CEIC Meeting 2016

From notes by Patrick D F Ion — `pion@umich.edu`

February 20–21 2016, Chicago, IL USA

1 Business

1.1 Members present

JD: James Davenport [Chair]; ID: Ingrid Daubechies [EC liaison]; TC: Tim Cole [Local organizer]; OC: Olga Caprotti; CH: Carol Hutchins; RV: Ravi Vakil; PI: Patrick Ion; GT: Gerhard Telschow [IMU technician, by Skype].

1.2 Apologies

Thierry Bouche, Victoria Stodden & Masakazu Suzuki

1.3 Thanks

JD thanked TC very much for organising this meeting. Also GT for Skypeing in so late.

2 CEIC & IMU website

2.1 Website

JD The last news item on the Website is about 2014, more or less the last thing done. The minutes of last meeting weren't on site. How best to get material on there? GT said to mail him.

JD Put there the minutes of last as soon as agreed then. **JD** Now we know we should do more. How much of GDML effort to be put there? E.g., slides and videos at Fields. PI reported that slides and videos from the SRMKW at Fields are now on the Fields site at `http://www.fields.utoronto.ca/video-archive/event/2053`.

JD Under News it points just to the Newsletter archive. The room's question is whether there should be a more detailed link in?

There was then a lengthy discussion about the structure of the website and access, indexing and formatting tools . This was in the context of the decision to do a fundamental review of the whole IMU website by external consultants. It was agreed that JD should write to HH with our concerns.   **JD**

2.2   Analytics

The question had been brought up in the previous discussion about the usage of the website, and IMU usage in general.

GT I could ask for such data. We have 3 years of analytics. ID asked how we could look at the analytics. GT thought that OC could.

OC I didn't look at the analytics recently. I did check last year. I'll look up the matter now. I did work on the Task assigned me of investigating the archiving of the Twitter feed relative to the Congress: I did my homework and can say that PI sent a link he had to a Google script already doing this sort of thing. I was able to set it up with a cron job. It searches Twitter over the most recent 7 days (presumably using their old API) and it logs the results in a spreadsheet. There are also visualisers and a data explorer that look at hashtags and timelines.

However people have already started using the hashtag #ICM2018, so for historical reasons we should be moving fast. Therefore install it immediately since it only looks back 7 days each time. The ICM2018 organisers have already put out a newsletter. I can show an experiment. Even with the Advanced Search I get empty results for longer periods. Twitter itself does have the archive, but they seem gone for us for 2014.

JD I see three issues over this Twitter matter

- liaise with OC and harvest the ICM2018 tweets daily (which OC is doing already)
- CEIC recommends thinking about archiving as you (the IMU) introduce a hashtag
- researching the Tweets from ICM2014 as to their availability for IMU archiving

OC If I make you, Gerhard, a co-author on Google Docs then you'll be able to do it all.

It was agreed that JD should add analytics to his letter to HH   **JD**

3 EC Matters

ID There's not much to report. We have to give respectful advice on some issues. There's more helping of Helge to be done. Then there's the whole matter of the web site, which we've just understood is to be changed again.

4 GDML

4.1 Ingrid Daubechies

GDML restarted when I was President of IMU, and now asked to shepherd. Patrick Ion is point person here.

4.2 PI's report

The WG still consists of the original 8 members who have continued to participate with enthusiasm, each with appropriate interruptions from ordinary other commitments, since the WG's creation.

Telecons The WG has conducted its extensive discussions largely by telecons. These are typically on Fridays 10:00–11:00 EST (UTC -5) which suits the normal spread of member times from UTC +1 to UTC -8. [`http://www.timeanddate.com/time/map/#!cities=37,308,791,319,416,1268,784,1203,907`] The telecons are held over a telephone bridge generously provided by Wolfram Research Inc. We have actually held 57 minuted telecons, so a lot has been discussed in detail and some matters repeatedly.

GDML-wiki The WG has a private wiki at `http://www.mathontheweb.org/gdml-wg/` where draft documents and some pages of resources such as links and snippets are kept. In addition to the WG presently Ingrid Daubechies and Peter Olver have logons. [This can be extended to the CEIC as required.] The wiki has not been prepared for public viewing but is rather a collection of bits of information and trial balloons that help clarify discussions. There are about a hundred pages or so. In addition, the WG has made use of DropBox to exchange materials and of GitLab. The use of Git has as yet made more for overhead than simplification. Not all WG members have the same comfort levels with all our technological aids.

