

Related to other papers in this special issue	16 (p158); 13 (p131); 15 (p151)
Addressing FAIR principles	F, A, I, R

# Growing the FAIR Community at the Intersection of the Geosciences and Pure and Applied Chemistry

Shelley Stall<sup>1†</sup>, Leah McEwen<sup>2†</sup>, Lesley Wyborn<sup>3</sup>, Nancy Hoebelheinrich<sup>4</sup> & Ian Bruno<sup>5</sup>

<sup>1</sup>American Geophysical Union, 2000 Florida Ave. NW, Washington DC 20009-1277, USA

<sup>2</sup>Cornell University, Ithaca NY 14853-0001, USA

<sup>3</sup>Australian National University, Canberra, Australian Capital Territory 0200, Australia

<sup>4</sup>Knowledge Motifs, San Mateo, CA 94401, USA

<sup>5</sup>The Cambridge Crystallographic Data Centre, Cambridge, Cambridgeshire CB2 1EZ, UK

**Keywords:** Geoscience; Geophysics; Chemistry; Crystallography; Collaboration

Citation: S. Stall, L. McEwen, L. Wyborn, N. Hoebelheinrich & I. Bruno. Growing the FAIR community at the intersection of the geosciences and pure and applied chemistry. *Data Intelligence* 2(2020), 139–150. doi: 10.1162/dint\_a\_00036

---

## ABSTRACT

The geoscience and chemistry communities have numerous common practices and dependency on data standards. Recent efforts from the International Union on Pure and Applied Chemistry (IUPAC) and the American Geophysical Union (AGU) are to explore and collaborate on approaches and sharing lessons learned on efforts to implement the FAIR Guiding Principles as they apply to data in their respective communities. This paper summarizes their efforts-to-date highlighting the importance of existing communities, Scientific Unions, standards bodies and societies in taking deliberate steps to move and encourage researcher adoption of the FAIR tenets.

---

<sup>†</sup> Corresponding author: Shelley Stall (E-mail: SStall@agu.org, ORCID: 0000-0003-2926-8353); Leah McEwen (E-mail: lrm1@cornell.edu, ORCID: 0000-0003-2968-1674).

## **1. SCIENTIFIC DATA ARE VALUABLE RESEARCH OBJECTS**

Researchers in the geosciences and chemistry often find data valuable to their work by identifying relevant articles and gaining access through the supplementary information, or by requesting the data from the authors. Few researchers place their data in a trusted repository, openly sharing and linking it to their article using a citation. Data in the supplement tend neither to be well-documented nor indexed for discovery. In a recent study of papers in *Science*, the authors analyzed 180 papers to determine accessibility of data and code [1]. Only 13% of the papers included the data and code necessary to attempt reproducing the research. It is important to note that the data policy for *Science* during the period reviewed required that the authors make their data and code available to requestors. With some level of work, only 36% of all the authors of the papers in the study shared the data and code that supported their research. This represents a fraction of the authors being compliant with the requirements of the journal's data policy.

These numbers indicate that our scientific data are at high risk of being unavailable or lost, with the probability increasing over time. The data and code that underpin research and the scholarly record are very important for the integrity, transparency and reproducibility of science. Our data must be treated as a valuable research object. It must be FAIR for humans and machines, where FAIR stands for Findable, Accessible, Interoperable and Reusable.

Cross-domain research teams are challenged to find data from domains other than their area of expertise to meet the needs of their research. The FAIR Data Guidelines [2] provide guidance for all domains on how data can be better managed and preserved to maximize state-of-the-art automated workflow to support the scientific record more accurately at a larger scale. The desired result is that data are easier to discover, well-documented, and reusable in future research, for both researchers and expert systems.

The data needed to address such complex interdisciplinary scientific questions and the problems of the future will benefit from common guidelines and best practices that all researchers follow to help each other navigate the complexity of our world through data. This includes standards that are well-adopted, well-implemented and managed, and endorsed through bodies such as the long-standing Scientific Unions and other professional organizations. Increasingly they have a new role in recommending the authoritative source of this information and in collaborating with other Unions on vocabularies and best practices that can be used by multiple domains [3].

