

# FuncNet

a distributed platform for  
high-throughput  
protein function analysis

Andrew Clegg  
University College London

*[funcnet.eu](http://funcnet.eu)*



# Outline of talk

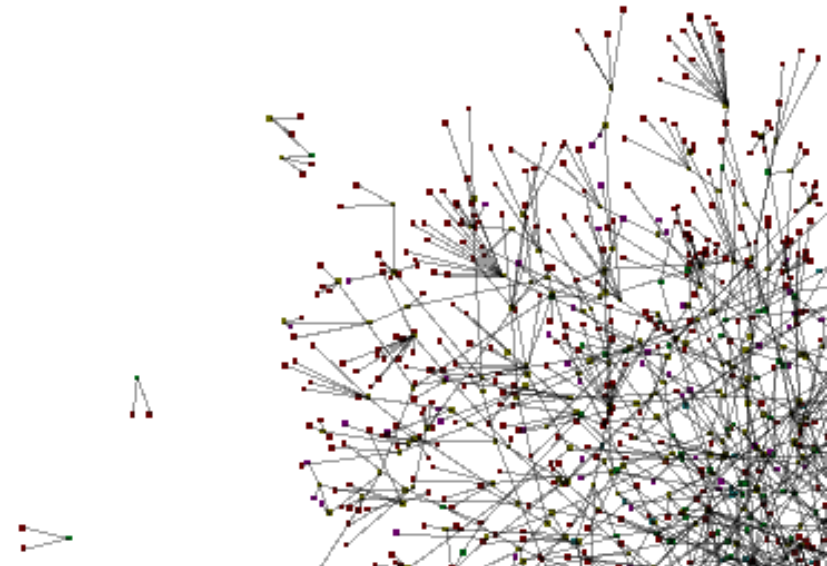
- ▶ Introduction and background
- ▶ Interactive web interface
- ▶ SOAP web services
- ▶ Additional features



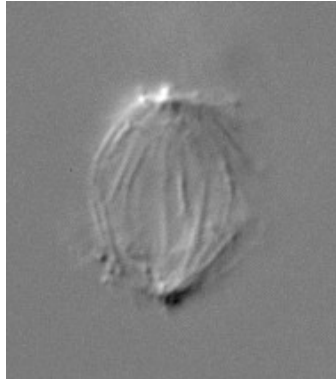
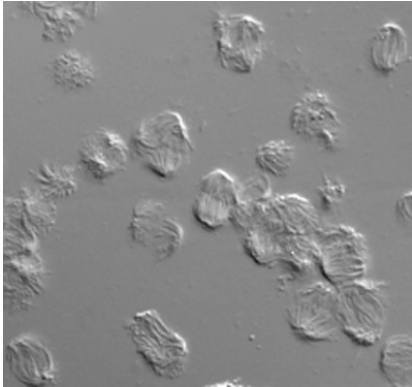
# Aims of FuncNet

*Given one set of proteins which are known to share a particular biological function...*

*... which of these other proteins also share that function?*



# Usage scenario



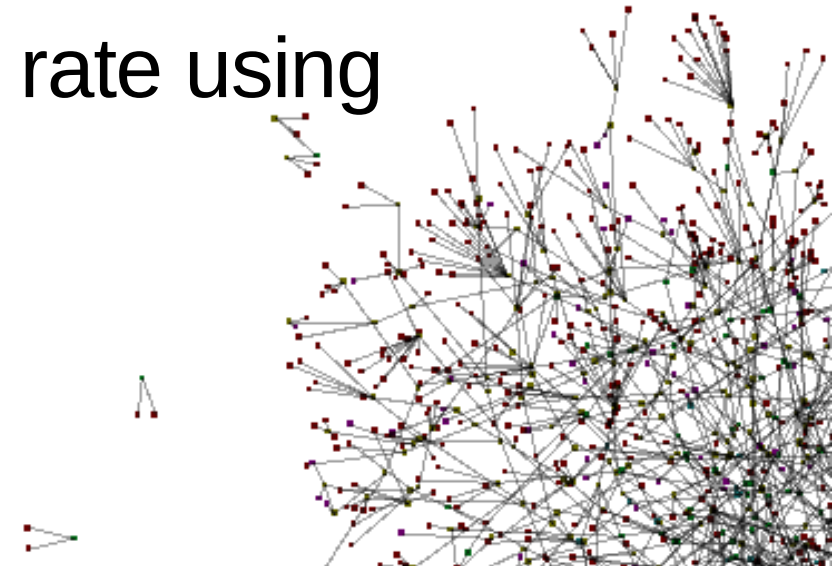
Erich Nigg – Max Planck Inst. – proteomics study.

