Semi-Annual Progress Report for the DDE-Funded Project

Reporting Period: January 1, 2025, to June 30, 2025
Prepared by: Robert M. Hazen

1. Project Title: Deep-Time Data-Driven Discovery in Earth and Planetary Materials

   [Research Network Node Project (KY-C03-RNN-2025-01)]

2. Project Deliverables:

Deliverables for the first 6 months of this program include 4 peer-reviewed papers published or newly in press, 10 public lectures and departmental seminars, and 8 related podcasts and Youtube videos. We have organized 2 DDE-related conference sessions, while coordinating activities of the International Mineralogical Association (IMA) Working Group on Mineral Informatics (which we lead) with DDE interests. Numerous additional manuscripts are in preparation, while at least 25 public lectures and departmental seminars on mineral informatics themes have been scheduled in 6 countries.

Our research primarily falls under 12 specific activities, as outlined in the original proposal:

- Activity #1: Establish and expand open-access data resources in Earth and planetary materials.
- Activity #2: Expand the open-access Mineral Evolution Database (MED).
- Activity #3: Coordinate with the Rruff.info/ima data resource.
- Activity #4: Link the IMA Mineral Informatics Working Group with DDE.
- Activity #5: Perform cluster analysis on the beryl group of minerals.
- Activity #6: Initiate cluster analysis of tourmaline group minerals.
- Activity #7: Investigate the evolving chemical and structural complexity of Earth's minerals.
- Activity #8: Study roles of specific chemical elements in Earth's mineral evolution.
- Activity #9: Continue to explore trends of mineral co-occurrence, especially as related to igneous and metamorphic rocks.
- Activity #10: Apply mineral evolution data to understanding the evolution of Earth's near-surface redox state, for example using Mn minerals.
- Activity #11: Investigate the evolution of mineral/microbe interactions.
- Activity #12: Employ machine learning methods with mineral data resources to help in the discovery of critical mineral resources.

In addition to PI Robert Hazen, the original proposal listed 12 Participants in 6 countries. All but one of the 12 Participants have been active during the past 6 months. However, this report will focus primarily on activities based at the Carnegie Institution's Earth and Planets Laboratory in Washington DC.

3. Summary of Work Completed During the Reporting Period:

In terms of the 12 activities listed above, work completed is as follows:

Activity #1: Establish and expand open-access data resources in Earth and planetary materials. Our team, notably Hazen, Ma, Morrison, and Prabhu, contribute regularly to large and growing data resources. Recent contributions published by our group prior to this grant include:

- Chiama, K., Morgan Gabor, Isabella Lupini, Randolph Rutledge, Julia Ann Nord, Shuang Zhang, Asmaa Boujibar, Emma S. Bullock, Michael J. Walter, Kerstin Lehnert, Frank Spear, Shaunna M. Morrison, and Robert M. Hazen (2023) The secret life of garnets: A comprehensive, standardized dataset of garnet geochemical analyses integrating localities and petrogenesis. *Earth Systems Science Data.* https://doi.org/10.5194/essd-15-4235-2023
- Ma, X., Ralph, J., Que, X., Prabhu, A., Morrison, S.M., Hazen, R.M., Wyborn, L., and Lehnert, K. (2024) OpenMindat: Open and FAIR mineralogy data from the Mindat database. *Geoscience Data Journal,* **11**, 94-104. http://doi.org/10.1002/gdj3.204 [Winner of GSIS Best Paper of 2024 Award]
- Ralph, J., Von Bargen, D., Martynov, P., Zhang, J., Que, X., Prabhu, A., Morrison, S. M., Li, W., Chen, W., & Ma, X. (2024). Mindat.org – the open access mineralogy database to accelerate data-intensive geoscience research. *American Mineralogist, in press*. https://doi.org/10.2138/am-2024-9486

Activity #2: Expand the open-access Mineral Evolution Database (MED). The expansion of the MED is on hold until the platform is moved to https://rruff.net from its former site, https://rruff.info/evolution.

Activity #3: Coordinate with the https://rruff.info/ima data resource. We are actively working with Prof. Robert Downs (University of Arizona) to find a permanent home and stable funding for "rruff."

Activity #4: Link the IMA Mineral Informatics Working Group with DDE. Fortunately, the IMA Working Group on mineral informatics includes several collaborators on this DDE grant: Hazen, Krivovichev, Li, Lu, Morrison, and Prabhu. We are ensuring a close relationship, for example by proposing DDE-endorsed/sponsored mineral informatics sessions at the July 2026 IMA meeting in Nanjing, China.

