Demonstrating MinMod: A Large-scale Knowledge Graph of Historical Mining Data

Craig A. Knoblock[1], Binh Vu[1], Basel Shbita[1,2], Pothula Punith Krishna[1] and Namrata Sharma[1]

[1]USC Information Sciences Institute, Marina del Rey, CA, 90292, USA
[2]IBM Research, San Jose, CA, USA

Abstract

We present MinMod, one of the largest knowledge graphs of historical mining data. MinMod is built using scalable machine learning methods to extract hundreds of thousands of records from mining reports, databases, and tabular data in articles, and to normalize and integrate the results into a unified knowledge graph. It also provides tools that enable end-users to explore, curate, and leverage the data to support predictions of new sources of critical minerals. In this demo, we walk through the process of extracting and integrating data into MinMod, demonstrate data exploration and curation, and showcase tools such as the Grade & Tonnage model that assist scientists in mineral assessments.

Keywords

knowledge graphs, critical minerals, large language models, data integration

1. Introduction

Given the rising global demand for critical minerals in key industries such as microelectronics and electric cars, it is crucial to the assessment process to provide actionable information to decision-makers in a timely and reproducible manner. However, the well-established methodology used by the U.S. Geological Survey (USGS) for mineral resource assessment is too time-consuming because it relies on a slow manual process of data gathering, preparation, and spatial analysis.

To address this problem, we created MinMod, a knowledge graph of mineral data extracted from mining reports, databases, and tables from articles. MinMod also infers the deposit types of the mineral sites and consolidates duplicated mineral sites to produce de-duplicated normalized data for mineral prospectivity mapping [1] and critical mineral assessments conducted by scientists at the United States Geological Survey (USGS). The resulting knowledge graph is one of the largest data repositories with 680,000 mineral sites covering 190 commodities.

In this demo, we present MinMod, demonstrating how users can quickly add new tables of mineral data, explore, and edit mineral site data for a given commodity, and produce grade and tonnage models. This demo covers the work described in the ISWC 2025 In-Use Paper on MinMod [2] and the ISWC 2025 Research Paper on SAND [3].

2. Overview of MinMod

MinMod is a comprehensive system comprising multiple components for extracting data from mining reports, tables in articles, and databases, as well as predicting deposit types and linking duplicate mineral sites. Since one of the primary user groups of MinMod is geologists conducting critical mineral assessments, MinMod also includes a web-based user interface that allows users to interactively explore and curate the extracted data, ensuring the highest possible data quality.

In the first part of the demo, we will present an overview of MinMod using the MinMod dashboard (Figure 1) and the interactive map view of mineral sites (Figure 2). When users visit https://minmod.isi.edu, they are greeted with a dashboard summarizing the number of extracted mineral sites, their

ISWC 2025 Companion Volume, November 2–6, 2025, Nara, Japan

Figure 1: MinMod dashboard showing extraction coverage by data type and commodity.

inventories (e.g., estimated ore amounts, grades, or historical production), and the number of data sources integrated into MinMod. The dashboard also provides a breakdown by commodity, allowing users to hover over a commodity to quickly see the number of associated mineral sites.

Users can explore mineral sites using the interactive map by clicking the `Map View` tab on the navigation bar. Figure 2 shows the map of copper mines currently available in MinMod across the globe. Users can zoom in and out, toggle the satellite layer to examine surrounding areas, and click on a site to open a webpage containing all data about that site extracted from multiple sources by MinMod.

