Integrating text-mining efforts for pathway curation requires output standards.

Andra Waagmeester

Maastricht University

March 22, 2011

Maastricht University

Introduction

Pathway Curation

Pathway Loom demo

Text mining

Text mining standard

How to expose text-mining results?

- Nifty website?

How to expose text-mining results?

- Nifty website?
- Sentences with "gold nuggets"?

How to expose text-mining results?

- Nifty website?
- Sentences with "gold nuggets"?
- Precision/Recall

How to expose text-mining results?

- Nifty website?
- Sentences with "gold nuggets"?
- Precision/Recall
- SOAP
- REST
- csv/tsv
- Unix pipes (ie: cat textfile — makelist.sh — sort — uniq -c)

Finding novel pathway parts on a pathway (1)

OMICS A Journal of Integrative Biology
Volume 13, Number 5, 2009

DOI: 10.1089/omi.2009.0029

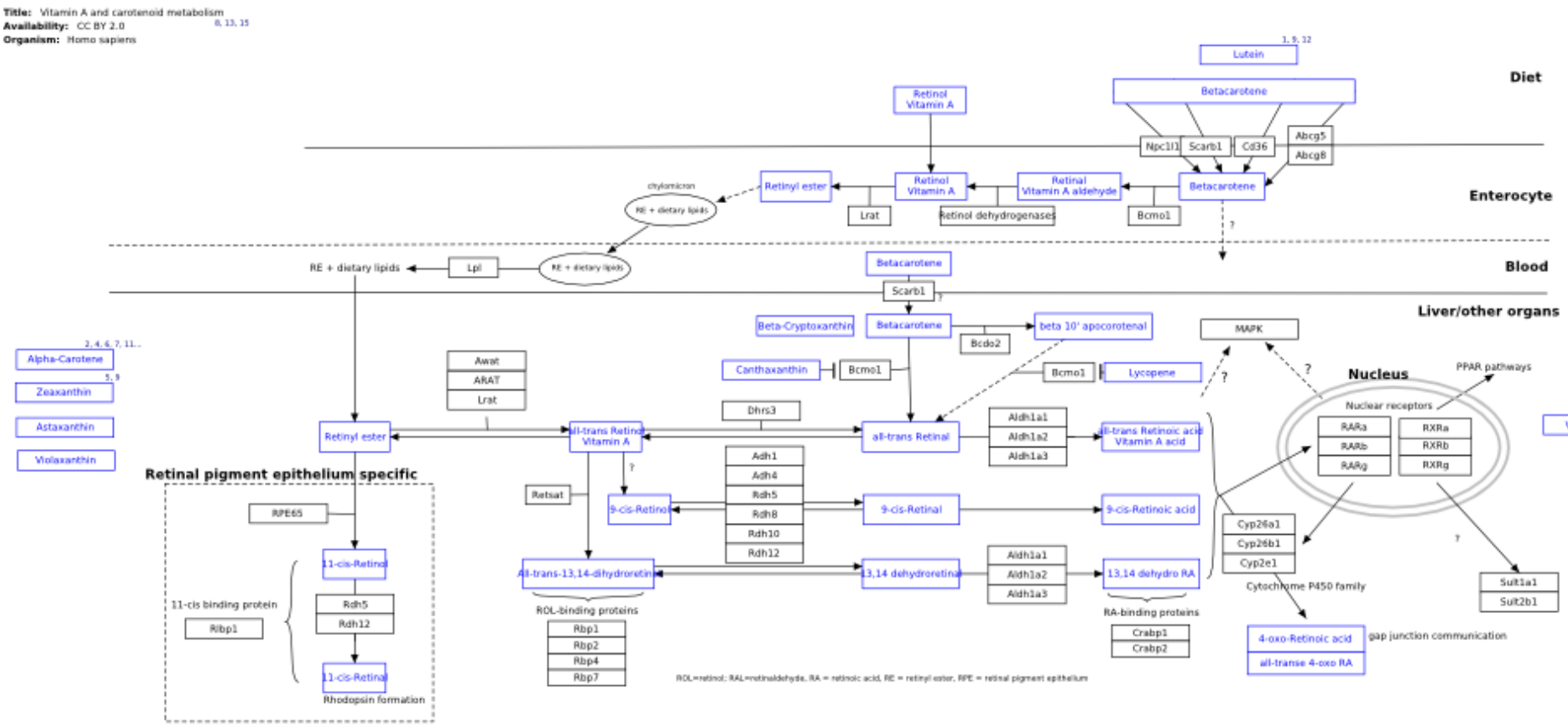
Original Article

Pathway Enrichment Based on Text Mining and Its Validation on Carotenoid and Vitamin A Metabolism