150 known spindle proteins

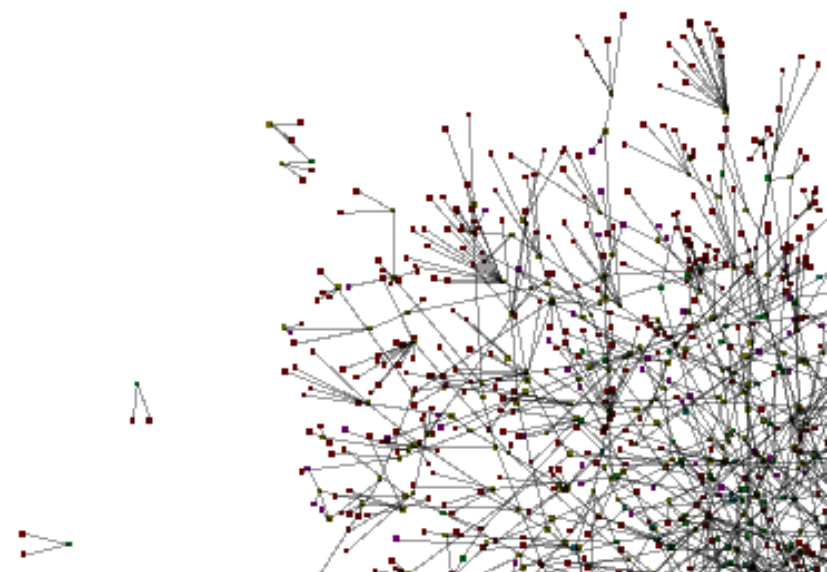
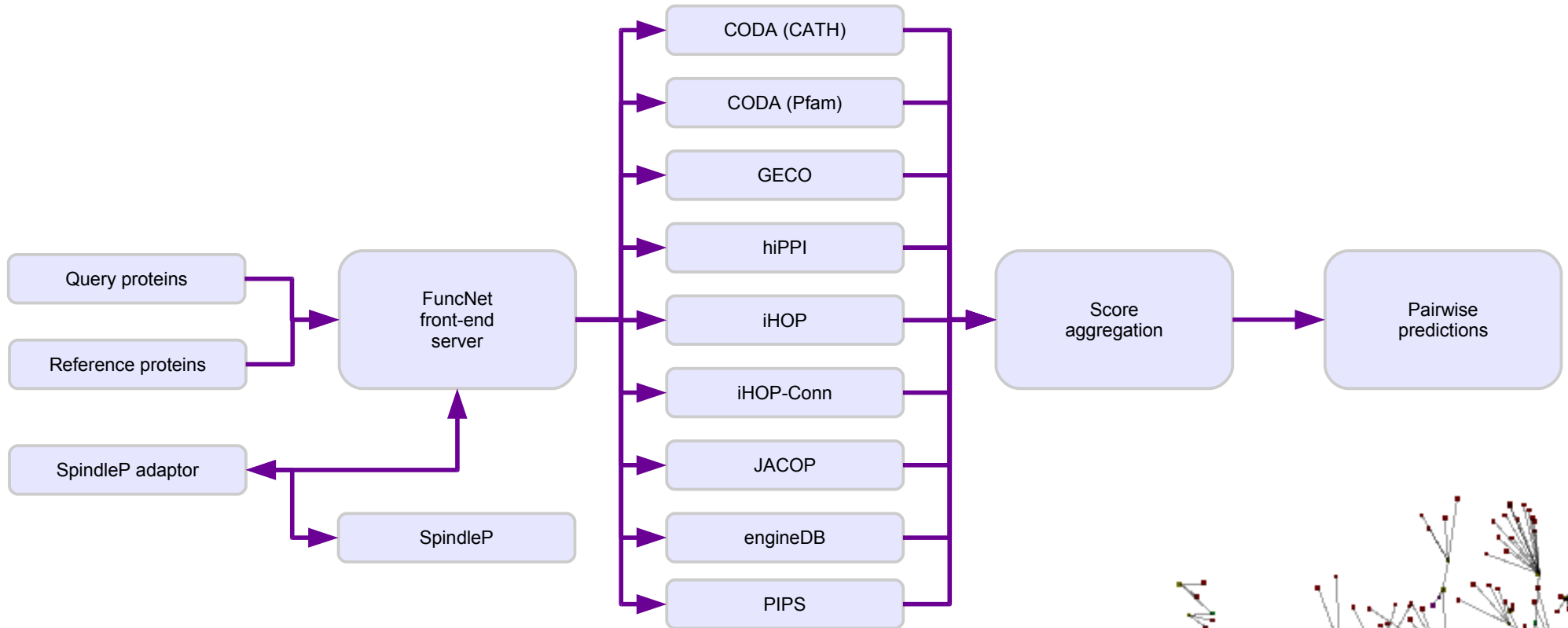
> 600 unknown/novel proteins

How can we help identify *true* spindle proteins?

Experimenters *doubled* success rate using FuncNet-style approach.