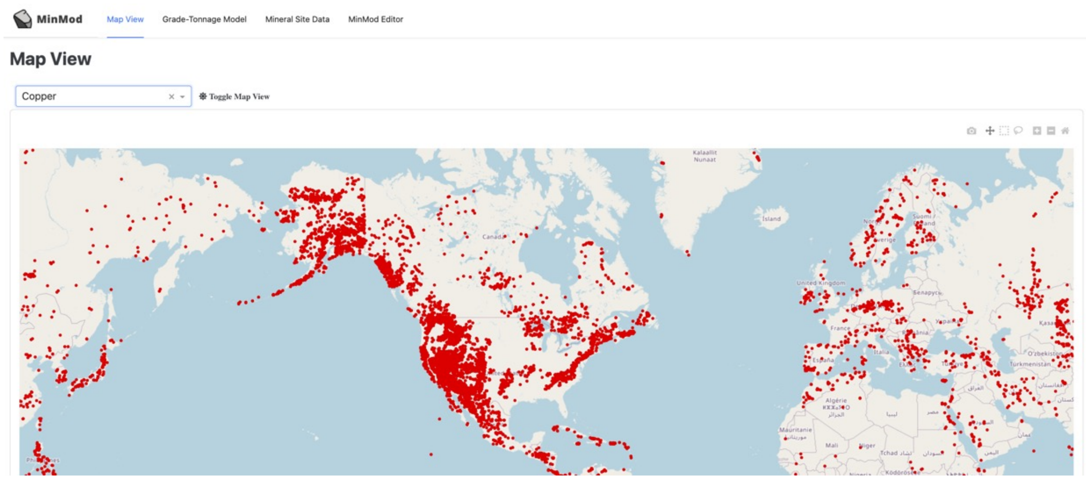
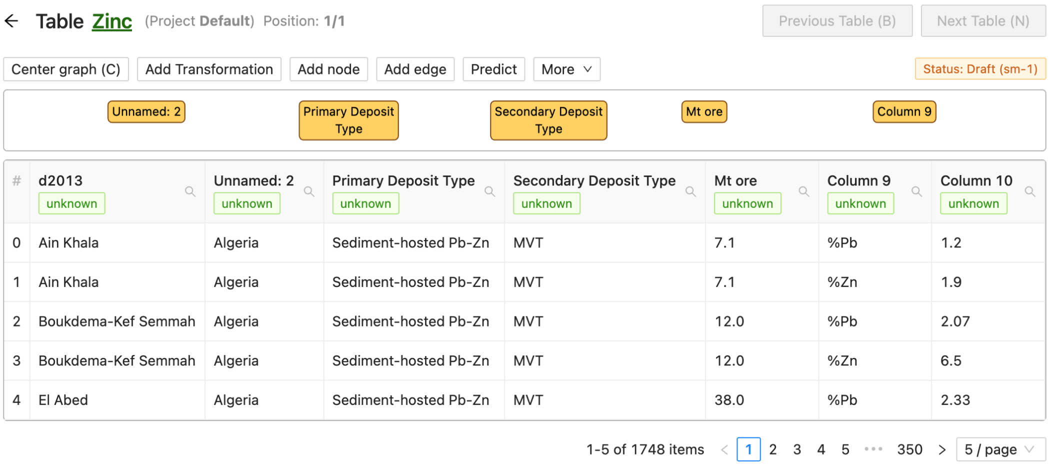


Figure 2: Interactive map view of mineral sites filtered by copper deposits.

3. Adding Data to MinMod

In this part of the demo, we demonstrate how users can import mineral data stored in tables from articles into MinMod. For this task, we use SAND [4], a semi-automated interactive semantic modeling tool, because of its flexibility to be easily configured to use the MinMod knowledge graph and a domain-independent approach [3] to automatically map tables to MinMod without any code modifications.

