INTERNATIONAL UNION OF
PURE AND APPLIED CHEMISTRY

Committee on Publications and Cheminformatics Data Standards (CPCDS)

Report to the Bureau, April 2020

Leah McEwen, Committee Chair

I. Executive Summary

The IUPAC Committee on Publications and Cheminformatics Data Standards (CPCDS) advises on issues related to dissemination of information, primarily of IUPAC outputs. The portfolio is expanding and diversifying, in types of content, modes of publication and potential readership. Publication of IUPAC recommendations, technical reports and other information resources remains at the core of IUPAC dissemination activity and that of CPCDS. Increasingly the committee also focuses on the development and dissemination of chemical data and information standards to facilitate robust communication in the digital environment. Further details on these two priority areas are provided in Appendix A (Subcommittee on Publications: Report to CPCDS) and Appendix B (CPCDS Chair's Statement: Towards a "Digital IUPAC").

The 2018-2019 biennium brought stability to key systems and workflows including those supporting IUPAC's flagship journal *Pure and Applied Chemistry (PAC)* and online access to the *Compendium of Chemical Terminology* (a.k.a., the *"Gold Book"*). A number of collaborative symposia and workshops with strategic partners including CODATA and the GO FAIR Chemistry Implementation Network (ChIN) surfaced use cases and infrastructure needs for digital science among diverse community stakeholders. These activities lay the groundwork for developing a robust and systematic program for Digital IUPAC and "the creation of a consistent and interoperable global framework for human and machine-readable chemical information," as articulated in the CPCDS Terms of Reference. This vision will be a critical component of success for IUPAC's contribution towards the United Nations Sustainable Development Goals.

Critical among community needs are standard IUPAC definitions that can be parsed accurately and consistently for use in the many computer applications that support chemical data collection and analysis. There are very few *digital* standards for chemical information that enable unambiguous, consistent and interoperable expression and even fewer venues/approaches to develop such standards that facilitate broad input and consensus across sectors and jurisdictions in the manner in which other IUPAC standards are developed for use globally. Developing digital standards represents much more than building software to distribute IUPAC information; programmers in any chemistry organization or context need standard specifications and conventions to ensure that the encoding, expression and exporting of chemical information in these systems supports a common language. This is especially critical for those applications that combine input from multiple disciplines and the Open Science agenda. As we are learning in the

current social environment of the global coronavirus pandemic, the rules of communication and navigation are different online than onsite.

Advising on principles and practices for supporting a common language and standardization of processes and procedures in the digital environment is a primary goal for CPCDS in the coming biennia. Establishing IUPAC's role in digital standards will involve looking under the hood of how we structure the definitions that convey the authoritative meaning of standard nomenclature, terminology, measurement methodology and critical evaluation of data as robustly as we do for trained chemists. Collectively we will need to review metadata used to name/label/enumerate data and information for machine parsing, consider what constitutes acceptable scenarios for machine access to express in standard licenses, and devise sustainable business models for the platforms necessary to curate and disseminate IUPAC standards. This effort will also entail looking at how the Union manages issues of documentation, intellectual property and community engagement as it advises on supporting these critical functions in the digital environment. Broad collaboration across Divisions, other Standing Committees and the Secretariat, as well as active partnership/engagement with other international bodies, initiatives and disciplines will be critical in realizing these objectives. The outcomes will align the highly valuable intellectual scientific work of IUPAC with mechanisms for digital science, facilitate communication worldwide, and better position the Union to contribute to critical efforts for global health and sustainability.

For the 2020-2021 biennium, the priority areas of work are as follows and described in more detail in section II.

- digital dissemination of IUPAC publications
- alignment of IUPAC outputs with FAIR data principles
- best practices, procedures and policies for digital project work
- participation in community meetings
- establishing and strengthening partnerships
- archiving IUPAC's digital work and online presence

II. Plans and priorities for this biennium, and beyond

CPCDS is implementing a multi-pronged strategy to prioritize the above issues in the 2020-21 biennium. Formulation of machine-processable technical descriptions that build on authoritative IUPAC scientific definitions and content is progressing through a number of projects recently launched and in the pipeline in conjunction with Divisions as well as active user groups. Many of these focus on standardization of the metadata that describe the components of critically evaluated data or terminology entries and the establishment of validation criteria to enable systems to check for interoperable representation based on the standards. Projects are also looking to develop more efficient mechanisms to augment curation practices that adhere to the FAIR Data Principles and preserve provenance links to IUPAC authority. Recent projects include:

- The Phase 2 project for the digital Gold Book is reviewing requirements for an online term management system to facilitate the process of updating the content with current approved IUPAC terminology and broader dissemination with accurate provenance.

