

# Query language and collation

When specifying the lexical search type for term matching there is a need to specify the collation used, and to specify the default collation for the language in which the terms are to be matched are represented.

Examples based on mysql collation behavior:

"AAO" matches "ÄÄÖ" in utf8\_generic\_ci and utf8\_unicode\_ci (and utf8\_german2\_ci) but not in utf8\_swedish\_ci collation.

"Aâa" matches "aaa" in utf8\_generic\_ci and utf8\_swedish\_ci but not in utf8\_bin collation (i.e. case insensitive vs. sensitive, sometimes you need case sensitivity when searching...).

Similar behavior can be implemented e.g. by java.text.Collator in java or by the collection.find() or cursor.collation() method in MongoDB.