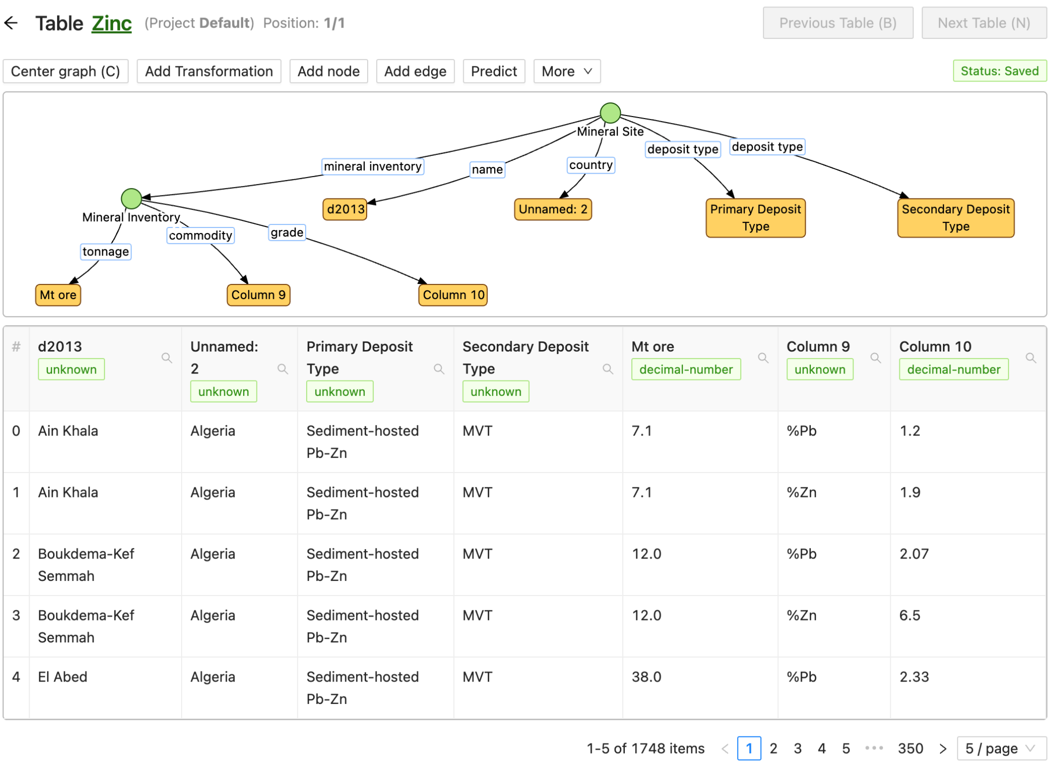
In particular, we showcase a typical workflow for mapping a table of zinc deposits, which includes deposit names, locations, deposit types, and their grades and tonnages. Figure 3 shows the table after it has been imported into SAND. By clicking the `Predict` button, the user can automatically generate a semantic description of the table, which annotates the column types and relationships between them. The user can then manually correct any errors in the description. Figure 4 illustrates the corrected semantic description after user edits. In this demo, the only required update is changing the property of



#	d2013 (unknown)	Unnamed: 2 (unknown)	Primary Deposit Type (unknown)	Secondary Deposit Type (unknown)	Mt ore (unknown)	Column 9 (unknown)	Column 10 (unknown)
0	Ain Khala	Algeria	Sediment-hosted Pb-Zn	MVT	7.1	%Pb	1.2
1	Ain Khala	Algeria	Sediment-hosted Pb-Zn	MVT	7.1	%Zn	1.9
2	Boukdema-Kef Semmah	Algeria	Sediment-hosted Pb-Zn	MVT	12.0	%Pb	2.07
3	Boukdema-Kef Semmah	Algeria	Sediment-hosted Pb-Zn	MVT	12.0	%Zn	6.5
4	El Abed	Algeria	Sediment-hosted Pb-Zn	MVT	38.0	%Pb	2.33

Figure 3: Zinc deposit table with names, locations, deposit types, grades, and tonnages

`Column 10` from `tonnage` to `grade`. Finally, the user can extract the data by clicking the `More` button and selecting the `Export` function to publish data into MinMod.



