

Thank you for the chance to see this revision.

I remain positive about the potential for this work, but I still regard this manuscript as too long. There is so much detail here that I find it difficult to appreciate the story. I'm sorry if this seems harsh, and possibly I'm in the minority in this feeling, but the purpose here is to get your research into the world so that people can read and appreciate it, not to document every aspect of the work.

So while I can see that the MS has been reduced by about 20%, I believe it could stand being half its current length. Most of what has been culled so far was the appendices, as far as I can tell. But I think there's scope for more brevity in the paper itself.

The good news is that I don't think it's difficult to make this paper a lot more readable. For example:

- The figure I picked on last time, Figure 9, is still repeated in its own caption, and with everything written out in Figure 3.2.1 and Figure 3.2.2.
- Have a look at the caption to Figure 3. I think it could be reduced to "Figure 3. Lithology Hierarchical Thesaurus showing the 7 major Lithology_Groups." I think there are a few pieces of information in the caption which are not simply repetition of what's in the diagram — but they are drowned out by everything else.
- As I also mentioned before, you do not need to provide the equations for standard statistics like precision and recall.

Turning to the code, I see you have addressed the licensing issues I mentioned (e.g. with QDriller). There are others however — I'm afraid my licence review was not that thorough. For example, you're depending on `fuzzywuzzy` — another GPL'd library. Again, this must be replaced with something compatible, or you must change your licence. I just had a quick look and couldn't see any others, but you should do a thorough audit here and be careful about using code from places like StackOverflow, most of which is also copyleft (CC BY-SA). (A side-note about your `requirements.txt` — you don't need to put things from the standard library in there, like `math`, `sys`, `bisect`, etc.)

I will mention that I do still have some reservations about the overall code quality (no docstrings, no tests, non-Pythonic idioms), but as I said before it's not clear to me how the journal feels about this sort of thing being part of a peer review.

I sincerely hope you persevere with this piece of work, because I do believe it deserves to be published.