# Overview of workflow

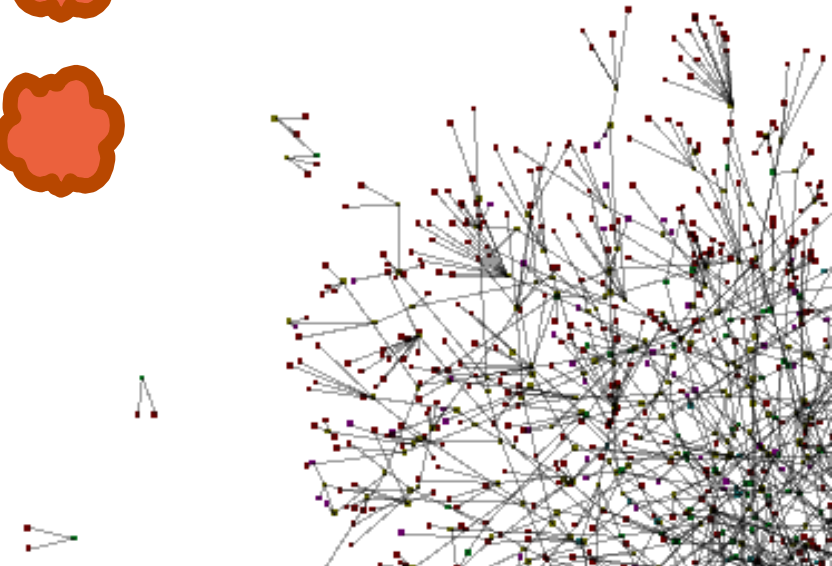


# Inputs

Reference proteins:  
known function



Query proteins:  
unknown function



# Outputs

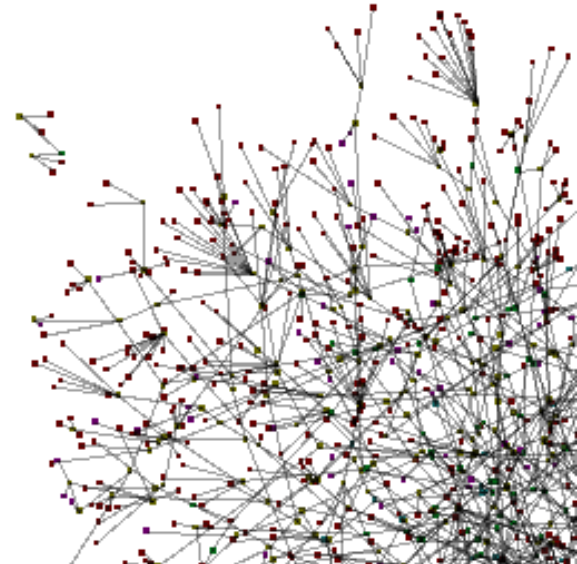
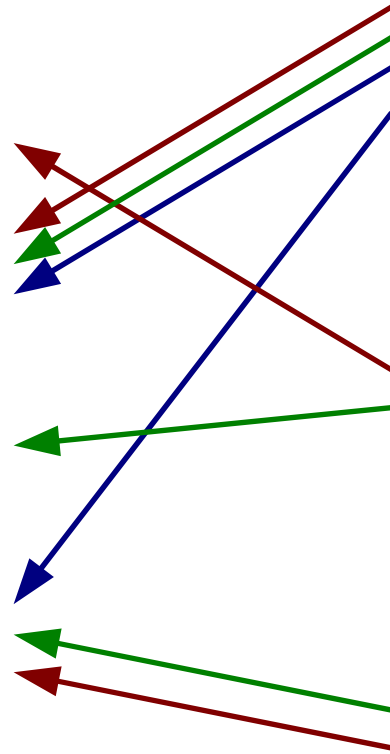
Reference proteins:  
known function



Query proteins:  
unknown function



predictor\_1  
predictor\_2  
predictor\_3

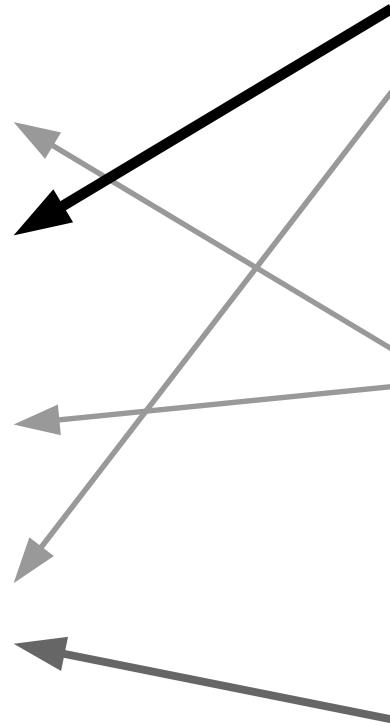


# Outputs

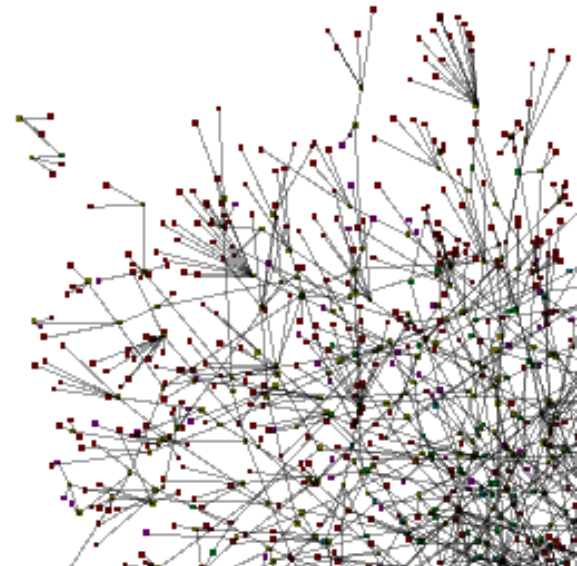
Reference proteins:  
known function



Query proteins:  
unknown function

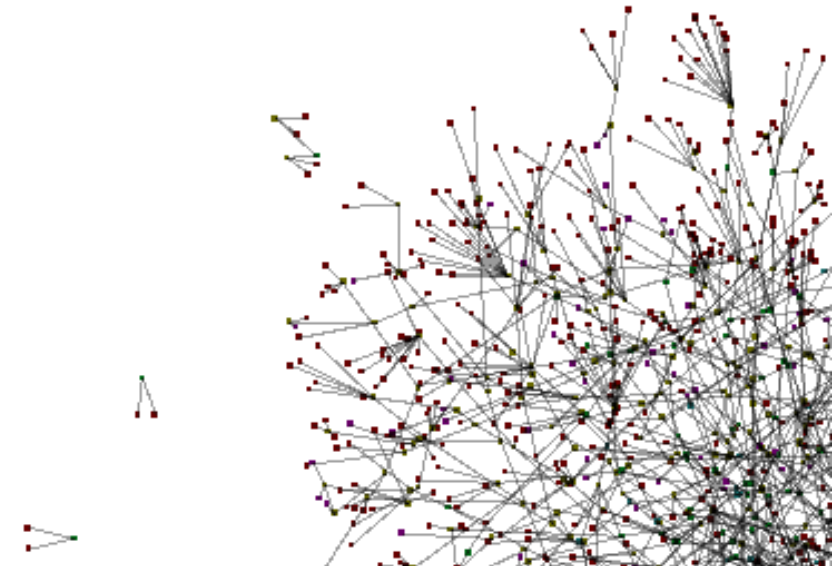


$$Fisher\ score = -2 \sum_{i=1}^k \log_e(p_i)$$



# Outline of talk

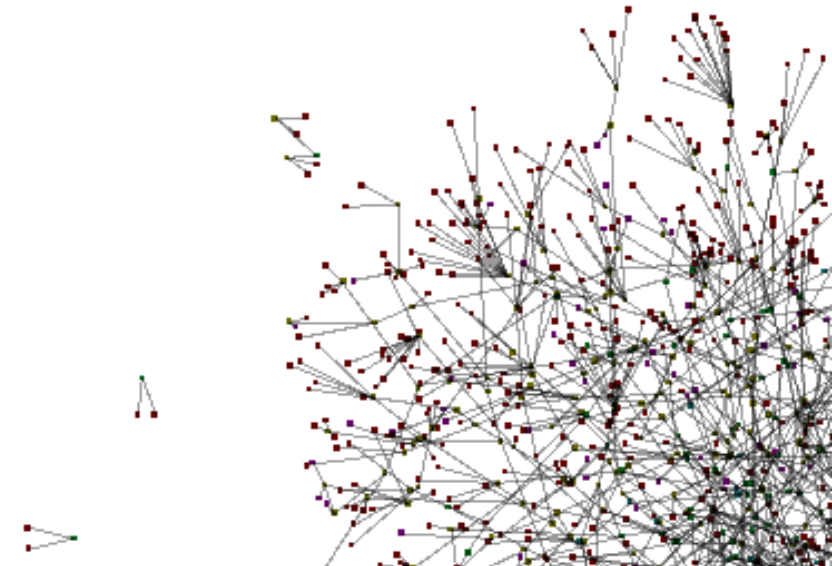
- ▶ Introduction and background
- ▶ **Interactive web interface**
- ▶ SOAP web services
- ▶ Additional features



