Project Documentation

RI Annotation Package :

This package contains

Directories :

* AnnotatedData : contains all the manually annotated xml files
* Batches : Each batch with the original unannotated documents
* MAE\_tool : contains the tools for annotating

Modules :

* convert\_xml : utilities to transform the original xml documents into the format used for the annotation task, and prepare batches
* agreementEvaluation : functions to compute inter-annotator agreement
* main\_annotations : applies functions from both previous modules

Tables :

* annotators\_documents : a table which associates the document names with their annotator and the batch they were part of

Normalization package :

Directories :

* DataTables : contains dataframes (in excel format), with i2b2 annotations, as well as the result of the first filtering of time expressions.

Modules :

* analysis\_original\_data : performs analysis on the annotations provided by Weiyi
* extract\_annotations : extracts i2b2 original xml annotations to create the data tables found in DataTables
* filter\_ri\_timexes : performs the first filtering of relative and imprecise time expressions (before the annotation step)
* svm\_anchoring : contains all the functions to train svm classifier on the original anchor problem
* map\_custom\_annotations : utilities to transform our annotations into a format that can be used to train svm classifiers
* main\_normalization : executes functions from the other scripts