been rather well understood and is already implemented in a quite satisfactory manner in existing retrodigitisation programs—however, the long-term sustainability status of these programs is unknown. The second one, digital preservation, is, and will remain, a vastly open question. Indeed, it is only in case of failure, in the future, of the current preservation plans that we shall know for sure whether they were robust enough to succeed. The last one, acquisition of new electronic content and its transfer to virtual shelves for access and preservation, is in a very problematic state. Some promising projects such as Cornell university library's Euclid do provide the feature, but they typically go as far as hosting the online edition of the journals in order to be able to get enough data to archive them: this gives no hint on how we could convince commercial publishers to feed a stable and ever growing DML.

Nevertheless, we hope it will be useful to draw a sketch of the picture we envision, and to detail the implementation of a prefiguration of some of its components in the very specific French context.

3. The French actions

"What is important:

- *Make things easily usable, even better, make them simple!*
- *Long-term funding and not short-term projects.*
- *Sustainable activities are more important than hype killer apps.*
- *International and interdisciplinary cooperation."*

Martin Grötschel [7]

3.1. **The French ecosystem.** We have in France a long tradition of mathematical publishing, which yields the current situation where about 20 journals are alive. We shall distinguish those categories:

(1) Independent journals (usually published by a university department): *Annales de l'institut Fourier, Annales mathématiques Blaise-Pascal, Annales de la faculté des sciences de Toulouse, Journal de théorie des nombres de Bordeaux, Cahiers de topologie et géométrie différentielle catégoriques.*
(2) Learned societies publications: *Bulletin et mémoires de la SMF, Astérisque, Revue d'histoire des mathématiques* (SMF), *Revue de statistiques appliquées* and *Journal de la SFdS* (SFdS).
(3) Journals published by a commercial publisher under the auspices of an academic institution or a learned society who owns the title: *Annales de l'institut Henri Poincaré* (IHP/Elsevier), *Annales scientifiques de l'École normale supérieure* (ENS/Elsevier), *ESAIM – Control, Optimization and Calculus of variations, ESAIM – Mathematical modelling and Numerical analysis, ESAIM – Probability and Statistics* and *RAIRO: RO* (SMAI/EDPS), *Journal of the Institute of Mathematics of Jussieu* (IMJ/CUP), *Publications mathématiques de l'IHÉS* (IHÉS/Springer).
(4) The unique case of a public service operated by contract by a commercial publisher! *Comptes rendus de l'académie des sciences, série A mathématiques* (Académie/Elsevier).
(5) Journal titles currently owned by a commercial publisher : *Bulletin des sciences mathématiques* and *Journal de mathématiques pures et appliquées* (Elsevier).

Up to my knowledge, most journals of categories 1-3 accept exchanges. It might be useful to notice that one finds the most scientifically respected journals under no. 3, which is also the most populated number, while each category holds one or more internationally respected journal: scientific reputation is rather independent of the organisational structure. Most of these journals are supported by the CNRS, which is a public institution of the French state (the support might be to provide grants as well as permanent technical staff, employed by the CNRS, to the journal).

3.2. **Digitisation.** In France, in our context, digitisation is spelled NUMDAM (aka *numérisation de documents anciens mathématiques:* old mathematical documents digitisation). This might change, as some publishers and librarians could decide to digitise themselves their collections, but this is not expected soon. We could also mention projects which were not designed specifically for mathematical journals but deal with some of them, like Gallica, or a project which is similar to (in fact partly inspired by) NUMDAM in the fields of social sciences and humanities: PERSÉE. What all these programs have in common is that they are

managed by public institutions of the French republic, and meant as a public service, hence most of the collections are freely accessible.

Outline. NUMDAM was initiated just before year 2000 by MathDoc's directors of the time: Pierre Bérard and Laurent Guillopé, thanks to generous support from the ministry of research. The *Cellule de coordination documentaire nationale pour les mathématiques* (aka MathDoc) is a small service unit belonging to both the CNRS and the University Joseph-Fourier, located in Grenoble. It is dedicated to providing bibliographical services at a national level to the mathematical community, which is why it is headed by mathematicians, and employs a balanced staff of librarians and computer scientists.

The actual industrial production is outsourced: MathDoc concentrates on inspecting and preparing the collections, controlling the output of the production, archiving and maintaining the online access at NUMDAM website.

We have only consumed public funds so far, adding to our initial funding specific actions from various sources. Although we consider ourselves a part of the DML, we'd like to emphasise that MathDoc is not a library, and NUMDAM is not associated to any library in particular: MathDoc owns no documents itself. NUMDAM is meant as a service to other parties, under the auspices of the French mathematical community. Similarly, NUMDAM posting is not considered a new edition: just the faithful copy of the paper edition, in another medium. MathDoc acts as an agent for existing publishers, it is not a new one.

NUMDAM's design was a long task, and we must acknowledge the support that we received from our funding partners who didn't urge us for a quick output, but left us think a while how to do it best: the first online posting happened in 2002 on December 20, when 2 journals appeared. At the time of writing, 5 years later, we offer 20 serials, 1 series of proceedings, 27 important seminars. We first started with the obvious, highly regarded and currently alive, pure mathematical journals, but we now provide a much more comprehensive landscape, with journals in applied mathematics, mathematical physics, statistics, summing up 26 000 articles spanning over 560 000 pages. These include the first ever mathematics journal (*Annales de Gergonne*, 1810-1831) and our first non-French participating journal (*Compositio Mathematica*), thanks to an agreement with the Foudation Compositio and the London mathematical society.