Two organizations working hard to leverage each other's good work are the International Union of Pure and Applied Chemistry (IUPAC) and the American Geophysical Union (AGU). Both are celebrating their centennial in 2019 and have been actively engaging over the last year to explore their common challenges and support each other in taking steps toward more open and FAIR data. These celebrations are timely opportunities to establish goals that acknowledge the future by providing even better support to researchers and the many types of research products that constitute the scientific record, not least of which is data.

IUPAC was established in 1919 as a neutral and objective international scientific organization to formulate a common language for chemistry and provide expert guidance on community processes and procedures

for standards development. They are a recognized member of the International Science Council (ISC). The IUPAC vision is to be “an indispensable resource for chemistry”, through its mission, “to provide objective scientific expertise and develop the essential tools for the application and communication of chemical knowledge for the benefit of humankind and the world”<sup>Ⓛ</sup>. Standard recommendations and technical guidance are published in IUPAC’s flagship journal, *Pure and Applied Chemistry*, along with a number of specific terminologies in various sub-disciplines through the Color Book program, as well as a number of evaluated data collections<sup>Ⓜ</sup>.

AGU is both a professional society with 60,000 members in the Earth, space, and environmental sciences, and a society publisher with 22 peer-reviewed journals. The AGU vision is to “galvanize a community of Earth and space scientists that collaboratively advances and communicates science and its power to ensure a sustainable future”. In AGU’s 2019 centennial celebration and programming the organization is focused on supporting the advancement of Earth and space science while providing a platform to broaden and deepen engagement within and outside the Earth and space science community.

## **2. OPEN AND FAIR DATA**

AGU has recently convened a community effort to make data open and FAIR through the Enabling FAIR Data project [4] funded by Arnold Ventures [formally the Laura and John Arnold Foundation]. In a separate effort, IUPAC has been working toward similar goals since 2014.

Researchers from the two communities share expertise in thermodynamics, petroleum, geochemistry, solubility, toxicology and element signatures to name just a few. They are both challenged with how best to establish common vocabularies, metadata best practices, and formats with formal structures that sustainably support these areas of expertise. They are further challenged with limitations on the number of repositories and the amount of curation support available to preserve the large and complex body of supporting data as part of the scientific record. Researchers in both communities face the burden of nonexistent or unaligned guidelines from funders, publishers, and others on what is required for data management and preservation in these disciplines.

## **3. ESTABLISHING A FAIR COLLABORATION**

AGU and the Enabling FAIR Data community have put into place a Commitment Statement [5] that designates for members of different stakeholder communities, such as journals, repositories, and researchers, their role in enabling open and FAIR data. The primary goal is to move data that supports research out of the supplementary information and into a trusted repository where it can be discovered and well-documented separately from the article. In this effort the focus was mostly on the “F” and “A” of FAIR with additional

---

<sup>Ⓛ</sup> <https://iupac.org/who-we-are>.

<sup>Ⓜ</sup> <https://iupac.org/what-we-do>.

work needed to firm up the approach to “I” and “R” working with GO FAIR and other science communities with intersections in the geosciences.

In 2014, IUPAC identified the need to help facilitate a consistent global framework for Human AND Machine-readable (and “understandable”) chemical information in collaboration with other science communities, industries and governments. This vision was articulated by one member as “Digital IUPAC” [6], and was incorporated into the Terms of Reference of the IUPAC Committee on Publications and Cheminformatics Data Standards, to “advise [IUPAC] on all aspects of the design and implementation of publications and data-sharing, ... and to promote the compatibility of the electronic transmission, storage, and management of digital content through the development of standards...” [7]. A subcommittee was established in 2016 to explore the cheminformatics data standards needs of the chemistry community, coordinate expertise within IUPAC, and prioritize international activity through collaborative efforts with the Research Data Alliance (RDA) and the Committee on Data (CODATA-ISC), among others. In 2018, IUPAC worked with community members to establish a nascent Chemistry Implementation Network (ChIN)<sup>®</sup> within the GO FAIR Initiative, and officially endorsed the manifesto in early 2019 [8].

In the collaboration efforts between AGU, IUPAC, and their respective broader communities, we can begin to agree how our vocabularies and standards are related. For example, in the field of geochemistry, vocabularies and standards that describe a rock or mineral that is chemically analyzed can come from the International Union of Geological Sciences (IUGS), whilst the chemical properties can come from data initiatives associated with the Periodic Table stewarded by IUPAC [9].