#	d2013 (unknown)	Unnamed: 2 (unknown)	Primary Deposit Type (unknown)	Secondary Deposit Type (unknown)	Mt ore (decimal-number)	Column 9 (unknown)	Column 10 (decimal-number)
0	Ain Khala	Algeria	Sediment-hosted Pb-Zn	MVT	7.1	%Pb	1.2
1	Ain Khala	Algeria	Sediment-hosted Pb-Zn	MVT	7.1	%Zn	1.9
2	Boukdema-Kef Semmah	Algeria	Sediment-hosted Pb-Zn	MVT	12.0	%Pb	2.07
3	Boukdema-Kef Semmah	Algeria	Sediment-hosted Pb-Zn	MVT	12.0	%Zn	6.5
4	El Abed	Algeria	Sediment-hosted Pb-Zn	MVT	38.0	%Pb	2.33

Figure 4: The semantic description of the table aligned to the MinMod ontology.

4. Assessing the Potential Grade and Tonnage of Undiscovered Deposits

One of the essential tools used by USGS scientists for critical mineral assessments is plotting the grade, tonnage, and deposit types of mineral sites in a chart. This plot, commonly referred to as the Grade & Tonnage Model (GTM), can be generated by clicking the `Grade-Tonnage Model` tab on the navigation

Figure 5: Grade & Tonnage Model of 16 rare-earth elements

bar and selecting a commodity to visualize.

In this part of the demo, we show how to generate a GTM for rare-earth elements (REEs) (Figure 5). Users can select "Rare Earth Elements" from the `Select Commodity` dropdown menu and click Generate. Since REE is a group of 17 metallic elements, the interface automatically expands it into individual elements and plots all of them (except Promethium, which has a very short half-life and is extremely rare in nature) in a single chart. In the plot, the x-axis represents tonnage (in millions of metric tons), while the y-axis represents grade (in percentage). The right-hand panel displays a list of deposit types in which the selected elements have been found or associated, sorted in descending order by the number of mineral sites. Users can click on any of the deposit types to filter and focus on the types of interest.

The diagonal lines in the chart represent sites with the same order of magnitude of contained metal, calculated as grade multiplied by tonnage. These lines help USGS scientists quickly identify which deposit types are likely to be the most productive when searching for a specific commodity or set of commodities in the case of REEs.

5. Exploring and Curating Mineral Site Data

To ensure the highest quality data in MinMod for training machine learning models for mineral prospectivity mapping [1], we provide an editor that allows users to modify any data in the knowledge graph. The editor can be accessed by clicking the `MinMod Editor` tab in the navigation bar. Users can then search for mineral sites by commodity, deposit type, location, or perform text-based searches on other properties such as name, type, and rank. Figure 6 shows a table of Graphite deposits currently stored in MinMod. Each row represents a de-duplicated mineral site, which may include entries extracted from multiple sources. Users can click the Edit button to view and modify grouped sites identified as the same. For example, the Cranston Mine in the figure includes data extracted from the MRDS database, a USGS article, and manual user updates. To edit any property of a mineral site, users can click the pencil icon on the corresponding column header. User-edited properties are highlighted in green.

Users can also refine the grouping of same sites by selecting the checkboxes at the beginning of rows, allowing them to merge entries referring to the same site or split those that are incorrectly grouped. Finally, the data can be downloaded as a CSV file by clicking the `Download` button at the top right of the page.

