Response to Ignacio Fuentes

The breakdown of replies to the individual comments are below:

Abstract, Line 23: what is an extraction rate of 865? What units? Or is it a typo and should be 86.5%?
→ This should be 86%, this was revised in the updated version.

Introduction, Line 43: I'm not sure if lithological drill core logging is "inevitably" subjective. In my impression, a lack of standardization in the lithological descriptions makes it subjective, but the subjectivity might be reduced through a standard procedure of description. However, it is not clear that such standardization is what we want.
→ The information and level of detail contained in logs is highly dependent on the purpose of the study, this already makes geological logging subjective. This subjectivity is also influenced by the lack of standards between project and/or companies combined with the personal biases of the logging geologists.

Materials and methods, Line 146: "The module was re-written into python to be make it more compact"... The grammar there sounds funny.
→ This was revised in the updated version.

Materials and methods, Line 195: Figure 2 seems to be wrongly enumerated in the text. The study area is referring to Fig 2, but it corresponds to Fig 3.
→ This was revised in the updated version.

Materials and methods, Figure 4: Regarding volcaniclastic rocks, they are classified as igneous and sedimentary. Is it ok to have the same subgroup in two lithological groups?
→ The matching is done at the Detailed_Lithology level, thus not causing confusion in the Subgroup and Group level. Volcaniclastics are present in both lithological groups as although volcaniclastics are volcanic in origin and are categorized as igneous rocks, ambiguous lithologies such as "metavolcaniclastic_sandstone" is more sedimentary than igneous.

Materials and methods, Line 325: it should be EPSG:4326 for WGS84
→ This will be revised in the updated version.

Materials and methods, Line 326: Relative level with respect to the sea level? Does the relative level refers to any reference level or is it a standard level? Because if it refers to any reference level, there is no way to know the real location unless it is corrected using a DEM and assuming the collar at the surface of the terrain.
→ We use RL here to refer to elevations of survey points with reference to the mean sea level. This definition of RL is equivalent to the elevation values used in DEMs. It is possible to cross-check the values extracted to the values in the DEM. However, it is important to note that drillholes could be drilled from underground, thus would not have a collar location at the surface of the terrain. There are entries with positive survey dip/inclination values that suggest underground holes drilled upwards.

Results, Line 498: you specified 820,612 entries for lithology, and 273,684 matched records with the thesaurus. What happened with the remaining 546,819 entries (66.6% of the total entries)? Can you give a simple example of entries not matched?
→ We did not obtain a match with a score greater than 80. Example of unmatched entries in Table 2. Added this information the manuscript.

Results, Line 507: Does it mean that in about 40,000 records you had a ratio() score from the fuzzy string matching) lower than 80? Maybe you could be more explicit in this? Additionally, you defined

the score threshold based on the exact match. But, might it be a kind of balance between the number of matched records and the exact match percentage? I'm just wondering because it seems that by defining that threshold you lose a lot of entries to be matched (83.5%). Could you give a simple example of records not matched?

→ Yes, we expounded on this in the updated version. It is a balance, as you mention. In this case study, we selected a cut-off score of 80 since this is where the # of exact matches obtained plateaus. A lower cut-off score could be used, depending on the familiarity to the data and/or purpose of drillhole processing. For our case, we wanted to be as conservative as possible without being too stringent (cut-off score 100). Example of unmatched record in Table 2.

Results, Lines 542 - 560: These are more materials and methods than results.
→ This was revised in the updated version.

Results, Line 549: Couldn't you get the rest of the hierarchical categories based on the lower hierarchy defined?
→This is what we do, the Subgroup and Group level matched is based on the detailed lithology match.

Results, Table 2. It gives an example of unmatched cases, so disregard that part of previous comments.

Results, Lines 663 - 678. Accuracy metrics should be included in the materials and methods section and not in the results.
→ This was revised in the updated version.

Discussion, Lines 812: Did you tried to replace the "same as above" with the descriptions?
→ This was not included in the scope of this work. Replacing "same as above" requires building a dictionary for all possible permutations to refer to this (blank, ""). It also makes some assumptions that the rock type is exactly the same.

Discussion, Lines 824: Is there any way to automatise the building of a thesaurus given new advances in NLP and machine learning?
→This should be possible. However, there is a need to first understand the syntax in which geological data is captured. The thesaurus provided by dh2loop provides a starting point/training set for this.

Discussion, Line 831: "extraction rate of 16% from the Comments is not bad at all" This sounds too subjective, how do you define what is a good or a bad extraction rate?
→This was revised in the updated version by describing further that although it is a low extraction rate, there is value in being able to obtain 16% more data that was previously deemed "unusable". We also stated the number of records that amount to 16%.

Discussion, Lines 852-853: "For sedimentary rocks, the lack of a standard syntax as to how comments are recorded impacts the classification. " You see, imagine if such standardization is achieved, wouldn't it reduce the subjectivity
→ Standardization will definitely reduce subjectivity. However, achieving a standard is not an easy task, nor is it ours. This is something that is for the geological surveys to decide and implement. It is also important to note that a "standard" would be tricky to achieve as the information and level of detail contained in logs is highly dependent on the purpose of the study. This study provides a basis for creating a pre-standard. Not so much providing a guide of practice but highlighting what shouldn't be done and what practices create ambiguity.

Discussion, Line 863: grammar error " Soils are technically are not rocks"
→ This was revised in the updated version.