#### **4. IMPORTANCE OF SCIENTIFIC COMMUNITIES, UNIONS AND SOCIETIES**

The scientific ecosystem has many stakeholders that must work together to make incremental changes toward significant goals. The tenets defined in the FAIR Guiding Principles have been a tool for convergence by most, if not all, of the stakeholders throughout the history of scientific research. Communities, Unions and Societies recognize the importance of their roles in promoting improvements in scientific communication and being drivers behind the changes needed.

AGU has a long-standing partnership with both Earth Science Information Partners (ESIP)<sup>®</sup>, and more recently the RDA<sup>®</sup> to further improve how scientific data are managed. Through ESIP and RDA research communities bring their ideas to collaborate across the sciences in an international setting.

The geoscience standards bodies include the Open Geospatial Consortium (OGC), the IUGS, and more broadly, the International Organization for Standardization (ISO). Standards bodies endorse vocabularies,

---

<sup>®</sup> <https://www.go-fair.org/implementation-networks/overview/chemistryin/>.

<sup>®</sup> <https://www.esipfed.org>.

<sup>®</sup> <https://www.rd-alliance.org>.

definitions and other supporting information necessary for well-governed geostandards. Through their leadership, the way we document our interactions with our research has a common language for better, more-accurate understanding.

IUPAC is regarded by the chemistry community as the world authority on chemical nomenclature and terminology, standardized methods for measurement, atomic weights and many other critically-evaluated data. This authority is established through the participation of National Adhering Organizations (NAOs) and companies in those bodies of the union responsible for formulating, ratifying and curating the standard recommendations, including many chemical societies and national standards agencies.

The scientific unions belong to the ISC and many participate in the ISC's CODATA. CODATA is an advocate for the FAIR Guiding Principles through recent efforts by Simon Hodson, CODATA Executive Director, who chaired the European Commission's Expert Group on FAIR Data<sup>®</sup> and published, *Turning FAIR into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR Data* [10].

## **5. HOW "I" BRINGS US TOGETHER**

This collaboration between chemistry and the geosciences has led to a better understanding of our strengths in data management and an opportunity to share our experiences and approaches. We are engaged at two levels. First, the general implementation of FAIR, community organization, awareness and encouraging adoption. Second, a much deeper level of coordination and collaboration on the interoperability of data that is generated and used by both the chemistry and geoscience communities.

**Implementation of FAIR:** During a recent presentation at the National Meeting of the American Chemical Society (ACS) in April 2019, Leah McEwen and Shelley Stall described the approach taken by the Enabling FAIR Data project as compared to that of IUPAC in considering implementation of FAIR.

The Enabling FAIR Data project started from the premise that to achieve findable data, we need common guidelines for all scholarly journals, scientific repositories and funders. We also need our researchers to deposit their data in trusted repositories that support the FAIR principles. With certifications like CoreTrustSeal<sup>®</sup> and the use of persistent identifiers, this is an area where our repository communities can adopt existing best practices. Further, to encourage and ensure this behavior takes place, we need our journals and funders to implement common guidelines for our research data. The Enabling FAIR Data project is working with journals on adoption of the project's common author guidelines. Some funders such as Wellcome Trust [11] and the Belmont Forum [12] have put into place clear guidance on their requirements for open data and other digital objects.

---

<sup>®</sup> <http://www.codata.org/working-groups/fair-data-expert-group>.

<sup>®</sup> <https://www.coretrustseal.org>.

As a long-standing standards body, IUPAC has been supporting the function of a common language and standard practices for communicating chemical information for a century. As we assess these outputs and how they can be translated toward FAIR practices, we appreciate that IUPAC efforts toward authoritative community-wide standards have always been grounded in the needs of the greater scientific community to be able to reuse, compile, and analyze interoperable chemical data of measurable quality. Looking forward towards a more “Digital IUPAC”, we have aspirations for improving accessibility and findability of chemical data across the globe, across sectors and across disciplines. Some mechanisms developed to support machine-accessible interoperability across the community, such as the International Chemical Identifier (InChI)<sup>®</sup>, can be further applied toward metadata schema to improve discovery of chemical data by other systems more broadly. However, we are most challenged in realizing wider dissemination of chemical data through lack of sustainable and scalable technical expertise and infrastructure to manage these processes.

