MetaMap's Strict and Relaxed Models

For efficiency, MetaMap partitions its UMLS data into Strict and Relaxed Models. MetaMap by default uses the Strict Data Model; the Relaxed Model is a proper superset of the Strict Model, and typically contains strings containing internal syntactic structure, such as conjunction. Each data model includes the Base, USAbase, and NLM Data Versions.

The following strings are just a few that appear in the relaxed model only:

- not fasting (C2585491)
- \bullet arms and legs (C0015385)
- \bullet asks questions (C0566217)
- abdomen not distended (C0424826)
- reaction time reduced (C0542189)
- ullet has difficulty with speech (C0233715)
- disease or syndrome of heart (C0018799)

In order to map text to such concepts, use the --relaxed_model (-C) command-line option.

We recommend using the Relaxed Model in conjunction with Term Processing.

Both MetaMap data models (Strict and Relaxed) include only the English strings from the UMLS sources (i.e., LAT == "ENG"). Users needing to include sources in languages other than English in order to process non-English text should create their own UMLS-like database using our Data File Builder suite, available here.