The recent additions concern 1 French journal at the boundary of mathematics and social sciences (*Mathématiques et sciences humaines*), 2 Italian journals (*Annali della classe di scienze, Scuola normale superiore di Pisa*; *Rendiconti del seminario matematico della università di Padova*), seminar proceedings from the École polytechnique (around Laurent Schwartz) and Collège de France (around Jean Leray), the Bourbaki seminar... As the collections keep growing steadily, we should reach the million of digitised pages at the end of the current phase of the program.

Goals. The main keywords that led to the decision of raising funds for the digitisation of mathematical literature published in France were: long-term archiving, visibility and availability.

Nevertheless, the inception of NUMDAM had its roots in ongoing discussions at French and European level with inspiration from JSTOR, CEIC, DIEPER, while Gallica was already viewed as an old project.

As with so many projects in this field, the main motivation was the commercial policy of big publishers: rising prices, bundling, licensing model for the electronic access which didn't make it clear what would be available to the average researcher if the publisher was ever to change policy, or to get out of business. So the motivations were very much bound to dated interrogations (see § 1.2) about big publishers—especially those who are in such a dominant position in their domain that they can rule the market.

One of our goals was thus to build a safe archive of the 'French' mathematical cultural heritage, that would be maintained by some public institution, and would not depend on the internal archiving policy of the current publishers, or even of its future destiny. As a side note, we observe that, when some of these publishers decided to digitise themselves their own collections, they had to query libraries because they had no stock and, anyway, much of those came from the acquisition of other publishing houses which they had dismantled right away.

The other, companion, goal was to endow independent and society journals with the same Internet presence, so that they could compete in functionnality, expose their rich history—they didn't need to compete on the scientific excellence. This logic has naturally been followed up in the CEDRAM project, which proposes to the same stakeholders to rethink and upgrade their born digital production to current best practices, see § 3.3.

There is an obvious competitive advantage for visibility in being an early mover. Being more visible surely enhances usage, hence impact. As the medium, and much of the paradigm, move, one can bet on a redistribution of cards. It is thus natural that national funds be spent in digitisation as a strategy to get more return on previous investment in this new era.

Finally, we'd like to recall here shortly one issue in our brave new world of scientific publishing: the mathematical literature is highly multilingual in essence, current mathematics are the result of a worldwide effort of thousands of human beings accross centuries and boundaries. This is both a barrier for the current mathematician to access primary writings in Arabic, Chinese, Greek, Japanese, Russian or most European languages, and the opportunity to read some great mathematician's thoughts in their own words. As of today, the French idiom is the last one, beside English, used for publishing major works in international journals. Each other country with a strong tradition gave up during the last century. As the DML gets populated, it gets more multilingual, and it could reimpulse linguistic pride, thus raise higher the quality of the language used in communicating mathematics. We hope that NUMDAM website's bilinguism will sustain the continuing effort of the French community in writing important articles in French.

Copyright model. Our copyright model can be summarized by the following items. The underlying idea is that our work should modify the previous state of affairs as little as possible.

- The electronic version is under the journal title's owner's control (usually an academic institution or a society), as it happened for the paper version.
- Authors are asked to transfer their (exclusive) electronic copyright to them if they are still alive.

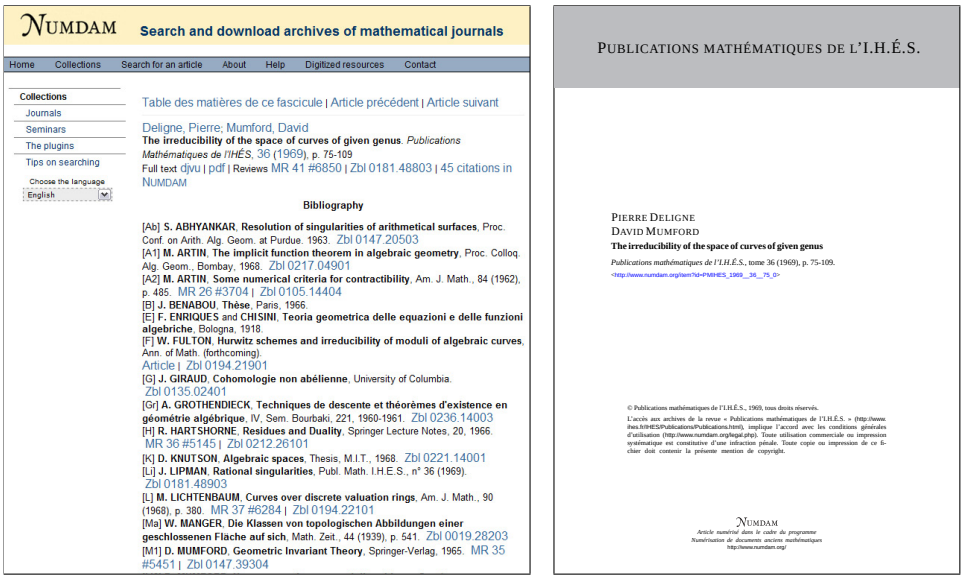


FIGURE 3. An article from *Publications mathématiques de l'IHÉS* at NUMDAM.

- We contract with the copyright owner to allow online access at NUMDAM website.
- We define a moving wall on a journalwide basis (the most frequent value is of 5 years, there are few free access journals). We view free access after the moving wall period as a fair counterpart to public funding.

