annotate

Table of contents

1. Tools for annotations

- Annotation.org provides open-source resources for linguistic annotation. These include tools for human annotation, automatic computer annotation, and managing annotation projects.

- TIMS (Tag Information Management System) - University of Tokyo
  - allows user to perform/view tagging on a particular document
  - tag information is stored separately from original documents and managed using an external database software (various different types of tags for the same document can be added)
  - keeps track of the Audit Trail or History, i.e., the date and time, the user or system that performed the tagging etc
  - exports a document from TIMS (all the tag information will be converted to XML format and embedded within the document for portability)

- Annotate

Annotate is a tool for efficient semi-automatic annotation of corpus data. It facilitates the generation of context-free structures and additionally allows crossing edges. Functions for the manipulation of such structures are provided. Terminal nodes, non-terminal nodes, and edges are labeled. In the NEGRA project, these labels are used for parts-of-speech and morphology (terminal nodes), phrase categories (non-terminal nodes), and grammatical functions (edges). Type and number of labels are defined by the user. Annotated corpora are stored in a relational database. Annotate has a specified interface for communication with external taggers and parsers.