Face to Face The group has only had face-to-face meetings of subsets of the members, with others calling in remotely if possible. These were at the events CICM 2015 and SRMKW described below.

Website There was an early stub website produced at `http://mathontheweb.org/gdml/` but though some found it interesting the WG concluded that it was too clearly a Wikipedia ripoff and this was not continued. Another stub for the impending IMKT is at `http://imkt.org`.

IMU WDML pages The WDML presence on the IMU web site has not been enhanced as was certainly suggested long ago. The WG has been taken up with trying to make forward progress toward the big goal. The overhead felt in designing and submitting material at the IMU, which essentially has to be painstakingly produced in professionally perfect form before it can be submitted, since there is no easy way to correct an oversight quickly just proved to great. The idea of a blog that might carry along an interested community is an easy one to formulate. It is actually quite difficult to put into practice as it requires the attention of a suitable personality with a good deal of time to devote to it. The WG could not find that.

IMKT Since the inception of this work it has been the idea that a legal entity dealing with efforts toward a WDML should be set up. The initial effort emphasizes the 'World' aspect since the IMU is, like mathematics, very much an international affair. So the WG finally decided that Canada will be a good place to do this. It is not the US but there are possibilities for funding from Foundations in the US. The WG has prepared a proposed Charter which has been circulated to the CEIC members. The next step is submission of a grant proposal. The presently intended target is the Alfred P. Sloan Foundation which has made encouraging noises, for instance at the recent workshop they funded in Toronto at Fields [see below].

Project creation The WG has been considering many aspects of getting toward a Digital Library in its many discussions. However, a particular one that took much attention has been the creation of a semantic capture markup language. This matter has also been a particular interest of the well-known figure Stephen Wolfram. He has been very helpful, and sometimes impatient, with the rate of progress. The idea of a Workshop to work out problems associated with the design of a semantic capture language was floated early in the GDML's existence and encouraged by Wolfram who had previously done a project Funded by Sloan on capturing mathematical knowledge from the continued fraction literature, a project called eCF. The GDML WG collaborated with Wolfram and initially it was hoped that a workshop could be held the Sloan Foundation's headquarters in New York and involve 25-30 experts. Extensive background material was prepared by Michael Trott of Wolfram with the participation of Eric Weisstein, a WG member. The final preparation of an application for funds and sorting out of details took a long time, but was done mostly by Trott, Weisstein and Ion for the WG, who later effectively formed the organizing committee. Eventually an application describing a slightly more ambitious Semantic Representation of Mathematical Knowledge Workshop (SRMKW) at the Fields Institute, with as Principal Investigators Michael Trott (Wolfram), Ian Hambleton (Fields Director) and Ingrid Daubechies (Duke, for the GDML,) was submitted and generously funded by the Alfred P. Sloan Foundation. The workshop took place 3-6 February at Fields with 40 participants from a dozen countries. It was a very successful event. The local organization by Fields and the Wolfram

events staff provided an excellent infrastructure for an intense scientific program of presentations and discussions, and the weather was unseasonably pleasant. Fields put up a workshop web site where now there are videos and slides from almost all the talks, and there is background information public on the Wolfram Foundation Computable Mathematics site:

- http://www.fields.utoronto.ca/programs/scientific/15-16/semantic/

The SRMKW is also to put put out a White Paper from the meeting. This is in progress but may take a month or so. As an indication of the mathematical heft of the workshop, we had 2 Fields medallists, two Hilbert Problem solvers, the authors of the three biggest machine checked mathematical proofs (Four Colors, Odd Order Groups, Kepler-Hales) and other notable contributors including representatives of the two best known computer algebra platforms, and of the best-known theorem proving approaches.

Since Buchberger, Ion, Kohlhase, Watt, and Weisstein (5/8 of the GDML WG) were at the workshop they held sessions the evening after and the following morning going over IMKT documents rather faster than can be done in a telecon.

OutReach Though it is better yet to produce artefacts or services that serve the mathematical community, it is of value to the GDML effort to ensure the community's benevolent awareness and a positive image of it.

OPSF13A - NIST PI attended the 13th Conference on Orthogonal Polynomials, Special Functions and their Applications (the first such held in the US) June 1–5, 2015, at NIST, Gaithersburg MD. He presented a paper in their digital libraries session. He had hoped to be able to get a group of experts on the subject begin collaborating on making a concordance of their material as represented by DLMF (NIST), DDMF (INRIA), Mathematica and Wolfram Functions, Maple (all represented there) and others. Individual conversations would produce interest and a general willingness to collaborate, almost without exception. However, Wolfram, a key player, was represented not as expected by Eric Weisstein but by a (mathematically very well qualified) colleague whose capacity for annoying others was unrivaled at the event. As a result the hoped-for coalition building did not happen. Nonetheless, there is still hope that this community will be early contributors to enhancement of semantic representation of mathematical knowledge borne out by further contacts at JMM 2016 and SRMKW.