During our ongoing collaboration, we continue to discuss how we can help each other achieve mutual goals using the FAIR Guidelines as the framework. Approaching FAIR with somewhat different emphasis and priority suggests several areas where our progress on these principles can complement and build on each other, furthering our collective efforts towards multi-domain interoperability. A brief summary:

**Findable** – More guidance is needed on implementing FAIR metadata and other criteria at all levels of describing data objects, from DOIs to trusted repositories. The efforts within the Enabling FAIR Data project to recommend the CoreTrustSeal certification for repositories in the geosciences could inform similar reviews of repositories in chemistry.

**Accessible** – Data are made accessible through the services provided by the selected trusted repository. The value of a trusted repository to the FAIR principles is paramount, as demonstrated in the Enabling FAIR Data project. While there are some specialized data repositories in chemistry, very few options that provide appropriate services exist for chemical data more generally. Outside the US there are few specialized data repositories for geoscience data making adequate curation of data difficult.

**Interoperable** – Chemistry, and specifically Crystallography, could be considered important examples of interoperable data. These communities have developed exemplar data information formats specific to their data types [13]. Within the geosciences, the IUGS has led global interoperability of geoscience data since 2004. Similarly, in geophysics there have been global standards [14] for more than 30 years. This provides solid ground work for building interoperability within and across chemistry and the geosciences.

**Reusable** – The ability of researchers and other stakeholders to reuse data depends on many factors, including adequately-documented provenance, domain specific metadata, and licensing information. Several stakeholders may be involved in contributing to this guidance, including standards bodies, publishers and trusted repositories.

---

<sup>®</sup> <https://www.inchi-trust.org>.

**Coordination and Collaboration on Interoperability:** Within FAIR, “I” or Interoperability is complex. Agreement on what that means needs to happen in local communities, across a domain, internationally by domain, and then cross-domain. One example of a domain working hard on being FAIR and interoperable is that of the ocean sciences [15]. Researchers in this domain have a driving need to use data collected from the many funded scientific vessel cruises to use in larger research efforts. They are implementing semantic standards and have a growing worldwide network. Their next step is to move toward cross-domain interoperability that will require a well-defined and described vocabulary that is inclusive of what is currently adopted by research communities and mapped to other relevant vocabularies. Practices need to be put into place that encourages the use of common vocabularies to maintain their value and the authority of their use. Cross-domain interoperability needs common formats for data so that it can easily be pulled into tools used by all communities.

The chemical information community has been striving toward this reality through various efforts to “translate” and “harmonize” chemical nomenclature, other terminology and chemical data reporting standards into digital formats. Building on IUPAC’s authoritative scientific definitions, their goal is to develop machine readable technical descriptions to facilitate accurate reporting and the exchange of chemical data from system to system and further scientific analysis and informatics processing. See the Gold Book Compendium of Chemical Terminology development project for an early phase example [16]. As with communicating data between human experts, common units of meaning and modes of expression are necessary to accurately define the context of chemical data for use in expert computer systems. Understanding how other disciplines refer and describe chemicals in their research to help build these bridges across domains and use cases is an emerging goal for IUPAC in the coming year.

Broader adoption of community-used vocabularies is needed, but even more so, we need the entire research process to be well intertwined with good data management (Figure 1). Each of these elements needs to be aligned and there are very few examples of domains that are doing this well. This includes data management plans, field notes, instruments, lab notebooks, data and sample preparation, analysis, modeling, data transformations, visualizations, archiving and preservation.

