

IMU CEIC

Notes by J.H. Davenport

7–8 May 2011

1 Agenda Items

Present Peter Olver (Chair), Thierry Bouche, Olga Caprotti, James Davenport, Carol Hutchins, László Lovász, Ravi Vakil.

Posting (6a) LL raised the EC wish for us to discuss an IMU (not just CEIC) website, advising people about (in favour of) posting papers on the Internet: see section 2. OC mentioned that her EU project had a requirement (imposed as part of the programme) for Open Access. Others likened this to the U.S. NIH requirement. OC pointed out that this requirement forces her project to decide which conferences they would go to. It was agreed that the key requirement was the ability of authors to post the

unformatted but corrected (in particular post-referee reports) version.

It was argued that publishers supplied archiving, but this was debatable. JHD pointed out that *Journal of Symbolic Computation*, which had been Academic Press until they were taken over by Elsevier, suffered from the fact that the early Academic Press years were not available electronically at all, although they are now.

CH pointed out that no single body could provide secure archiving: entities such as Portico were important here. PO pointed out that the Library of Alexandria was a classic example of the dangers of single-source archiving.

LL noted that in Hungary, as a typical example of a smaller country, there is a national agreement. From that point of view, AMS (and CUP) are small players, and it was difficult to get even MathSciNet included in the agreement. CH noted that there was pressure in the U.S. against secret bundling agreements. TB pointed out that France was going the same way, but the Mathematics branch of CNRS was trying to include mathematics in these agreements.

It was also noted that, unlike a paper copy, a lapsed subscription may no longer allow access to back issues. TB said that Springer's contract with France included the back file, but not necessarily tools to access it.

PO pointed out that one journal had explicitly stated that it adhered to “best practices”, and that IMU/CEIC could publish a list of adhering journals. It wasn’t clear how this list would be ‘policed’ — responding to complaints was probably the only option. What happened if there were a large number of “borderline” complaints. It would occasionally happen that that editors would struggle to find referees who actually responded. CH pointed out that many mathematics journals did not adhere to the, fairly basic, requirement of publishing editorial policy. She had had a lengthy correspondence, which included having to cancel the journal, with one journal which had, apparently, an undocumented statement of electronic-only subscription availability and lacked license or terms of use statement. RV noted that we could arrange a public blog on the subject.

PO said that a key question was how much work was involved in policing and moderating any such initiative. CH pointed out <http://retractionwatch.wordpress.com/>, apparently set up by a couple of individuals. It was noted that there two retractions linked to *Applied Mathematical Letters*, a journal which had been positively rated by the Australian system.

JHD pointed out that a list of journals which had explicitly told us (e.g. <mailto:journals@mathunion.org>) that they adhered to Best Practices was feasible. We could also do random sampling of the list (guaranteeing a full audit might be more troublesome). Policing the list (in terms of responding to complaints etc.) was more difficult, but could be set up. There was an issue of how we told journals that we were setting up such a list, but JHD thought we should contact them via, say, MathSciNet, and this would probably be effective in practice. LL pointed out that we probably would need EC approval for such a move. PO thought was we should probably have a two-phase approach — public comments on a Facebook (or Facebook-like) site, and a more formal complaints list.

The Committee noted that coverage of mathematics-specific publishers (specifically the London Mathematical Society) was patchy in sherpa.ac.uk, and this should be chased up.

JHD

CEIC Web (3) Following from the item 6a discussion, TB asked whether we had a concise statement of “Best Practices”. PO pointed out that this was part of the whole issue of CEIC web pages, which contained a large amount of obsolete information. He would like to streamline the web pages, possibly moving a lot of material to an archive. CEIC web pages had been housed at Dalhousie, but have now been moved to Berlin¹ JHD proposed that we start from a blank sheet, rather than asking what was obsolete, and this met with general agreement.

PO/all

The following items should probably be included (see also Appendix B) .

¹Regrettably the Berlin site seemed to be inaccessible at precisely the time of the CEIC Meeting.

