The BioMeXT group specializes in Information Extraction (IE), from the biomedical literature, as well as from other textual sources. Information extraction consists in the task of automatically extracting structured information from textual documents, and it is an important component of Text Mining systems. We focus in particular on the extraction of domain-specific entities (such as genes, proteins, drugs, diseases), and their semantic relations (e.g. protein-protein interactions, gene-disease associations). Our tools are often evaluated through participation in community-run evaluation challenges (e.g. BioCreAtiVe: Critical Assessment of Information Extraction systems in Biology). Additionally, we provide an environment for Assisted Curation (ODIN), which is currently being used in the curation pipeline of the RegulonDB database in a project funded by the US National Institutes of Health.

The Bio Term Hub (BTH) is a combined terminological resource created by dynamically sourcing entity names and their identifiers from reference databases. A web interface allows a user to access selected resources and download them in an uniform format. http://www.ontogene.org/resources/termdb

The OntoGene's Biomedical Entity Recogniser (OGER) is a RESTful web service implemented on top of the BTH which allows a remote user to batch annotate a collection of documents. Recently, we have participated in a community-organized evaluation of Bio Text Mining services (BioCreative/TIPS), in which our system obtained the best results according to several of the evaluation metrics. http://www.ontogene.org/resources/oger

Funding

The OntoGene/BioMeXT group is partially supported by the Swiss National Science Foundation (SNF): grants 100014-118396/1, 105315-130558/1, CR30I1_162758 (F. Rinaldi). We are also involved in an SNF-NRP74 project, and in a large project funded by the US-NIH. Additionally we received funding from the COGITO foundation and the Federal Food Safety and Veterinary Office (FSVO). Recently a new project proposal has been approved by the Commission for Technology and Innovation (CTI).