# FuncNet web interface

- Upload protein lists from your PC
- Monitor progress of job
- Download results as CSV or XML

<http://funcnet.eu/gwt-client/>



FuncNet job caffdubya/-124554:1227e0eebda:-6849 - current status - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address [http://funcnet.eu/gwt-client/info.cathdb.funcnet.gwt\\_client.FuncClient/FuncClient.html#caffdubya/-124554:1227e0eebda:-6849?anonymous-web-user@funcnet.eu](http://funcnet.eu/gwt-client/info.cathdb.funcnet.gwt_client.FuncClient/FuncClient.html#caffdubya/-124554:1227e0eebda:-6849?anonymous-web-user@funcnet.eu) Go Links »


Start New Session Go to FuncNet Website Send us a Mail

## FuncNet session status report.

Large queries may take some time (minutes or hours) to complete on all predictors. This page refreshes every 60secs.

You can bookmark this page to return to this session later, or take a note of the job ID and the email address you used.

User: anonymous-web-user@funcnet.eu Job ID: caffdubya/-124554:1227e0eebda:-6849



CodaCathService	WORKING
CodaPfamService	WORKING
GecoService	WORKING
HippiService	COMPLETE
JacopService	WORKING
PIPsPredictorService	COMPLETE
iHOPFuncNetConnService	WORKING
iHOPFuncNetService	WORKING



Results for FuncNet job caffdubya/-124554:1227e0eebda:-683d - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address [http://funcnet.eu/gwt-client/info.cathdb.funcnet.gwt\\_client.FuncClient/FuncClient.html#caffdubya/-124554:1227e0eebda:-683d](http://funcnet.eu/gwt-client/info.cathdb.funcnet.gwt_client.FuncClient/FuncClient.html#caffdubya/-124554:1227e0eebda:-683d) anonymous-web-user@funcnet.eu

Start New Session Go to FuncNet Website Send us a Mail

## FuncNet job results.

The following table shows the FuncNet scores for each pair of possibly interacting proteins, where a higher score indicates a better prediction. Also included are the prediction p-values for each predictor. The lower the p-value, the more likely that this is a genuine interaction, in the view of one particular predictor. Statistically significant predictions are highlighted in green, and predictors that failed due to server-side or network errors are show in red. Click a protein to see its UniProt page.

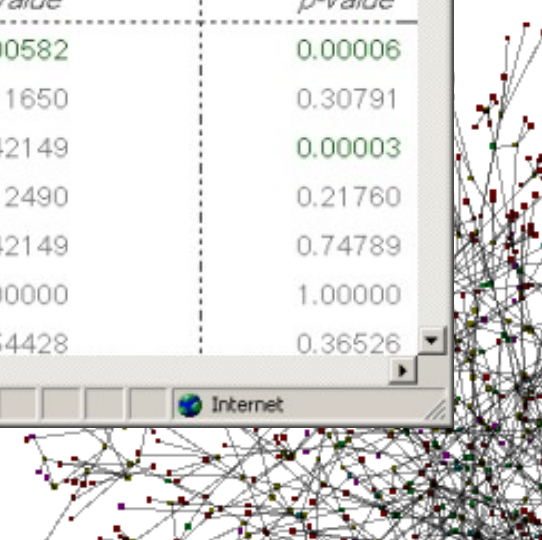
You can bookmark this page to return to for up to a week.

User: anonymous-web-user@funcnet.eu

**Results for job caffdubya/-124554:1227e0eebda:-683d**

[Click here to download the results in XML format.](#) [Click here to download the results in CSV format.](#)

Query protein	Reference protein	FuncNet score	CodaCathService p-value	CodaPfamService p-value	GecoService p-value	HippiService p-value	iHOPFuncNetConnService p-value	iHOPFuncNetSe p-value
<a href="#">Q99728</a>	<a href="#">P38398</a>	62.48168	1.00000	1.00000	1.00000	1.00000	0.00582	0.00006
<a href="#">Q99728</a>	<a href="#">P51587</a>	25.65159	1.00000	1.00000	1.00000	1.00000	0.11650	0.30791
<a href="#">Q9P0W2</a>	<a href="#">P51587</a>	22.92018	1.00000	1.00000	1.00000	1.00000	0.42149	0.00003
<a href="#">Q9Y6A5</a>	<a href="#">O75410</a>	20.97225	1.00000	1.00000	1.00000	1.00000	0.12490	0.21760
<a href="#">Q9UM11</a>	<a href="#">P38398</a>	19.85311	1.00000	1.00000	1.00000	1.00000	0.42149	0.74789
<a href="#">Q9Y6A5</a>	<a href="#">P43487</a>	17.19031	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
<a href="#">Q9Y6A5</a>	<a href="#">O95359</a>	16.62138	1.00000	1.00000	1.00000	1.00000	0.54428	0.36526