Standards. Because we're not trying to create a new edition, we won't perform new editorial choices, thus we respect any aspect of the paper publication: we deal with full journals backruns, we scan every page at high resolution (600 dpi black & white for text pages, grey or colour when appropriate). We retain the original page format in the delivered files, which provide thus a faithful image of the printed material. Besides, we consider that we are responsible for the metadata generated within the project, links, and other edits, hence we keep a clean separation between articles as images and added metadata.

We capture detailed structured metadata allowing to search over basic bibliographical data plus plain-text and cited references. This means that we have a metadata model that allows to describe journal runs, made of physical (bound) volumes, which are essentially made of articles. Among the metadata describing an article are its typical bibliographic reference, which is fielded, its abstract, its OCRed full-text, and its bibliography, as it is printed, with important bibliographical elements tagged.

We provide a clear identification of the originating journal on all supports through the visible mention of the origin and links to the website of the publisher on the individual articles covers.

Our user interface to the archive is dual, it is made of both the article metadata exposed on our website, and the user version of the scanned article itself. We consider that it is good practice to associate to any article a stable or persistent URL, which should not point to a graphical file, but to the nevralgic center where

all the relevant metadata is collected and exposed using a plain HTML format. These are the vertices of the DML graph.

One important reason for this dual design is that metadata features are subject to change, while the article's text won't be modified. An obvious example is that, as the DML grows, new citation links will be activated, another one is the correction of errors in metadata.

We have thus endowed each full-text file with an added first page which is similar to a reprint cover. This provides all the necessary information for the users to clarify the origin of the file, the allowed usage, and its persistent URL, which is an actual link from the article file, where its full record is available, with all navigation features up-to-date.

Links. A hyperlink network places the article in proper context: we add as many meaningful links as possible to each article. For instance, some items in its bibliographical reference are clickable: author gives the list of this author's papers (using a local authority list of authors), issue number leads to the table of contents of that issue, etc. There is of course a link to the full text (in both PDF and DjVu formats) if freely accessible, or to the article's location at its publisher website if applicable. If the article has an erratum that we are aware of, then we provide a link. If the article has been reviewed in one of the reviewing databases that we know (*Jahrbuch über die Fortschritte der Mathematik* (JFM), *Zentralblatt MATH* (ZM), *Mathematical Reviews* (MR)) and we can match the review there, then we add links as well. In some occasions, we have more special relations: one old journal had the special feature that some of its longer articles were spread over various volumes, printed in arbitrary page ranges with no consideration for logic, so that a sentence could be running at the end of a page, and you had to wait for the next volume to know how it would finish, and this resulted in many chunks of the same article in our database: we added links to the list of all the parts from each of them. In the case of the Probability seminar of Strasbourg, we had an already running database of third-party written abstract of all the talks out of the initial period: we linked these abstracts to the texts, and added links from the articles similar to those to other reviewing databases.

But most of our links concern our bibliographical references cited from individual articles: those quotations have been matched to the extent possible to JFM, ZM and MR reviews. Moreover, when a quoted article happens to be part of NUMDAM, we provide the direct link. We started to add similar outer links as a byproduct of the mini-DML that will be discussed below. Without taking into account their respective coverage periods, we have matched 75% of our references to a ZM review, 68% to MR, 4% to JFM, 9% to NUMDAM itself, and 12% to other digitisation centres thanks to the mini-DML facility. This means that we provide around 21% of cited references with a direct link to the cited article's full-text, and still more through the reviewing databases.

Collections. We summarize in the tables below the serials that have been handled by NUMDAM.

Currently online (Febr. 2008)

Title	Period	Owner	Volumes	Pages	Articles
Ann. Fac. Sci. Toulouse	1887-2002	University	207	36 270	1 041
Ann. Gergonne†	1810-1831		22	8 700	1 083
Ann. inst. Fourier	1949-2000	Assoc. A.I.F.	156	51 054	1 811
Ann. I.H.P.†	1930-1964	I.H.P.	71	5 860	147
Ann. I.H.P. sér. A, B, C	1964-2000	I.H.P.*	119	55 000	2 400
Ann. math. Blaise-Pascal	1994-2002	Labo/UBP	19	2 554	165
Ann. Sci. École norm. sup.	1864-2000	É.N.S.*	295	68 898	1 867
Ann. univ. Grenoble†	1945-1948	UJF	3	1 006	47
Ann. Scuola Norm. Sup. Pisa	1871-2001	SNS	253	46 062	1 726
Bull. Soc. math. France	1872-2000	S.M.F.	167	45 774	2 608
Mém. Soc. math. France	1964-2000	S.M.F.	134	18 118	396
Compositio Math.	1935-1996	Found. Comp.	275	39 500	1 934
J. théor. nombres Bordeaux	1989-2003	IMB	33	7 700	431
Journées É.D.P.	1974-2003	C.N.R.S.	31	5 976	514
Math. et Sciences humaines	1962-2000	EHESS	148	12 600	763
Publ. math. I.H.É.S.	1959-2000	I.H.É.S.*	92	17 424	344
Rend. Sem. Mat. Padova	1930-2000	U. Pad.	108	33 300	1 980
Rev. Stat. appl.	1953-2004	SFdS	207	22 150	1 676
Sém. Bourbaki	1948-2000	Assoc. N. B.*	45	18 200	920
Sém. Leray†	1961-1977	Coll. de F.	38	3 200	129
Sém. Paris†	1953-1985	Secr. math.	146	37 700	2 647
Sém. Polytechnique	1969-1996	X	37	12 800	910
Sém. Proba Strasbourg	1967-2002	IRMA*	37	17 352	1 254
				567 200	26 800

* Contract with a commercial publisher. † Dead serial.

