The Effort to Realize a Global Digital Mathematics Library

Patrick Ion(✉)

AMS, University of Michigan, Ann Arbor, MI, USA
`ion@ams.org` `pion@umich.edu`
`http://www-personal.umich.edu/~pion/`

Abstract. A decade after a resolution in 2006 by the International Mathematical Union endorsing the notion of a global digital mathematics library, and following a thorough report on possibilities written under the auspices of the US National Research Council in 2012, an 8-person Working Group, set up in 2014, is still working toward implementations of some of the ideas. There are difficulties with mobilizing the mathematical community toward building worthwhile infrastructure in times that are both perilous and well off, depending on where you stand. But progress continues.

1 Introduction and History

Vignettes from history help to emphasize the way that members of the mathematics community have long wished access to more of the world's mathematics to have a better understanding of it. The name of Giuseppe Peano (1858–1932) is universally well known among students of mathematics as the author of an axiom system for the natural numbers. That this remarkably inventive and productive Italian mathematician, who has recently been the subject of renewed historical interest [Dolecki&Greco:2016, Dolecki&Greco:2010a, Dolecki&Greco:2010b] also was a leading light of the push for the international auxiliary language Interlingua is less familiar [Peano:1903]. This can be seen in the mathematical context as feeding into the efforts at pasigraphy, a writing system where each symbol represents a concept. Indeed it was in connection with such efforts at representing mathematics by well-designed formulas that Peano developed his natural number axioms [Peano:1894].

At the ICM in 1897 there was a session under the chairmanship of Peano on how to encode mathematical knowledge. It remains striking to me that Ernst Schröder, the algebraist and logician from Karlsruhe, delivered a plenary address "On pasigraphy, its current status and the pasigraphic movement in Italy" [Schröder:1897]. By the last phrase he meant the work of Peano and his followers. Schröder disagreed with the distinguished chairman and suggested that the new system he had developed with a small number of basic symbols, somewhat in the vein that had been worked on by the American C. S. Peirce, was what was wanted. He began his remarks with the ringing statement that if there

G.-M. Greuel et al. (Eds.): ICMS 2016, LNCS 9725, pp. 458–466, 2016.
DOI: 10.1007/978-3-319-42432-3_59

were any topic that really belonged at an International Conference of Mathematicians, then it was pasigraphy. He was sure that pasigraphy would take its rightful place on the agenda of all succeeding such conferences.[1]

At the 1928 ICM in Bologna there was active discussion of how to provide comprehensive bibliographic resources for mathematics to everyone, especially in regard to publication of the comprehensive catalogue of mathematical work prepared by the German Georg Valentin (which was later bombed out of existence at Unter den Linden in Berlin in February 1942) [Valentin:1900, GöbelSperber:2010]. From 1931 on there were publications such as Zentralblatt [ZM], Mathematical Reviews [MR] and Referativny Zhurnal [RZ] indexing and abstracting the mathematical literature, but full text could only be had in the traditional way. In 2006 the IMU saw the possibility of realizing the imagined Global Digital Mathematical Library, or World Digital Mathematical Library, which had been discussed again [BallBorwein:2005,IMU/CEIC:2006, Jackson:2003], and put out a resolution to this effect [IMU:2006]. But that was all.

The adjective 'digital' is important here as it is the new digital technologies that allow better access to the resources of mathematical knowledge than ever before. Surely in 16th century Europe, or even earlier in 14th century Korea, when printing from metal type was a brand-new technology, the possibilities of new forms of books for the recording and dissemination of knowledge were welcomed by some. That they were right we all now know. That sort of opportunity is open to us again now. Of course, printing revolutionized in other ways: in Europe printing of indulgences was a labor-saving device for the church that may have changed religion, and printed money changed economics in China.

The adjective 'global' is important too. We all think the truths of our subject to be global, independent of location in this world.[2,3] Think of the IMU's membership from over 120 countries. A GDML can have a global reach as a result of digital technology, particularly the internet. It should be a shared global good. We are looking for global support for the idea and expect that there can be contributions to a GDML from all over the world. The benefits will be felt all over the world.