Andra Waagmeester,[1,2,3] Piotr Pezik,[1] Susan Coort,[3] Franck Tourniaire,[4] Chris Evelo,[3] and Dietrich Rebholz-Schuhmann[1]

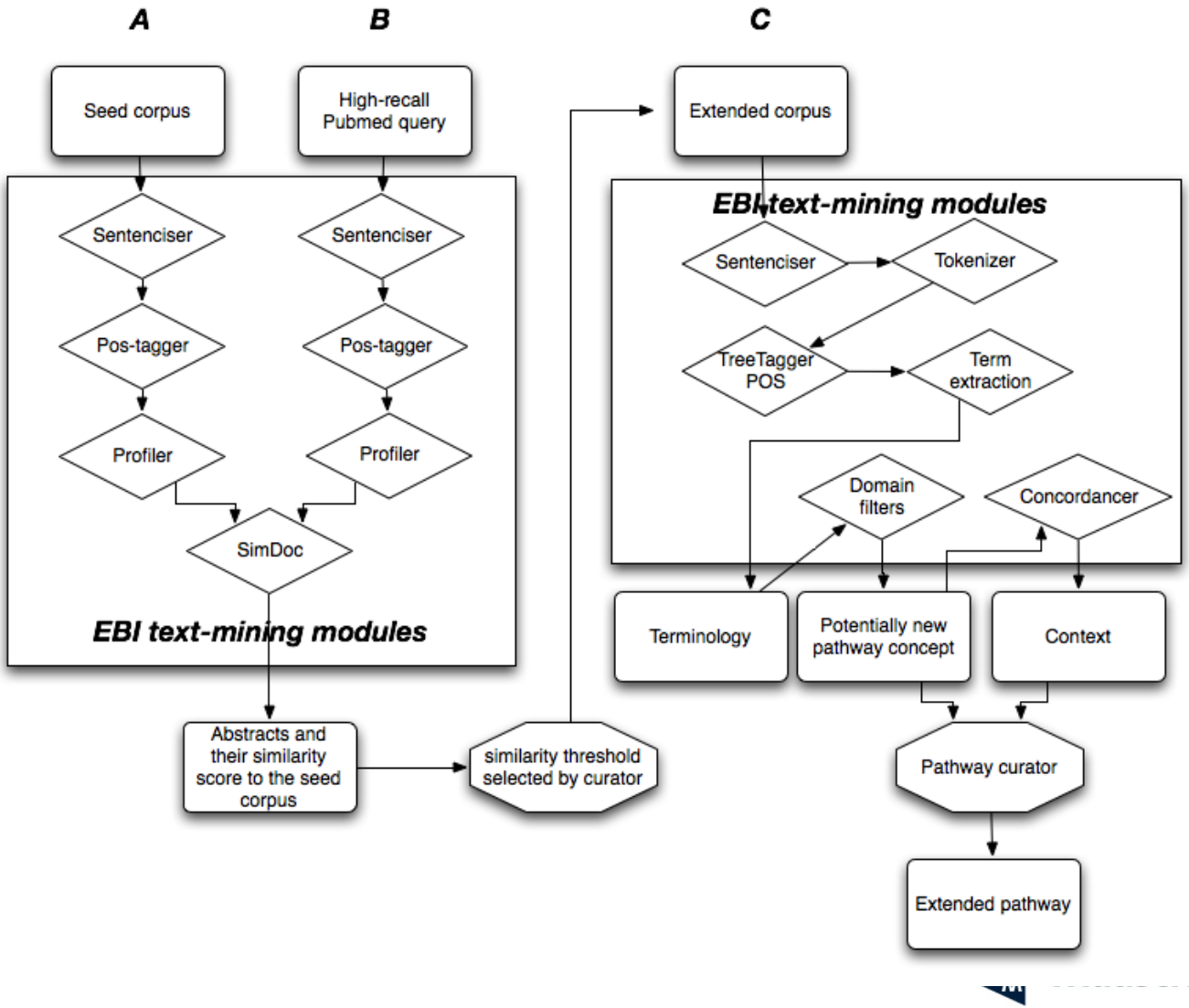
Maastricht University

Finding novel pathway parts on a pathway (2)