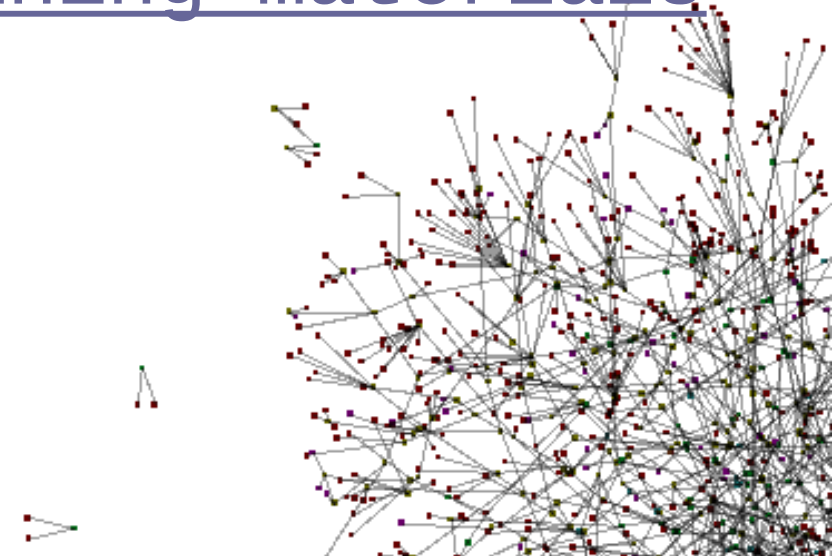


# Task 1 Instructions

This task will involve running a FuncNet session through the web interface, downloading the results as a CSV file and viewing them in Excel.

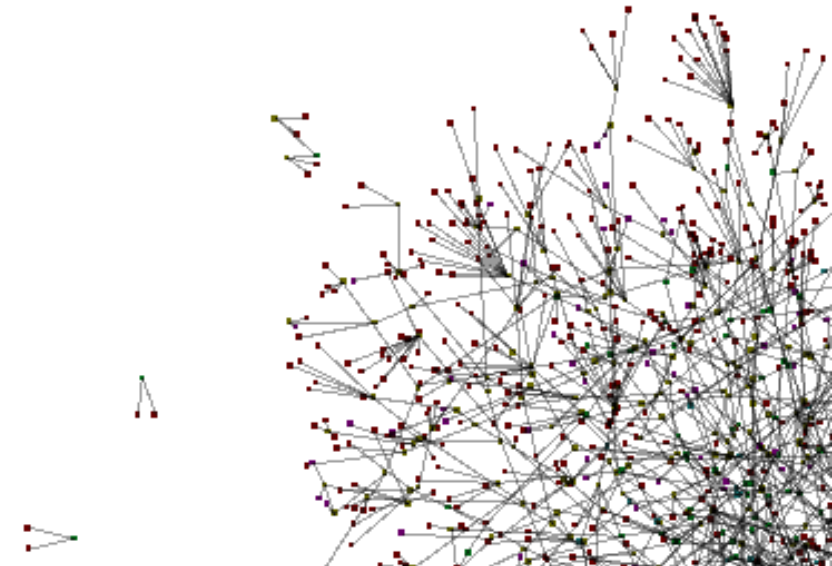
See section 1 of

<http://funcnet.eu/training-materials>



# Outline of talk

- ▶ Introduction and background
- ▶ Interactive web interface
- ▶ **SOAP web services**
- ▶ Additional features



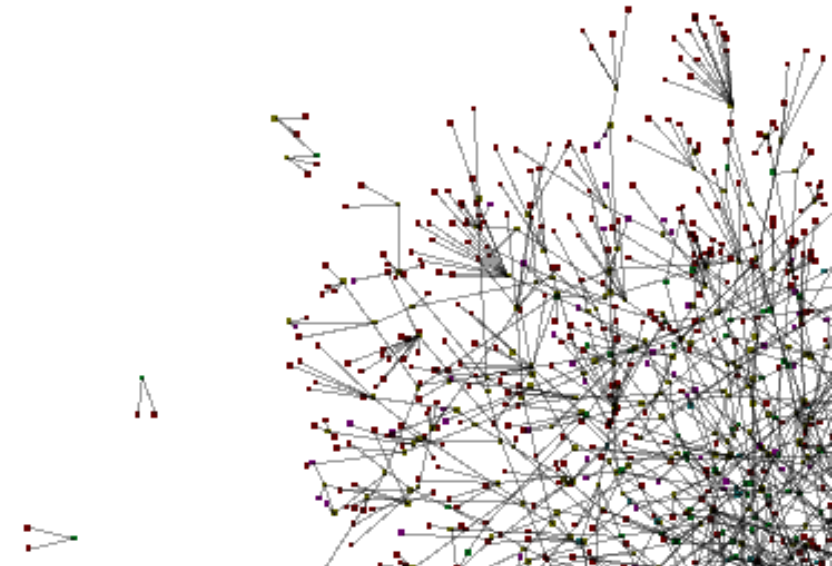
# Introduction to SOAP

- Protocol for sending XML between applications
- Usually uses standard HTTP (web) requests

variables/objects → serialization → XML

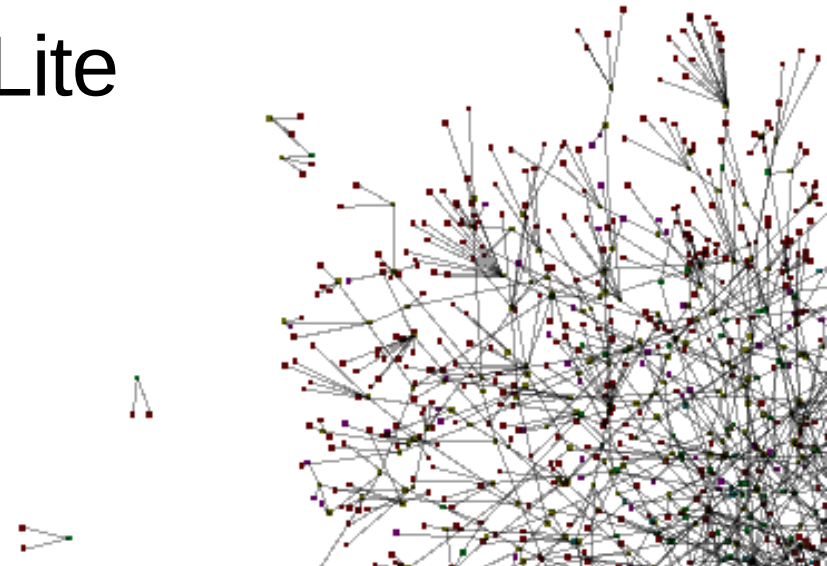
XML → deserialization → variables/objects

