Followup Question on ICD10 Dictionary and Concept Matching Quality

---

me to dev@ctakes.apache.org   10:49 AM

Hi cTAKES Community, Sean, Tim,

As you recall, I recently put together the YouTube video "cTAKES: How to Create an ICD10 Dictionary" (https://www.youtube.com/watch?v=4aOnafv-NQs) to improve the documentation of using the NIH/UMLS and cTAKES tooling to generate custom dictionaries. It is working well and Dr. Tim Miller's cTAKES Docker repository will soon have further documentation/support for using this approach in establishing multiple custom cTAKES dictionaries.

However, I am writing to because while everything with ICD10 is working, I'm finding the concept matching to be inadequate as compared to SNOMED/RXNORM (I'm sure I'm just missing something :)). For example, if I type in "Type 2 Diabetes", there is no concept match. However, when I type "Type 2 Diabetes Mellitus", there is a match. Is there a way I can better "train" or configure the dictionary to have matching at parity with that of the SNOMED/RXNORM dictionaries?

For reference, here is the ICD10 configuration I'm using:

```
<dictionary>
  <name>icd10Terms</name>
  <implementationName>org.apache.ctakes.dictionary.lookup2.dictionary.UmlsJdbcRareWordDictionary</implementationName>
  <properties>
        <property key="jdbcDriver" value="org.hsqldb.jdbcDriver" />
        <property key="jdbcUrl" value="jdbc:hsqldb:file:resources/org/apache/ctakes/dictionary/lookup/fast/icd10/icd10" />
        <property key="jdbcUser" value="sa" />
        <property key="jdbcPass" value="" />
        <property key="rareWordTable" value="cui_terms" />
        <property key="umlsUrl" value="https://uts-ws.nlm.nih.gov/restful/isValidUMLSUser" />
        <property key="umlsVendor" value="NLM-6515182895" />
        <property key="umlsUser" value="CHANGE_ME" />
        <property key="umlsPass" value="CHANGE_ME" />
  </properties>
</dictionary>

<conceptFactory>
  <name>icd10Concepts</name>
  <implementationName>org.apache.ctakes.dictionary.lookup2.concept.UmlsJdbcConceptFactory</implementationName>
  <properties>
        <property key="jdbcDriver" value="org.hsqldb.jdbcDriver" />
        <property key="jdbcUrl" value="jdbc:hsqldb:file:resources/org/apache/ctakes/dictionary/lookup/fast/icd10/icd10" />
        <property key="jdbcUser" value="sa" />
        <property key="jdbcPass" value="" />
        <property key="umlsUrl" value="https://uts-ws.nlm.nih.gov/restful/isValidUMLSUser" />
        <property key="umlsVendor" value="NLM-6515182895" />
        <property key="umlsUser" value="CHANGE_ME" />
        <property key="umlsPass" value="CHANGE_ME" />
        <property key="tuiTable" value="tui" />
        <property key="prefTermTable" value="prefTerm" />
        <property key="icd10cmTable" value="text" />
        <property key="icd10amaeTable" value="text" />
        <property key="icd10amTable" value="text" />
        <property key="srcTable" value="text" />
        <property key="icd10Table" value="text" />
        <property key="icd10aeTable" value="text" />
        <property key="icd10pcsTable" value="text" />
  </properties>
</conceptFactory>
```

As always, I appreciate the hard work this community has done and, based on the feedback from this thread, I will do my best to improve the documentation for others.

Thanks,

Matthew Vita
www.matthewvita.com