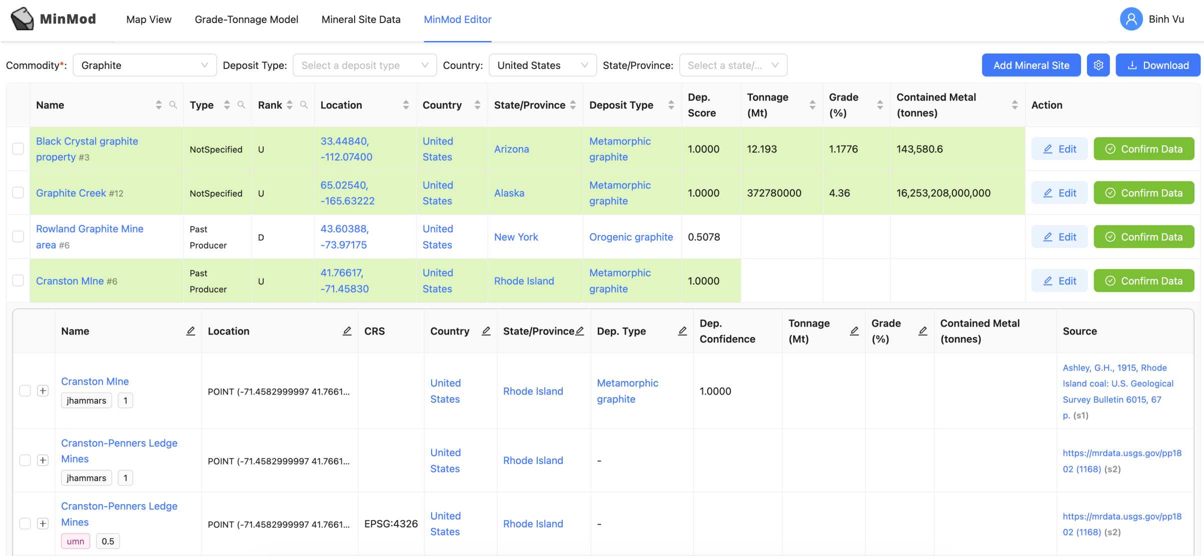


Figure 6: Exploring and editing Graphite deposits in MinMod. Properties that have been validated by an experts are highlighted in green.

6. Conclusion

In this paper, we demonstrate the MinMod knowledge graph of mineral resources, along with the related tools such as Grade and Tonnage models that enable scientists to explore and analyze data in MinMod for mineral assessments. We also show how tables of mineral data can be added into MinMod using SAND.

MinMod has been deployed internally at the U.S. Geological Survey (USGS), where feedback has been overwhelmingly positive. It has significantly reduced the time required to manually extract and collect data from diverse sources such as NI 43-101 reports, scientific articles, and databases.

In this demo, we did not showcase other tools involved in building MinMod, such as text data extraction from reports, deposit type prediction, or linking duplicated mineral sites. Readers are encouraged to refer to the full paper [2] for a detailed discussion of these components.

7. Acknowledgments

This material is based on work supported by the Defense Advanced Research Projects Agency (DARPA) under Agreement No. HR00112390132, and Contract No. 140D0423C0093. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Defense Advanced Research Projects Agency (DARPA), or its Contracting Agent, the U.S. Department of the Interior.

The authors thank Graham Lederer, Garth Graham and Jane Hammarstrom of the US Geological Survey and Simon Jowitt of the Nevada Bureau of Mines and Geology for their advice and feedback on the technology and tools.

Declaration on Generative AI The content of this paper was written by the authors. Generative AI tools were used solely for grammar checking, sentence polishing, and word choice refinement. The authors carefully reviewed and verified all AI-assisted outputs to ensure that the intent and originality of the work were fully preserved.

References

[1] A. Daruna, V. Zadorozhnyy, G. Lukoczki, H.-P. Chiu, GFM4MPM: Towards geospatial foundation models for mineral prospectivity mapping, in: Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems, SIGSPATIAL '24, Association for Computing Machinery, New York, NY, USA, 2024, p. 565–568. doi:10.1145/3678717.3691268.
[2] C. A. Knoblock, B. Vu, B. Shbita, Y.-Y. Chiang, P. Krishna, X. Lin, G. Muric, J. Pyo, A. Trejo-Sheu, M. Ye, Exploiting LLMs and semantic technologies to build a knowledge graph of historical mining data, in: International Semantic Web Conference, Springer, 2025.
[3] B. Vu, C. A. Knoblock, F. Lin, A domain-independent approach for semantic table interpretation, in: International Semantic Web Conference, Springer, 2025.
[4] B. Vu, C. A. Knoblock, SAND: A tool for creating semantic descriptions of tabular sources, in: European Semantic Web Conference, Springer, 2022, pp. 63–67.