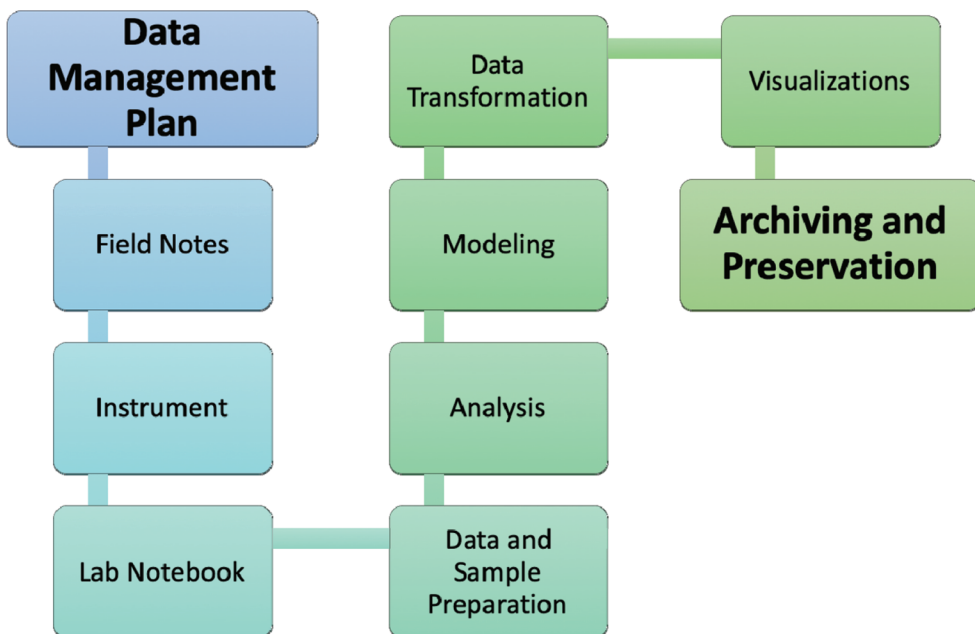


Figure 1. Research processes that need to be intertwined with good data management practices.

There are different entities responsible for nearly each step of this simplified process including commercial software, instrument manufacturers and industry. Further, the process in reality is not linear, but more iterative.

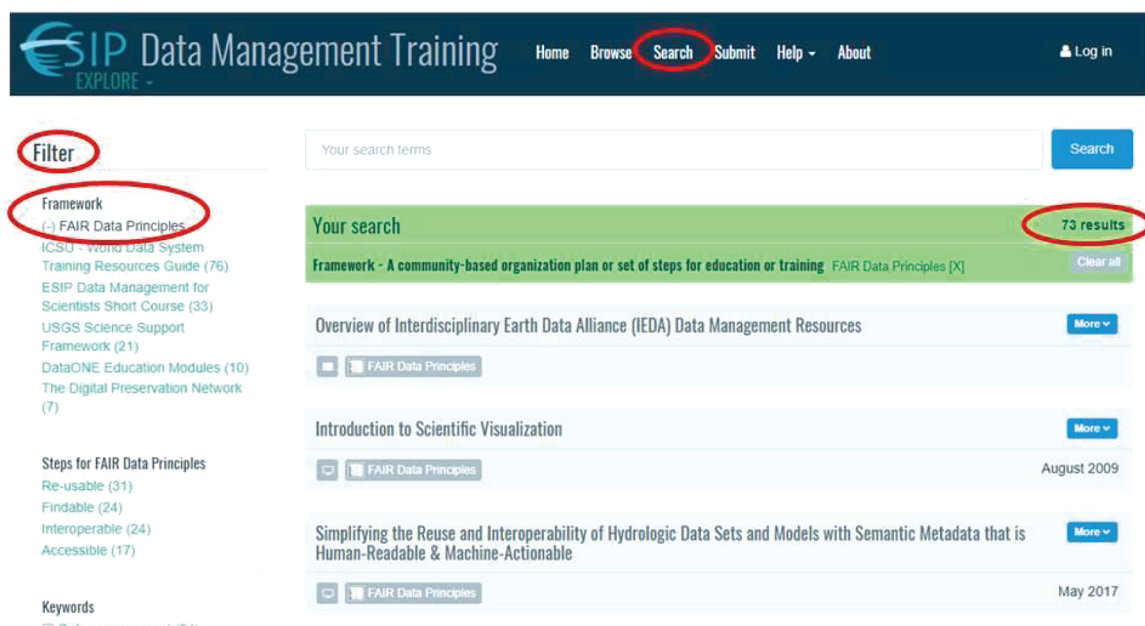
## 6. PROGRESS IN IMPLEMENTING AND SUPPORTING FAIR

**Outreach and Awareness:** Communication of new recommendations and guidelines is important for awareness and adoption. For the Enabling FAIR Data project, papers on the outcomes of each stakeholder meeting [4, 17, 18] were published, along with a useful guide for reviewers and editors on the author guidelines data [19].