---

[1] Zur Diskussion auf einem internationalen Mathematiker-Kongresse dürfte sich kaum ein Thema in höherem Maße empfehlen als das der Pasigraphie, und ich bin berzeugt, daß der Gegenstand von der Tagesordnung künftiger Kongresse nicht mehr verschwinden wird.

Ist doch das Ziel dieser neuen Disziplin kein geringeres als: die endgültige Festlegung einer wissenschaftlichen Universal-Sprache, die völlig frei von nationalen Eigentümlichkeiten und bestimmt ist, durch ihre Konstruktion die Grundlage zur wahren, nämlich exakten Philosophie zu liefern!

[2] Naturally there are ethnomathematical considerations to be discovered, and different developments in mathematics to be related to sociology. Also Marcus and Watt point out that there are linguistic differences even in what is an equation [MarcusWatt:2012].

[3] See also a discussion by Gray of natural languages used for the expression of mathematical ideas [Gray:2002].

In 2010, IMU President Ingrid Daubechies and Peter Olver, chair of the IMU's Committee on Electronic Information and Communication (CEIC) took the initiative to work toward a WDML or GDML through consultations with a broad-based expert group [WDML Blog]. This culminated in a comprehensive report from a Workshop at the US National Academy of Sciences [NAS:2012]. At the 2014 ICM a small working group of 8 persons was given the task of working toward realizing some of the dreams [Seoul CEIC Panels, Seoul DML Panel]. The resources need to be found to begin building the GDML. It will not be an easy task [Bouche:2014, PitmanLynch:2014].

2 Challenges

One can distinguish four facets of the GDML initiative:

- The mathematical community
- The mathematical literature
- Mathematical knowledge management
- Management of the enterprise

They are all discussed in the NRC report [NRC/USA:2014]. Some parts of the GDML require work that is understood, or already done in part, but that just takes much time and effort to complete. Other parts require serious investigation and prototyping which will take time, although the general ideas and development paths may seem clear. But it turns out that each facet, though many would think their meanings pretty clear, leads a number of people to hold differences of opinion as to how each should be understood. It's in those details that the sticking points lie in realizing the public good we crave. Nonetheless, the WG and its successors will keep moving forward, probably with progress showing stick-slip characteristics that are hard to describe.

3 Goals

The goals of the GDML effort are still of considerable generality. To achieve them is going to be a longer process. Setting up any new organization for the public good which is to be sustained over a long period is a slow business. But we need to do it since the openness of mathematical knowledge to all who need it is very important.

To this end the GDML WG is founding an International Mathematical Knowledge Trust [IMKT], which it is hoped will function eventually as an organizing center for the knowledge base that will be the GDML. Setting out a charter for this proposed trust, and getting it endorsed by the IMU has been one of the tasks of the WG. It has proposed the present goals:

The purpose of the International Mathematical Knowledge Trust, IMKT, is to establish a mathematical knowledge commons — a public resource consisting of mathematical knowledge represented in non-proprietary, machine-readable formats and an international network of knowledge providers, information systems, and semantic services based on it, that is, a global digital mathematical library.

The mission of the IMKT is to construct a mathematical knowledge commons as a global public good, an effective knowledge base of open mathematical knowledge data, encompassing the world's mathematics through collaborations deploying both present and future technology, and to foster a supporting community. In particular, the IMKT should work to

- enhance accessibility of all mathematical knowledge world-wide, present, past and future,
- serve people in research mathematics, education, and applications of mathematics,
- promote the creation of open standards and best practices for management of mathematical data, and encourage the use of such standards,
- facilitate the development of open source tools and open mathematical data repositories,
- facilitate creation, dissemination and open archiving of semantically rich forms of mathematical data,
- encourage the collaborative development of open services based on representations of mathematical knowledge.

The mathematical knowledge commons resulting from these efforts of the IMKT and affiliated organizations should be a truly global resource, which matches the highest possible standards of independence, of reliability, and of data protection.