CICM 2015 - Washington DC This group of conferences, where much of the community in mathematical knowledge management regularly meets was attended 7–12 July 2015 by Pitman (an Invited Plenary Speaker), Bouche, Kohlhase, Watt, Weisstein and Ion, i.e. 6/8 of the

GDML WG, so face-to-face sessions were held on the Sunday evening preceding the conference proper, with attempts to get others in the phone.

JMM 2016 - Seattle WA Teschke, Watt and Ion, acting for the GDML WG, organized an AMS Special Session on "Mathematical Information in the Digital Age of Science" at the Joint Mathematics Meetings, 6–9 Jan 2016. The program was very well received and well attended. There were three periods totaling 11 hours over two days with 18 presenters (another 3 who could not be fitted in were scheduled in a later General Session). Tim Gowers, who was also the JMM's Colloquium Lecturer and had his own Special Session led off to a packed hall. David Mumford unfortunately had sustained a neck injury that prevented his attendance but provided a half-hour video presentation that began the second period with a stimulating discussion. The third session was started by Tom Hales. An idea of the interest of the talks can be had from the program listing at `http://jointmathematicsmeetings.org/meetings/national/jmm2016/2181_program_ss65.html#title` with links to abstracts. Watt and Teschke presented, but Ion was relegated to the General Session (by Ion who set the schedule).

The GDML expects to be visible at the following in the future, and PI also expects to attend.

ICMS 2016 - Berlin A track is being organized by Watt and Kohlhase both GDML WG members.

ECM7 - Berlin Teschke and Ion got accepted a short special session where the speakers will be Ingrid Daubechies, Bruno Buchberger, James Davenport and Jir Rakosnik.

CICM 2016 - Białystok Kohlhase is organizing a session there. The DML and MKM communities both make up a significant part of the CICM base.

The WG hopes therefore the CEIC will feel able to

1. Endorse the proposed IMKT Charter (after revision if needed)

2. Approve the GDML WG's activity so far

3. Welcome its continued effort to get GDML going, in particular by founding the IMKT which should replace the WG in leading such efforts

PI explained that IMKT wants to be at arm's length from the IMU but wants the IMU's endorsement.

4.3 Discussion

The following points emerged.

1. IMKT needed to be able to accept donations in several tax jurisdictions, and donations in kind as well.

2. Endorsing is important. Clearly the IMU will continue to appraise its relationship. It should want regular reports about how things are as long as endorsement is given.

4.4 Resolution

TC proposed, and the following resolutions was agreed

> Having discussed at some length the problem of benefits and its relation to the IMU as laid out in the charter, CEIC welcomes publication of the proposed Charter and recommends it to the EC as something that can be endorsed as in line with the vision expressed in the 2006 resolution.

JD will send such a note after PI has amended the proposed Charter appropriately.

Post meeting note. A revised IMKT Charter proposal was later prepared by PI and sent to ID who forwarded it to HH. PI also checked the GDML WG was OK with the tweaks.

4.5 Related Matters

CEIC congratulated the WG the work to date and looked forward to its continuation through an IMKT.

OC The IMU is an Associate member of the WDS and can participate in its Member Forum. That group is now moving on from Big Data to Broader Data, which is a form of Linked Data.

CH Quality control: is it dominated by certain disciplines? Discussions dominated by GeoData.

OC WDS does much work in publishing data. It's working on establishing workflow rules for data curation. This is a step forward for it's not yet clear how to do this. They're looking into procedures for certifying data repositories. This raises the question whether GDML is part of a bigger general picture or not? Does it impact what it means to be a WDS [Associate] Member?

There was a further discussion on WDS matters.

5   Item 5: Accessibility

JD Masakazu Suzuki can't be here but sent an email you've seen with some general points and a link to a new paper on this front.

PI Since Masakazu really just mentions the piece by Cervone, Krautzberger and Sorge[1] as a new thing on the accessibility front, I disagree. The paper is really a planned work programme.

JD
- I feel the IMU is doing nothing in this are and should be
  - This probably belongs with the CEIC
  - We should be prepared to recognize it's our problem.

RV Our opportunity.

JD Obligation even.