Forthcoming (Febr. 2008)

Title	Period	Owner	Volumes	Pages	Articles
Ann. Fac. sci. Univ. Clermont	1962-1993	UBP	37	5 050	328
Cah. topol. géom. différ. catég.	1966-2006	A. E.	184	16 300	691
Nouvelles annales	1842-1927		94	50 700	8 648
Bull. Darboux	1870-1884		27	9 400	609
SMAI journals (prior to ESAIM series), 1914-1945 French thesis, ...					

Future plans. At the time of writing, we can consider that we have completed our initial task: digitising important French mathematical journals and making their mathematical content widely available. We've even gone somewhat

beyond this goal, as we've been dealing with older material, seminars and European journals. Our main failure is the fact that the titles currently owned by Elsevier Science, as a result of their acquisition of Gauthier-Villars, have not been digitally archived at all, except for the early years at Gallica, with lower standards than ours. Some other important resources have yet to be digitised because their overcautious owners have not yet settled a clear policy regarding archiving or access.

We shall concentrate many of our next efforts on the improvement of the website's features (MathML versions for the OCR and metadata, better full-text search engine, including math expressions, and ranking hits according to relevance), and better integration with other projects. An important feature of the archive is that it should stay alive, not only by active maintenance of the whole system, which is something MathDoc is committed to, but also by growing with the addition of the new articles, and this is a domain where there is still a lot to be done.

Our policy is to store metadata of post-digitisation articles as soon as they are published, and to post the full texts when they are freed from the moving wall. We implemented this policy with three journals handled by Elsevier (the *Annales* of, resp. ENS and IHP) in 2007 and 4 by CEDRAM. We are willing to update the NUMDAM database with newly published articles' metadata on a regular basis from all suppliers, adding links from NUMDAM to the actual article's location at its publisher's site, so that the archive also drives audience to the journals we support. For this to occur, we still need support from more publishers...

MathDoc has had some collaborations, in parallel to NUMDAM:

- The Bourbaki archive project has digitised all the manuscripts, drafts, letters, internal publications from the early days of the association. We built a database on top of this, and provide web access for researchers and the curious.
- We are using our scanned articles in an SMF edition of the collected works of Laurent Schwartz, which will be a partnership between SMF, Polytechnique, and MathDoc.

3.3. **Redesign of journal production.** There are many ways to measure whether a journal is successful or not: how many paid subscriptions? how many exchanges? how much profit generated? how much subsidy consumed? How many paper submissions? How many papers published? Among those, how many are good mathematical papers? Would *you* be proud or ashamed to publish there? etc.

If you are some institution providing funding to various journals, you might wonder whether it's your name/logo that gives value to the cover, or the other way round. For instance small universities tend to support a laboratory's journal and accept happily exchanges all over the world in the hope that they will increase their visibility. This is also to some degree the rationale which is behind the official support to Open access by funding agencies worldwide (Budapest open access initiative, Berlin declaration...): if you finance research, you prefer it to be most visible, used, referred to, so you prefer the derived publications to be universally accessible.

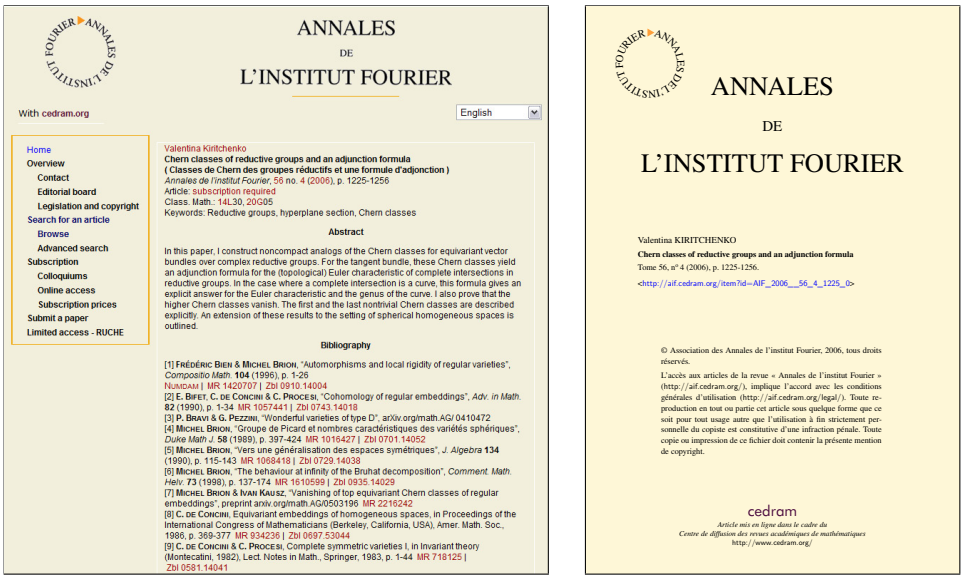


FIGURE 4. An article from *Annales de l'institut Fourier* at CEDRAM.

During the first years of this century, the math division of the CNRS had conducted a survey on the math journals it supported, and the results were puzzling as for the ratio cost-efficiency, whatever measure for efficiency you pick in the above list. We were then in the strange situation where the NUMDAM website was acclaimed for its quality, usefulness and visibility, where every individual journal's fame and legacy papers added weight to the whole lot, while they were divided on the current front, relying on much amateur or unpaid voluntary work, thus being successful in very different areas. As this is essentially a management issue, this is not related to the quality of published articles. It was then discussed inside the community whether some synergies could be brought out, and this ended up in a seamingly simple motto: "Let's make a French math journals portal!". It was also decided that MathDoc would undertake the effort, Yves Laurent being its director.

