We are interested in evaluating completeness of metadata collections in different dialects with respect to a recommendation made in a single dialect. Our approach is illustrated in Figure 2 which shows two dialects, a conceptual recommendation with two levels (L1 and L2) in Dialect 1, implementations of the recommendation in dialects 1 and 2, and two metadata collections in each dialect.

Typically, recommendations are associated with a native dialect, as illustrated in Figure 1 with R1-5 and D1, so they include an implementation in that dialect. The first step in our analysis is to map those implementations (H-N) to dialect-independent documentation concepts (A-G). For example, the recommendation might recommend that the metadata include an XML element <title> that holds a dataset title and an element <pointOfContact> that holds the name of a point of contact. These two elements could be mapped to the documentation concepts “Resource Title” and “Resource Contact”. These mappings are identified by open, bi-directional arrows in Figure 2. Note that all the recommended concepts can be mapped to implementations in the native dialect, as communities do not recommend concepts that do not exist in their implementations. In the LTER case, the recommendations were originally described as documentation concepts, and we mapped them to implementations as the first step in our process.

Once the implementations are determined, the metadata evaluation is straightforward. We examine collections of metadata records to determine which of the concepts they include. We simplify the illustration here by considering only two concepts (A and E). Figure 2 includes two collections in Dialect 1. Implementation H of concept A is included in all four records in the first collection (indicated by filled arrows) and in two of the three records in the second. Implementation L of concept E is included in two of the four records in collection 1 and all three of the records in collection 2. The “occurrence completeness” of concept H in the first collection is 100% and of concept E is 50%. In collection 2 these are 66% and 100% respectively.

In many cases, we identify groups of metadata records that include, and therefore are missing, the same concepts. Collection 1 includes two such groups. The first two records are missing concept E and the second two records are not missing either A or E. We term these *signature groups* and identify them by the number of concepts that they are missing in each level of the recommendation. The signature of the first group in collection one is “0 1” as these records are missing zero concepts from L1 and one concept from L2. The signature of the second group is “0 0” as they are missing zero concepts from L1 and L2. Note that low numbers are better in these signatures, so “0 0” indicates a complete record and the sum of the signature group is the total number of concepts missing from the records in the group. Collection 2 also has two groups. The first two records are in group “0 0” and the third is in group “1 0”.

The process for applying the recommendation from Dialect 1 to metadata that are in dialect 2 differs slightly from this process because some of the concepts in the recommendation may not exist in dialect 2. This can happen for many reasons. Our understanding of documentation needs evolves as a function of time and newer dialects may include concepts that emerge from that evolution. For example, the importance of unique identifiers in metadata records has become clear in the last decade, but dialects created prior to that time do not include that concept. Dialect differences may also reflect differences between data types in different communities. In Figure 2, concepts A and B exist in dialect 1 but are not included in dialect 2. This effectively decreases the maximum possible score for L1 in dialect 2 from four to two. We refer to this new maximum as the *dialect max* for L1 in dialect 2.

The metadata evaluation step is the same in dialect two and leads to the identification of signature groups in the same way as the previous analysis. In this case, four groups are identified. Note that these signature groups are described as counts of missing elements, so smaller sums still indicate more compete records. This also makes it possible to quantify the amount of improvement needed for complete collections. Each signature group is missing (number of missing elements) X (number of records) total elements. This number gives provides a starting metric that can guide and measure the improvement process.