ID IMKT should be aware of this issue as well.

JD It's very largely a semantics extraction problem.

OC The main issue is there are many approaches. One thing clear is that reducing the cognitive load is a good thing. Conversely, deeply nested structures are hard to comprehend. It's an important problem.

Several Understandability!

JD It does say "This paper aims to stimulate discussion" and it did. This is clearly an important issue on which we do not have a consensus.

6   Item 6: IMU-Net

CH Citation Statistics continue to be discussed. I have talked to Martin Hubert Raussen. There's nothing more we have not discussed already.

7   Item 7: DOIs, handles and identifiers

TC It is something useful to have (for IMU documents) DOIs or handles.

OC We give no guidelines on how to cite etc.

CH Citation statistics of WordPress is often cited; cf. Web of Science and see that.

JD That we can settle overnight. Anyway in time for the March EC Meeting.

---

[1] JHD: presumably http://workshop.sciaccess.net/DEIMS2016/articles/p15_Volker_ARIA.pdf.

OC A Webinar for WDS: Herbert van de Sompel from Los Alamos [see, e.g.,`http://public.lanl.gov/herbertv/home/`, `http://lanlsource.lanl.gov/hello` https://www.youtube.com/watch?v=xYZEMTBo02E]

8 ICM2006 Web site archiving

TC I'm happy to say the Spanish ICM2006 web site is now archived in its entirety at the IMU thanks to the efforts of GT. This was not trivial for some of the linked material, like videos and so on, had first to be found.

CH A possible plan for the next 6 months might be to enhance its Accessibility (to the visually impaired , that is).

9 IMU-Net matters

Admin General discussion.

JD Content for this doesn't always have to be articles. Little bits of information would also be good.

CH 3 brief squibs.

JD Say for the May issue

- Announce the White Paper from the Fields SRM Workshop
- Announce the Accessibility Sate of the Art piece [post-meeting note: this did not happen: see conclusion to section 5
- Announce Minutes of the CEIC meeting

10 ICSU/WDS

OC We've discussed this already but there's one thing more: my replacement.

JD We have VS for this.

OC New this year is Member Access to the web site. Associate members can submit some things. I can pass on the credentials and login arrangements. I used it for uploading the poster.

11 Copyright etc.

JD We decided that the Wilfrid Hodges et al. paper should be updated and revised. This was to be done by VS and Ulf Rehmann. This has not happened. VS is unhappy to do a rewrite without the addition of a lawyer to the team.

PI I thought she was a lawyer (has a law degree).

JD It's a serious change to have a lawyer write this. A lawyer is always embedded in a jurisdiction and sensitive to peculiarities of that region. The idea was general guidelines and things for the publishing mathematician to consider in the process of publication. The need for a non-Anglo-Saxon perspective on this (e.g. from a legal system with Napoleonic history) was to be met by adding Ulf Rehmann in. We may have to restart with a new mandate.

CH The original report from Wilfrid Hodges, Van de Poorten et al. was really good for its time. Henry Cohn of Microsoft Research [NB, not Henri Cohen] is really very interested in copyright. Perhaps we could discuss with him, for another viewpoint, what we might do.

JD We could ask his advice. What we want is a decent job by a non-lawyer.

TC Yes a lawyer will tend to a conservative approach and may not suggest authors investigate some possibilities that others see.

ID Ask Thierry for a French perspective.

JD to contact VS.   **JD**

12   ICM 2018 Ideas

12.1   Formal Proofs

ID nice idea; but HH says outside remit but matter of program committee. We should send respectful note about such good things like the panels that worked in 2014 we'd be happy to work with PC or local organizing committee; do not wait until they think. of it; this is not interfering; Chair of PC is Janos Kollar; put it into context.   **JD**

12.2   IMKT

ID IMKT as in introduction of it is a second topic.

PI at SRMKW we had a very good encounter.

12.3   Other ideas

ID free software was suggested as a panel; there are legal issues: typically making available what shouldn't be.

JD I've just been elected a fellow of `software.ac.uk` that concerns itself with sustainability of software etc.

PI so no IMU Warez.

RV who runs ICTP math?

QQ Fernando Rodriguez Villegas `http://users.ictp.it/~villegas`.

ID could ask for a suggestion.

13 Closing

JD expressed the Committee's thanks to OC and CH, saying "you have to work another 10 months, but thanks a lot".

ID glad to have come and got some points covered.

JD thanks to Tim for local arrangements; ACCLAIM.

Next Meeting: around the same time next year, to be arranged by e-mail.