- WSDL file describes services

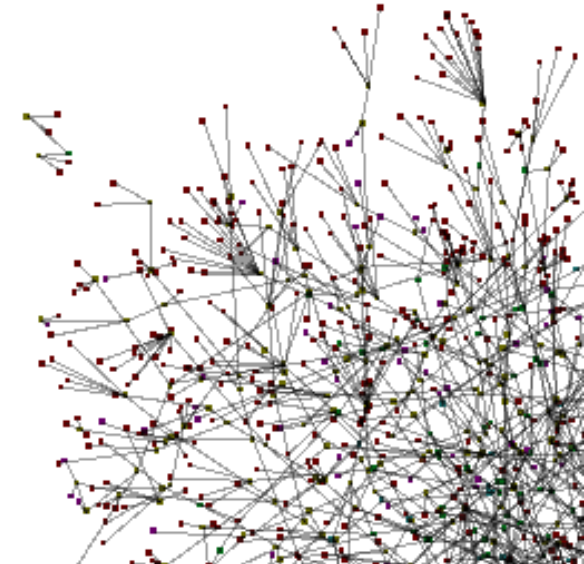
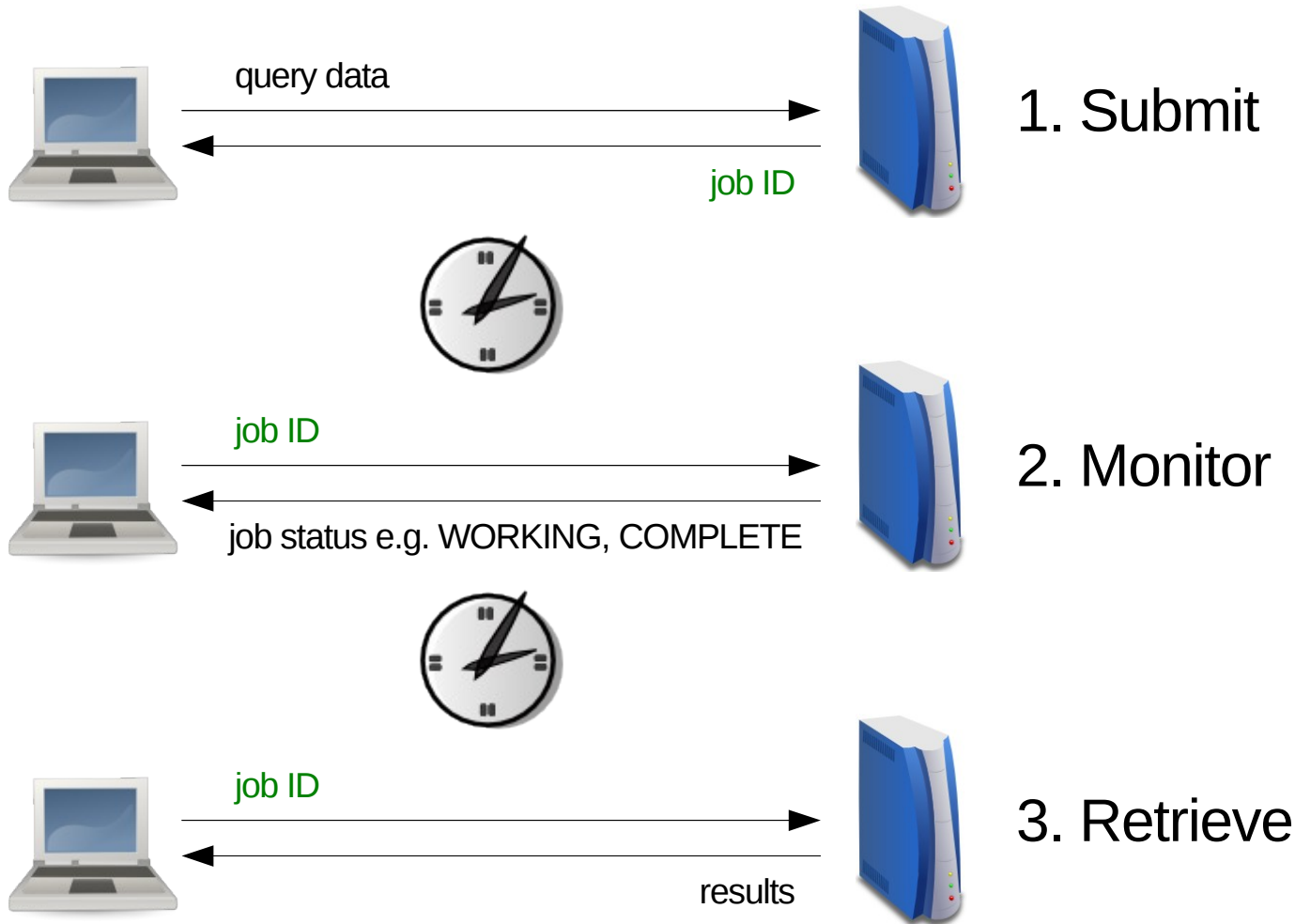


# SOAP client libraries

- Build [most] client-side code for you from WSDL
- Handle serialization and deserialization
- Make web requests over HTTP
  
- Java – CXF, Metro, Axis2, etc. etc.
- Perl – XML::Compile, SOAP::Lite
- PHP – SoapClient, NuSOAP
- Python – Zolera (ZSI)