- Membership and Terms of References
- (current) Best Practices (certainly Journals, probably Digitization, and some others)
- Minutes etc.
- Endorsed links, such as `www.sherpa.ac.uk`.
- News of some variety, e.g. “IMU on the web” or its replacement, possibly blogs.
- (Pointer to) archives and past documents.

It was generally felt that the “directory of mathematicians” items (FWDM, EWDM etc.) were “past their sell-by date”.

WDML was a more complicated question. The most recent “news” on the website is 2006. It was noted that it was the result of an unfunded vision of a top-down project. In fact, there has been ‘bottom-up’ evolution of archives. TB noted that the situation was comparable to journals: we could have a list of WDML-compliant archives. The EuDML project was trying to organise such a list at the European level. The meaning of “library” was itself a challenge. CH noted that the library community tended to refer to such initiatives as “aggregators”. RV thought that we could have a page of

- Vision, best practices etc.
- Compliant archives
- Other archives
- Commercial sites, such as JSTOR, ScienceDirect etc.
- A mechanism to add others.

TB noted that the question was “where to stop”, e.g. individual universities, departments, private home pages? LL wondered whether suitable metadata was a partial answer — TB noted that Google Scholar worked this way, as did Mendeley (but a different mechanism) and others.

What do we mean by the “reference version”? It should be capable of supporting unambiguously

The first theorem on page 10.

RV noted that the arXiv supported this *provided* one said “archive version 7” or whatever. Of course, the version(s) in arXiv is not the journal version. DOI is a unique identifier *where applicable*.

The Committee reminded itself that Ulf Rehmann’s private list² was essentially the most comprehensive, even if not easy to search/use. RV pointed out that having a link to this would be far better than doing nothing.

²http://www.mathematik.uni-bielefeld.de/~rehmann/DML/dml_links.html.

The general conclusion was that, from the “start from scratch” CEIC page, there should be a “start from scratch” WDML-like page, including at least such links. PO said that he would start the initiative of “start from scratch” web pages, and circulate a non-public link to the committee.**PO**

Websites We noted MG’s comment that the website should be moving from ZIB. He had raised the question of redesigning the IMU website. LL noted that the pages had improved *in content* over the last few years, but were still not “attractive”. The phrase “lacked humanity” was used. Why didn’t they link to, for example, the blogs of Fields Medallists (the set is non-empty: Tao and Gowers).

Security LL commented that he had been on the Abel Committee, which used a central server, and the Fields, which use ordinary e-mail. JHD noted that the archiving committee had suggested secure e-mail, but MG’s experience was that he found this sufficiently difficult to use that MG felt that it was a non-starter. JHD’s view was that the only viable method for secure discussion available today was via a central server. RV noted that Google Wave apparently did all this. Though this was being discontinued, it was an existence proof that such technology exists.

The question had also been asked about the use of electronic voting for the IMU itself (i.e. by the constituent societies). JHD argued that the issue was the electorate, not the mechanism, and this was a human issue not a technological one. The Committee shared this view. **MG**

Committee e-mail This seemed to be working. CH thought there had been *one* spam, but others weren’t sure. JHD felt that it was probably time to check the archiving, and others agreed (that he should). **JHD**

World Data System Waiting for them to get back to us. **WDS**

Mathoverflow RV demonstrated www.mathoverflow.net. This opened in 2009. Questions are tagged out of a mixture of arXiv terms and free-text. The system of “badges” does seem to work for those who are interested in such things³, and possessing more badges does increase “rights”. Self-policing does seem to work here: probably more so than Wikipedia. The software is provided by a company, but they do *not* own the data, and provide a dump every month. It is possible to be anonymous, or to have an account (which can use Gmail details). As with the earlier days of the arXiv, usage is not uniform across fields of mathematics.

³RV noted the links with gaming sociology.