Activity #5: Perform cluster analysis on the beryl group of minerals. This task was accomplished by Dr. Rhiana Henry in the second half of 2024.

Activity #6: Initiate cluster analysis of tourmaline group minerals. This task is on hold because we have learned that scientists at the University of Louisiana have built a large tourmaline data resource. They may decide to apply cluster analysis themselves, or they may choose to collaborate with our team.

Activity #7: Investigate the evolving chemical and structural complexity of Earth's minerals. Significant ongoing advances are being made by Prof. Sergey Krivovichev, Marko Bermanec, and other colleagues. Three papers on this topic by our group were published in late 2024, shortly before the start date of this grant:

- Hazen, R.M., Burns, P.C., Cleaves, H.J.II, Downs, R.T., Krivovichev, S.V., and Wong, M.L. (2024) Molecular assembly indices of minerals: Some mineral structures are as complex as large biomolecules. *Journal of the Royal Society Interface, 20230632*. https://dx.doi.org/10.1098/rsif.2023.0632
- Bermanec, M., Vidović , N., Ma, X., and Hazen, R.M. (2024) The average symmetry index of minerals co-varies with their hydrogen content, rarity, and paragenetic mode. *Minerals,* **14**, 387. DOI: 10.3390/min14040387
- Hazen, R.M., and Wong, M.L. (2024) Open-ended versus bounded evolution: Mineral evolution as a case study. *PNAS Nexus*, in press. https://doi.org/10.1093/pnasnexus/pgae248

Please note that DDE support is acknowledged in those 3 papers.

Activity #8: Study roles of specific chemical elements in Earth's mineral evolution. We have begun a major study on the role of hydrogen in minerals, led by Marko Bermanec. In addition, we were approached by Mr. Ezra Bekker, a graduate student at Eastern Carolina University, who developed a large database on lithium minerals. We have completed a manuscript on lithium mineral evolution, which will be submitted soon. The associated open-access data resource will prove especially valuable in efforts to apply machine learning to the discovery of critical minerals (see Activity #12).

Activity #9: Continue to explore trends of mineral co-occurrence, especially as related to igneous and metamorphic rocks. Our major papers on the evolution of igneous and metamorphic minerals have now been published.

- Hazen, R.M., Morrison, S.M., Prabhu, A., Walter, M.J., and Williams, J.R. (2023) An evolutionary system of mineralogy, Part VII: The evolution of the igneous minerals (> 2500 Ma). American Mineralogist, 108, 1620-1641. https://doi.org/10.2138/am-2022-8539
- Morrison, S.M., Prabhu, A., and Hazen, R.M. (2024) An evolutionary system of mineralogy. Part VIII. The evolution of metamorphic minerals. *American Mineralogist*, **109**, 1760-1784. https://doi.org/10.2138/am-2023-9004

Note that both papers acknowledge the support of DDE.

Activity #10: Apply mineral evolution data to understanding the evolution of Earth's near-surface redox state, for example using Mn minerals. Prof. Yan Li has spearheaded an impressive effort to enumerate thousands of manganese mineral localities, coupled with their ages and redox states, to elucidate the evolution of atmospheric oxygen. Subsequent work with additional redox-sensitive elements will greatly expand this work.

Activity #11: Investigate the evolution of mineral/microbe interactions. Work on mineral/microbe interactions is being led by Hazen, Dong, and Giovannelli, who are investigating the various modes of microbial biomineralization. We are systematizing three types of mineralization—directed (such as magnetite to make magnetosomes), indirect (minerals formed because of redox metabolisms), and environmental (minerals formed to concentrate and remove toxic metals).

Activity #12: Employ machine learning methods with mineral data resources to help in the discovery of critical mineral resources. A rapidly expanding application of mineral informatics is the use of machine learning and large mineral data resources to locate new economic deposits of critical resources. Work of Prabhu, Morrison, and Hazen set the stage, with much more opportunity ahead. Note Activity #8 above with respect to lithium. Closely related papers have been recently published by members of our team.

- Morrison, S.M., Prabhu, A., Eleish, A., Hazen, R.M., Golden, J.J., Downs, R.T., Perry, S., Burns, P.C., Ralph, J., and Fox, P. (2023) Machine learning approaches for predictive mineralogy in Earth and evolution of planetary science: A study in mineral association analysis. *Proceedings of the National Academy of Sciences Nexus*, 2, pgad110. https://doi.org/10.1093/pnasnexus/pgad110
- Que, X., Huang, J., Zhang, J., Prabhu, A., Morrison, S.M., Hazen, R.M., and Ma, X. (2024) Using adjacency matrix to explore remarkable associations in big and small mineral data. *Geoscience Frontiers, in press*. https://doi.org/10.1016/j.gsf.2024.101823
- Zhang, J., Que, X., Madhikarmi, B., Hazen, R.M., Ralph, J. Prabhu, A., Morrison, S.M., and Ma, X. (2024) Using a 3D heat map to explore the diverse correlations among elements and mineral species. *Applied Computing and Geosciences*, 21, Article #154. https://doi.org/10.1016/j.acags.2024.100154.

