

#Reviewer1 (Report #2)

This is my second review. The authors did a very good job of addressing both reviewers' comments. I found it much easier to understand and follow and think it will be the same for many readers.

I only have very minor wording suggestions for the authors' consideration.

L 11: "had been" -> was

Reply: Revised as suggested. Please see Line 11.

L45: constraints: parameters

Reply: Revised as suggested. Please see Line 46.

L63: Would the platform be R Studio, and RMarkdown the framework? Just suggesting something more in line with/parallel to Jupyter. Not sure

Reply: Thank you for your comment. We replaced the RMarkdown with R Studio. Please see Line 63.

Table 1: use of "volume" is a little awkward. Since there is no unit, I would prefer "number," but this is just a preference.

Reply: Thank you for your comment. The title was changed as suggestion. Please see Lines 72-75.

Table 1: some definitions are not meaningful definitions to me ("occurrence records" and "literature references")

Reply: Thank you for your comment. We revised the description of occurrences and remove the row of "literature references". Please see Lines 73-75.

L 75: highlighted phrase does not make sense to me. "Currently, it provides a separate access endpoint for each data subject."
- I don't know the significance of that. Maybe explain in a sentence or leave out.

Reply: Thank you for your comment. As the number of subjects and records in the crowdsourced Mindat database increases, the API server may open or update its data access endpoints. Some new features of our package will be updated according to the endpoints provided by the server. We have added some descriptions in the revised manuscript. Please see Lines 77-79.

L86: The description of the hierarchy and what is at the "top" or "broadest" is a bit confusing to me. Somehow I thought top of the hierarchy would be most detailed. Not sure if it can be reworded to reduce confusion, but it is okay the way it is.

Reply: Thank you for your comment. The top level of locality (0) is defined by the mindat.org database, and we just follow its principles here to keep the concepts consistent. Please see Lines 88-89.

L87: plate = tectonic plate?

Reply: Thank you for your comment. Yes. We used “tectonic plate” in our revised manuscript. Please see Line 89.

L95: strange to call CSV a specified format for these particular cases since it is the most general format, but I guess I get the point. Might be better to list the different formats for and say they are for different use cases.

Reply: Thank you for your comment. We have made changes based on your suggestions. Please see Lines 97-98.

Figure 1: typo in cotain = contain

Reply: Thank you for pointing this out. We revised it and updated the Figure 1. Please see Line 99.

Figure 1: better, but would be even more improved if each step was labeled 1-4 to show the sequence of events.

Reply: Thank you for your comment. We revised it and updated the Figure 1 as your suggestion. Please see Line 99.

L 103: remove "and it"

Reply: Thank you for your comment. Changed as suggested. Please see Line 106.

L 105: "records"?? better word?

Reply: Thank you for your comment. We use "consists of" instead of "records". Please see Line 108.

Table 3: typo in filename "wildcar" vs. "wildcard"

Reply: Thank you for your comment. We have revised the filename and updated Table 3. Please see Line 141.

L 139: says "three tasks" but there are two examples. Maybe the two examples do three things, but it is confusing, reword to reduce confusion.

Reply: Thank you for pointing this out. We only showed two examples and have revised the relevant statements accordingly. Please see Line 142.

L164: write out the element As after the abbreviation

Reply: Thank you for your comment. Changed as suggested. Please see Line 167.

L173: "the record retrieval" -> "retrieval of records"

Reply: Thank you for your comment. Changed as suggested. Please see Line 176.

L288: potentials -> potential

Reply: Thank you for pointing this out. Changed as suggested. Please see Line 294.

L289: this line seems like it got jumbled and needs to be reworked "the current version does not friendly support queries involving mineral occurrences." maybe just swap the two words friendly and support. or "the current version does not support user-friendly queries..."

Reply: Thank you for your comment. Changed as suggested. Please see Line 295.

#Reviewer2 (Report #1)

reviewer: Dominik Hezel

ms received: 05.02.2025

review completed: 04.03.2025

The authors greatly improved their manuscript compared to its first version, and I find this an exciting, highly important, and interesting read – but even more so, a fantastic R package. I only have some minor comments below, and am now looking forward to finally see this paper published.

Detailed comments

1. 59: I was excited to read there might also be a python package, but this is only mentioned here. So either delete or make some more statement here or later that there is also a python package in the works, and provide some hint to when this might be released.

Reply: Thank you for your comment. We did develop a Python package to meet the needs of different development users, but considering its implementation mechanism is different; to avoid misunderstanding, we deleted the description here. Please see Lines 59-60.

2. Table 1: Why is the meteoritics volume number in red?

Reply: Thank you for pointing this out. We merged and checked the data and forgot to change it back to black. It has now been changed to black. Please see Line 74.

3. 84: I find the term ›address‹ confusing. Could this be simply replaced by e.g., ›hierarchy‹? A hierarchy can have levels, so this might be more descriptive than ›address‹.

Reply: Thank you for your comment. Following the principle of locality hierarchies in Mindat, we will replace the word "address" with "locality" and explain its hierarchical structure. This should make it clearer for readers. Please see Lines 88-89.

4. 87: ›tectonic plate‹?

Reply: Yes. We changed it to “tectonic plate”. Please see Line 89.

5. 88ff: The bracket starting with ›(Studies of ...)‹ is confusing. Maybe this could be made as a proper sentence that has a better connection to its previous sentence.

Reply: Thank you for your valuable comment. We removed the brackets and used them as an example to explain the previous statement. Please see Lines 91-93.

6. 114: Delete ›above‹

Reply: Thank you for your comment. Changed as suggested. Please see Line 117.

7. 115: Delete ›us‹

Reply: Thank you for your comment. Changed as suggested. Please see Line 118.

8. 119: Delete extra space before ›it supports‹

Reply: Thank you for your comment. Changed as suggested. Please see Line 121.

9. 173: ›This function category ...‹? cf. your Table 6

Reply: Thank you for pointing this out. We have revised the content of the function category in Table 6 to make it more accurate. Please see Line 183 (Table 6).

10. 184: ›... few lines of code ...‹ would be more idiomatic

Reply: Thank you for your comment. Changed as suggested. Please see Line 187.

11. 224: ›... code are shared ...‹

Reply: Thank you for your comment. Changed as suggested. Please see Line 225.

12. 231: ›... section, is also shared ...‹

Reply: Thank you for your comment. Changed as suggested. Please see Line 234.

13. 233: Delete ›those‹

Reply: Thank you for your comment. Changed as suggested. Please see Line 236.

14. 235: I don't think the tool is reproducible, but it allows for reproducible data query.

Rephrase accordingly.

Reply: Thank you for your comment. Changed as suggested. Please see Lines 238-239.

15. 239: ›interactions‹ instead of ›interventions‹

Reply: Thank you for pointing this out. Changed as suggested. Please see Line 242.

16. 239 ›webpages‹ or rather ›webpage‹?

Reply: Thank you for your comment. Changed as suggested. Please see Line 242.

17. 242: ›... programming skills, but require large datasets ...‹

Reply: Thank you for your comment. Changed as suggested. Please see Line 245.

118. 250: The FAIR principles and open science are two different things. An e.g., database can be FAIR but not open, i.e., there is e.g., a paywall. This is not overly important in this context, but regarding these principles and concepts are still new to many, I would suggest to make it clear that these are different things.

Reply: Thank you for your valuable comment. According to your suggestions, we have added relevant statements to make the manuscript clearer and more complete. Please see Lines 250-255.

19. 251: ›... is increasingly demanded in the ...‹

Reply: Thank you for your comment. Changed as suggested. Please see Line 257.