This led to the CEDRAM project (an acronym for *Centre de diffusion de revues académiques de mathématiques*, meaning something like "Diffusion centre for academic mathematical journals"), which is viewed as a combined effort by the participating journals and MathDoc, with advising partners, including the French mathematical societies SMF and SMAI, making up a steering committee whose president is Claude Sabbah [2].

Currently, CEDRAM is mostly a set of internal tools meant for the production of mathematical journals with full-featured paper as well as electronic editions, it is also a platform for hosting their electronic edition. The visible part of the project is at the CEDRAM website, which is the awaited portal: it provides links to all French mathematical journals supported by the CNRS, and a combined search engine over all the articles of the currently hosted journals.

The main design decision for this platform was that it should provide NUMDAM features as a minimum, plus access control to full texts, and a wide versatility as to preserve a distinct look-and-feel for each journal.

Each journal remains completely independent, regarding scientific content as well as financial balance. The synergies are thus currently limited to sharing a common underlying platform, and be simultaneously present on the portal. More options like bundled subscriptions will probably be introduced in the course of the project.

Journals. As a matter of fact, the founding journal partners of CEDRAM were all belonging to the first type, according to my taxonomy in § 3.1. One can even remark that all of them were published by mathematics laboratories from non-Parisian universities. The project was not designed exclusively for them, but it seems that they are those for which there is no doubt that the benefits exceed the burden of conforming to some production standards, and of changing some production steps. This has mildly changed as we help the SMAI (French society of applied mathematics) launch a new journal in 2008: *MathematicS in Action*.

Three journals are online since year 2006, opening the service:

- *Annales de la faculté des sciences de Toulouse, mathématiques*;
- *Annales de l'institut Fourier* (Grenoble);
- *Annales mathématiques Blaise Pascal* (Clermont-Ferrand);
- *Journal de théorie des nombres de Bordeaux*.

In 2007-2008, news concern the addition of three series of seminar proceedings, and the above mentioned SMAI journal.

Architecture. Of course, some of the partners had a French version of project Euclid in mind while the CEDRAM was under development. But this is not what devised *in fine*: while project Euclid is an electronic edition *back-end* that will accept whatever the participating journals produce, and build the electronic offer on top of that, we conceived CEDRAM as a set of modular *plug-ins* that operate at various crucial steps in the production process, so that end-user features are uniform.

One of the main reasons for this option was that we wanted to be upward compatible with NUMDAM, so that the CEDRAM production could be archived by converting the metadata to NUMDAM format, which would then provide long-term access (notice that all CEDRAM journals have agreed on the moving wall model for eventual open access). Another reason was that we also wanted to provide preservation for the source files so as to have two different archiving strategies: NUMDAM for long-term accessibility to user files, some black-box system for long-term preservation of the production system, in order to keep the sources nearest to the mathematical meaning published, rather than its representation in some ephemeral format.

So far, the existing modules that are already in use are:

(1) RUCHE, a software for managing the editorial process: from paper submission and refereeing through a web-based interface, to the preparation of published volumes [9].
(2) Cedram.cls, a LaTeX+BibTeX driven production environment [4] that automates whatever can be (page numbers, metadata generation...). The paper edition is a by-product of the whole process when using this system.
(3) An environment that prepares all required data in the *ad hoc* formats for updating the electronic edition.

(4) website development with a content management system that allows each journal to maintain the static pages of its website.
(5) EDBM, a database, indexing, matching, linking, interactive searching software suite that takes care of the user interface to the collections (essentially the same system as in NUMDAM, which has been made more customisable).

Step 3 has been profoundly redesigned in 2007 thanks to the second generation of the CEDRAM environment. It outputs, directly from LaTeX source, XML metadata ready for generating dynamic content in XHTML+MathML, thus getting rid of our legacy Latex2html-style images for mathematical expressions in titles and abstracts.

3.4. **Registries, lists and catalogues.** We won't repeat here what has already been written elsewhere [16, 3] about the exponential growth of the mathematical literature, and the induced need for adapted discovery mechanisms, going from bibliographical lists and catalogues, to systematic reviewing journals, then ending with relational databases.

The latest challenge is, as there are now so many concurrent databases, each with its own set of metadata, user interface, and document coverage, that people started all over again to write lists of databases, and we're now under the pressure of maintaining either a database of databases (registry) or unified access points to the content of those databases using the greatest common denominator of their metadata set, which is of very low granularity.

We give here three examples that maybe only add to the whole entropy of the databases universe.

Gallica-Math. The BNF's server Gallica has a huge amount of valuable mathematics that are somewhat hidden by weak metadata policy. It is mostly based on a paper library catalogue, which means that only physical volumes possess a record, but users deal with logical units, which might be one volume monographs, as well as multivolume works or single contributions published inside a journal, or proceedings volume, e.g.

MathDoc has built a user front-end to Gallica's resources where the available volumes are indexed at the article level. Two collections are concerned so far:

- Some important mathematician's collected works.
- The early years of Liouville's *Journal*, for which Gallica supplied HTML tables of contents.

Each logical item inherits an identifier and a full record so that third parties (including our mini-DML, see below) can link to them.