As part of its outreach efforts to monitor cheminformatics needs and raise community awareness, IUPAC is engaging in a series of symposia and workshops worldwide in collaboration with other chemistry and data organizations, including the ACS, the Royal Society of Chemistry (RSC), the International Union of Crystallography (IUCr), RDA, CODATA and several additional Chemical Societies and organizations in Europe and on the Pacific Rim, among others. Outcomes have appeared in a number of reports and articles [20] and have led to the identification of a number of key areas of activity to support human and machine-readable exchange of chemical data [21].



**Training:** Educational and training resources are becoming available on the FAIR Data Principles generally, but also on how the FAIR Data Principles can be applied to the data from different domains, and on various tools that can be used to help make your data FAIR. The ESIP-hosted Data Management Training Clearinghouse (DMTC) [22] is a continually growing, curated registry of information about existing educational and training resources ranging from full online courses for credit to short online tutorials, video presentations, and learning activities. The educational and training resources in the DMTC can be accessed by means of both search and browse functionality, including the use of filters or facets that can help you find appropriate resources for your needs more precisely. As an example, a key access point into the training resources available on the FAIR Data Principles is to use the “framework” filter to see the list of resources currently available that discuss in whole or in part, one or more of the FAIR Data Principles (Figure 2). Creators of educational and training resources are encouraged to submit their resources to the DMTC for publication using the “Submit” button on the menu bar and/or the landing page.



The screenshot shows the ESIP Data Management Training website interface. The top navigation bar includes 'Home', 'Browse', 'Search', 'Submit', 'Help', and 'About'. A search bar is present with the text 'Your search terms' and a 'Search' button. On the left side, there is a 'Filter' section with a dropdown menu set to 'Framework'. Below this, a list of filter categories is shown, including 'FAIR Data Principles' (76), 'World Data System Training Resources Guide' (76), 'ESIP Data Management for Scientists Short Course' (33), 'USGS Science Support Framework' (21), 'DataONE Education Modules' (10), and 'The Digital Preservation Network' (7). The main content area displays search results for 'FAIR Data Principles' with a total of 73 results. The first result is 'Framework - A community-based organization plan or set of steps for education or training' with a 'Clear all' button. Other results include 'Overview of Interdisciplinary Earth Data Alliance (IEDA) Data Management Resources', 'Introduction to Scientific Visualization', and 'Simplifying the Reuse and Interoperability of Hydrologic Data Sets and Models with Semantic Metadata that is Human-Readable & Machine-Actionable'.

**Figure 2.** The Data Management Training Clearinghouse website demonstrating the filter alignment with the FAIR Guiding Principles.

**Adoption:** Awareness needs to be coupled with adoption for change across our broad ecosystem to occur. For the Enabling FAIR Data project we track the number of signatories and continue to engage with those that are working toward becoming a signatory making sure that all their questions and concerns are addressed. The Enabling FAIR Data project has over 150 signatories [23] with strong participation from the Earth, space and environmental science journal and repository communities.

The most recent chemistry data community workshop focused on devising “FAIR Publishing Guidelines for Spectral Data and Chemical Structures”, funded by the National Science Foundation (NSF) and held in conjunction with the Spring 2019 ACS meeting in Orlando, Florida [24]. A preliminary survey of current chemical data publishing requirements revealed a mix of digital and analog practices with generally little guidance on preparing data [25]. The workshop brought together domain publishers, databases, repositories, software developers, researchers, librarians, standards organizations and data initiatives to draft practical workflows and community-wide value propositions for publishing these common chemistry data types in a more FAIR enabled manner. Planning for a pilot is underway among several chemistry publishers.

## **7. OPPORTUNITIES FOR COLLABORATION**

Chemistry is one of the fundamental physical sciences and has branches in many related scientific disciplines. By creating a foundational set of interoperable vocabularies and standards in the Periodic Table and other fundamental chemistry standards, groups like the geosciences can leverage such standards and vocabularies in their own fields [9]. Chemistry standards can be used in combination with relevant standards that describe the geological samples that were analyzed (e.g., controlled vocabularies from the Commission for the Management and Application of Geoscience Information (CGI) of the IUGS [26] on lithology, composition, alteration, etc., or the International Mineralogical Association (IMA) list of standardized names of mineral species [27]).