Maastricht University

Finding novel pathway parts on a carotenoid pathways (3)



ht University

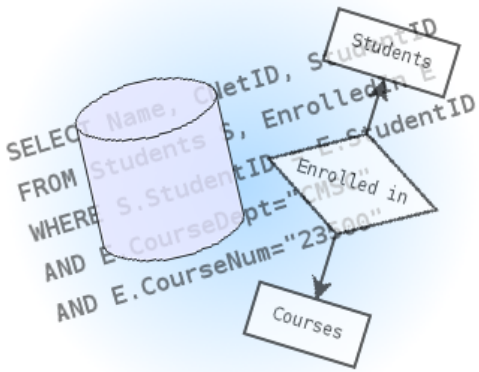
Finding novel pathway parts on a carotenoid pathways (4)

We were able to retrieve 13 novel pathway parts from Pubmed abstracts:

- Vitamin D3
- ADH3
- alpha-carotene
- astaxantin
- beta-cryptoxanthin
- Canthaxanthin
- lutein
- lycopene
- zeaxanthin
- PEPCK
- cryptoxanthin
- triglyceride
- lipase

Pathway curation

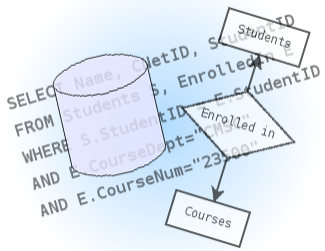
Knowledge is hidden in:



Maastricht University

Computers

× Can't listen

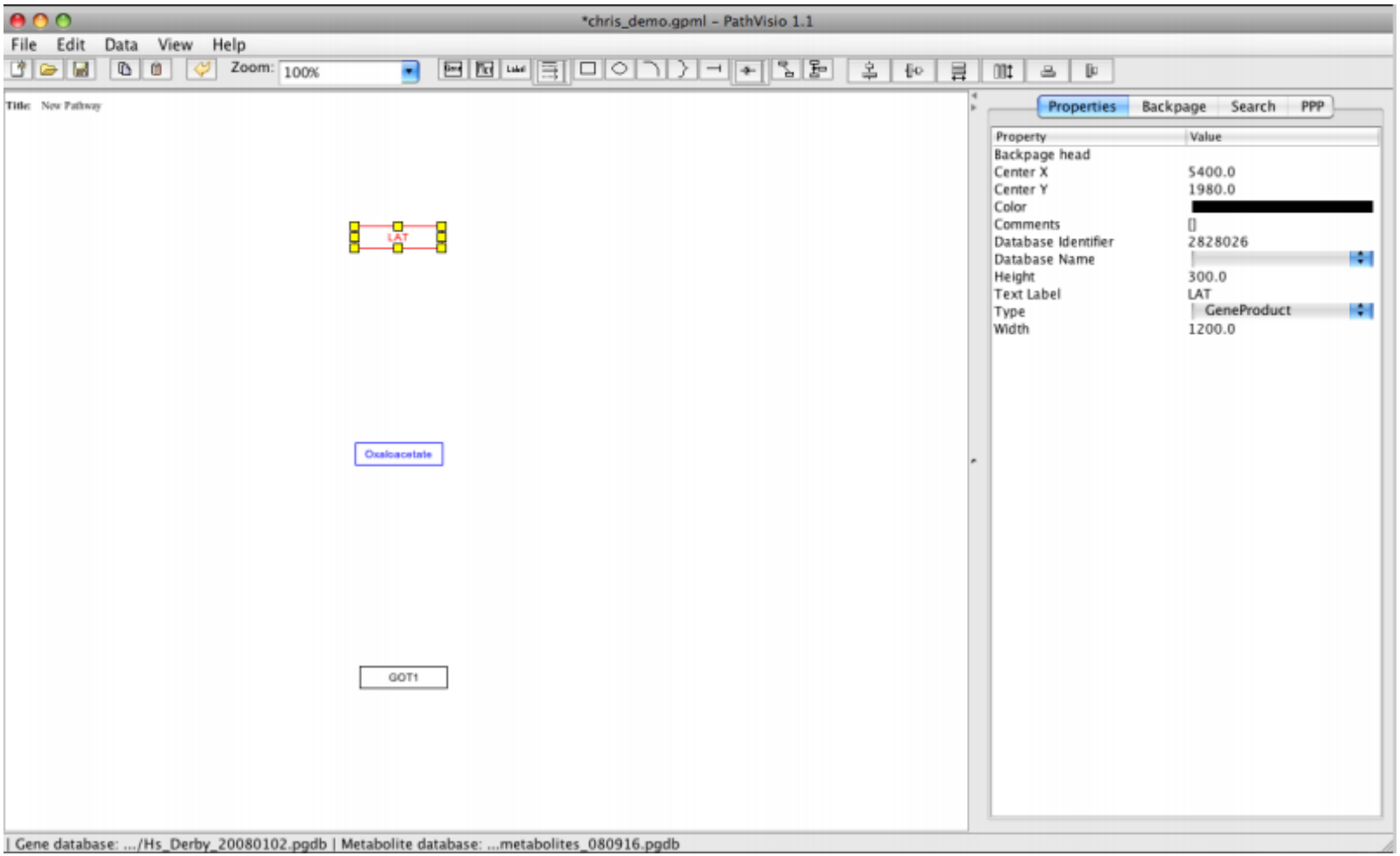


√ Can Query

× Can't read

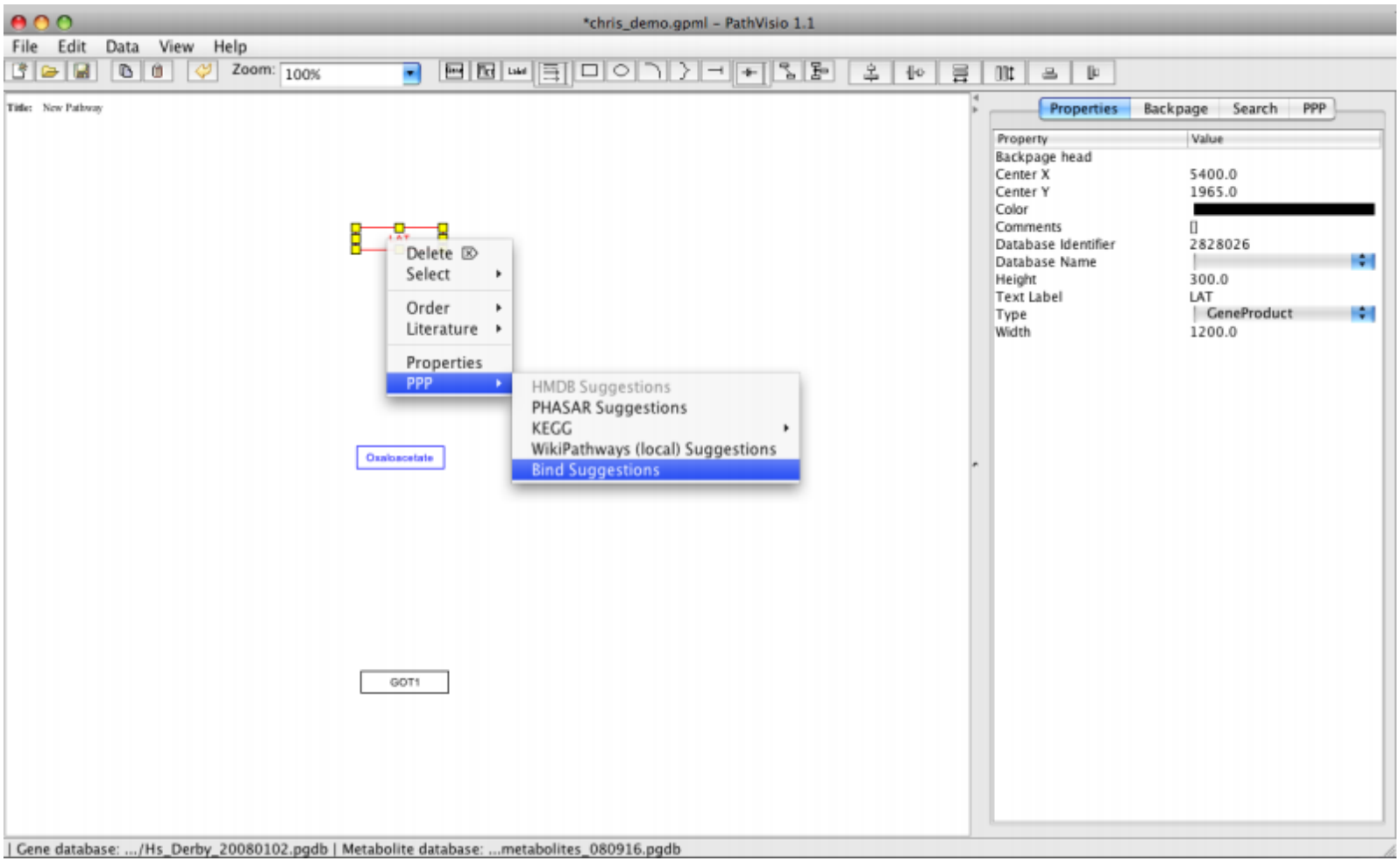
Maastricht University

Pathway Loom demo (1)