Work from the past 6 months is represented by 4 peer-reviewed publications that were published or in press since January 1, 2025.

1. Boucher, C.M., Eleish, A.M., Hazen, R.M., and Morrison, S.M. (2025) Network analysis of Virginia's mineralogical systems. *Virginia Journal of Science,* 75, 1-38. https://doi.org/10.25778/ajzd-qp38. This paper uses data in mindat.org coupled with network analysis to describe the diversity and distribution of minerals in the state of Virginia. The work was led by high school student, Ms. Cadence Boucher, who work with the Carnegie mineral informatics team.
2. Hazen, R. M. (2025) 2025 Andrei Borisovich Vistelius Research Award: Anirudh Prabhu. *Mathematical Geosciences, in press*. This contribution details the impressive achievements of Carnegie data scientist Dr. Anirudh Prabhu, who won the early-career Vistelius Prize for his pioneering work in mineral informatics.
3. Bermanec, M., Zhang, J., Downs, R.T., Kunić, D., Ma, X., Morrison, S.M., Prabhu, A., and Hazen, R.M. (2025) On the nature and distribution of mineral-forming ions: A mineral informatics approach. *Mineralogical Magazine*, *in press*. This paper is the Hallimond Lecture Prize paper of the Mineralogical Society of Great Britain and Ireland. The first author, Marko Bermanec, is a PhD student at the University of Graz, Austria.
4. Nanda, V., Siess, J., Jagilinki, B., Hazen, R., Kamerlin, S.C.L., Koder, R.L.,

Noy, Dror, Silberg, J.J., Tezcan, F.A., Ulijn, R., Yee, N., and Falkowski, P.G. (2025) On the emergence of metabolism: The evolution of proteins that powered life. *Philosophical Transactions, in press*. This work, in which Robert Hazen joined a group of NASA-sponsored biochemists, examines the roles of minerals in the evolution of proteins.

In addition to these published contributions, between January and June of 2025 our team presented 10 public lectures, departmental seminars, and conference talks, and contributed several podcasts, Youtube videos, TV recordings, and other media activities (see Item 7, "Outreach," below).

4. Description of the tasks undertaken by individuals involved in the project

- Principal Investigator Robert Hazen continues to contribute to all parts of the mineral informatics effort, compiling data, analyzing data, writing manuscripts, and delivering numerous presentations.
- Co-Investigator Dr. Anirudh Prabhu continues to produce network graphs and other analytical and visualization aspects of mineral informatics.
- Co-investigator Prof. Shaunna Morrison continues to contribute to all aspects of mineral informatics, with a special focus on planetary materials.
- Graduate student Marko Bermanec is the lead author on a range of mineral evolution and mineral informatics projects, with a focus on mineral/water interactions.
- Prof. Hailiang Dong continues to collaborate on research related to mineral/microbe interactions.
- Prof. Donato Giovannelli continues to collaborate on research related to mineral/microbe interactions.
- Prof. Sergey Krivovichev continues to collaborate on aspects of the structural and chemical complexity of minerals using a mineral informatics approach.
- Prof. Yan Li continues to collaborate on applying mineral evolution data to understand the rise of atmospheric oxygen on Earth.
- Prof. Anhui Lu continues to lead mineral research and education in China.
- Prof. Xiaogang Ma continues to contribute to the development and application of data resources devoted to Earth and planetary materials.
- Prof. Dietmar Muller continues to contribute pioneering research that links Earth materials data resources to studies of evolving tectonic and environmental processes on Earth.
- Dr. Michael Walter continues to apply mineralogical and petrological data resources to fundamental problems in planetary evolution.
- Dr. Rhiana Henry completed her research on beryl group minerals (Activity #5), including database development and cluster analysis, in December 2024.
- In addition, Postdoctoral Fellow Dr. Yanzhang Li will arrive at Carnegie from the China University of Geosciences, Beijing, in August. He will engage in a range of mineral informatics projects.

5. Changes in staff at the Co-PI level.

Marko Bermanec is now a PhD candidate at the University of Graz, Austria. Dr. Rhiana Henry has completed her tenure at the Gemological Institute of America.

We are excited that Dr. Yanzhang Li will arrive from the China University of Geosciences, Beijing, in August to begin a 2-year postdoctoral fellowship in mineral informatics.