Facebook LL reminded us that there had been a Facebook account started for the ICM of 2010. See ID's mail of `Mon, 18Apr201123:01:21-0400`. **ID/MG**
OC is following ICMI and MAA Facebooks. These are clearly active (\geq one post/day). RSS feeds, such as UR's on recent digitizations, can be pushed to Facebook. She was in favour of the IMU's following suit, as it reaches a different community.

Twitter It was felt that it would make sense for the same people who handle Facebook to handle Twitter, as the media are very similar, and it is possible to 'push' tweets to Facebook. **ID/MG**
It was noted that "replying" doesn't necessarily have to be done. It is possible for many people to share an "IMU" account. Twitter is used for disseminating information about debates, rather than being used for debating itself. Twitter was likened to a quick press release. CH pointed out that there are widgets which copy from Twitter or Facebook to the home page. The key question is who is responsible for IMU communication to the outside world. **EC**
While on the subject of "communication", it was noted the Wikipedia article on IMU is pretty sparse and should be updated. **MG**

IMU-Net/Notices TB noted that he is on the Editorial Board of the publishing column of Notices of the AMS. IMU-Net is a collection of comments from EC, CEIC and CDC. The CEIC component has historically been named "IMU on the Web" (despite the fact that it is *not* on the web as such). PO noted that it would be nice to get statistics on *who* is receiving IMU-Net, e.g. by geographical distribution. **technician**

MCM's (editor of IMU-Net) view is that contributions have to be very concise, which has made more technical contributions difficult. For example Terry Tao's speech had to be shortened for the July 2010 IMU-Net (No. 42).

CH noted that we have run out of the previous list of suggested topics, and that she herself is running dry. IMU on the Web's rôle could be seen as

interesting things/developments in CEIC's remit.

- Kuperberg's idea that a "refereed subset of arXiv" could replace journals as they current exist. It was noted that this was controversial.
- It was noted that, while the original remit was to introduce mathematicians to the web, new tools (and hardware, e.g. Kindle) are appearing all the time, and should be described.
- OC had mentioned one example of this as `paper.li` (producing a newspaper from Twitter feeds). See also `news.me`.

- CH: finding appropriate (and reasonably priced) textbooks. In particular previous permissions to redistribute paper copies no longer apply to the electronic age. JHD/TB noted that laws differ substantially, but even an anglo-saxon article would be useful. It was agreed that CH should tell the story, noting that implications might be different in different legal systems.
- Polymath. We should ask Gowers. LL
- MathJAX (especially because of the accessibility issues). OC

- CH recalled that MG had asked about an organised approach to improving the mathematics in Wikipedia. JHD was dubious about an organised approach, noting that
 - Wikipedia is generally pretty good;
 - if it's not, he fixes it, and this is a fairly general ethos.
 - In his view, contributing to Wikipedia falls into the same “service” category as refereeing, conference organising, etc.
 - It was noted that the Wikipedia page devoted to the IMU is rather sparse and in need of an editor.

A dialogue should take place. JHD/RV

- CH raised the issue of “projects”. We could maybe use an “interview” format. LL mentioned Bill Barton’s project on updating Klein’s *Elementary Mathematics from an Advanced Standpoint* — kleinproject.org.
- Mendeley, OC noted that it has useful tools, such as BiBTeX support. CH knows the founder and CEO. CH
- CH raised Google Book Search. There is still on-gong litigation.

PO noted that coordination with *Notices* is still unresolved. Should this be systematic or *ad hoc*. *Notices* length requirement is c. 1200 words, whereas IMU-Net is under 600.

CH drew attention to the e-mail from Christiane Rousseau. Some concern was expressed that *Notices* wanted first publication rights. It was noted that it didn’t make much sense to publish the longer version first. PO should engage in more correspondence with *Notices*. PO

Web publication LL reminded the Committee that EC wanted CEIC to produce an IMU document on the importance of posting all new mathematics on the Internet. There is an existing CEIC recommendation in this area (see Appendix A and http://www.mathunion.org/ceic/Publications/Recommendations/6_call.shtml).