4 Achievements

Aside from working toward the problems of global existence and governance of an IMKT and associated international efforts, the GDML WG has been paying attention to building wider support for a GDML in the mathematical community and to planning and beginning projects that start constructing tools and structures to underpin a GDML. An activity of the support-building type was a Special Session "Mathematical Information in the Digital Age of Science" at the Joint Mathematics Meetings in Seattle Jan. 6–9, 2016 with 18 speakers over 11 h [JMM:2016]. This is a large gathering with this year about 6,300 registrants. These meetings are organized by the Mathematical Association of America (MAA) and the American Mathematical Society (AMS), and host additional sessions for the Association for Symbolic Logic (ASL), the Association for Women in Mathematics (AWM), the National Association for Mathematicians (NAM), and the Society for Industrial and Applied Mathematics (SIAM).

It was an excellent venue for speaking to the mathematical community on matters related to GDML efforts.

The materials collection efforts for a digital library, such as a world-wide extension of what EuDML did [EuDML], and building on their experience is being considered, but as yet have not been proposed to any funding agency. A European consortium has, however, begun the process of applying for significant funds under a European Union program.

The GDML WG helped organize a workshop to start on on of the infrastructure developments with long-range promise for a GDML. The Semantic Representation of Mathematical Knowledge Workshop was held 3–5 February 2016 at the Fields Institute, located at the University of Toronto, Toronto, ON CA [Fields:2016]. The support of a grant from the Alfred P. Sloan Foundation, staff and resources provided by the Wolfram Foundation, and staff and resources for local arrangements provided by the Fields Institute made possible a very successful workshop. The workshop was organized by the Wolfram Foundation, represented by Michael Trott and Eric Weisstein, by the Fields Institute and its Director Ian Hambleton, and by the GDML WG.

To paraphrase the application for support to the Alfred P. Sloan Foundation, the goals of the workshop were to pool the knowledge and experience of a group of experts to agree on design principles leading the way toward implementation of a semantic capture language applicable to all mathematics. Such a semantic encoding is expected to help realize one of the goals of a Global Digital Mathematical Library.

Through a program of talks and discussions the workshop was to work toward consensuses enabling the creation of a semantic language both for mathematics as a whole and for its sub-disciplines. The workshop was intended to produce (1) a white paper outlining the structure of the proposed semantic language, (2) a concrete plan for creating an explicit semantic language that will be used to mark up results in a specific area of mathematics, and (3) internet publications for all to see.

A great number of opinions were offered and a good deal of intense discussion ensued amongst the 40 participants of the workshop. Part 3 of the intentions is well covered by materials on the Fields Institute and Wolfram Foundation websites. However, production of a white paper and of definite plans proved difficult. Though discussions were started it turned out there was more to learn than expected. The first difficulty was the fact that there were several meanings of the word semantics in use, which could be seen as contradictory and were often confusing to one participant or another. The resulting short white paper is online, and linked to a wiki for further discussion [SRMwiki].

After term distinctions there are, for instance, possibilities for different views of the literature of mathematical interest, different levels of formalization, different kinds of semantically explicit mathematics, different audiences for mathematical communications, and even different levels of readability required.

When it comes to semantic capture language design there are language design issues, use cases to list and satisfy, issues of organizing design when a large vocabulary is presumably involved, and the whole matter of the relationship of semantic levels to formal proofs which can be machine checked.

There were some ideas for possible projects exploring the semantic capture aspect of recording the mathematical literature, whether old or newly created. One can try and accord various mechanized reasoning styles with each other, and tie them to more conventional computer algebra systems. There is lots of scope for exploring particular subjects and trying to capture, to various degrees, the special peculiarities of, say, algebra versus analysis, or geometry and probability.

One particularly attractive opportunity is the area of orthogonal polynomials and special functions (OPSF). This is a classical subject which is still being explored and has fascinating connections with other parts of mathematics. However, the basics of the subject are generally thought to be well understood. The fact that current computer algebra systems do not always agree on definitions, and so can produce strongly contrasting results to a fairly simple calculation, shows this is not so. The need for a concordance of special functions seems clear, and will be a good test case for proposed techniques of capturing mathematical semantics.

