

Report from the IMA Mineral Informatics Working Group

August 2024

The IMA Mineral Informatics Working Group has been tasked with exploring needs and opportunities in the exciting direction of data science applied to mineralogy. The past year has seen an ever-increasing number of creative and original publications that apply association analysis, network analysis, and a range of machine learning methods to provide insights to our understanding of the mineral kingdom. Our ambition is to provide a forum for mineralogists to learn about and share relevant data resources, analytical methods, and visualization options.

A full-page feature in the December 2022 issue of *Elements* (see Figure), coupled with more than 100 emails sent to individuals, has resulted in a growing distribution list of more than 100 scientists from more than a dozen countries. Our 2023 survey to identify the most critical needs and opportunities related to the development and analysis of open-access mineral data resources pointed to four important ideas:

- 1) Build open-access mineral data resources based on FAIR principles [Ma et al. (2024) *Geoscience Data Journal* **11**, 94-104; <http://doi.org/10.1002/gdj3.204>]. We are lucky to have core websites, notably: <https://mindat.org>; <https://ruff.info/ima>; <https://earthchemn.org>. But we still need large databases for compositions and properties of large mineral groups such as a new data resource with ~100,000 garnet chemical analyses [Chiama et al. (2023) *Earth Systems Science Data* **15**, 4235-4259; <https://doi.org/10.5194/essd-15-4235-2023>].
- 2) Foster a new culture of data sharing among mineralogists, some of whom have built “private” data resources. Provide robust incentives for data sharing, for example by expanding the use of citable DOI’s. In the process, track down unpublished “dark data” and expand existing data resources. This situation has proven to be the most intractable aspect of evolving the culture of mineralogical data management.
- 3) Create a library or public platform for analytical and visualization methods tailored to mineralogical data, especially those linked to existing data resources. We are gratified that new code for analysis and visualization are appearing on GitHub (<https://github.com>) every month.
- 4) Advance mineral informatics training and engagement by organizing workshops and summer schools related to mineral informatics, and organizing special sections or issues in *American Mineralogist*, *Elements*, and other venues. We are pleased that special sessions on mineral and planetary materials informatics have been organized in the past year at the European Geological Union (2024), American Geophysical Union (2023), Geological Society of America (2023 and 2024), European Mineralogical Association (2024), International Mineralogical Association (2023), and International Geological Congress (2024).
- 5) Utilize open-access mineral data to create data-driven scientific discoveries. These novel findings refresh our understanding of mineral and environment co-evolution [e.g., Morrison et al. (2023) *PNAS Nexus* **2**, <https://doi.org/10.1093/pnasnexus/pgad110>], mineral properties [e.g., Bermanec et al. (2023) *Minerals* **10**, 139,

<https://doi.org/10.3390/min13081089>], Li et al. (2023) *Nature Communications*, in review], and many other aspects related to minerals.

In addition to the above efforts, we are pleased to see a growing collaboration between IMA and the new international Deep-time Digital Earth (DDE) program, which is recognized as the first “big science” program of the International Geological Union. IMA is now an official DDE affiliate organization.

Submitted by Robert M. Hazen (Chair), Sergey Krivovichev, Yan Li, & Shaunna M. Morrison

SOCIETY NEWS



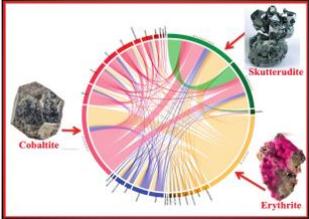
International Mineralogical Association

www.ima-mineralogy.org

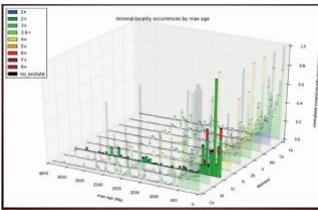
NEWLY ANNOUNCED: THE IMA WORKING GROUP ON MINERAL INFORMATICS

Informatics, or information science, focuses on all aspects of extracting information from data. The objectives of informatics include assembling and providing access to well-curated data resources, developing and applying advanced analytical and visualization methods, and the interpretation of results after applying these methods. Open and reliable data resources that conform to FAIR (Findable, Accessible, Interoperable, and Reusable) practices are an essential pillar of scientific advances through informatics. Mineralogists have long benefited from open-access data resources such as mindat.org, rruf.info, and earthchem.org, but a significant amount of published and unpublished data on mineral occurrences, compositions, physical properties, and other attributes are not yet available on any open-access platform. Most mineralogical publications do not require new data to be deposited in an open-access form, nor are there uniform standards for reporting such data.

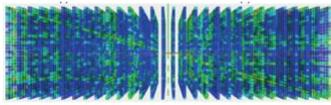
Chord diagram illustrating co-existing pairs of the 30 most abundant cobalt-bearing minerals. These data reveal significant correlations between cobalt and arsenic mineralization.



Accordingly, significant opportunities exist to improve the accessibility and reliability of a wide range of mineralogical data, as well as to develop and disseminate analytical and visualization methods to advance mineralogical research. We hope to start a conversation among engaged members of the community to identify needs and opportunities, to formulate best practices, to encourage a culture of data sharing among members of the Earth and planetary materials community, and to develop and share new resources.

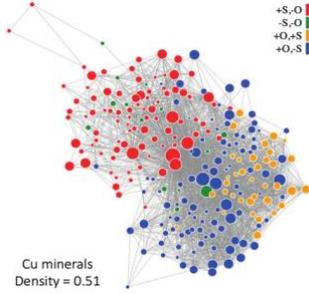


“Skyline diagram” of the temporal distribution of minerals containing 10 first-row transition elements colored according to the oxidation state. These data display episodic mineralization associated with the supercontinent cycle and changes in oxidation states associated with increases in atmospheric oxygenation.

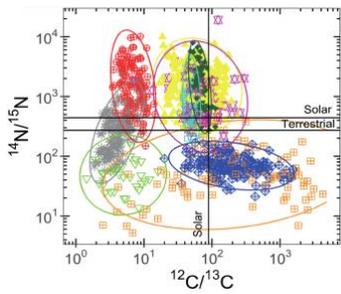


Three-dimensional 30 x 30 x 30 (27,000-matrix-element) Klee diagram revealing the relative abundances of minerals with three different co-existing chemical elements. Brighter colors indicate three-element combinations that are more common than predicted by crustal abundances.

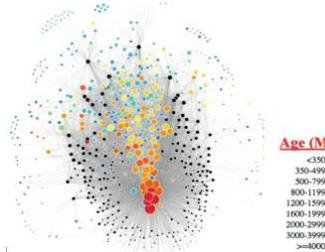
Accordingly, Sergey Krivovichev, Shaunna Morrison, Yan Li, and Robert Hazen have been asked to lead a new International Mineralogical Association called the “Mineral Informatics Working Group.” As a first step, we are developing a list of interested mineralogical community members. If you would like to receive future notices, and perhaps participate in this effort, please email Robert Hazen at rhazen@ciw.edu.



Cu minerals
Density = 0.51



Cluster analyses of stellar silicon carbide grains, based on isotope ratios, revealing various types of parent stars, including AGB stars and supernovas.



Bipartite network of carbon-bearing minerals (colored circles, sized according to their abundances) illustrating links to their localities (black circles, sized according to the number of carbon minerals). Colors indicate relative ages, which decrease from the center of the graph to the periphery.