An extensible web interface for databases and its application to storing biochemical data

Nicos Angelopoulos and Paul Taylor

nicos@gmx.co.uk

Department of Biological Sciences,

University of Edinburgh, Edinburgh, Scotland

overview

- motivation
- generic features
- storing biochemical data
- R interface
- demo ?

motivation

Prolog provides high level of programming. Term structures are naturally translating to html code. Relational databases are a restricted form of LP.

Request and DB driven page construction fit well to the ‘closed world’ Prolog model

Swi provides:

- html <-> terms
- http server
- odbc interface

context

c. elegans intervention experiments

Utilising mass-spectrometry to identify effect of 3 different ligands (small chemical structures known to interact with c.elegans proteins).

4 labs: chemistry, c.elegans growth, spectrometry and analysis

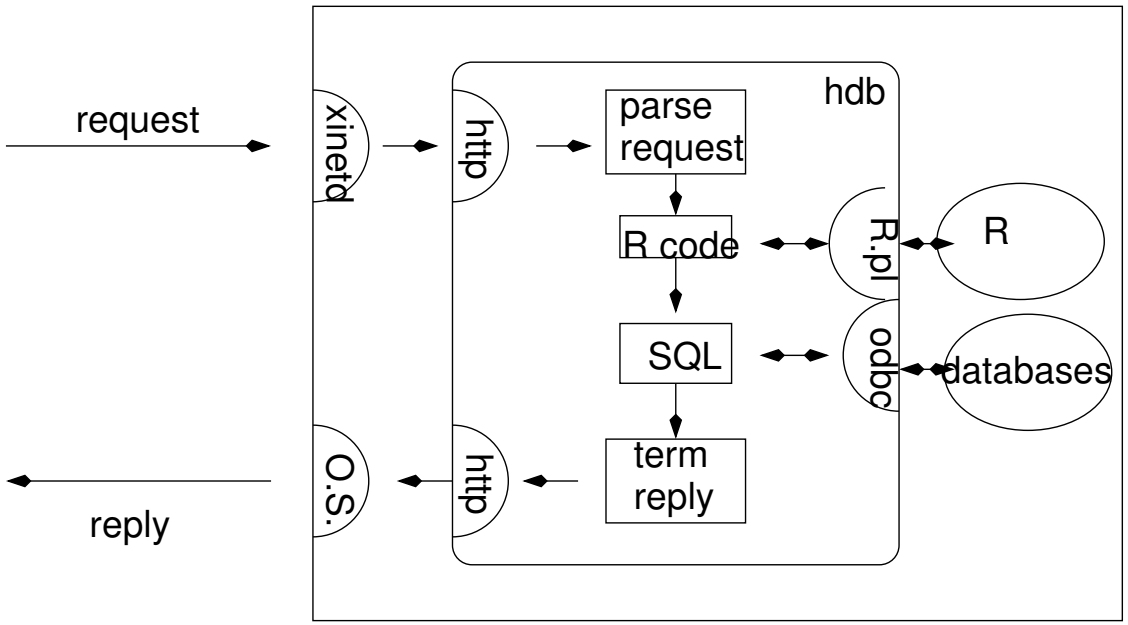
envisaged applications

Small intelligent interfaces, where relational databases and computation can be naturally combined.

Scientific databases, collaborative bibtex databases.

Full blown LIMS ?

architecture



Start server either from terminal (development) or via xinetd upon request on assigned port :

`http://scibsn1.bch.ed.ac.uk:8080`

odbc and schema interrogation

```
odbc_current_connection(-Connection,-DSN)

odbc_get_connection(+Connection,database_name(-DB))

odbc_current_table(+Connection,-Table)

odbc_table_column(+Connection,+Table,-Column)
```

hooks and handlers

hooks - ‘system plugs’

- default values- such as dates
- storing derived entries
- live links for html output
- file pointers

handlers - ‘user code’

- displaying molecules via Jmol applet
- statistical analysis of samples via R

standard extension operations

- repertoire of standard operations bridging the gap between database theory and user intuition