# Aynchronous services



# FuncNet client tools

- We supply Java and Perl clients
- Use as-is or adapt to your needs

<http://funcnet.eu/source-code/>

- Workflows: Taverna, Bioclipse, Pipeline Pilot
- EnVision2

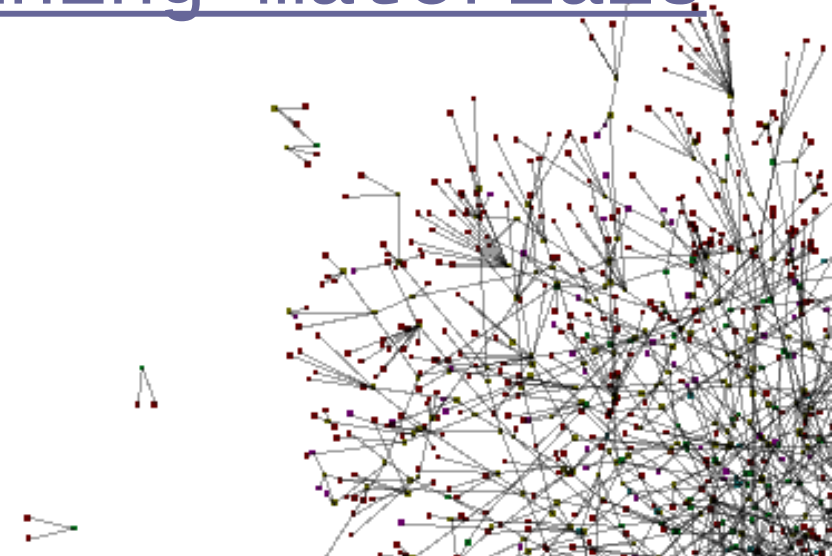


# Task 2 Instructions

This task will involve downloading and installing the Java command-line client, and sending a query to FuncNet.

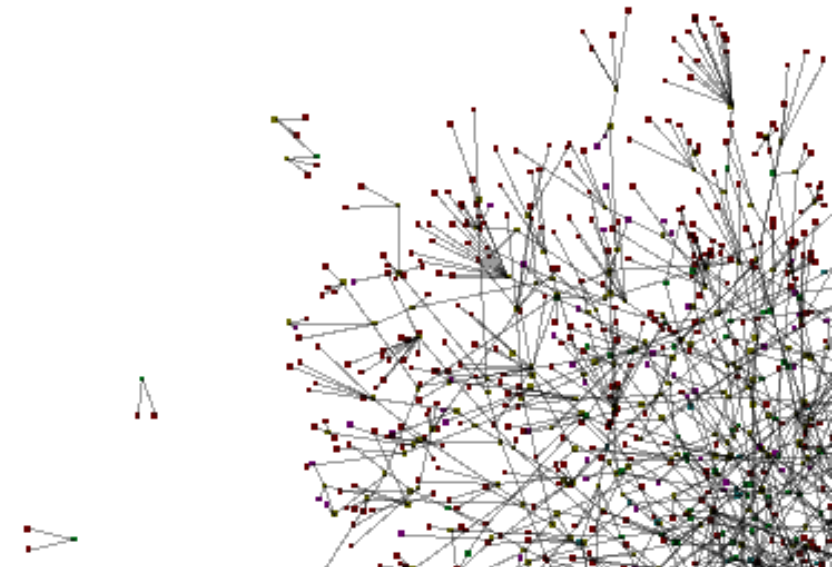
See section 2 of

<http://funcnet.eu/training-materials>



# Outline of talk

- ▶ Introduction and background
- ▶ Interactive web interface
- ▶ SOAP web services
- ▶ **Additional features**



# FuncNet

a distributed protein function analysis network

## Current status

 Search

FuncNet Service Status Monitor

Service Identifier	Status	Response Time	Checked On
<a href="#">FrontEndService</a>	OK	72.30s	April 30, 2009 10:30:05 AM GMT+01:00
<a href="#">GecoService</a>	OK	2.48s	April 30, 2009 10:31:18 AM GMT+01:00
<a href="#">HippiService</a>	OK	0.17s	April 30, 2009 10:31:20 AM GMT+01:00
<a href="#">CodaCathService</a>	OK	2.55s	April 30, 2009 10:31:21 AM GMT+01:00
			April 30, 2009

### Pages

- » [About FuncNet](#)
- » [Current status](#)
- » [Getting started](#)
- » [News](#)
- » [Partners and predictors](#)
- » [Publications](#)
- » [Source code](#)
  - » [Perl examples](#)
- » [WSDLs](#)
  - » [Example queries](#)