2019 is the International Year of the Periodic Table and an area of joint need for the geoscience and chemistry communities is access to the evaluated elemental data that underlies the table [28]. The IUPAC Commission on Isotopic Abundances and Atomic Weights (CIAAW) is dedicated to making these data as open and accessible as possible [29, 30]. A preliminary implementation pilot is underway to disseminate these data along with other authoritative agency sources in machine accessible form through the PubChem data framework at the US National Institutes of Health [31]. Additional initiatives are also underway in IUPAC to standardize more types of digital data formats and descriptive metadata for chemical characterization, such as spectroscopic data.

Sharing common chemical standards across these multiple disciplines will ultimately facilitate interdisciplinary and transdisciplinary science. As the recognized authority in chemical nomenclature and representation, IUPAC is keen to gain an understanding of the how and where other disciplines are representing chemicals in their data workflows and metadata schema. To further map and analyze the chemical landscape more broadly, a challenge area for IUPAC, ChIN, the RDA Chemistry Interest Group, and other collaborators, will be to survey the collective data space in many fields for references to chemical information. The geoscience data community holds a wealth of data with diverse interests in chemical data standards and vocabularies – shall we start a joint initiative on “What is a Chemical”?

## 8. CONCLUSION

The aim of this article has been to articulate how we see current data sharing practices within these communities, and to compare and contrast with a view toward highlighting similarities, differences and shared challenges. Hopefully this might provide a launching pad for newer initiatives to identify how they can enrich and complement existing community activities and ensure their endorsement by the relevant International Science Union and/or appropriate professional body.

## AUTHOR CONTRIBUTIONS

Stall (sstall@agu.org) and L. McEwen (lrm1@cornell.edu) contributed equally to the design and writing of the article. L. Wyborn (lesley.wyborn@anu.edu.au) provided significant content on efforts within the geoscience unions and communities on data standards and contributions supporting FAIR data. N. Hoebelheinrich (nhoebel@kmotifs.com) developed the section highlighting the Data Management Training Clearing House. I. Bruno (bruno@ccdc.cam.ac.uk) supported the design of the article.

## REFERENCES

- [1] V. Stodden, J. Seiler & Z. Ma. An empirical analysis of journal policy effectiveness for computational reproducibility. *Proceedings of the National Academy of Sciences of the United States of America* 115(11)(2018), 2584–2589. doi: 10.1073/pnas.1708290115.
- [2] M.D. Wilkinson, M. Dumontier, I.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, ... & B. Mons. The FAIR guiding principles for scientific data management and stewardship. *Scientific Data* 3(2016), Article No. 160018. doi: 10.1038/sdata.2016.18.
- [3] L. Wyborn, N. Norris, R. Francis, K. Finney & J. Hunte. A new role for CODATA in harmonizing the data standards activities of the International Science Bodies (2010). In: *The 22nd International Conference, Cape Town South Africa*. Available at: <http://codata.info/10Conf/abstracts-presentations/High%20level%20F2/L.%20Wyborn.pdf>.
- [4] S. Stall, L.R. Yarmey, R. Boehm, H. Cousijn, P. Cruse, J. Cutcher-Gershenfeld, ... & L. Wyborn. Advancing FAIR data in Earth, space, and environmental science. *Eos*, 99(2018). doi: 10.1029/2018EO109301.
- [5] Enabling FAIR data project commitment statement in the Earth, space, and environmental sciences. Available at: <https://copdess.org/enabling-fair-data-project/commitment-statement-in-the-earth-space-and-environmental-sciences/>.
- [6] J. Frey. Digital IUPAC. *Chemistry International – Newsmagazine for IUPAC*, 36(1)(2014), 14–16. doi: 10.1515/ci.2014.36.1.14.
- [7] IUPAC Committee on Publications AND Cheminformatics Data Standards. Available at: <https://iupac.org/body/024>.
- [8] IUPAC Endorses the Chemistry GO FAIR Manifesto statement. Available at: <https://iupac.org/iupac-endorses-the-chemistry-go-fair-manifesto>.
- [9] Wyborn, L.A.I, Cox, S.J.D., Lehnert, K.A., Klump, J. A modular approach to developing interdisciplinary, interoperable standards for geochemical data based on the Semantic Sensor Network (SSN) Ontology. In: *SciDataCon 2018*. Available at: <https://www.scidatacon.org/IDW2018/sessions/223/paper/659/>.