  - joint key input (e.g. input experiment with specific type)
  - multiple rows entry for single foreign key (e.g. purchase)

storing biochemical data

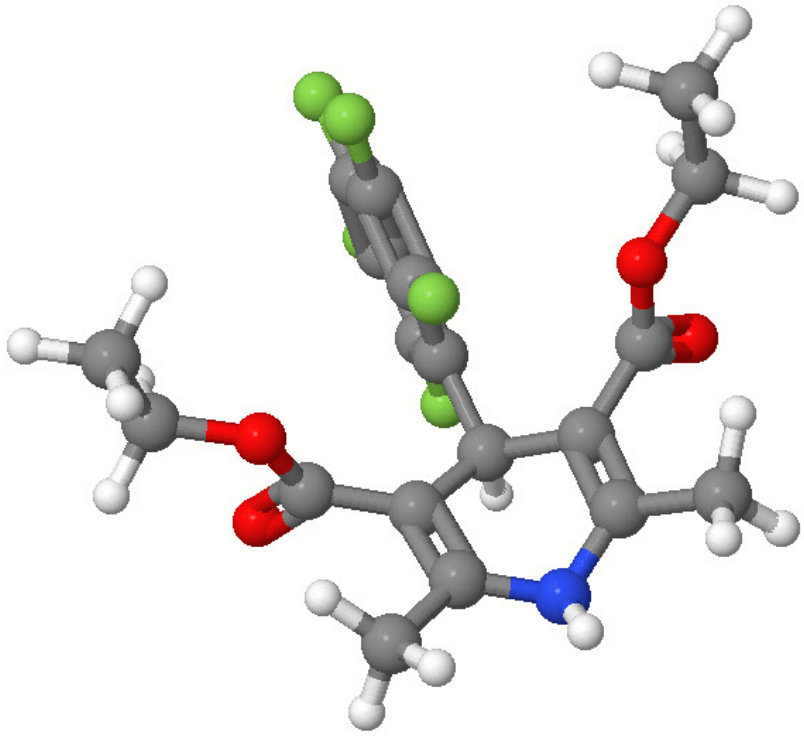
2 main databases: molecules + experiments

intuitive interface

store statistical analysis

display chemical molecules

visualising molecules



Create .sdf file from molecules db.

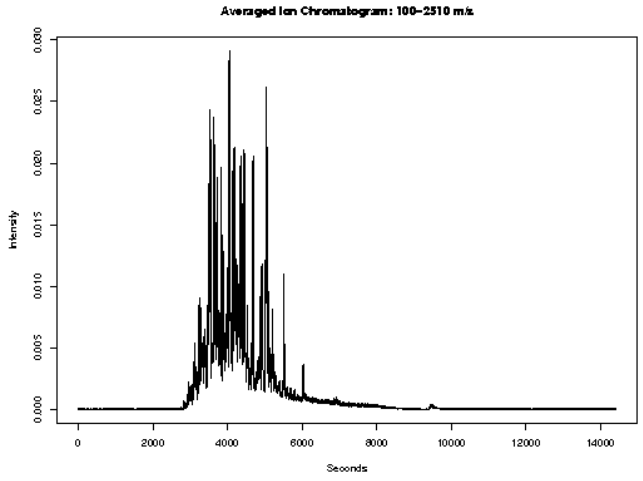
Create .html page with references to Jmol applet.

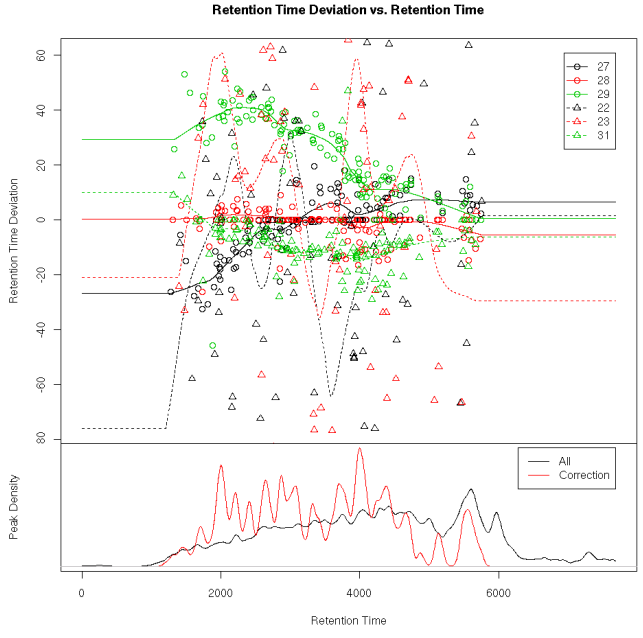
Serve redirection page.