6. Compliance with the Work Plan

PART/TASK	2024	25/I	25/II	26/I	26/II	27/I	27/II
Activity #1: Open-access data							
Expand existing data resources		X	X	X	X	X	X
Activity #2: Mineral Evolution Database							
Expand existing data resources		X	X	X	X	X	X
Activity #3: Coordinate with Rruff.info/ima							
Conversion to rruff.net			X				
Find stable funding				X	X		
Activity #4: IMA and DDE							
Propose IMA/DDE session			X				
Convene IMA/DDE session					X		
Activity #5: Beryl cluster analysis							
Research	X						
Manuscript			X	X			
Activity #6: Tourmaline cluster analysis							
Create database	X						
Cluster analysis				?			
Activity #7: Evolving mineral complexity							
Role of hydrogen—research		X	X				
Role of hydrogen—manuscript			X	X			
Activity #8: Specific chemical elements							
Li mineral evolution data	X	X					
Li mineral evolution manuscript			X	X			
H mineral evolution data			X	X			
H mineral evolution manuscript				X	X		

Activity #9: Igneous/metamorphic minerals							
Manuscripts	X						
Activity #10: Earth's redox state							
Mn research/manuscript	X						
Other elements research			X	X			
Other elements manuscript					X	X	
Activity #11: Mineral/microbe interactions.							
Microbe biomineral data	X	X	X				
Microbe biomineral manuscript			X	X	X		
GOE data			X	X			
GOE manuscript					X	X	
Activity #12: ML for critical mineral resources.							
Li minerals data	X	X					
Li minerals ML manuscript			X	X			
Other anticipated manuscripts							
Paragenetic modes revisited			X	X	X	X	
Water/mineral interactions 1			X	X	X		
Water/mineral interactions 2				X	X	X	
Water/mineral interactions 3					X	X	X
Phanerozoic biomineralization						X	X
DDE-Themed Conference Sessions	**GSA;AGU**	**EMC;IGC**	**GC;GSA;AGU**	**IMA;GC**	**GSA;AGU**	**GC**	**GSA;AGU**

7. Outreach

Robert Hazen and colleagues continue to present numerous public lectures, departmental seminars, conference talks, and media events (podcasts; TV; video) related to mineral informatics. Selected speaking events in the first 6 months of 2025 have included:

1. Hazen: February 20, 2025: "Mineral evolution and the search for critical resources, life's origins, and time's second arrow." Caltech GeoClub, Pasadena CA.
2. Hazen: March 4, 2025: "Mineral evolution and a missing law of nature." Island of Knowledge Workshop, Tuscany, Italy.
3. Hazen: April 11, 2025: "Mineral evolution and a missing law of nature." York University, Toronto, Canada.
4. Hazen: April 12, 2025: "The evolutionary system of minerals." Royal Ontario Museum, Toronto, Canada.
5. Hazen: May 4, 2025: "Mineral evolution and mineral informatics." Canadian Micro Mineral Association, Brock University, Canada.
6. Hazen: May 5, 2025: "The evolutionary system of minerals." Canadian Micro Mineral Association, Brock University, Canada.

7. Hazen: May 14, 2025: "The evolutionary system of minerals." GAC-MAC, Ottawa, Canada.
8. Prabhu: June 2025: "Predictive Mineralogy." Ma'aden (Saudi Arabian Mining Co.)
9. Prabhu: June 2025: "Detecting Biosignatures from Pyrolysis-GCMS Data Using Machine Learning." Gordon Research Conference on Chemical and Biological Defense (2025)
10. Prabhu: May 2025: "The Next Decade of Data-driven Discoveries: Perspectives from a Data User." NIH Data Repository & Knowledgebase Program Network Meeting.

Public Media Presentations:

Big Think Templeton Podcasts (Released January 2025)

1. Mineral evolution: https://www.youtube.com/watch?v=XwA9wEQ8DxM
2. Devil's advocate: https://www.youtube.com/watch?v=VQNAEec9W38
3. Missing Law: https://www.youtube.com/watch?v=lepxTr9zKDc&t=105s
4. Minerals and Alien Life: https://www.youtube.com/watch?v=es6Cs5V-M7A
5. Earth's Colors: https://www.youtube.com/watch?v=LBh758B92nw

Quanta Youtube Interviews (June 2025)

Release pending

NOVA (TV) Youtube Interview for the new series, "Particles of Thought" (June 2025)

Release pending

Wisconsin Public Radio, "To the Best of Our Knowledge" podcast (released April 2025)

https://www.ttbook.org/interview/why-we-need-new-theory-life

8. Benefits to DDE

All publications, talks, and other contributions prominently feature acknowledgement of DDE, including use of the DDE logo where appropriate. These outputs of this grant will reach many thousands of readers and audience members, including both scientists and members of the general public.