Another useful beginning will be to take a subject area and to try and capture all the significant theorems in it with a full set of definitions. A good prototype area might be geometry, since geometrical thinking can be contrasted with algebraic although much of classical geometry can be done by algebraic calculation; and there are sometimes insights to be gained into algebra from a geometrical view.

This leads on naturally to the matter of ontology creation. In the medical and life sciences use of mechanizable ontologies has proved useful. A lot of resources and attention have gone into the Gene Ontology, for instance. However, the use of the term ontology is also one fraught with possibilities for incomprehension, it seems. Indeed opposition to the idea that an ontology could be useful can provoke the sort of heat that political opinions are more known for. Nonetheless mathematics may be able to benefit from the practical experience and developments in other fields, where researchers have developed software that, *mutatis mutandis* could be useful for mathematics.

5 Prognosis

I expect to see the IMKT founded, and that it will work toward spreading the ideas of cooperation to achieve a GDML. We should see other regional mathematical Knowledge Trusts formed. There will come a wider awareness of what well-organized mathematical knowledge resources can bring both rich and poor communities. With better communication about how our subject's knowledge can be managed there will be a chance that it will not be lost to most people, as could happen for plausible commercial reasons. If the mathematical community can organize itself a little better, it can hope to avoid duplication of effort, and to achieve more. To a perhaps surprising extent many of the problems in implementing a GDML are social ones, though there are intellectual problems enough in trying to clarify what mathematical knowledge is and to make machinery to help us with it.

References

[BallBorwein:2005] Ball, J., Borwein, J., ACCESS: Who gets what access, when and how? In: MSRI Digitizing Mathematics Workshop (2005). http://www.mathunion.org/fileadmin/CEIC/Publications/MSRI.pdf

[Bouche:2014] Bouche, T.: The digital mathematics library as of 2014. Not. Am. Math. Soc. **61**(2014), 1085–1088 (2014)

[Dolecki&Greco:2010a] Dolecki, S., Greco, G.H.: Tangency vis-à-vis differentiability by Peano, Severi and Guareschi. J. Convex Anal. (2010). http://arXiv.org/abs/1003.1332

[Dolecki&Greco:2010b] Dolecki, S., Greco, G.H.: Towards historical roots of necessary conditions of optimality. Regula of Peano. http://arxiv.org/abs/1002.45811002.4581

[Dolecki&Greco:2016] Dolecki, S., Greco, G.H.: The astonishing oblivion of Peano's mathematical legacy (I). Youthful achievements, foundations, arithmetic, vector spaces, (13 pages). http://dolecki.perso.math.cnrs.fr/dolecki_greco_pro_Peano_I.pdf (Submitted). Dolecki, S., Greco, G.H.: The astonishing oblivion of Peano's mathematical legacy (II). Analysis and geometry, (18 pages). http://dolecki.perso.math.cnrs.fr/dolecki_greco_pro_Peano_II.pdf (Submitted). Dolecki, S., Greco, G.H.: The astonishing oblivion of Peano's mathematical legacy (III). Measure theory and topology, (15 pages). http://dolecki.perso.math.cnrs.fr/dolecki_greco_pro_Peano_III.pdf (Submitted)

[EuDML] EuDML - The European Digital Mathematics Library. https://eudml.org/

[Fields:2016] Semantic Representation of Mathematical Knowledge Workshop Videos. http://www.fields.utoronto.ca/video-archive/event/2053, http://www.fields.utoronto.ca/activities/workshops/semantic-representation-mathematical-knowledge-workshop

[GöbelSperber:2010] Göbel, S., Sperber, W.: Bibliographische Klassifikationen in der Mathematik : Werkzeuge der inhaltlichen Erschlieung und für das Retrieval. Forum der Berliner Mathematischen Gesellschaft. **12**, 77–99 (2010)

[Gray:2002] Gray, J.J.: Languages for mathematics and the language of mathematics in a world of nations. In: Parshall, K.H., Rice, A.C. (eds.) Mathematics Unbound: The Evolution of an International Mathematical Research, pp. 201–228. Providence, Rhode Island, American Mathematical Society and London Mathematical Society, London (2002)