- The Machine-Accessible Periodic Table (MAPT) is a joint project of CPCDS and the CIAAW to develop robust machine-readable specification of the values, associated uncertainties and other descriptive information to ensure accurate reuse of these critically evaluated data in computation and analysis.
- The recently approved FAIR Spectroscopic Data project is building on a long history of IUPAC activity to apply the FAIR Data Principles to standards for consistent expression and dissemination of spectroscopic data in online networks.
- The SMILES+ project is seeking to formalize a standard interpretation of the commonly used SMILES notation by chemistry software and cheminformatics toolkits, building on a former open community-based initiative and leveraging the GitHub environment for development work.
- The InChI Open Education Resource project is a collaborative project with Division VIII and the InChI Trust to encourage broader awareness and use of the IUPAC InChI identifier in the scientific community.

To fulfill its remit, CPCDS is establishing a number of functional groups in the current biennium to ensure regular support for ongoing IUPAC activities. Subcommittees include: Publications, Records & Archives, and the Gold Book, along with the existing Subcommittee on Cheminformatics Data Standards (SCDS).

- The Subcommittee on Publications provides support for the Secretariat regarding publication contracts, the Editorial Boards of *PAC*, *Chemistry International* (*CI)* and *Chemistry Teacher International* (CTI) to ensure strong content pipelines and liaise with IUPAC Divisions and Committees on issues that impact publication.
- The Subcommittee on Records & Archives supports the Secretariat in matters of archiving the outcomes and documentation of IUPAC activity, including both print and digital records such as correspondence, web-based materials, social media, datasets and digital cheminformatics artifacts.
- The Gold Book Subcommittee provides stewardship for the digital format of the Gold Book, including administration of the website and DOI registrations, site security, web accessibility and technical documentation, and coordinates outreach and promotion in conjunction with other IUPAC Divisions and Committees.
- SCDS will continue to focus on the dissemination of digital standards through engagement with the cheminformatics community and planning for the Cheminformatics Color Book. Dissemination of digital standards and supporting materials through the well-known Color Book brand will establish the authority of IUPAC in describing chemical data in the digital environment.
- Several CPCDS members also participate in the Interdivisional Subcommittee for Critical Evaluation of Data (ISCED) to facilitate improved dissemination and curation practices.

CPCDS is actively working to address a number of topical issues that have emerged around publication and digital technologies. Goals include providing the Union with more clear guidance around critical questions of intellectual property and accessibility of digital outputs, and to provide the broader community with guidance on IUPAC's role for supporting the chemical enterprise in the digital era. The following task forces have been convened to address these topics:

- The Task Force on Intellectual Property is working with the Secretariat to develop policies and guidelines that will protect IUPAC's investment and that of its volunteers in the intellectual property that is created by IUPAC in the fulfillment of its mission.

- The FAIR Data Framework Task Force is reviewing IUPAC assets relative to the FAIR criteria to be Findable, Accessible, Interoperable and Reusable by both humans and machines, including a gap analysis across sectors, business considerations necessary for IUPAC to ensure sustainable support digital standards, and coordination with key partners including CODATA, the GO FAIR ChIN, the Research Data Alliance (RDA) and other scientific groups.
- The Task Force on Critical Issues in Scientific Communication is coordinating preparation of a series of white papers focusing on emerging technologies, new areas of science, and current issues in global chemistry to serve as information pieces for IUPAC's NAOs and the broader community. Expert groups are being convened on Blockchain and AI as initial topics.
- The PacifiChem Task Force is organizing a two-day symposium on "Chemistry on the Global Stage: Data, Standards, Infrastructure, and Challenges for the Future" to be held at the quinquennial meeting of the Pacific Rim chemistry societies in December 2020, with a goal to promote IUPAC activities and form a Pacific-based CPCDS subcommittee to facilitate engagement in cheminformatics data standards in these regions.

III. Report of Committee activities and achievements during the 2018-2019 biennium

CPCDS activities support the IUPAC Mission, particularly "the development of the essential tools for the application and communication of chemical knowledge" through support of IUPAC's publications, databases, and the development of standards for the management, storage, and sharing of digital chemical information. CPCDS' objective is to help build a technical infrastructure that will facilitate the maximum dissemination and usage of IUPAC's content in a digital environment in support of IUPAC's Vision to be an indispensable worldwide resource for chemistry. CPCDS activities and achievements during the 2018-2019 biennium are primarily related to the following goals.

Goal #2: **Increasing the value of our products and services** (publications, databases, and other intellectual property), and the following strategic objectives:

1) Brand IUPAC in the minds of stakeholders,
2) Improve the quality and frequency of communication with stakeholders
3) Increase revenue and improve long-term financial stability, and
4) Enhance interdivisional interaction and collaboration

IUPAC publications and databases serve two main purposes: 1) promotion of IUPAC as a source of indispensable information for the global chemistry community; and 2) generation of revenue to provide financial support for IUPAC activities.

