**Format**

DANFEVER contains three SQLite databases (SQLite Consortium, 2000); da_fever.db, da_wikipedia.db and den_store_danske.db.

*da_wikipedia.db* contains article data from Danish Wikipedia. It contains an id-field, which is the articles’ curid from Wikipedia. It also contain the text and title of each article, as well as the url to that article.

*den_store_danske.db* contains the same information from Den Store Danske, but the article ID is a simple enumeration, and the article texts are not included due to data-right (evidence extracts are included though as described below).

The *da_fever.db* database contain the annotated claims. Each row in the database contain a claim and a unique id. With each claims comes the labels verifiable (Verifiable and NotVerifiable) and label (Supported, Refuted and NotEnoughInfo). The evidence column contain information about what articles were used to create and annotate the claim, and is composed by a comma-separated string, with IDs referring to the articles. The ID-format is $Y_X$ where $Y$ is either wiki or dsd to indicate whether the article comes from Danish Wikipedia or Den Store Danske, and $X$ is the numerical id from that data-source. Finally the claims has a related evidence_extract that can be used to train inference models. Note that there may be some character-level incongruence between the original articles and the evidence_extract, due to formatting and scraping. The evidence extracts for Verifiable claims are the text-snippets used to create and annotate the claims. The evidence extract for NotEnoughInformation samples are based on a chrF1 search from the related article, in order to get relevant text.

All three databases are also provided in TSV-format.