interface to R

ScanID	bigint(20)	NO	PRI	autoinc	autoinc
ExperID	bigint(20)	YES			
ScanDate	date	YES			
LabFilename	text	YES			
ScanLoc	text	YES			Browse...
PinType	enum(2)	YES		glass	
Ligand	tinytext	YES			
FeedID	bigint(20)	YES			
Source	tinytext	YES			
Fraction	tinytext	YES			
ScanAICLoc	text	NO			
ScanIMGLoc	text	NO			
SpectraNof	int(4)	YES			
TimeMin	float	YES			
TimeMax	float	YES			
IntensityMin	float	YES			
IntensityMax	float	YES			
MassMin	float	YES			
MassMax	float	YES			
PrfMethod	tinytext	YES			
PrfStep	tinyint(4)	YES			
Note	text	YES			

input





R library

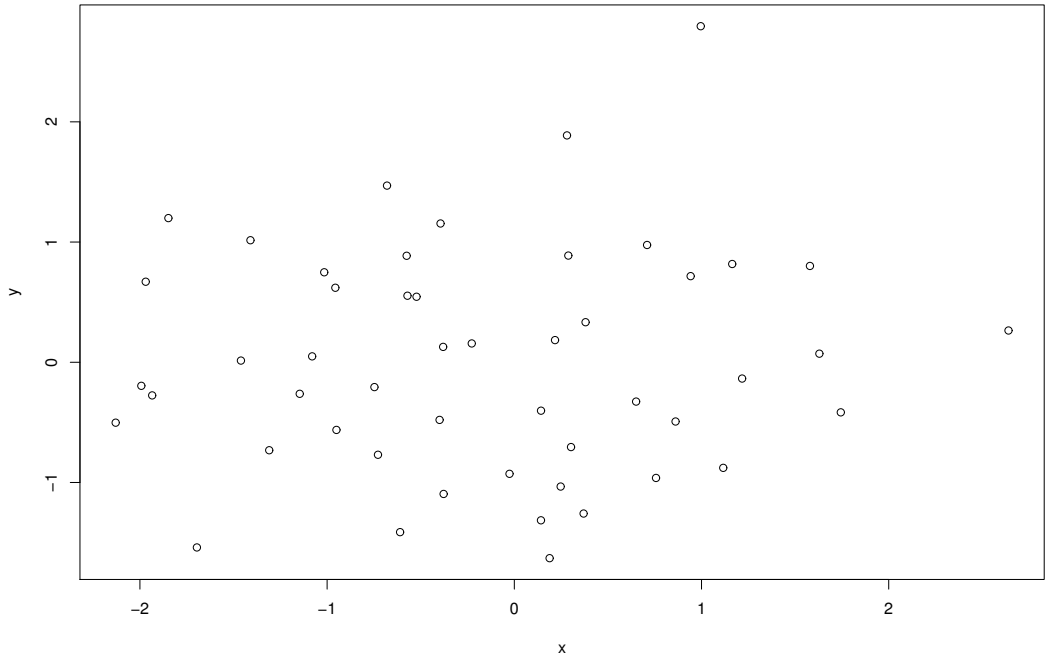
Session management

- r_bin(?Rbin)
- r_open(+Opts)
- r_close(+R)

Communicating

- r_in(+Rcmd)
- r_out(+Rcmd,-Lines)
- r_print(+Obj)

R examples



r_in(y <- rnorm(50))

r_print(y)

r_in(x <- rnorm(y))

r_in(x11(width=5,height=3.5))

r_in(plot(x,y))

X = [1,2,3,4,5]

Y = [6,6.5,7,8,8.5]

r_in('x11()')

r_in(plot(X,Y))

future work

Theory and clean implementation of basic operations.

Ability to select table/report rows.

Application driven development of the code.

`http://scibsfs.bch.ed.ac.uk/~nicos/sware/hdb`
`http://scibsn1.bch.ed.ac.uk:8080`