A. Publications

Pure and Applied Chemistry (*PAC*)
The content pipeline for PAC has been strengthened through regular production meetings. Special issues are published periodically to broaden the journal's appeal, including: Chemistry and Cultural Heritage (March 2018), Distinguished Women in Chemistry and Chemical Engineering (February 2019), and IUPAC Distinguished Women in Science in 2021 to honor the women who received the award in Paris during the 2019 World Chemistry Congress. Potential

cancellation of IUPAC-endorsed conferences due to COVID-19 could impact future production and more special topic issues are being considered

IUPAC and DeGruyter are celebrating the 60th anniversary of PAC in 2020 with a special journal cover and virtual compilations of important PAC articles from the past. De Gruyter is also planning special events at their booth at both the fall meeting of the American Chemical Society (ACS) in San Francisco (August 2020) and the PacifiChem meeting in Honolulu (December 2020).

Chemistry International (*CI*)
CI is currently produced four times/year in print with an online version of the print content hosted on De Gruyter's website. A new initiative to identify the top ten emerging technologies in chemistry each year was launched in 2018. The first annual selection was featured in the April 2019 issue of CI and presented by former CPCDS Chair Bonnie Lawlor at the IUPAC Centenary celebration symposium at the ACS meeting in Orlando (April 2019). Proposals for the 2020 Top Ten Emerging Technologies in Chemistry are under review and will be highlighted in the October 2020 issue of CI.

CPCDS continues to explore options for *Digital CI* with DeGruyter as they have recently switched to a new publishing platform. We are looking at a variety of new potential business models for CI based on De Gruyter's experience with digital journals and reviewing background documentation just provided.

Chemistry Teacher International (CTI)
CTI was launched in July 2019 as part of an IUPAC project grant and two issues have been published so far (https://www.degruyter.com/view/journals/cti/cti-overview.xml). The journal is open access and Article Processing Charges (APC's) are being reviewed to support sustainable publication. CTI was put forth as a trial journal and an assessment should be made at some point in the future if it should continue.

DeGruyter Partnership
An authentication process to facilitate access from the IUPAC website to IUPAC Publications on the De Gruyter website by IUPAC members and other authorized users was implemented in mid-2018. The IUPAC/De Gruyter contracts for both PAC and CI have officially been extended from December 31, 2020 to December 31, 2021 and both organizations are looking at how to best move forward.

B. Databases

IUPAC Standards Online
The IUPAC Standards Online database was originally launched in March 2016 as a searchable online interface to IUPAC Standards and Recommendations. De Gruyter is offering free access during 2020 as part of the PAC 60th Anniversary celebration. People will need to register so that information can be gathered for future marketing and development purposes. At the end of the year usage will be analyzed to see who is using the database, how it is being used, and what content is of most interest, and also how best to market and position the content.

Compendium of Chemical Terminology (Gold Book)

The new Gold Book website launched in July 2019. This project enabled deployment of a stable, modern version of the current Gold Book website, a downloadable vocabulary of Gold Book terms and a simple Application Programming Interface (API) for programmatic access to individual terms. This first phase focused on website infrastructure to ensure a secure and stable online environment. Updating content from published Recommendations and supporting ongoing curation of terms is the focus of a new project that will engage ICTNS and the Divisions, as described below.

Goal #3: Improve the vitality, effectiveness and efficiency of our Union, and the following strategic objectives:

1) Brand IUPAC in the minds of stakeholders,
2) Improve the quality and frequency of communication with stakeholders, and
3) Enhance interdivisional interaction and collaboration.

CPCDS initiatives that involve technical enhancements and the development of standards aim to serve several purposes for IUPAC and the broader community: 1) facilitate management of quality chemical data globally; 2) expand the utility of authoritative IUPAC information more broadly across disciplines; and 3) support best practices for IUPAC curation of robust machine-readable content.

Standards for the Storage, Management, and Sharing of Digital Content

The CPCDS Subcommittee on Cheminformatics Data Standards (SCDS) was established in 2016 to engage with interested IUPAC Divisions and Committees, as well as external organizations with interest in quality chemical data to collaborate on the resolution of the pain points in the storage, management, sharing, and usage of digital chemical information. SCDS initiatives in the 2018-2019 biennium are described below, several of which led to successful IUPAC projects.

Cheminformatics Color Book
The first phase of the Cheminformatics Color Book project was to explore the needs for disseminating machine-readable chemical data standards for automated processes by engaging in outreach discussions worldwide at meetings of the American Chemical Society (ACS), the Royal Society of Chemistry, the Beilstein Institute, the InChI Trust, the European Bioinformatics Institute, the U.S. National Library of Medicine Center for Bioinformatics, the Research Data Alliance, and CODATA, taking into account input from scientific disciplines beyond chemistry and the principles of the GO FAIR movement which IUPAC now supports. A framework for an IUPAC Cheminformatics Color Book was articulated over several discussions that will inform the second phase to implement a digital resource to coordinate and manage source reference material for IUPAC outputs related to cheminformatics and data standards.

