o   file name: LDT_lexicon.csv

o   file contents: data

o   file format: csv (comme-separated-values)

o   size of file (in bytes): 1.215 KB

o   other relevant dimensions:

        o   number of records: 11186

        o   number of fields: 9
            (author,subdoc,span,verb,voice,pattern,pattern_morph,pattern_fillers)

Documentation:

The file contains a database table. Each record (row) corresponds to a verb occcurrence in the Latin Dependency Treebank. The fields are the following:
1. author: author of the work where the verb occurrence appears;
2. subdoc: section of the work where the verb occurrence appears;
3. span: span of text where the verb occurrence appears;
4. verb: verb lemma of the verb occurrence;
5. voice: voice of the verb occurrence ('A' for active, 'P' for passive);
6. pattern: syntactic pattern containing the syntactic labels of the arguments of the verb ('Sb' for subject, 'Ocomp' for object complement, 'Pnom' for predicate nominal, 'Obj' for all other arguments);
7. pattern_morph: syntactic pattern containing the syntactic labels of the arguments of the verb plus morphological information on the case (for pronouns, nouns, and adjectives) or mood (for verbs) of the arguments;
8. pattern_fillers: syntactic pattern containing the syntactic labels of the arguments of the verb plus morphological information on the case or mood of the arguments, and lemmas of the arguments.

The metadata (author, subdoc, and span) are those used in the Latin Dependency Treebank (LDT 1.5). The documentation for LDT 1.5 is available at nlp.perseus.tufts.edu/syntax/treebank/ldt/1.5/docs/README.txt and contains the list of works and authors, together with the editions of the texts.
The syntactic annotation guidelines for the Latin Dependency Treebank are available at http://nlp.perseus.tufts.edu/syntax/treebank/ldt/1.5/docs/guidelines.pdf.