LiNum: Livres numérisés mathématiques. This is a consolidated database registering 2 679 freely accessible books, and 651 digitised but copyrighted, provided by Gallica (Paris), *Digital Math Books Collection* (Cornell), *Historical Math Collection* (Ann Arbor), *Mathematica* (Göttingen), *Biblioteka Wirtualna Matematyki* (Warsaw), and smaller digitisation centers. It was meant to be updated on a regular basis, but only Gallica provides us with notification for its recent additions.

RBSM: the Répertoire bibliographique des sciences mathématiques. This is a very special list of mathematical items [16] that was asembled at the turn of 20th

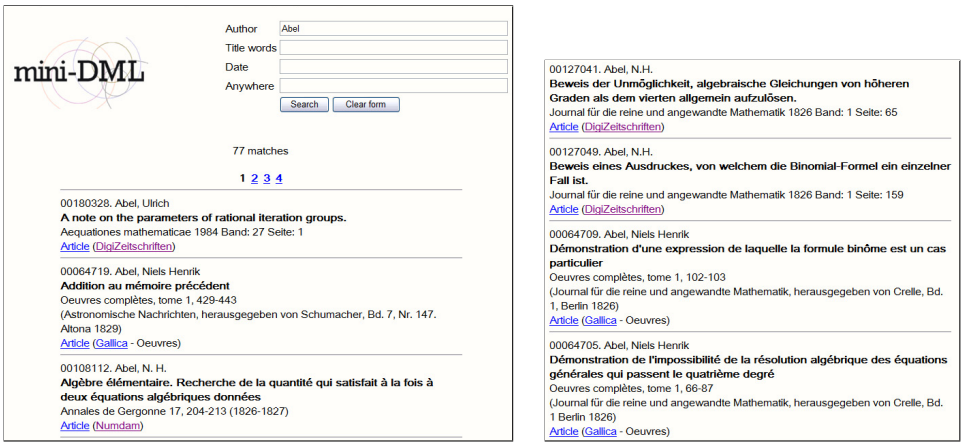


FIGURE 5. Some articles authored by Niels Henrik Abel, from various sources in the mini-DML.

century (1894-1912). It was meant as the list of articles published during the 19th century that were expected to be valuable to future research. Papers were grouped into files, sorted according to a specific classification scheme.

The MathDoc version, which amounts currently to half of the whole document, is a searchable database, which might be useful for various historical studies. Moreover, links to the actual articles have been added when available, possibly achieving the dream of Henri Poincaré: the RBSM will ultimately act as a gateway to the selected articles.

This work is a collaboration between Gallica, Paris (scan of the files), the *laboratoire de philosophie et d'histoire des sciences*, Nancy (structured keyboarding of the cards) and MathDoc (database, indexing, online interface).

3.5. **Integrating bits and pieces: the mini-DML.** We have seen that the comprehensive DML is not here, and won't be a reality before long. There are many reasons, some of them obvious, others quite obscure, that prevent us from building today a navigable web of the mathematical knowledge. However, to the naïve mathematician's eye, it seems that we can't expect more than a world where each stakeholder takes care of the collections it is or feels responsible for: publishers provide the best access and navigational features to their current offer, librarians and archivists populate secure and stable repositories, reviewing databases collect all relevant metadata, agregators of any sort provide instanciated views and personalised discovery tools over the collections. We can't believe that there will be some day one central access point to such a disorganised system, we even might be led to think that it is not the ideal situation when the actual usability of a fundamental resource depends upon a too restricted group (we could compare the robustness of the Internet, thanks to its web structure, with the fragility of the DNS system).

Since the announcement of the prototype for a mini-DML in 2004 [3], we have developed both the internal structure and the content so that we can consider it already as a useful service: one of those foundations upon which the actual DML will be built. This can be defined as unified indexation of articles available in digital format, taking advantage on the general dissemination of OAI-PMH

Bérard-Bergery, Lionel
Quelques exemples de variétés riemanniennes où toutes les géodésiques issues d'un point sont fermées et de même longueur suivis de quelques résultats sur leur topologie. *Annales de l'institut Fourier*, 27 no. 1 (1977), p. 231-249
Texte intégral djvu | pdf | Analyses MR 55 #9124 | Zbl 0309.53038 | 1 citation dans NUMDAM