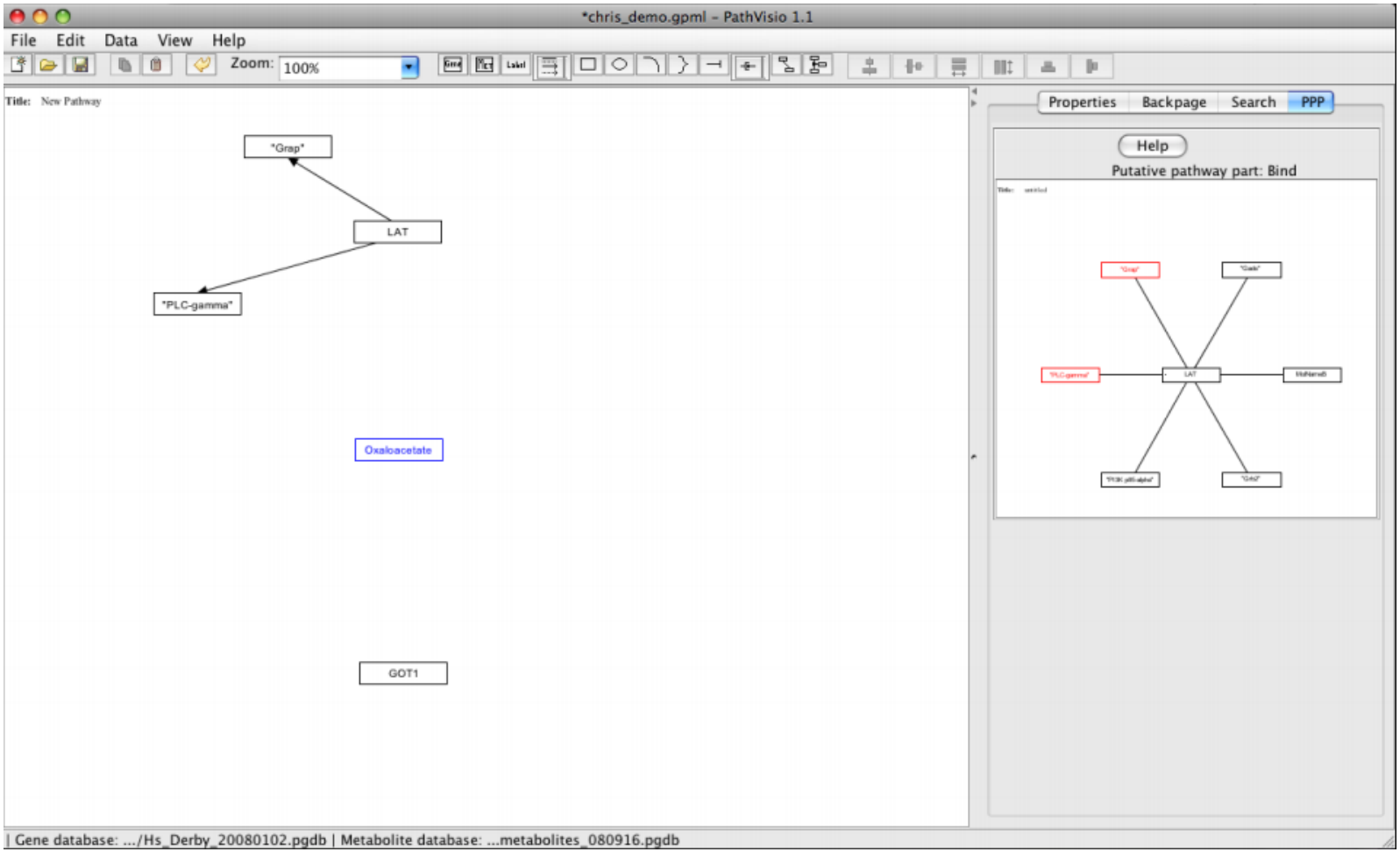
Maastricht University

Pathway Loom demo (2)

