Supplementary Note 3: Type-wise error analysis of the node scoring model

The scorer shows larger prediction error on broader lexical/category nodes, especially Culture-loaded Word and upper-/lower-level cultural content classification, than on provenance-oriented nodes such as Source, Chapter/Content Unit, Person, and Time Period. This pattern is consistent with the greater semantic ambiguity of broad lexical/category nodes, whereas provenance-oriented nodes usually play more stable evidential roles under question-conditioned labeling.