Chemical Representation for Computers
The IUPAC International Chemical Identifier (InChI) is a critical tool for cheminformatics and digital dissemination of chemical information and data. The InChI descriptor generated by a canonical algorithm facilitates the accurate matchup of chemical data and information records when linking between different computer systems. SCDS has contributed to several InChI workshops since 2017 and several CPCDS projects are incorporating InChI as a core feature of the metadata to facilitate interoperability.

The SMILES (Simplified Molecular-Input Line-Entry System) family of chemical notation is also implemented in many cheminformatics toolkits and databases and complements the canonical identifier function of InChI. SMILES is designed for automated retrieval of structural information and supports substructure searching, molecular pattern matching and parsing of reaction transforms. SMILES is no longer being updated and concern in the cheminformatics community prompted a project to develop open reference documentation that articulates a standard interpretation of SMILES (https://iupac.org/project/2019-002-2-024). Formalizing the SMILES specification will also enhance the accuracy of input used to generate canonical InChIs.

Standards for Spectral Data and Repositories
SCDS members participated in a CPCDS Task Force to consider the feasibility of supporting a repository for experimental spectral data. SCDS members have a range of expertise in working with spectral data, including the development of the widely used IUPAC JCAMP-DX file format (jcamp-dx.org). As a result of the feasibility study, a joint workshop was held with CODATA in Amsterdam in July 2018 to flesh out the needs of researchers and NMR equipment vendors. This has led to the successful project proposal for development of a standard for FAIR management of spectroscopic data. The recently approved project will focus on metadata for data publication as well as validation tools and workflows.

FAIR Guidelines for Chemical Data
Since the FAIR Data Principles were published in 2016, SCDS has organized and participated in a number of symposia and workshops to review the criteria and consider what is involved to share data more effectively across disciplines. Many organizations in research and applied sectors are incorporating the principles into both internal processes and outward-facing products and services. As IUPAC evolves its work practices to become more virtual, adopting the FAIR principles will enable the Union to tap into many motifs for digital exchange emerging in the data sciences and informatics expert communities. IUPAC FAIR data and information standards can also play a key role in facilitating reproducibility and replicability of scientific research and communication of authoritative information.

Systems for Curation of Digitally Enabled IUPAC Content

Machine-Accessible Periodic Table (MAPT)
Access to authoritative machine-readable data for the chemical elements was raised at a number of SCDS sponsored symposia, including by scientists in other disciplines as Geochemistry. CPCDS launched a joint project with CIAAW in 2019 to honor the International Year of the Periodic Table by developing machine readable specification of the values, associated uncertainties and other descriptive information associated with the chemical elements. Ensuring these data can be accurately parsed by machine systems will facilitate more accurate computation, maintain links to provenance, and expose this content more broadly. This project is supporting CIAAW to align their curation practices with the FAIR Data Principles and exemplifies the complementary role of CPCDS to enhance the viability and value of authoritative IUPAC output.

Gold Book Term Management System
The improved availability of the Gold Book through the new website has generated interest in extending machine representation of chemical concepts to enable new capabilities. The second of

three planned phases will identify the requirements for a term management system that supports curation and dissemination of terminology, nomenclature, and symbols for chemistry commensurate with the digital environment. This project will engage all the Divisions to develop a sustainable process for promulgating and reviewing terms and provision them through a system that will more sustainably enable accurate connections into many digital venues such as Wikipedia and online textbooks. Future initiatives will refine the machine representation to support development of chemical ontologies and the semantic web.

IV. Tabular material

Symposia/Workshops/Webinars

15-20 Dec. 2020:
Chemistry on the Global Stage: Data, Standards, Infrastructure, and Challenges for the Future, a symposium organized by the IUPAC Committee on Publications and Cheminformatics Data Standards to be held at PacifiChem 2020: A Creative Vision for the Future. The meeting is jointly organized by members of the American Chemical Society, the New Zealand Chemical Society, and the Chemical Society of Japan.

16-20 Aug. 2020 (final symposium schedule TBD):
Making Chemistry FAIRer, a symposium jointly organized by the IUPAC Committee on Publications and Cheminformatics Data Standards and the GO FAIR Chemistry Implementation Network to be held at Fall 2020 ACS Meeting in San Francisco: Making Chemistry FAIRer. The outputs of this symposium are anticipated to feed into ongoing and future activities in IUPAC and CODATA/GO FAIR.

23-24 Mar. 2020 (*cancelled*):
The Current State of FAIR Chemical Data, a two-day symposium jointly organized by the ACS Division of Chemical Information, the IUPAC Committee on Publications and Cheminformatics Data Standards and the GO FAIR Chemistry Implementation Network was to be held at the Spring 2020 ACS meeting in Philadelphia. The meeting was cancelled due to the COVID-19 pandemic. Planning is underway to address various themes of FAIR and chemistry data across a number of future symposia in which IUPAC CPCDS will also take part.

25-29 Aug. 2019:
Chemical Nomenclature and Representation: Past, Present and Future, a symposium organized by the IUPAC Committee on Publications and Cheminformatics Data Standards jointly with the CSA Trust, ICTNS, and the ACS Committee on Chemical Nomenclature under the auspices of the Division of Chemical Information of the American Chemical Society (ACS) at the ACS meeting on San Diego, CA.