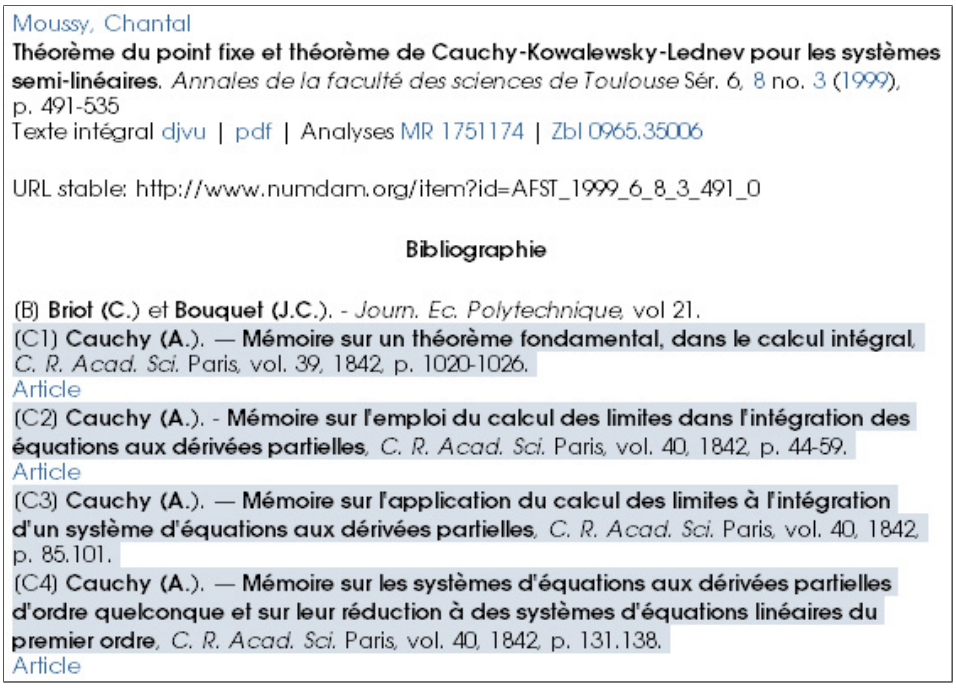
Bibliographie

(1) **R. BOTT**, **On manifolds all of whose geodesics are closed**, Ann. of Math., 60 (1954), 375-382.
Article | MR 17,521a | Zbl 0058.15604
(2) **J. EELLS** and **N.H. KUIPER**, **Manifolds which are like projective planes**, Publ. Math. I.H.E.S., 14 (1962), 181-222.
NUMDAM | MR 26 #3075 | Zbl 0109.15701
(3) **R. HERMANN**, **A sufficient condition that a mapping of Riemannian manifolds be a fiber bundle**, Proc. A.M.S., 11 (1960), 236-242. MR 22 #3006 | Zbl 0112.13701
(4) **W. KLINGENBERG**, **Manifolds with restricted conjugate locus**, Ann. of Math., 78 (1963), 527-547.
Article | MR 28 #2506 | Zbl 0117.38701
(4)• **W. KLINGENBERG**, **Lectures on closed geodesics**, (Livre en préparation) Zbl 0397.58018
(5) **S. LÓPEZ de MEDRANO**, **Involutions on manifolds**, Springer-Verlag, Berlin Heidelberg New York 1971.
MR 45 #7747 | Zbl 0214.22501
(6) **K.H. MAYER**, **Fixpunktfreie Involutionen von 7-sphären**, Math. Ann., 185 (1970), 250-258. MR 41 #6234 | Zbl 0181.27002
(7) **D. MONTGOMERY** and **C.T. YANG**, **Free differentiable actions on homotopy spheres**, Proc. of Conf. on transformation groups, New Orleans 1967, Springer-Verlag, Berlin Heidelberg New York 1968.
Zbl 0198.56801
(8) **H. NAKAGAWA**, **A note on theorems of BOTT and SAMELSON**, J. Math. Kyoto Univ., 7-2 (1967), 205-220. MR 37 #933 | Zbl 0162.53502
(9) **H. SAMELSON**, **On manifolds with many closed geodesics**, Portugaliae Mathematica, 22 (1963), 193-196.
Article | MR 37 #5819 | Zbl 0134.17704
(10) **S. SMALE**, **On the structure of manifolds**, Amer. J. of Math., 84 (1962), 387-399. MR 27 #2991 | Zbl 0109.41103

FIGURE 6. An example from NUMDAM, with direct links to JSTOR, NUMDAM, and Portugal's national library.

(Open access initiative protocol for metadata harvesting) technology. A simple search interface is up and running, providing seamless > access to more than 232 000 articles from 11 sources.

The mini-DML is not a competitor to the traditional reviewing databases, it is a register of the basic bibliographic metadata associated to mathematical articles in digital form, as provided by their publishers. It allows for one-stop searching over multiple collections: for instance, one can find the articles of the *Annals of mathematics* that have been spread at JSTOR, arXiv when it was an overlay journal, and project Euclid which now incorporates also the arXiv-EMIS period; one can at the same time find Euler articles from its collected works available at Gallica and the current English translation ongoing at arXiv; in fact, it is the only place where one can search *Portugalia mathematicæ* and obtain links to the article's full texts digitised by Portugal's national library. It provides also matching of bibliographical references, thus making it possible to add direct links from a reference to an electronic version of the text itself. This last feature is not yet public, but it is already in use at NUMDAM, and provides more than 36 000 direct article external links, which is slightly less than the number of direct NUMDAM links, so that about 21% of our references have already a digital article associated to them.


Moussy, Chantal
Théorème du point fixe et théorème de Cauchy-Kowalewsky-Lednev pour les systèmes semi-linéaires. *Annales de la faculté des sciences de Toulouse* Sér. 6, 8 no. 3 (1999), p. 491-535
Texte intégral djvu | pdf | Analyses MR 1751174 | Zbl 0965.35006

URL stable: http://www.numdam.org/item?id=AFST_1999_6_8_3_491_0

Bibliographie

(B) **Briot (C.)** et **Bouquet (J.C.)**. - *Journ. Ec. Polytechnique*, vol 21.
(C1) **Cauchy (A.)**. — **Mémoire sur un théorème fondamental, dans le calcul intégral**, *C. R. Acad. Sci.* Paris, vol. 39, 1842, p. 1020-1026.
Article
(C2) **Cauchy (A.)**. - **Mémoire sur l'emploi du calcul des limites dans l'intégration des équations aux dérivées partielles**, *C. R. Acad. Sci.* Paris, vol. 40, 1842, p. 44-59.
Article
(C3) **Cauchy (A.)**. — **Mémoire sur l'application du calcul des limites à l'intégration d'un système d'équations aux dérivées partielles**, *C. R. Acad. Sci.* Paris, vol. 40, 1842, p. 85.101.
(C4) **Cauchy (A.)**. — **Mémoire sur les systèmes d'équations aux dérivées partielles d'ordre quelconque et sur leur réduction à des systèmes d'équations linéaires du premier ordre**, *C. R. Acad. Sci.* Paris, vol. 40, 1842, p. 131.138.
Article

FIGURE 7. An example from NUMDAM, with direct links to Cauchy's collected works at Gallica-MATH.