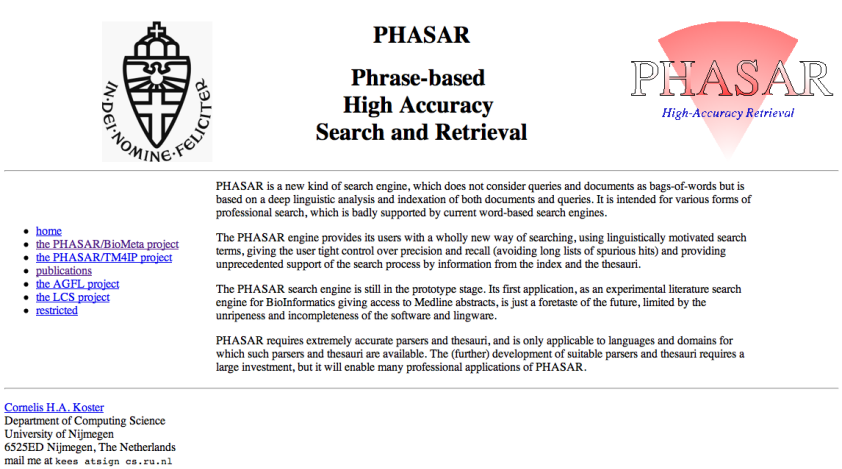


Maastricht University

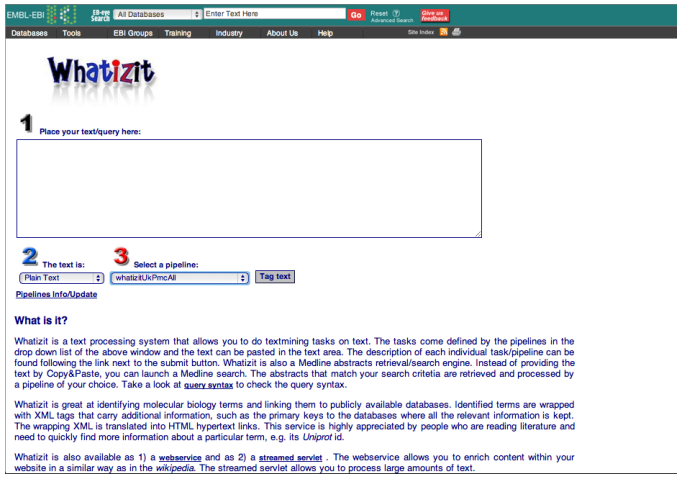
Pathway Loom demo (3)



Maastricht University



(g) PHASAR



(h) WHATIZIT (http://www.ebi.ac.uk/webservices/whatizit/info.jsf

Maastricht University

Andra Waagmeester 488 ●9 | log out | faq search

BioStar

Questions Tags Users Badges Unanswered Ask Ques

Which are stable text mining solution for biologists.

4

I am working on integrating pathway relations into a pathway diagram. I have two text mining solutions that I use to find interactions. I am looking for new suggestions. The idea is that I provide a pathway entity (ie. gene/protein, metabolite) and all relations based on the literature are returned.

text-mining interaction pathway

2 edit | rollback | close | delete | flag

edited Mar 1 at 7:28
Lars Juhl Jensen ♦
4,441 ●4 ●15

asked Feb 26 at 21:08
Andra Waagmeester
488 ●9

Could you please tell us which are the two text mining solutions that you already have? – Lars Juhl Jensen ♦ Feb 27 at 7:57

I am using Whatizit: ebi.ac.uk/webservices/whatizit/info.jsf and Phasar: phasar.cs.ru.nl. The first works decently, the second has issues with performance and accessibility. – Andra Waagmeester Feb 27 at 9:41

add comment

tagged
pathway × 11
text-mining × 10
interaction × 8

asked
15 days ago

viewed
205 times

latest activity
11 days ago

Spread the word!
Tell a colleague about BioS
http://www.biostars.org
Subscribe: RSS feed
Admin group: Biostar Cent

BioStar content is licensed u
Creative Commons Attribut
ShareAlike 3.0 Unported Lic

