

IUPAC Database Enhancements

Teleconference
10:07am - 11:00pm EDST
October 29, 2015

Draft Minutes

IUPAC Attendees: Colin Humphris, IUPAC Secretary General, Lynn Soby, IUPAC Executive Director, Bonnie Lawlor, Chair, IUPAC Committee on Publications and Cheminformatics Data Standards (CPCDS)

De Gruyter Attendees: Katharina Butsch, Karin Sora

The meeting opened at 10:07am EDST. The objective of today's teleconference was to discuss possible enhancements to the IUPAC Standards and Recommendations database. The enhancements will be introduced in 2017, but Karin Sora needs to budget for the preparation work that will need to be done in 2016.

Status of Phase One

Katharina Busch said that the contractor who is responsible for the conversion of articles from PDF to XML will have everything turned over by the end of November. So far the content looks good. Colin Humphris asked if the font conversion problem discussed in Busan, Korea with regards to *Pure and Applied Chemistry* will also impact the database. Katharina said that they have looked into that issue and the problem will not impact the database because there will not be a font conversion. Bonnie Lawlor said that Ron Weir had been brought into the discussion and he agreed with Katharina. At present, everything is on schedule for phase one.

Database Enhancements

Bonnie Lawlor noted that this meeting was called to identify potential database enhancements that would be released in 2017. The work on those enhancements would begin in 2016 and therefore De Gruyter needs to budget for the work required. Bonnie expressed concern that without customer feedback it is difficult to determine the additional content/features/functionalities that users will want to see implemented. However, she went through all of the database planning meeting minutes and pulled out the potential enhancements that CPCDS members discussed. These are as follows:

- Properties (meeting of 12/8/14: p.3)
- Methodologies (meeting of 12/8/14: p.3)
- Standard identifiers and links (meeting of 12/8/14: p.3)
- Tracking of the use of IUPAC Standards (meeting of 12/8/14: p.1)
- Retrieval of tables/captions/charts within the context of the article (meeting of 12/8/14: p.2)
- Substances (should be connected to InChi-Codes and can be considered as a long term option)
- Thesaurus (user searches for a term (s)he is used to & is redirected to the IUPAC term in the database, if necessary; meeting of 12/8/14: p.1 and meeting of 5/13/15, p. 5)
- Cross linking between database entries ((meeting of 1/22/15: p.3)
- Citation linking (CrossRef; meeting of 8/11/14, p. 3)

- Clickable browsing structure for each entry (meeting of 12/8/14: p.1)
- Element search by a clickable periodic table (meeting of 8/14/14, p.2)

Karin Sora said that they are very interested in ongoing enhancements to the database thesaurus. Any changes will be invisible to the user and simply work in the background to improve the search function. The second enhancement that they would like to add is substance information; e.g. InChI codes, CAS registry numbers, etc. They need IUPAC approval to do so and would like input on the sources from which such content can be reliably retrieved. They are also interested in having a periodic table that could be used to retrieve information on specific elements.

Bonnie said that from her personal perspective all three of those enhancements would be valuable, but she does not know offhand from whom InChI codes and CAS registry numbers can be obtained for free. Karin Sora said that originally Springer retrieved InChI codes from the Internet, but they have since acquired InfoChem (Peter Loew's company) and can now use their InChI codes. Karin said that Peter is willing to help, but she is concerned about the cost because Springer is indirect competition with De Gruyter. Bonnie said that perhaps the InChI Foundation could help and Colin Humphris noted that IUPAC has representation on the Foundation's Board. He said that Richard Hartshorn who will be the IUPAC Secretary General starting in 2016 is the IUPAC rep to their Board and perhaps he can help. Colin will talk to Richard since he will see him at the IUPAC Executive Committee meeting this weekend.

ACTION: *Colin Humphris* will talk to Richard Hartshorn to see if he can suggest a no/low cost source of InChI Codes.