5-12 Jul. 2019:
Digital Chemistry and the Lab of the Future, a symposium organized by the IUPAC Committee on Publications and Cheminformatics Data Standards to be held at the 2019 IUPAC World Chemistry Congress in Paris, France.

1 Apr. 2019:

Creating a Common Language for Chemistry: IUPAC's Role - Past, Present, and Future, a symposium organized by the IUPAC Committee on Publications and Cheminformatics Data Standards in cooperation with the ACS Division of Chemical Information at the Spring 2020 ACS meeting in Orlando, FL.

30-31 Mar. 2019:
FAIR Publishing Guidelines for Spectral Data and Chemical Structures, a workshop sponsored by NSF and jointly organized by members of the CPCDS Subcommittee on Cheminformatics Data Standards and the ACS Division of Chemical Information, in conjunction with the Spring 2019 ACS meeting in Orlando, FL.

5-8 Nov. 2018:
Data Interoperability in Chemistry, Biology, and Crystallography, a joint symposium/workshop organized by the CPCDS Subcommittee on Cheminformatics Data Standards and the International Union of Crystallography held at the 2018 SciDataCon/International Data Week held in Gaborone, Botswana.

16-17 Jul. 2018:
Supporting FAIR Exchange of Chemical Data through Standards Development, a joint workshop organized by the CPCDS Subcommittee on Cheminformatics Data Standards and CODATA held at the University of Amsterdam, Science Park, The Netherlands.

16 Feb. 2018:
Chemistry Data Standards, a joint webinar from the CPCDS Subcommittee on Cheminformatics Data Standards, the RDA Chemistry Research Data Interest Group, and the GO FAIR Chemistry Implementation Network.

24 Jan. 2018:
Repurposing the IUPAC Gold Book, a webinar presented by Dr. Stuart Chalk on behalf of the IUPAC Committee on Publications and Cheminformatics Data Standards through the auspices of the American Chemical Society.

Publications

Coles, S. J.; Frey, J. G.; Willighagan, E. L.; Chalk, S. J. Taking FAIR on the ChIN: The Chemistry Implementation Network. *Data Intel*. **2020**, *2* (1-2), 131-138. https://doi.org/10.1162/dint_a_00035

Stall, S.; McEwen, L.; Wyborn, L.; Hoebelheinrich, N.; Bruno, I. Growing the FAIR Community at the Intersection of the Geosciences and Pure and Applied Chemistry. *Data Intel*. **2020**, *2* (1-2), 139-150. https://doi.org/10.1162/dint_a_00036

Chalk, S. The IUPAC Gold Book. *IUPAC 100 Stories*. Jul. 29, 2019, story 0.13. https://iupac.org/100/stories/the-iupac-gold-book

Hepler-Smith, E.; McEwen, L. A Century of Nomenclature for Chemists and Machines. *Chem Int.* **2019**, *41* (3), 46-49. https://doi.org/10.1515/ci-2019-0315

Clark, A. M.; McEwen, L.R.; Gedeck, P.; Bunin, B. A. Capturing mixture composition: an open machine-readable format for representing mixed substances. *J. Cheminform*. **2019**, *11*, article 33. https://doi.org/10.1186/s13321-019-0357-4

Lawlor, B. A Common Language for Chemistry and More. *IUPAC 100 Stories*. Mar. 21, 2019, story 0.9. https://iupac.org/100/stories/a-common-language-for-chemistry

Lawlor, B. Preprints and Scholarly Communication in Chemistry: A Look at ChemRxiv. *Chem. Int.* **2018**, *40* (4), 18-21. https://doi.org/10.1515/ci-2018-0406

McEwen, L. InChI'ng forward: Community Engagement in IUPAC's Digital Chemical identifier. *Chem. Int.* **2018**, *40* (1), 27-31. https://doi.org/10.1515/ci-2018-0109

Projects

INCHI OPEN EDUCATION RESOURCE (OER)
https://iupac.org/project/2018-012-3-024
Chair: Robert Belford

IUPAC SMILES+ SPECIFICATION
https://iupac.org/project/2019-002-2-024
*Chair: Vincent Scalfan*i

MACHINE-ACCESSIBLE PERIODIC TABLE
https://iupac.org/project/2019-020-2-024
Chairs: Leah McEwen, Juris Meija

DEVELOPMENT OF A STANDARD FOR FAIR DATA MANAGEMENT OF SPECTROSCOPIC DATA
https://iupac.org/project/2019-031-1-024
Chairs: Robert Hanson, Damien Jeannerat

DEVELOPMENT OF AN IUPAC RECOMMENDED TERM MANAGEMENT SYSTEM FOR EXPANSION OF THE COVERAGE OF THE IUPAC COMPENDIUM ON CHEMICAL TERMINOLOGY
https://iupac.org/project/2019-032-1-024
Chair: Stuart Chalk