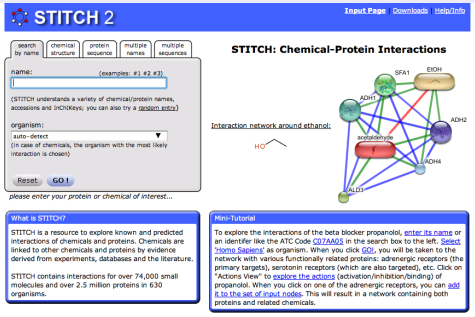
4 Answers

oldest newest votes

5

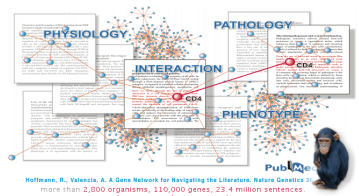
You might want to take a look at STITCH. It is a database that allows you to query with a gene/protein or a small molecule (e.g. a metabolite) and retrieve the interaction partners. A very large part of the evidence in STITCH comes from text mining, and if you really want to, you can turn off all the other evidence types to *exclusively* get the results obtained through text mining.

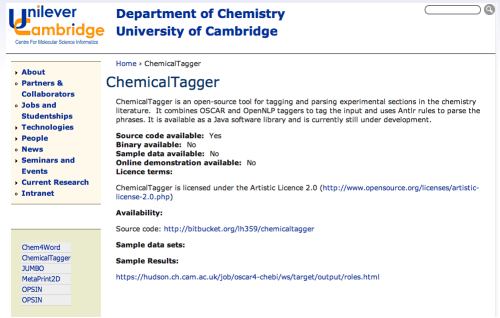
Nature Precedings : doi:10.1038/npre.2011.5850.1 : Posted 25 Mar