WDML While the discussion above covered the IMU/CEIC pages describing WDML, LL noted that we should still write a document explaining archiving of web publications etc., and its relationship with tools such as Google Scholar.⁴ TB noted that NumDam had to create a special interface to Google to let it crawl their resources, but this had been obsoleted by a change of policy at Google. It is now the case that all NumDam papers have, as “abstract”, the NumDam copyright statement. TB noted that initial estimates, 10 years ago, were that there were 2.5M mathematical articles: the estimate today is 3.3M, with almost all of the new ones “born digital” — all too often “born badly digital”, both as regards the text (bad pdf \TeX) and the metadata (or lack thereof).

OC asked whether it was possible to automate production of metadata, e.g. via a \LaTeX style file. TB has one, but it is based on TRALICS (their \TeX →XML system). CEDRAM uses an extension of the `amsart` style: in particular it splits up the author name as first/last. LL pointed out that his pre-publication versions (obviously) do not include the publication data, so he would need to re-edit the source to include this. Hence the work-flow is not obvious.

mini-DML is a federated service with 356678 articles from 15 sources. This doesn’t include JSTOR, basically as a consequence of their anti-deep-linking policy.

TB communicated estimates for the cost of long-term archiving, which mean that “born digital” was much cheaper than digitized.

Social Media RV used this phrase to cover Facebook, Twitter, Wikipedia etc.

In general organisations give this to secretaries and staff members, and the result is “corporate bland”. Instead of this, let’s appoint the equivalent of a “poet laureate”, and some-one for a year (say) to talk about (not formally represent!) the IMU in social media. LL noted that the ICM 2010-related Facebook *had* used the official IMU logos, which he thought should be reserved for “official” statements. Perhaps we need an “IMU-light” logo, e.g. the original with a section removed. This was generally thought of as a good idea.

There were questions on how all this would relate to the wider IMU publicity, and in particular ICM publicity. CH felt that PR/journalism devoid of social media was not the way forward. It was noted that *much* of the publicity for the ICM is local/national, but not all. Hence the conclusion was that

IMU needs both a social media person (say one-year rotating)
and an overall communications head (4-year elected, possibly
with a Head Office assistant).

⁴JHD’s notes. For example, author names in Google Scholar often have extra letters affixed which are in fact references to footnotes. There are also problems with the version of pdf \LaTeX , which may well generate non-standard representations of ligatures etc.

EC

President's business ID had noted, following her stay in Madagascar, that access to equipment and the internet is very unevenly distributed, and had suggested a joint CDC/CEIC working group. It was not clear what the IMU, or, say, AMS, could do by itself, but possibly an approach to charities such as the Gates Foundation would be more profitable. CDC should probably lead on this, but CEIC should provide technical advice. JHD has had experience of such initiatives in the former Soviet Union, and would be willing to assist.

EC/CDC

It was also noted that ICTP has a strong mandate in this area, and CIMPA in France. These organisation are represented on CDC.

Software availability and information is also unevenly distributed. It was noted that <http://www.ams.org/mathweb/mi-software.html> is a reasonable list, though incomplete. JHD has mailed some suggestions, but other areas, e.g. optimization, need more work. There's also a list at <http://www.zib.de/en/services/software.html>.

Terms of Reference No action required.

ICM 2014 It was noted that CEIC had been notified about ICM 2010 very late in the day, by which time publication arrangements had been made. It was noted that traditionally local organisers had had autonomy in this area, but that this was less viable in the electronic age. CEIC believes that IMU should state its overall requirements, e.g.

- open access to the proceedings within a reasonable time frame (e.g. 2010's two year period)
- immediate broadcasting, and publishing, of plenary lectures
- archiving of these items by IMU as well as any local arrangements.

EC

Bulletin IMU It was noted that this is a formal publication, sent to member societies.

Miscellaneous It was noted that there will be another (fourth) DML workshop in Bertinoro(Italy) in July 2011. There is also a plan for a "WDML" conference in Berlin. Items such as these should be linked from the (new-style) web pages.