INTERNATIONAL UNION OF
PURE AND APPLIED CHEMISTRY

Subcommittee on Publications

Report to CPCDS, April 6, 2020

Bonnie Lawlor, Subcommittee Chair

The following is a brief status report on IUPAC Publications and Databases:

Pure and Applied Chemistry (*PAC*)

The first three *PAC* issues for 2020 have been published and the April issue is in process. Content is also in various stages of processing through the July issue. Of concern at the moment is the potential cancellation of IUPAC-endorsed conferences due to Covid-19 which could negatively impact 2021 production. Preventive action in the form of special topic issues will be discussed at the next *PAC* production meeting scheduled for April 8th. One special issue on the IUPAC Distinguished Women in Science award is already planned for 2021. It is based upon the women who received the award in Paris during the 2019 World Chemistry Congress.

The *PAC* IUPAC/De Gruyter production team continues to meet monthly to ensure that the publication has sufficient content and that the pipeline is robust. Note that *PAC* has a special cover for 2020 to highlight its 60th anniversary and that De Gruyter plans to have a special celebration at the fall meeting of the American Chemical Society (ACS) in San Francisco. The one planned for the spring Philadelphia ACS meeting was cancelled, but may take place at Pacifichem in December since IUPAC will be holding a symposium there and De Gruyter plans to exhibit. Virtual compilations of important *PAC* articles from the past will be posted to the De Gruyter site throughout this year.

Chemistry International (*CI*)

The first two issues of *CI* have been completed and content has been identified for the July and October issues - less for October than for July. However, the 2020 Top Ten Emerging Technologies in Chemistry call for proposals ended March 31st and most likely the article that will emerge from that project will be published in October.

The *CI* Editorial Board continues to meet approximately every six weeks. The next meeting is scheduled for April 17th.

Chemistry Standards and Recommendations Online Database

De Gruyter is offering the Standards online Database free to all for one year as part of the *PAC* 60th Anniversary celebration. People will need to register so that information can be gathered for future marketing and development purposes. At the very least user names, email addresses, organizational affiliations, and geographic location will be captured. At the end of the year usage will be analyzed to see

who is using the database, how it is being used, and what content is of most interest, and also how best to market and position the content.

Chemistry Teacher International (*CTI)*

CPCDS has not received regular reports on *CTI* and De Gruyter has been asked to provide regular reports at the quarterly business meetings. *CTI* is supposed to implement Article Processing Charges (APC's) this year or next. Originally, an APC of €150 was noted, but it was agreed in October 2019 that De Gruyter would do a financial analysis to determine if that number is still viable - does it provide a sustainable financial future, and is it a fee acceptable to and affordable by the education community who represents the pool of potential authors? *CTI* was put forth as a trial journal and an assessment should be made at some point in the future if it should continue. De Gruyter plans to do a financial and market analysis and report back.

Publication Contracts

The IUPAC/De Gruyter contracts for both *PAC* and *CI* have officially been extended from December 31, 2020 to December 31, 2021 and both organizations are looking at how to best move forward.

Plan S

Implementation of Plan S has been put off for a full year (see: https://www.coalition-s.org/rationale-for-the-revisions/). However, both IUPAC and De Gruyter continue to monitor its progress. Of more concern is the rumored "Executive Order" to eliminate all embargos on access to funder research (see: https://www.aip.org/fyi/2019/scientific-publishers-unite-oppose-potential-open-access-executive-order). At least two rounds of interagency reviews of the order have occurred, but there has been no word on when or whether it will be issued.

CI Digital Platform

The *CI* Digital Platform team met with De Gruyter on December 16th to learn about the features and functionalities of the new De Gruyter online system and how it will look and feel. The ultimate goal was to see if the new platform is an option to serve as a digital platform for *CI.* There was a lengthy discussion of what can and cannot be done both with the current online system and the new one. For both systems the major hurdles are: 1) the fact that *CI* is a hybrid journal (print/digital) and the work flow will be a challenge for both systems; 2) *CI* needs to be behind a firewall so that only members and paying subscribers can access it; and 3) the current log-on system for members/paying subscribers is less than ideal.

Lyndsey Dixon, who is now the De Gruyter contact person for the IUPAC account, said that from her perspective it appears that if *CI* were Open Access and in a digital-only format, that most of the challenges could be eliminated. She asked if it would be helpful if De Gruyter provided a financial picture on *CI* based upon a digital-only, open access version and the investment that would be required by both organizations. They could provide market information, trends, etc. A document was supplied in early

March, but upon review it needed to be revised due to some faulty assumptions. The updated document was received the week of March 30th and has not yet been reviewed.

