



INTERNATIONAL UNION OF
PURE AND APPLIED CHEMISTRY

(SCDS) Subcommittee on Cheminformatics Data Standards

Subcommittee on Cheminformatics Data Standards (SCDS)

(Subcommittee of the Committee on Publications and Cheminformatics Data Standards)

Inaugural Meeting

October 25, 2016, 11am - 12noon US-EDT

(Web conference)

Minutes

SCDS Members present: David Martinsen (co-chair), Leah McEwen (co-chair), Ian Bruno, Antony Davies, Robert Lancashire, Andrey Yerin

1. *Review of Subcommittee Aims:*

The Subcommittee on Cheminformatics Data Standards (SCDS) will communicate within IUPAC the cheminformatics data standards needs of the chemical community with the objective of prioritizing and efficiently meeting those needs through the collaborative efforts of relevant IUPAC Divisions, IUPAC Committees, and external scientific organizations, such as NIST, CODATA, and the Research Data Alliance. SCDS will widely promote this work of IUPAC both internally within IUPAC and externally to the global chemical community through symposia, workshops, presentations at relevant conferences, and articles in *Chemistry International (CI)* and other appropriate publications. To ensure broad communication, SCDS will make its reports available to all IUPAC Divisions and Committees as well as to its external collaborative partners.

2. *Meeting schedule*

The Subcommittee will plan to meet monthly at a similar time.

ACTION: *Leah McEwen* will send a doodle poll to set up the next meeting time.

3. *Review of SEDS, the former Subcommittee on Electronic Data Standards (SEDS):*

Robert and Tony discussed SEDS work on the JCAMP-DX standard. The page is still up on the old IUPAC site (<http://www.jcamp-dx.org/>), which gives a history of IUPAC's involvement starting in 1995 when they took over the DX standards suite from the Joint Committee on Atomic and Molecular Physical Data (JCAMP). SEDS was formed in 2003 to oversee IUPAC activities associated with JCAMP-DX and XML. JCAMP-DX includes standard reporting protocols for multiple spectra techniques, including EPR, NMR, IR and MS. The last earnest effort was made in 2005.



INTERNATIONAL UNION OF
PURE AND APPLIED CHEMISTRY

(SCDS) Subcommittee on Cheminformatics Data Standards

AnIML (Analytical Information Markup Language), is an XML template in development under the auspices of the ASTM International E13.15 Subcommittee, charged with development and implementation of standards relevant to analytical instrument data. The project was launched with a meeting at Shimadzu in Columbia, MD in September 2002, and was then formulated as a joint project of ASTM and IUPAC. The last joint meeting of the AnIML Working Group with IUPAC SEDS was at the IUPAC GA in Turin in 2007. Stuart shared a brief update on the project by email: they are still working on documentation, consistency checking the technique definitions, working on missing technique definitions, and preparing to ballot documents in ASTM in early 2017.

Reflecting on this history, one conclusion of the meeting was that this group needs to spend some time thinking about what constitutes a digital standard or recommendation today, and what process is needed to best support this development. For example, rather than specifying details about how numbers are stored (ASCII, 32-bit binary, 64-bit binary, etc; which is an aspect of both JCAMP and AnIML), there should be more focus on the semantics with a mechanism to simply record how numbers are stored. The rate of change is such that a standard needs to accommodate today's infrastructure as well as that in the future. It is important to garner input from stakeholders early and throughout the process for developing digital information standards.

ACTION: *Tony Davies* will share the review process considered by SEDS.

3. Liaising with the Research Data Alliance (RDA), Chemistry Research Data Interest Group (CRDIG):

CRDIG Co-Chair, Ian Bruno, described the RDA aims to support sharing of data, address infrastructure challenges and facilitate day-to-day research efforts associated with managing and sharing data. The CRDIG in particular is interested in acting as a community bridge among various data exchange initiatives relevant to chemistry, including IUPAC's contributions to chemistry data standards. The question of what are the goals and deliverables for developing digital standards is arising in multiple venues and organizations.

Respectfully submitted,

Leah McEwen, Co-Chair, Subcommittee on Cheminformatics Data Standards

December 20, 2016