Personalized Semantic Assistance
for the Curation of Biochemical Literature

Fedor Bakalov[1], Marie-Jean Meurs[2], Birgitta König-Ries[1],
Bahar Sateli[2], René Witte[2], Greg Butler[2] and Adrian Tsang[2]

seit 1558

[1] Friedrich Schiller University of Jena, Germany

[2] Concordia University, Montréal, QC, Canada

BIBM 2012   Oct. 6th, Philadelphia, PA

problem: CURATION OF BIOCHEMICHAL LITERATURE

finding relevant knowledge   →   of critical interest

  ▷ conducting research

  ▷ designing experiments

mining the literature   →   not that simple

  ▷ ever growing amount of scientific publications

  ▷ multiple repositories

  We are looking for needles in a huge haystack...

context: CROSS-DISCIPLINARY COLLABORATION

presented work

▷ computer scientists

user modeling

natural language processing (NLP)

software engineering

▷ biologists

genomics

molecular biology

users

▷ omics researchers

▷ biocurators

approach: WEB PLATFORM

web platform supporting researchers

    ▷ navigating literature

    ▷ curating relevant papers

single-point of access to abstracts

    ▷ publications harvested from multiple DBs
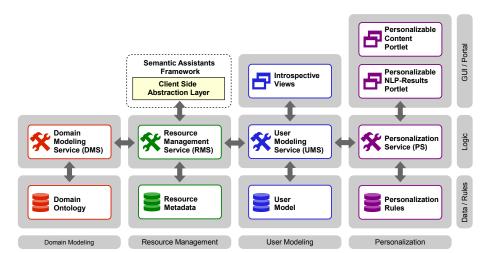
NLP services to process the literature

    ▷ various NLP pipelines

personalized view of the harvested papers

    ▷ editable according to the user's interests

Introduction
○○○

Architecture
●○○

User Interface
○○○○○○

Conclusion
○○

System Architecture - Overview

Introduction
ooo

Architecture
o●o

User Interface
oooooo

Conclusion
oo

System Architecture I

Domain Model Unit

▷ storing, accessing, managing the domain model

▷ domain model represented as an ontology (OWL)

Resource Management Unit

▷ harvesting literature

▷ annotating papers [↔SemanticAssistants↔NLP pipelines]

Introduction
000

Architecture
00●

User Interface
000000

Conclusion
00

System Architecture II

User Modeling Unit

▷ storing information about users' interests

▷ providing models, mechanisms, interfaces

▷ managing information about users required for adaptation

Personalization Unit

▷ storing personalization rules
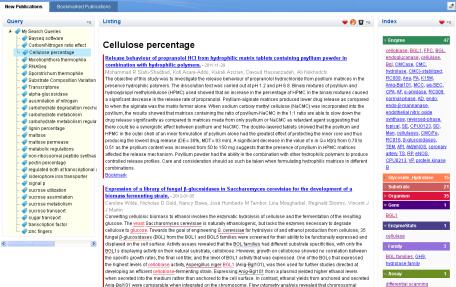
▷ providing mechanisms for performing personalization

Introduction
ooo

Architecture
ooo

User Interface
●oooooo

Conclusion
oo

Personalized Page

New Publications | Bookmarked Publications

Query

- My Search Queries
  - Bayseq software
  - CarbonNitrogen ratio effect
  - **Cellulose percentage**
  - Myceliophthora thermophila
  - RNASeq
  - Sporotrichum thermophile
  - Substrate Composition Variation
  - Transcriptome
  - alpha-glucosidase
  - assimilation of nitrogen
  - carbohydrate degradation mechs
  - carbohydrate metabolism
  - carbohydrate metabolism regula
  - lignin percentage
  - maltose
  - maltose permease
  - metabolic regulations
  - non-ribosomal peptide synthas
  - pectin percentage
  - regulated both at transcriptional a
  - siderophore iron transporter
  - signal p
  - sucrose utilization
  - sucrose assimilation
  - sucrose metabolism
  - sucrose transport
  - sugar transport
  - transcription factor
  - zinc fingers

Listing

Cellulose percentage

Release behaviour of propranolol HCl from hydrophilic matrix tablets containing psyllium powder in combination with hydrophilic polymers. - 2011-11-29

Mohammad R Siahi-Shadbad, Kofi Asare-Addo, Kakali Azizian, Davoud Hassanzadeh, Ali Nokhodchi

The objective of this study was to investigate the release behaviour of propranolol hydrochloride from psyllium matrices in the presence hydrophilic polymers. The dissolution test was carried out at pH 1.2 and pH 6.8. Binary mixtures of psyllium and hydroxypropyl methylcellulose (HPMC) used showed that an increase in the percentage of HPMC in the binary mixtures caused a significant decrease in the release rate of propranolol. Psyllium-alginate matrices produced lower drug release as compared to when the alginate was the matrix former alone. When sodium carboxy methyl cellulose (NaCMC) was incorporated into the psyllium, the results showed that matrices containing the ratio of psyllium-NaCMC in the 1:1 ratio are able to slow down the drug release significantly as compared to matrices made from only psyllium or NaCMC as retardant agent suggesting that there could be a synergistic effect between psyllium and NaCMC. The double-layered tablets showed that the psyllium and HPMC in the outer shell of an inner formulation of psyllium alone had the greatest effect of protecting the inner core and thus producing the lowest drug release (DE = 38%, MDT= 93 min). A significant decrease in the value of n in Q= k(t)n from 0.70 to 0.51 as the psyllium content was increased from 50 to 150 mg suggests that the presence of psyllium in HPMC matrices affected the release mechanism. Psyllium powder had the ability in the combination with other hydrophilic polymers to produce controlled release profiles. Care and consideration as such be taken when formulating hydrophilic matrices in different combinations.