- [10] Directorate-General for Research, and Innovation (2018). Turning FAIR into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR Data. Publications Office of the European Union. Available at: <https://publications.europa.eu/en/publication-detail/-/publication/7769a148-f1f6-11e8-9982-01aa75ed71a1>.
- [11] Wellcome Trust Policy on data, software and materials management and sharing. Available at: <https://wellcome.ac.uk/funding/guidance/policy-data-software-materials-management-and-sharing>.
- [12] Belmont Forum Open Data Policy and Principles. Available at: <https://www.belmontforum.org/about/open-data-policy-and-principles/>.
- [13] Crystallographic Information Framework. Available at: <https://www.iucr.org/resources/cif>.
- [14] Society of Exploration Geophysicists Technical Standards. Available at: <https://seg.org/Publications/SEG-Technical-Standards>.
- [15] T. Tanhua, S. Pouliquen, J. Hausman, K. O'Brien, P. Bricher, T. de Bruin, ... & Z. Zhao. Ocean FAIR data services. *Frontiers in Marine Science* 6(2019), 440. doi: 10.3389/fmars.2019.00440
- [16] Gold Book Compendium of Chemical Terminology version 3.0. Available at: <https://goldbook.iupac.org/>.
- [17] S. Stall, E. Robinson, L. Wyborn, L.R. Yarmey, M.A. Parsons, K. Lehnert, ... & B. Hanson. Enabling FAIR data across the Earth and space sciences, *Eos*, 98. doi: 10.1029/2017EO088425.
- [18] S. Stall, L. Yarmey, J. Cutcher-Gershenfeld, B. Hanson, K. Lehnert, B. Nosek, ... & L. Wyborn. Make scientific data FAIR. *Nature* 570(2019), 27–29. doi: 10.1038/d41586-019-01720-7.
- [19] S. Stall, P. Cruse, H. Cousijn, J. Cutcher-Gershenfeld, A. de Waard, B. Hanson, ... & L. Yarmey. Data sharing and citations: New author guidelines promoting open and FAIR data in the Earth, space, and environmental sciences. *Science Editor* 41(3)(2018), 83–87. Available at: <https://www.csescienceeditor.org/article/data-sharing-and-citations-new-author-guidelines-promoting-open-and-fair-data-in-the-earth-space-and-environmental-sciences/>.
- [20] IUPAC Subcommittee on Cheminformatics Data Standards. Available at: <https://iupac.org/body/036>.
- [21] Data Interest Group/Chemistry – A Vision for Chemical Data Standards. Available at: <https://iupac.org/digchem-a-vision-for-chemical-data-standards>.
- [22] ESIP Data Management Training Clearinghouse website. Available at: <http://dmtclearinghouse.esipfed.org>.
- [23] Enabling FAIR data project commitment statement in the Earth, space, and environmental sciences signatory list. Available at: <https://copdess.org/enabling-fair-data-project/commitment-statement-in-the-earth-space-and-environmental-sciences/signatories/>.
- [24] NSF OAC 2019 Workshop: FAIR Publishing Guidelines for Spectral Data and Chemical Structures. Available at: <https://osf.io/psq7k>.
- [25] Chemistry journal data submission and sharing policies checklist. Available at: [https://docs.google.com/spreadsheets/d/1P\\_7xGW5PEq8cVI6F-\\_YqupSG7VIR5uwHygZWMUMBT-w/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1P_7xGW5PEq8cVI6F-_YqupSG7VIR5uwHygZWMUMBT-w/edit?usp=sharing).
- [26] IUGS CGI geoscience vocabularies for linked data. Available at: <http://resource.geosciml.org/>.
- [27] IMA list of minerals. Available at: <https://www.ima-mineralogy.org/Minlist.htm>.
- [28] The International Year of the Periodic Table. Available at: <https://www.iypt2019.org>.
- [29] The Commission on Isotopic Abundances and Atomic Weights. Available at: <https://iupac.org/body/210>.
- [30] IUPAC Periodic Table of Elements. Available at: <https://iupac.org/what-we-do/periodic-table-of-elements>.
- [31] PubChem Periodic Table of Elements. Available at: <https://pubchem.ncbi.nlm.nih.gov/ptable>.