There are still ways to explore in order to make the system more robust and more useful: we will need to standardise the minimal scheme used by DML partners to share metadata (this is work in progress where Cornell and Göttingen libraries cooperate with MathDoc), to document what we did so far so that others can start other similar ventures from our experience rather than from zero. As the DML grows, and the data comes in, we will need something like selfmatching in order to group duplicates (we already have quite some duplicates, for instance when articles appear in digitised journals as well as in collected works).

We are expecting to add many references in a short term, but there are many valuable article repositories that don't make their metadata easily retrievable. We are still in the process of trying to convince stakeholders that it benefits them to have as many ways of accessing their scientific content as possible, and that the best way to achieve this nowadays is to deliver enough of their metadata so that third parties can generate useful links to their resources, which in turn adds visibility thus value to the content they host.

3.6. **MathDoc services' URLs.**

CEDRAM	`http://www.cedram.org/`
GALLICA-MATH	`http://portail.mathdoc.fr/GALLICA/`
mini-DML	`http://minidml.mathdoc.fr/`
NUMDAM	`http://www.numdam.org/`
RBSM	`http://math-doc.ujf-grenoble.fr/RBSM/`
MathDoc website	`http://www.mathdoc.fr/`
MathDoc portal	`http://portail.mathdoc.fr/`

REFERENCES

[1] Pierre Bérard, *Documentation issues for mathematics in the digital age*, 68th IFLA Council and General Conference, August 2002, `http://www.ifla.org/IV/ifla68/papers/095-112e.pdf`.

[2] Thierry Bouche, Yves Laurent, and Claude Sabbah, *L'édition sans drame*, Gazette des mathématiciens **108** (2006), 86–88.

[3] Thierry Bouche, *Introducing the mini-DML project*, New Developments in Electronic Publishing, AMS/SMM Special Session, Houston, May 2004 & ECM4 Satellite Conference, Stockholm, June 2004 (Hans Becker, Kari Stange, and Bernd Wegner, eds.), FIZ Karlsruhe, 2005, `http://www.emis.de/proceedings/Stockholm2004/bouche.pdf`, pp. 19–29.

[4] ———, *A pdfLaTeX-based automated journal production system*, TUGboat **27** (2006), no. 1, 45–50.

[5] Keith Dennis, G. O. Michler, G. Schneider, and M. Suzuki, *Automatic reference linking in distributed digital libraries*, Conference on Computer Vision and Pattern Recognition Workshop, vol. 3, June 2003, p. 26.

[6] John Ewing, *Twenty centuries of mathematics: Digitizing and disseminating the past mathematical literature*, Notices of the AMS **49** (2002), no. 7, 771–777.

[7] Martin Grötschel, *Electronic publishing, intellectual property, and open access in mathematics: The position of the international mathematical union*, Slides of a Berlin presentation available at `http://www.ceic.math.ca/Information/031022MPGOpenAccess.ppt`, October 2003.

[8] Derk Haank, *Is electronic publishing being used in the best interests of science? the publisher's view*, Proceedings of the Second ICSU/UNESCO International Conference on Electronic Publishing in Science (Paris, 20–23 February 2001), 2001, pp. 127–130.

[9] Philippe Jacquier-Roux, *RUCHE, an editorial flow management tool*, `http://ruchedemo.cedram.org/`, 2006.

[10] Donald E. Knuth, *Mathematical typography*, Bull. Amer. Math. Soc. (N.S.) **1** (1979), no. 2, 337–372.

[11] Gerhard O. Michler, *Report on the retrodigitization project "Archiv der Mathematik"*, Arch. Math. **77** (2001), 116–128.

[12] ———, *How to build a prototype for a distributed digital mathematics archive library*, Annals of Mathematics and Artificial Intelligence **38** (2003), 137–164.

[13] Committee on Electronic Information Communication of the International Mathematical Union, *Best current practices: Recommendations on electronic information communication*, Notices of the AMS **49** (2002), no. 8, 922–925.

[14] DML project, *Final report*, `http://www.library.cornell.edu/dmlib/DMLreport_final.pdf`, October 2004.

[15] Eugénio M. Rocha and José F. Rodrigues, *Mathematical communication in the digital era*, Bol. Soc. Port. Mat. (2005), no. 53, 1–21, an English translation is available at `http://www.ceic.math.ca/News/digitization.pdf`.

[16] Laurent Rollet and Philippe Nabonnand, *Une bibliographie mathématique idéale ? le* Répertoire bibliographique des sciences mathématiques, Gazette des mathématiciens **92** (2002), 11–26.

[17] Christian Rossi, *De la diffusion à la conservation des documents numériques*, Cahiers GUTenberg **49** (2007), 47–61, `http://www.gutenberg.eu.org/publications/`.

INSTITUT FOURIER & CELLULE MATHDOC, CNRS-UJF, GRENOBLE, FRANCE
E-mail address: `thierry.bouche@ujf-grenoble.fr`