Genozymes

Overview

In the Genozymes project, we investigated semantic technologies for scientists in biology, biochemistry, and genomics for the development of bioproducts and bioprocesses, in particular for second generation biofuel production.

The Biofuel Process

Biofuels produced from biomass are considered to be promising sustainable alternatives to fossil fuels. The conversion of lignocellulose into fermentable sugars for biofuels production requires the use of enzyme cocktails that can efficiently and economically hydrolyze lignocellulosic biomass. As many fungi naturally break down lignocellulose, the identification and characterization of the enzymes involved is a key challenge in the research and development of biomass-derived products and fuels. One approach to meeting this challenge is to mine the rapidly-expanding repertoire of microbial genomes for enzymes with the appropriate catalytic properties.
Integrating semantic support in curation, analysis and retrieval

Semantic technologies, including natural language processing, ontologies, semantic Web services and Web-based collaboration tools, promise to support users in handling complex data, thereby facilitating knowledge-intensive tasks. Trying to select the appropriate technologies and combine them in a coherent system that brings measurable improvements to the users, we work on the development of a semantic infrastructure in support of genomics-based lignocellulose research. Part of this effort is the automated curation of knowledge from information on fungal enzymes that is available in the literature and genome resources.

Text mining results from our OrganismTagger and mycoMINE systems displayed in Firefox through the Semantic Assistants plug-in
Working closely with fungal biology researchers who manually curate the existing literature, we develop ontological natural
language processing pipelines integrated in a Web-based interface to assist them in two main tasks: mining the literature for
relevant knowledge, and at the same time providing rich and semantically linked information.

Project Members

Project Supervision

- René Witte

Group Members

Marie-Jean Meurs
  Development of the mycoMINE text mining pipeline
Nona Naderi
  High-performance species name recognition system OrganismTagger
Bahar Sateli
  Semantic Wikis with NLP support for curation through IntelliGenWiki

Further Information

The following publications are currently available:


For information on the parent project, please visit the Genozymes Wiki.

Funding

The Genozymes project is funded by Genome Canada and Génome Québec. The Semantic Assistants project, the OrganismTagger system, and IntelliGenWiki are funded through an NSERC Discovery Grant.

Except where otherwise noted, all original content on this site is copyright by its author and licensed under a Creative Commons Attribution-Share Alike 2.5 Canada License.