Towards a "Digital IUPAC"

Leah R. McEwen, Cornell University
Chair, IUPAC Committee on Publications and Cheminformatics Data Standards

CPCDS Chair's Statement, *Chemistry International*, April 2020
[submitted 2020-02-10]

Communicating chemical knowledge is at the core of the IUPAC mission and underlies the success of the chemistry enterprise. In the global economy of the 21st century, this involves exchange among computer systems along with the expert scientists who use them. To enable the application of IUPAC outputs in the digital environment, IUPAC must augment its efforts to enable accessibility and interpretation by machines as well as humans. The Union must adapt to a digital work culture to engage in its mission of sustainable development, common language and free exchange of scientific information.[1]

The Committee on Publications and Cheminformatics Data Standards (https://iupac.org/body/024) is charged to develop standards that enable and "promote interoperable and consistent transmission, storage, and management of digital [chemical information] content." Since 2016, the CPCDS Subcommittee on Cheminformatics Data Standards has been tasked to explore the needs of the chemical community with the objective of coordinating the collective expertise of relevant IUPAC Divisions and Committees and external global organizations. A special issue of *Chemistry International* on "Research Data, Big Data and Chemistry" was edited by the Subcommittee for the 49th General Assembly in São Paulo. As demonstrated in related communities of practice such as crystallography, machine readable scientific definitions and standard data formats facilitate accurate reporting, further scientific analysis and processing of measurements. Collective sharing of data within a domain enables the generation of new insights that are applicable more broadly. The adoption of standard file formats and standard identifiers across the community and stakeholders greatly aids in workflows to accurately publish and share data in digital venues.[2]

Developing and disseminating digital representations of IUPAC intellectual assets is not simply a software problem. Criteria for machine readability needs to be robust, function consistently across many different computer systems and be based on accepted Internet protocols. The FAIR Data Principles describe high level criteria for enabling data and associated information to be Findable, Accessible, Interoperable and Re-usable for both humans and machines in a distributed digital environment.[3] These principles provide a good starting point for understanding what is required to enable data to be effectively shared and allow IUPAC to tap into many motifs for digital exchange emerging in the data sciences and informatics expert communities.

The goal of the committee in the coming biennium will be to formulate machine-processable technical descriptions that build on the authoritative scientific definitions developed by the

[1] [Frey 2014, https://doi.org/10.1515/ci.2014.36.1.14]
[2] [Bruno 2020, https://charlestonlibraryconference.com/here-come-the-data/]
[3] [Wilkinson et al. 2016, https://doi.org/10.1038/sdata.2016.18]

scientific Divisions of IUPAC. From an information perspective, IUPAC outputs may be classed into three pillars that support communication of chemical principles and knowledge: definitions of terms, names and symbols; critically evaluated standard data values; and specifications for chemical structures and other data representation. Describing the chemical world is too complex to accurately communicate through a single motif and there are different aspects of what to frame explicitly in machine depictions. The challenge of this work will be to break down this problem into discrete interoperable functionalities that are essential for accurate exchange of critical information and can enable broader utility collectively. CPCDS has been launching projects in conjunction with a number of Divisions as well as active user groups with a goal to show-case the application of IUPAC assets to global problems in digital science.

Digital Terminology

One of the most significant undertakings for CPCDS in collaboration with the Divisions is development and stewardship of the digital form of the IUPAC Compendium of Chemical Terminology. Colloquially known as the Gold Book after its first editor, Victor Gold, the electronic edition (https://goldbook.iupac.org) is a visible face of the significant investment that members of the IUPAC Divisions have made over the years to formally define many important chemical terms. A recent project has stabilized the content and provided the groundwork for more active curation and use of the terms (https://iupac.org/project/2016-046-1-024). Term definitions may now be downloaded, accessed through an Application Programming Interface (API), and cited with automatic links through Digital Object Identifiers (DOIs).

The improved availability of Gold Book terms for computer applications has generated interest in extending machine representation of chemical concepts to enable new capabilities. Through a newly formed project (https://iupac.org/project/2019-032-1-024), IUPAC seeks to support the development of terminology, nomenclature, and symbols for chemistry commensurate with the digital environment. This necessitates a more efficient mechanism for managing terms that supports rigorous articulation and approval processes and ensures this rigor and provenance in the digital space. This project will provide a secure system and engage all the Divisions to develop a sustainable process for promulgating and reviewing terms.

Machine readable critically evaluated data

The Periodic Table is one of the most well-known chemical information constructs, and a joint project of CPCDS with the IUPAC Commission on Isotopic Abundances and Atomic Weights (CIAAW) to develop machine readable specification of this resource exemplifies the complementary role of the standing committee in expanding the utility of authoritative IUPAC output (https://iupac.org/project/2019-020-2-024).

The CIAAW and has been streamlining the process for managing the speed and accuracy of the data evaluation and communication of updates to official IUPAC approved standard values through their website (https://ciaaw.org/). To ensure these data are accessed accurately by machine systems, the values, associated uncertainties and other descriptive information must be consistently expressed in formats that can be parsed without human interpretation or intervention. CPCDS is working with CIAAW to augment curation practices for digital

dissemination in adherence to the FAIR Data Principles that will facilitate more accurate computation, maintain links to provenance, and expose this content more broadly across disciplines.

