Semantic annotation is a machine-readable information about the meaning of texts (and other things which contain information: e.g., images, sounds, videos, etc.). Machines cannot understand what a text or an image refers to, but with the help of semantic annotation, they can process the meaning and carry out sophisticated tasks for users.

- Locating important facts reported in biomedical articles
- Searching images of a Buddhist statue

Ontology is a knowledge representation defined as “a specification of a conceptualization” (by Gruber 1993). An ontology usually represents relations between important concepts (such as is-a relation, part-of relation). Linking semantic annotations with classes/instances in ontology is important to get a logical description for reasoning with objects in texts.

Who does the annotation?

Manual annotation is a time-consuming and expensive task, thus researchers are seeking the way to make annotations automatically (or semi-automatically). Our group take the machine learning approach summarized as follows:

- Human annotators create a small corpus
- Machines learn how to make annotations from the corpus

Open Ontology Forge (OOF) is a software designed to have integrated functions to support human annotators to construct annotated corpora. This software provides a convenient environment for human annotators to do the tasks as follows:

- Ontology creation
- Text annotations based on the ontology
- Event annotation
- Image annotation
- Property/individual editing

In the adult, <NAME author="aichan" created="2003-12-01 11:23:16" expression_type="Name" id="1000001" modified="2003-12-01 11:23:16" ontology_id="protein" pool_id="C 000000" sure="True" term="True" text="FHL2">FHL2</NAME> is expressed in the myocardium of the heart and in the epithelial cells of the prostate, where <NAME author="aichan" created="2003-12-01 11:23:53" expression_type="Others" id="1000213" modified="2003-12-01 11:23:53" ontology_id="protein" pool_id="C 000000" sure="True" term="False" text="it">it</NAME> colocalizes with the AR in the nucleus.

An example of output in in-text XML format

<bibliography>