Bookmark

Expression of a library of fungal β-glucosidases in Saccharomyces cerevisiae for the development of a biomass fermenting strain. - 2012-01-05

Caroline Wilde, Nicholas D Gold, Nancy Bawa, José Humberto M Tambor, Lina Mougharbel, Reginald Storms, Vincent J J Martin

Converting cellulosic biomass to ethanol involves the enzymatic hydrolysis of cellulose and the fermentation of the resulting glucose. Towards the goal of engineering S. cerevisiae for hydrolysis of and ethanol production from cellulose, 35 fungal β-glucosidases (BGL) from the BGL1 and BGL5 families were screened for their ability to be functionally expressed and displayed on the cell surface. Activity assays revealed that the BGL families had different substrate specificities, with only the BGL1s displaying activity on their natural substrate, cellobiose. However, growth on cellobiose showed no correlation between the the specific growth rates, the final cell titer, and the level of BGL1 activity that was expressed. One of the BGLs that expressed the highest levels of cellobiase activity, Aspergillus niger BGL1 (Anig-Bgl101), was then used for further studies directed at developing an efficient cellobiose-fermenting strain. Expressing Anig-Bgl101 from a plasmid yielded higher ethanol levels when secreted into the medium rather than anchored to the cell surface. In contrast, ethanol yields from anchored and secreted Anig-Bgl101 were comparable when integrated on the chromosome. Flow cytometry analysis revealed that chromosomal integration of Anig-Bgl101 resulted in a higher percentage of the cell population that displayed the enzyme but with overall lower expression levels.

Remove Bookmark

Index

Enzyme — 47

cellobiose, BGL1, FPC, BGL, endoglucanase, cellulase, Bgl, CMCase, CMC, hydrolase, CMC)-stabilized, RC008, Ana, PA, K15M, Anig-Bgl101, MCC, aq-SEC, CPA, AF, α-amylase, RC009, normal-phase, AD, endo-, endo-β-mannanase, endothelial nitric oxide synthase, reversed-phase, topical, SE, CPU0123, SD, Man, cellulases, CMC/Fe, RC016, β-glucosidases, TEM, API, AktieNOS, coronary artery, TS, RP, eNOS, CPU0213, VP, protein kinase B

Glycoside_Hydrolase — 15
Substrate — 21
Organism — 35

Gene — 1
BGL1

EnzymeStats — 1
cellulase

Family — 3
BGL families, GH8, hydrolase family

Assay — 1
differential scanning calorimetry

OrganismStats — 1
Saccharomyces cerevisiae

Introduction
○○○

Architecture
○○○

User Interface
○●○○○○○

Conclusion
○○

Query Portlet

▷ list of user search queries

▷ used by the portal to retrieve publications

▷ allows users to add, edit, delete queries

▷ allows to organize queries hierarchically

Click on a query → display a list of matching publications in the Listing portlet.

Introduction
ooo

Architecture
ooo

User Interface
o●oooo

Conclusion
oo

Listing Portlet

▷ request various types of semantic assistance

▷ view a list of named entities extracted from the publications

▷ access Semantic Assistants menu

▷ decide how to display results

▷ access personalization options ♡

Personalization

personalization menu
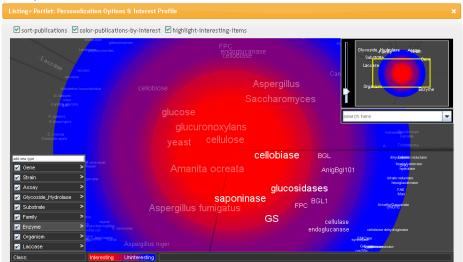
    ▷ switch between personalized and standard views

    ▷ display personalization options and users interest profile

    ▷ personalization options vary from portlet to portlet

listing portlet example: 3 personalization effects
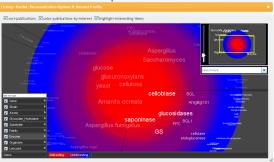
    ▷ sort publications according to user's interest profile

    ▷ highligh most interesting publications by a color marker

    ▷ highlight items from the user interest profile in the list

Visualisation

Introspective Views

Introspective Views

▷ each zone: items of certain interest degree

▷ each slice: items of a specific type

▷ hot zone: items that users are strongly interested in

▷ cold zone: items that users are not interested in

▷ getting overview, zooming, filtering, navigation, and search

▷ display relevant content upon request

▷ editing information in the model: adding/deleting items, changing interest degree, organizing items by type, defining user-specific types...

Introduction
000

Architecture
000

User Interface
000000

Conclusion
●○

Conclusion

What is done

 ▷ web platform supporting curation of biochemical literature

 ▷ single-point of access to publications

 ▷ process and analyze publications using semantic assistants

 ▷ personalized view of publications and NLP results

What is next

 ▷ **evaluation: user study**

 ▷ more publication sources (API?)

 ▷ integration of other textual contents

 ▷ components collaboration

 ▷ interaction user/semantic services

Acknowledgment