Semantic text mining support for lignocellulose research

Submitted by mj [1] on Tue, 2012-05-01 10:44

- Genozymes [2]
- Bioinformatics [3]
- NLP [4]
- Text Mining [5]

Title	Semantic text mining support for lignocellulose research
Publication Type	Journal Article
Year of Publication	2012
Authors	Meurs, M. - J. [6], C. Murphy [7], I. Morgenstern [8], G. Butler [9], J. Powlowski [10], A. Tsang [11], and R. Witte [12]
Journal	BMC Medical Informatics and Decision Making
Volume	12
Issue	Suppl 1
Pagination	S5
Date Published	04/2012
ISSN	1472-6947

Abstract

BACKGROUND
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.

RESULTS
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. An ongoing challenge is to select the appropriate technologies and combine them in a coherent system that brings measurable improvements to the users. We present our ongoing 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.

CONCLUSIONS
Working closely with fungal biology researchers who manually curate the existing literature, we developed 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.
URL	http://www.biomedcentral.com/1472-6947/12/S1/S5 [13]
DOI	10.1186/1472-6947-12-S1-S5 [14]
Copyright	
Impact Factor	2.23

Attachment	Size
BMCMIDM2012-1472-6947-12-S1-S5.pdf [16]	1.38 MB

Links:
[1] https://www.semanticsoftware.info/users/mj
[2] https://www.semanticsoftware.info/taxonomy/term/263
[3] https://www.semanticsoftware.info/category/topic/bioinformatics
[4] https://www.semanticsoftware.info/category/topic/nlp
[5] https://www.semanticsoftware.info/category/topic/text-mining
[6] https://www.semanticsoftware.info/biblio/author/79
[7] https://www.semanticsoftware.info/biblio/author/80
[8] https://www.semanticsoftware.info/biblio/author/81
[9] https://www.semanticsoftware.info/biblio/author/53
[10] https://www.semanticsoftware.info/biblio/author/82
[11] https://www.semanticsoftware.info/biblio/author/83
[12] https://www.semanticsoftware.info/biblio/author/1
[13] http://www.biomedcentral.com/1472-6947/12/S1/S5
[14] http://dx.doi.org/10.1186/1472-6947-12-S1-S5
[15] http://creativecommons.org/licenses/by/2.0
[16] https://www.semanticsoftware.info/system/files/1472-6947-12-S1-S5.pdf