[IMKT] International Mathematical Knowledge Trust. http://imkt.org/

[IMU:2006] Digital Mathematics Library: A Vision for the Future, International Mathematical Union (2006). http://www.mathunion.org/fileadmin/CEIC/Publications/dml_vision.pdf

[IMU/CEIC:2006] Some Best Practices for Retrodigization, International Mathematical Union (2006). http://www.mathunion.org/fileadmin/CEIC/Publications/retro_bestpractices.pdf

[Jackson:2003] Jackson, A.: The digital mathematics library. Not. Am. Math. Soc. **50**, 918–923 (2003)

[JMM:2016] AMS Special Session on Mathematical Information in the Digital Age of Science, 6-7 January 2016. http://jointmathematicsmeetings.org/meetings/national/jmm2016/2181_program_ss65.html

[MarcusWatt:2012] Marcus, S., Watt, S.M.: What is an equation? In: 2012 Proceedings of the 14th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC 2012, pp. 23–29. IEEE Computer Society, Washington, DC, USA (2012). ISBN: 978-0-7695-4934-7, doi:10.1109/SYNASC.2012.79, http://dl.acm.org/citation.cfm?id=2477708

[MR] Mathematical Reviews: MathSciNet. http://www.ams.org/mathscinet/

[NAS:2012] Symposium Wiki on The Future World Heritage Digital Mathematics Library. National Academy of Sciences (2012). http://ada00.math.uni-bielefeld.de/mediawiki-1.18.1/index.php/Main_Page

[NRC/USA:2014] National Research Council. Developing a 21st Century Global Library for Mathematics Research. The National Academies Press (2014)

[Peano:1894] Peano, G.: Formulaire de mathématiques. t. I-V. Turin, Bocca frères, Ch. Clausen (1894–1908)

[Peano:1903] Peano, G.: De Latino Sine Flexione. Lingua Auxiliare Internationale [1], Revista de Mathematica (Revue de Mathématiques), Tomo VIII, pp. 74–83. Fratres Bocca Editores: Torino (1903). https://en.wikipedia.org/wiki/Latino_sine_flexione

[PitmanLynch:2014] Pitman, J., Lynch, C.: Planning a 21st century global library for mathematics research. Not. Am. Math. Soc. **61**, 776–777 (2014)

[RZ] Referativny Zhurnal. http://www.viniti.ru/pro_referat.html

[Schröder:1897] Schröder, E.: Über Pasigraphie, ihren gegenwärtigen Stand und die pasigraphische Bewegung in Italien. In: Verhandlungen des ersten Internationalen Mathematiker-Kongresses in Zürich, pp. 147–162, vom 9. bis 11 August 1897. English translation in The Monist, vol. **9**, 44–62 (1899). (Corrigenda p. 320). http://www.mathunion.org/ICM/ICM1897/Main/icm1897.0147.0162.ocr.pdf

[SRMwiki] Semantic Representation of Mathematics White Paper and Wiki. http://www.fields.utoronto.ca//sites/default/files/whitepaper.pdf, http://www.wolframfoundation.org/programs/whitepaper.pdf, http://imkt.org/Activities/SemanticMathematics/Workshops/2016-02-03-Fields/whitepaper.pdf, http://imkt.org/Activities/SemanticMathematics/wiki/index.php?title=Main_Page

[Seoul CEIC Panels] http://www.mathunion.org/ceic/resources/icm-2014-panels/

[Seoul DML Panel] https://www.youtube.com/watch?v=OERXmv2oIyU

[Valentin:1900] Valentin, G.H.: Die Vorarbeiten für die allgemeine mathematische Bibliographie. Bibl. Math. **1**(3), S. 237–245 (1900) (JFM 31.0003.04). http://hdl.handle.net/2027/hvd.32044102938867

[WDML Blog] http://blog.wias-berlin.de/imu-icm-panel-wdml/

[ZM] Zentralblatt für Mathematik und ihre Grenzgebiete; zbMATH. http://zbmath.org