Colin said he has no idea where CAS registry numbers could be obtained for no/low cost (the American Chemical Society has a fee their CAS numbers). Karin said that De Gruyter can get them from the Internet using the IUPAC substance name, but would not do so without IUPAC's consent. Bonnie asked what Internet sources would be used and Karin said Open Access sources such as the U.S. National Library of Medicine's PubChem, the Royal Chemical Society's Chem Spider, etc. Bonnie said that her personal opinion is that those are fairly reliable sources.

Karin also noted that they believe that the ability for users to access "data sheets" that would contain substance information (molecular formulae, 3D structures, InChI codes, CAS registry numbers, etc.) would also be a nice feature. Other databases such as Springer Materials offer this (see: <http://materials.springer.com/periodictable>). We could discuss what kind of information IUPAC would want to include on such sheets (Karin noted that Springer has changed the look and feel of Springer Materials and it is not as good as in the past, but it still demonstrates what could be done).

ACTION: *Bonnie Lawlor* will ask members of CPCDS if they know of any reliable sources for CAS Registry numbers and InChI codes.

Lynn Soby asked about how De Gruyter sees the Periodic Table feature being used. Katharina said that the user would click on an element and retrieve all of the articles in which that element is mentioned. The user could alternatively do a keyword search, but retrieval would require that the element be spelled out in the article. They see the Periodic Table as a feature of the database, not as an open access source. Karin added, however, that the table could be "outside" the database as well, but that the user would retrieve limited information. They would have to subscribe to the database to get all of the relevant information from the articles in the database. The Periodic Table could also be used to access free information that is not in the database such as the data sheets mentioned above.

Lynn said that the Royal Society of Chemistry (RSC) has a very nice interactive table freely-available on their website and she uses it frequently to access information on elements (see: <http://www.rsc.org/periodic-table>).

Colin Humphris said that IUPAC has a special affiliation with the Periodic Table and is considering having one as part of its new website. Lynn said that in some ways the Periodic Table is an “inherent brand” for IUPAC. She agrees that it could be used as described by De Gruyter to enhance the searching of the Standards and Recommendations database in order to retrieve articles from that database, but any use beyond database searching would require a separate conversation. It would also be an innovative marketing tool for the database - giving people a “taste” of what they could retrieve if they subscribed.

Colin said that we probably need to give more thought on this to clarify/understand how the periodic table will be used by IUPAC on its website and how it will be used to access the database. Karin said that she will send some screenshots as well so that everyone can see how they envision the datasheets.

Colin then summarized the discussion regarding enhancements for 2017. IUPA supports:

- Thesaurus improvements
- The addition of substance information such as CAS registry numbers and InChI codes, but some work is needed to identify reliable sources
- The use of innovative visual tools such as the Periodic Table to enhance the search function, but we need to consider how IUPAC will want to utilize the Periodic Table on its own website

ACTION: *Katharina Butsh/Karin Sora* will forward screen shots to assist in the visualization of some of the ideas that they were discussing.

ACTION: *Karin Sora* will ask that Spencer McGrath forward the third quarter revenue figures and the projection for year-end figures for the IUPAC journals.

ACTION: A follow-up meeting will be scheduled before year-end (preferably in November) to continue today’s discussion.

Karin closed the meeting by reporting on the sad news that Katharina has resigned and will be leaving De Gruyter at the end of the year. Everyone expressed their dismay at Katharina’s departure, but expressed their best wishes for her future success. De Gruyter is in the process of seeking a replacement and Karin said that any recommendations for potential candidates would be most welcome. Bonnie said that she will send an email to her Chemical information peers in ACS to see if anyone is interested or knows of anyone who might be interested. Karin will send Bonnie the link to the job description.

There being no further discussion, the meeting was adjourned at 11:00am.

Respectfully Submitted,

Bonnie Lawlor, Chair
IUPAC Committee on Publications and Cheminformatics Data Standards (CPCDS)
November 23, 2015



INTERNATIONAL UNION OF
PURE AND APPLIED CHEMISTRY

(CPCDS) Committee on Publications and Cheminformatics Data Standards