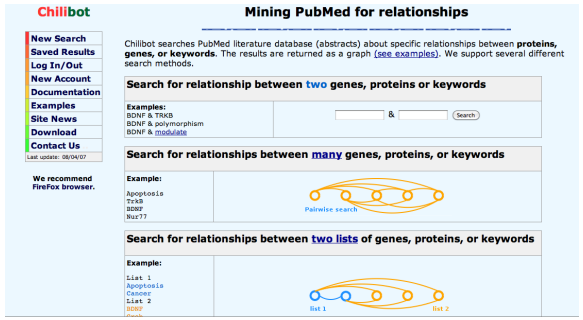
(i) STITCH

(j) iHOP

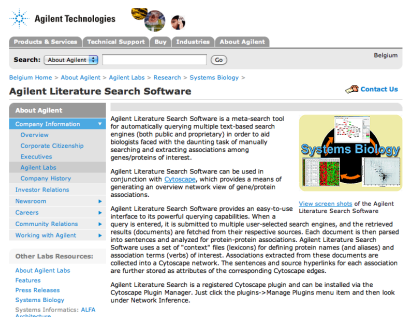




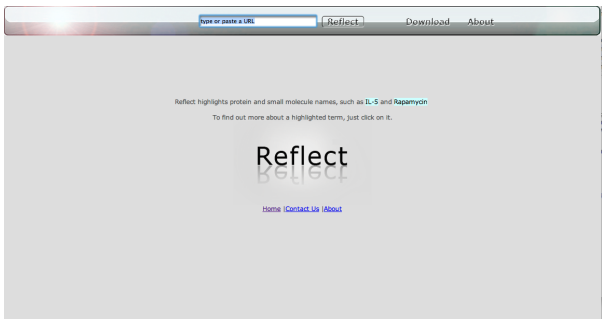
(k) CHEMICAL TAGGER



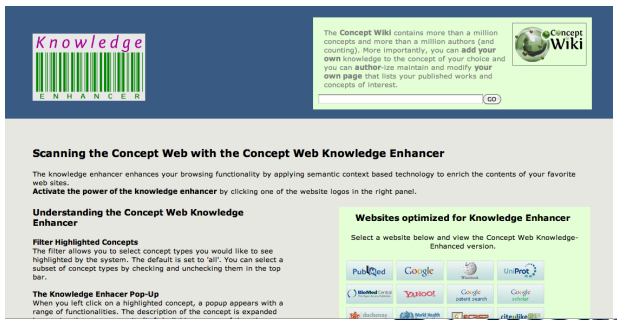
(l) CHILIBOT



(m) AGILENT



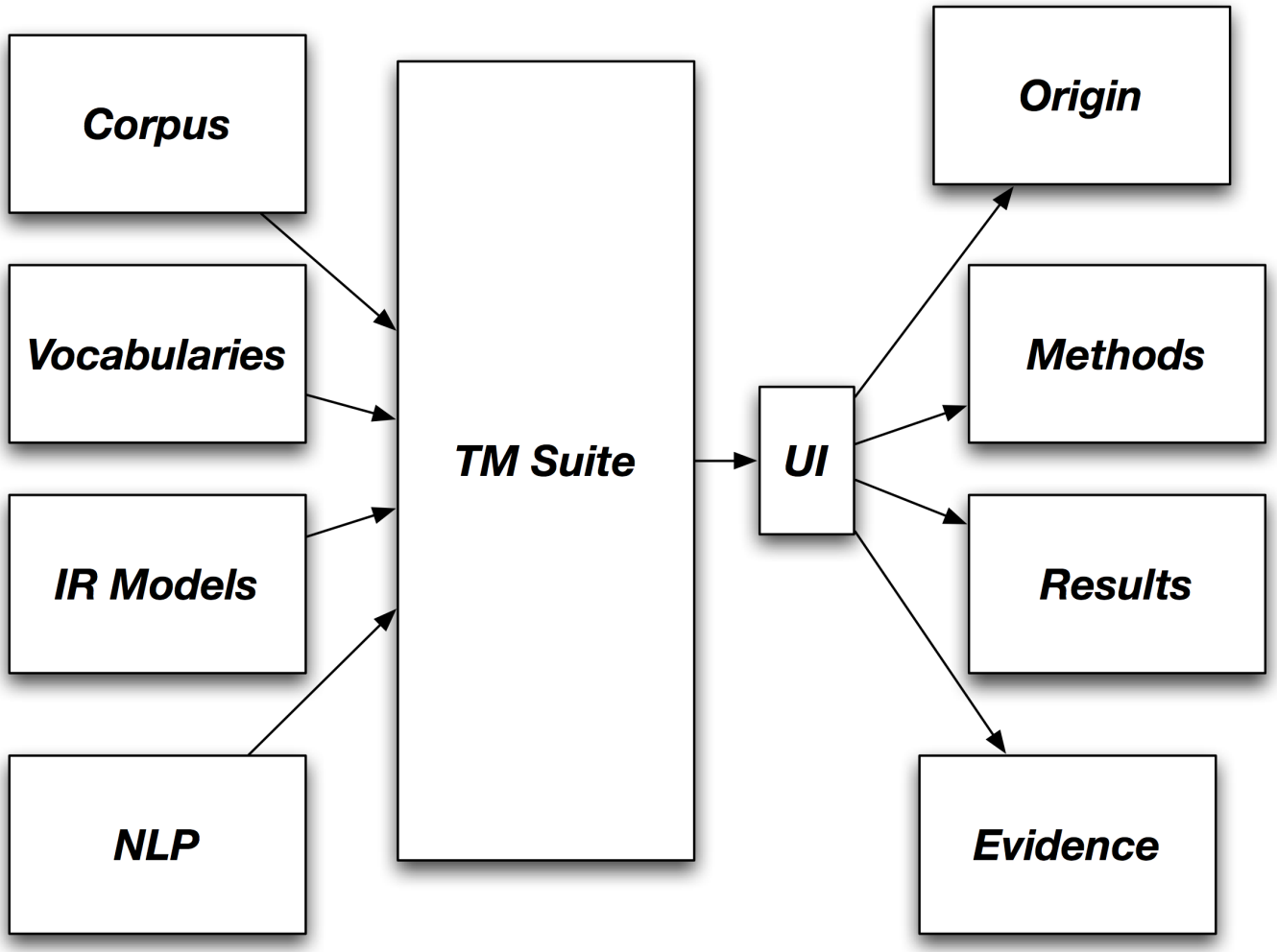
(n) REFLECT



(o) KNOWLEDGEENHANCER

Maastricht University

Text-mining standard for pathway curation



Origin

- Description
- Corpora
- Identifiers

Methods used

- Description
- Method name
- Vital statistics

Results

- Source
- Target
- Source Identifiers
- Target Identifiers

Evidence

- Relevant text parts

Awknowledgements

Maastricht University

- Chris Evelo
- Thomas Kelder
- Martina Kutmon
- Jahn Saito
- Martijn van Iersel

USC - SF

- Alexander Pico

BIOSTAR http://biostar.stackexchange.com

- Lars Juhl Jensen
- Casey Bergman
- Egon Willighagen
- Dominique Noel

nbic netherlands bioinformatics centre

Maastricht University