Proposed Evaluation for TA2

There are three outputs for TA2
1) Grade and tonnage models
2) Mineral site data
3) Mappable criteria (SRI should provide the details for this deliverable)

Since there are various components that go into creating each of these outputs, Graham Lederer proposed the following breakdown to evaluate #1 and #2:

A) Deposit type classifications
B) Mineral site locations
C) Grade and tonnage inventory

A&B can be composed to produce the mineral site data
A&B&C can be composed to produce the grade and tonnage models

Deposit Type Classification

For any given mining site, the challenges iss to assign one of the 189 deposit types from the CMMI list of deposit types. The data to do this classification can come from either or both of the mining reports and/or the information from the various USGS databases.

Since some deposit types can be quite close, this should be evaluated by producing the top N deposit types and evaluating how close to the top the correct deposit type is proposed. The metric that captures that is the Mean Reciprocal Rank (see https://en.wikipedia.org/wiki/Mean_reciprocal_rank). I would propose that we evaluate the top N where N=5.

The container USC delivered provides the top from ranks that SRI has published into the CDR and USC then loaded into the knowledge graph.

Mineral site locations

The mineral site locations come from a variety of sources: USGS MRDS, USGS USMIN, mining documents, tables from papers. The UMN team is working on the reconciliation of this data and determines which sites are the same sites. The grouping of these sites coming from multiple sources is then published as a set of links directly to the CDR, which can then be evaluated.

Given ground truth data of mineral sites (e.g., from assessment data), we can evaluate the performance of FuseMine using the standard Adjusted Rand Index (ARI) and Normalized Mutual Information (NMI). We will also evaluate the result using Accuracy, which is calculated by the number of correctly linked sites divided by the total number of linked sites in the ground truth data: $\frac{\{linked\ sites\ from\ FuseMine\} \cap \{linked\ sites\ in\ the\ groud\ truth\}}{\{linked\ sites\ in\ the\ groud\ truth\}}$.

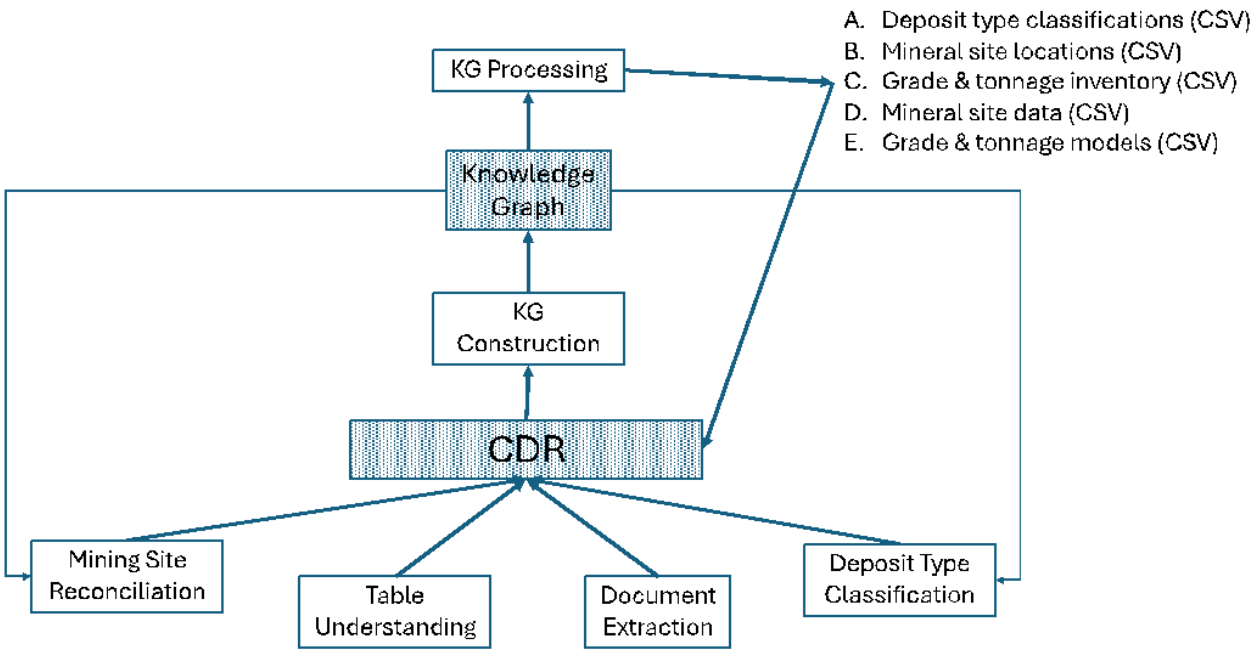
The container USC delivered provides the group sites that UMN has published into the CDR and USC then loaded into the knowledge graph.

Grade and tonnage inventory

The grade and tonnage inventory comes primarily from two sources: mining reports and tables of published data from papers on specific minerals. Since the papers contain significant detail, it requires some additional processing of the data to produce the final values used for grade and tonnage. The grade and tonnage inventory then is being compiled from the mining reports that Inferlink is processing and the tables of data that USC is compiling and then the postprocessing that USC is performing on top of this data to produce the final values. The data to evaluate is being published into the CDR by Inferlink and USC, but it must then be loaded into the knowledge graph to analyze and produce a final table of data. So the evaluation of the inventory should be done through a query on the knowledge graph.

The metrics to evaluate this are 1) recall to capture whether it was able to produce a grade and tonnage estimate for a given set of mines and 2) mean square error to measure how close the values are to the published values. We do expect some variance in the values since the estimates for a given mine can change over time and the important thing is to be in the right ballpark and not get a precise value. I would propose that we evaluate this by taking a published set of values for a given commodity and deposit type and evaluating how well we do in reproducing the values in that published report, measuring recall and mean square error.

The container USC delivered provides the grade and tonnage from both the Inferlink processed documents and the USC processed tables that are both published into the CDR and then loaded into the knowledge graph to produce the final result.


A. Deposit type classifications (CSV)
B. Mineral site locations (CSV)
C. Grade & tonnage inventory (CSV)
D. Mineral site data (CSV)
E. Grade & tonnage models (CSV)
KG Processing
Knowledge Graph
KG Construction
CDR
Mining Site Reconciliation
Table Understanding
Document Extraction
Deposit Type Classification