Next Meeting Roughly in a year's time. Berlin (Head Office) would be a good venue.

2 New Call for Availability of Research Mathematics

Open access to the mathematical literature throughout the world and irrespective of financial means is an important goal. Each of us can contribute to that goal by making available electronically as much of our own work as feasible.

CEIC believes that authors *should* post, and be free to post, the final author-generated version of their papers, taking into account any referees' comments, and possibly using the publically-available publisher style files. Most publishers accept the benefits of such postings, but CEIC notes that publishers' actual conditions vary. A fairly authoritative list of these can be found at <http://www.sherpa.ac.uk>. CEIC recommends that authors take into account publishers' policies on this when choosing their outlets. In particular, CEIC suggests that authors avoid publishers that do not allow authors to post versions taking account of referees' comments.

The lesson of the Library of Alexandria shows that posting one copy is not wise. We would recommend at least three copies: personal home page, institutional/ funding agency repository and an independent indexed archive such as <http://arxiv.org> or <http://hal.archives-ouvertes.fr>.

2.1 Way Forward

The IMU EC wishes this to become a joint IMU/ICIAM declaration.

now–22May CEIC to mull this over: comments at least to JHD.

23 May JHD to forward to ICIAM's nominated representatives: P.N. Srikanth (<mailto:Srikanth@math.tifrbng.res.in>) and Dinesh Singh (<mailto:dsingh@maths.du.ac.in>).

mid-June Circulate to IMU EC and within ICIAM.

September2011 Target date for publication.

A Original Call

Call to All Mathematicians to Make Publications Electronically Available

Endorsed by the IMU Executive Committee on May 15, 2001 in its 68th's session in Princeton, NJ.

Open access to the mathematical literature is an important goal. Each of us can contribute to that goal by making available electronically as much of our own work as feasible.

Our recent work is likely already in computer readable form and should be made available variously in TeX source, dvi, pdf (Adobe Acrobat), or PostScript form. Publications from the pre-TeX era can be scanned and/or digitally photographed. Retyping in TeX is not as unthinkable as first appears.

Our action will have greatly enlarged the reservoir of freely available primary mathematical material, particularly helping scientists working without adequate library access.

(http://www.mathunion.org/ceic/Publications/Recommendations/6_call.shtml)

B RV's Suggestions for CEIC Web pages

**** Proposed webpages for CEIC ****

*** Front page ***

Links to: each of the following pages
list: current membership of committee

*** Journals, and public access of research mathematics ***

(1) tell people to post (new and old), and link to some possibilities

(2) best practices for journals
link to sherpa

Subsection: Digital mathematics library
State the goal. Metadata. Give link

*** Links for people in under-resourced places ***

software repositories:
link to digital libraries
link to AMS site
link optimization site
link to latex users group
how to get pdf reader

*** Archives ***

Link to all old pages
List of meetings, minutes, etc.

C JHD's notes on web-site archiving

[We didn't discuss this, as the web site was down]

JHD's web site monitoring expert makes the following distinction.

Client-side archiving This is designed to show what the web-site *looks like*.

There are various ways of doing this — apparently Acrobat Acrobat (standard, not reader) can do this: http://help.adobe.com/en_US/Acrobat/9.0/Standard/WS58a04a822e3e50102bd615109794195ff-7f67.w.html.

The British Library will apparently do this for sites it considers to be “important British sites”, and the German equivalent (DNB) may well do the same. The advantage is that you get to preserve the *look* of the site: the disadvantage is that such archiving doesn’t descend into Content Management Systems (CMS) well.

Server-side archiving This is essentially archiving the site as a set of files, or the CMS. There are “open” (JHD is not sure how practical this is) standards for such archives which should mean that the web-site is re-creatable, but some of the presentation (contained in the CMS interface) may be lost.

Most professionals apparently recommend both. JHD suggests we ask the Head Office what their plans are, and check whether the DNB will archive mathunion.org.