We are also building an international community, for example by forging ties with a much wider range of scholars through the IMA Mineral Informatics Working Group, which has more than 100 followers.

We emphasize the empowerment of early-career scholars, with training of graduate students and postdoctoral fellows, as well as mentoring a diverse international group of undergraduate interns, both in person and virtually.

9. Articles published or in press in 2025 that acknowledge DDE support:

1. Boucher, C.M., Eleish, A.M., Hazen, R.M., and Morrison, S.M. (2025) Network analysis of Virginia's mineralogical systems. *Virginia Journal of Science,* 75, 1-38. https://doi.org/10.25778/ajzd-qp38. This paper uses data in mindat.org coupled with network analysis to describe the diversity and distribution of minerals in the state of Virginia. The work was led by high school student, Ms. Cadence Boucher, who work with the Carnegie mineral informatics team.
2. Bermanec, M., Zhang, J., Downs, R.T., Kunić, D., Ma, X., Morrison, S.M., Prabhu, A., and Hazen, R.M. (2025) On the nature and distribution of mineral-

forming ions: A mineral informatics approach. *Mineralogical Magazine, in press*. This paper is the Hallimond Lecture Prize paper of the Mineralogical Society of Great Britain and Ireland. The first author, Marko Bermanec, is a PhD student at the University of Graz, Austria.

Other papers related to this grant, but not citing DDE support:

3. Nanda, V., Siess, J., Jagilinki, B., Hazen, R., Kamerlin, S.C.L., Koder, R.L., Noy, Dror, Silberg, J.J., Tezcan, F.A., Ulijn, R., Yee, N., and Falkowski, P.G. (2025) On the emergence of metabolism: The evolution of proteins that powered life. *Philosophical Transactions, in press.*
4. Hazen, R. M. (2025) 2025 Andrei Borisovich Vistelius Research Award: Anirudh Prabhu. *Mathematical Geosciences, in press*. This contribution details the impressive achievements of Carnegie data scientist Dr. Anirudh Prabhu, who won the early-career Vistelius Prize for his pioneering work in mineral informatics.

10. Databases, software tools, platforms, services, standards, and other outputs generated through the project funding.

We continue to add to the major mineralogical data resources, including:

- https://mindat.org: the principal data resource for mineral localities, mineral associations, and associated references. Our group's recent work on "Open mindat" is making the resource more interoperable, including with DDE platforms.
- https://rruff.info/ima: the official IMA list of minerals, coupled with extensive data on mineral compositions, structures, and physical properties. This site is being updated at https://rruff.net.
- https://rruff.info/evolution: The Mineral Evolution Database is the largest repository of mineral/locality/age data. It is critical to developments in mineral evolution.

Many of our publications incorporate open-access code for mineral data analysis and visualizations. We will continue that practice and report any such contributions in future reports.

11. Problems/issues:

There have been no significant problems related to this work. The delay in the original proposed start data of July 1, 2024, to January 1, 2025, meant that some priorities changed. For example, Dr. Rhiana Henry completed her cluster analysis survey of beryl group minerals (Activity #5) in late 2024 and presented those results at the 2024 AGU meeting. She subsequently completed her tenure as postdoctoral fellow at GIA and has yet to obtain employment in the field. We hope to publish those results, but any advances are on hold while she searches for new employment. Our ambition to develop and exploit a tourmaline group mineral database (Activity #6) was modified when we learned that such a database was already being developed at the University of Louisiana, where they are likely to apply cluster analysis themselves (possibly in collaboration with us). At the same time, we have been approached by graduate student Ezra Becker at Eastern Carolina University regarding development of a very large lithium mineral database. This resource has opened opportunities to enhance our Activities #1, #2, and #12, using data resources and machine learning to search for critical lithium resources. An inevitable issue in this kind of broad and integrated program is that new opportunities constantly arise and priorities shift, and we will continue to be opportunistic in the best sense of the word.

12. Budget Utilization

We are on track to fully spend Year 1 funds on priority categories, as outlined in our budget. A significant part of the budget will be devoted to the salary of our new mineral informatics postdoctoral fellow, Dr. Yanzhang Li, who will arrive from China University of Geosciences, Beijing, in August.

Additional expenditures include publication page charges and travel to conferences with DDE sessions or DDE-related themes.

Project PI Signature

Signature:

Print Name: Robert M. Hazen

Date: July 30, 2025