FAIR description of measurement data

IUPAC has stewarded for many years a standard "Data Exchange" format for spectroscopic information originally developed by the Joint Committee on Atomic and Molecular Physical Data, known as JCAMP-DX.[4] A project building on IUPAC's extensive expertise is being formalized to apply the FAIR Data Principles through description of digital data objects that will facilitate the processing of raw and derived spectroscopic data from instruments through publication and review to further study and analysis. In addition to specification of a standard format for the metadata, the project will seek to formulate validation criteria to enable systems to check files for machine readable and interoperable representation based on the standard.

Communicating information about chemical structures

The IUPAC International Chemical Identifier (InChI) is a chemical structure descriptor generated by a canonical algorithm. It has become an essential standard for communicating chemistry in the Internet era. InChI facilitates the accurate matchup of chemical records for discrete compounds when linking and exchanging information across computer systems.[5] The InChI algorithm is jointly stewarded by Division VIII and the InChI Trust, an independent nonprofit charity established to promote use of the standard (https://www.inchi-trust.org/). Several CPCDS projects are incorporating InChI as a core feature of the metadata to facilitate interoperability.

The SMILES (Simplified molecular-input line-entry system) family of chemical representation notation is a common digital motif for automated retrieval of structural information that supports substructure searching, molecular patterns and reaction transforms. The continued ubiquitous use of dated SMILES documentation is limiting the accurate global exchange of chemical information and a project is underway to develop open reference documentation that articulates a standard interpretation of SMILES (https://iupac.org/project/2019-002-2-024). SMILES plays a complementary role to the standard InChI identifier in cheminformatics and formalizing the specification will enhance the accuracy of input used to generate canonical InChIs.

These projects represent some of the many opportunities in which CPCDS will engage over the course of the new biennium. CPCDS members are also participating in the Interdivisional Subcommittee on Critical Evaluation of Data (https://iupac.org/body/505) to harmonize formats for the archiving of both compiled and evaluated data. CPCDS is launching a task force to develop white papers that will focus on emerging technologies, new areas of science, and current issues in global chemistry, including current and proposed applications of Artificial Intelligence, Machine Learning and Blockchain Technologies in the chemical sciences.

[4] [Grasselli 1991, https://doi.org/10.1351/pac199163121781]
[5] [McEwen 2018, https://doi.org/10.1515/ci-2018-0109]

IUPAC is strategically placed to connect these collective efforts into open and FAIR data initiatives globally and across disciplines through participation on the International Science Council (ISC) Committee on Data (CODATA) (http://www.codata.org/). The Digital Revolution is one of the four domains of the ISC Action Plan to Advance Science as a Global Public Good (https://council.science/actionplan/). The Secretary General of IUPAC is a member of the CODATA Executive Committee and has highlighted on a number of occasions the strategic importance of Cheminformatics and of IUPAC work in developing the tools and standards that will be needed by chemists and all those who use chemical data in the world of big data. International IUPAC related collaborations were recently highlighted in a special issue of *Data Intelligence* on the FAIR Data Principles.

https://doi.org/10.1162/dint_a_00035
Taking FAIR on the ChIN: The Chemistry Implementation Network
Simon J. Coles, Jeremy G. Frey, Egon L. Willighagen and Stuart J. Chalk

https://doi.org/10.1162/dint_a_00036
Growing the FAIR Community at the Intersection of the Geosciences and Pure and Applied Chemistry
Shelley Stall, Leah McEwen, Lesley Wyborn, Nancy Hoebelheinrich and Ian Bruno

IUPAC Strategic Plan 2015

VISION

IUPAC is an indispensable worldwide resource for chemistry.

MISSION

The International Union of Pure and Applied Chemistry is the global organization that provides objective scientific expertise and develops the essential tools for the application and communication of chemical knowledge for the benefit of humankind and the world.

IUPAC accomplishes its mission by fostering sustainable development, providing a common language for chemistry, and advocating the free exchange of scientific information.

CORE VALUES

We serve humankind by advancing chemistry worldwide.
Scientific excellence and objectivity are the cornerstones of all our work.
We value collaboration and communication among all our stakeholders.
We strive for diversity and inclusiveness in all forms.
We respect each other and the Union.
We uphold the highest standards of transparent, responsible and ethical behavior.

GOALS (short term and long term)

Provide scientific expertise to address critical world needs.
Increase the value of our products and services.
Improve the vitality, effectiveness and efficiency of our Union.

OBJECTIVES (short term – substantial progress in coming biennium)

Brand IUPAC in the minds of stakeholders
Improve quality and frequency of communication with stakeholders
Increase revenue
Expand and retain Member and volunteer base with an emphasis on diversity and inclusion.
Enhance interdivisional interaction and collaboration
Emphasize multidisciplinary projects addressing critical global issues
Support chemistry education, particularly in developing countries