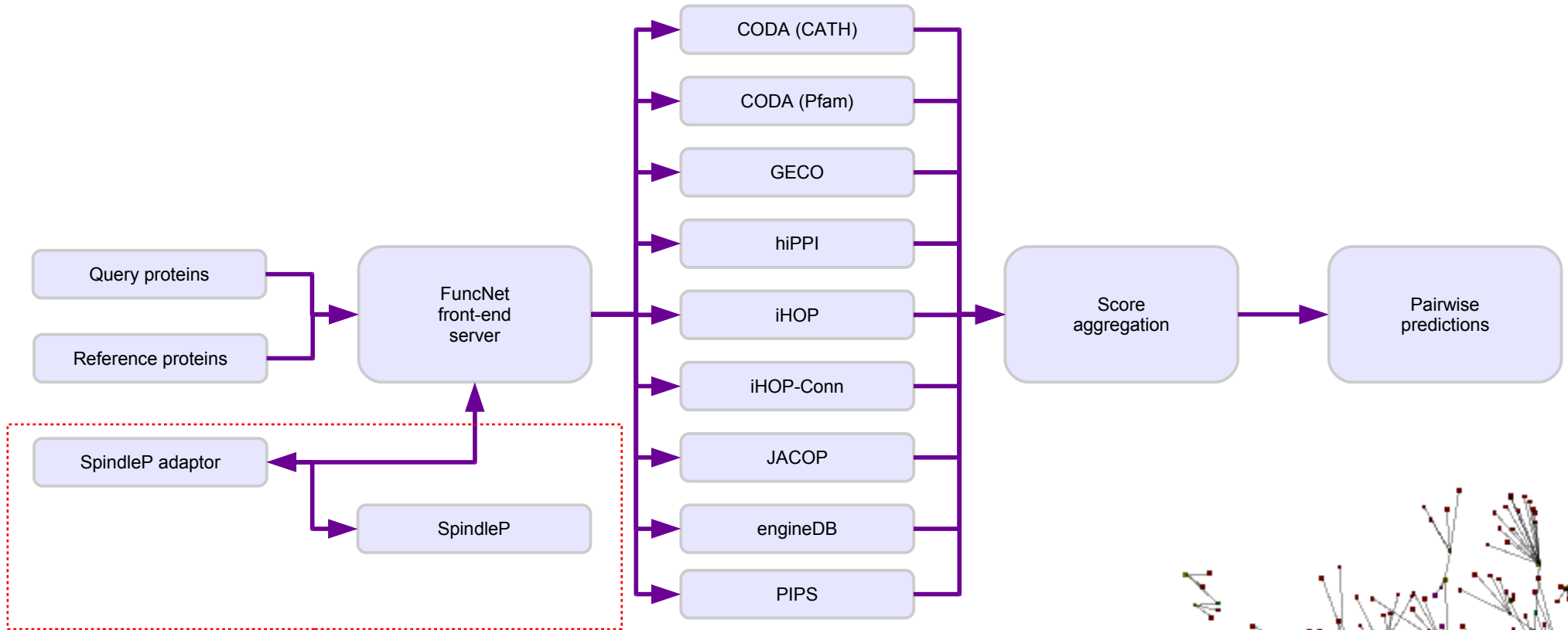
### News by Category

- » [Announcements \(5\)](#)
- » [Downtime \(2\)](#)

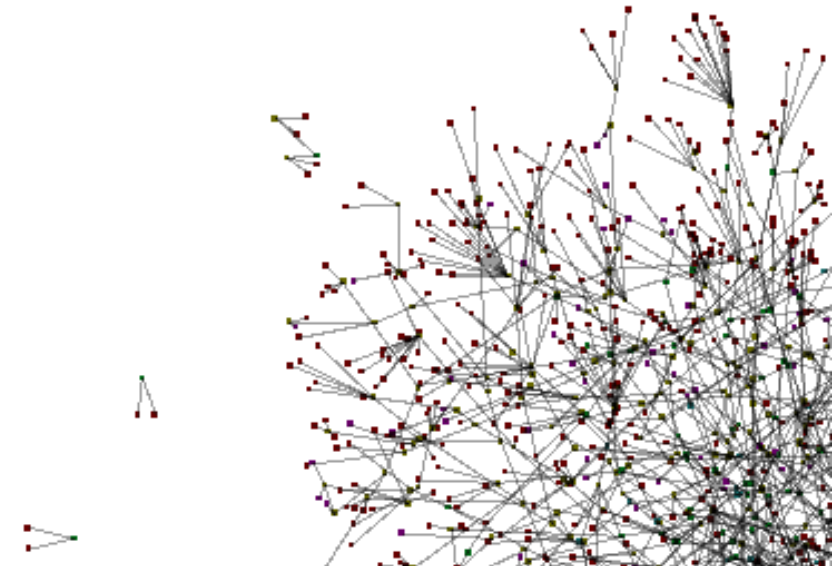
### News by Date

- » [April 2009 \(1\)](#)

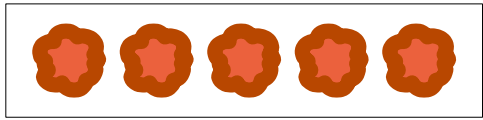
# Spindle protein prediction



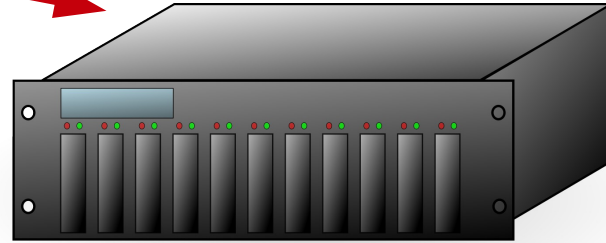
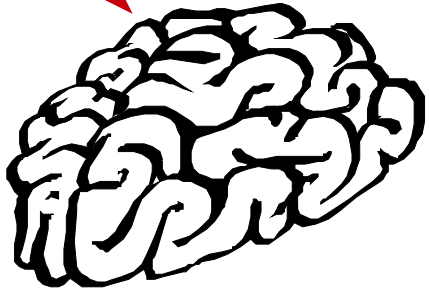
**Integrating with SpindleP**



Query set



Reference set = SpindleP training set



SpindleP neural network

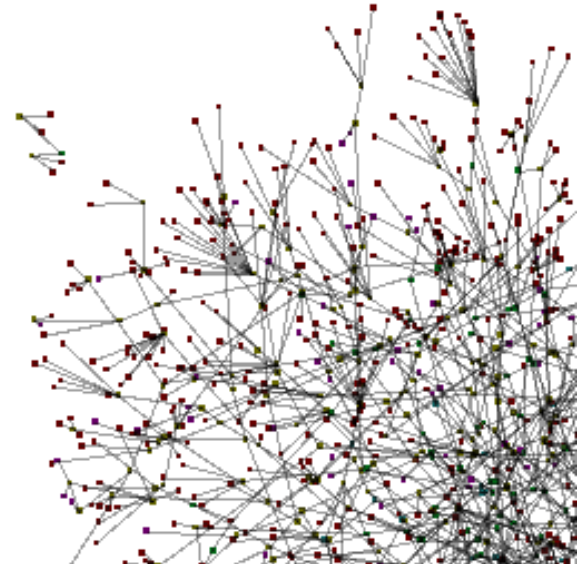
FuncNet front-end service

```

clegg@bsmlx47:~/export/people/clegg/data/EmbraceWorkspace/FuncNet-Spir
File Edit View Terminal Tabs Help
<pv>0.5297700703445435</pv>
</pr>
</s>
<s>
<p>Q8NFN7</p>
<rs>6.052770709804056</rs>
<pr i="1">
<pv>0.151909</pv>
</pr>
<pr i="3">
<pv>0.721078</pv>
</pr>
<pr i="4">
<pv>0.999999</pv>
</pr>
<pr i="6">
<pv>0.545864909294281</pv>
</pr>
<pr i="7">
<pv>0.810653860897854</pv>
</pr>
</s>
</results>

```

Score integration



# The FuncNet consortium

## *University College London*

- Christine Orengo
- Corin Yeats
- Jonathan Lees
- Ian Sillitoe
- Adam Reid

## *University of Malaga*

- Juan Ranea
- Ian Morilla

## *SIB, Lausanne*

- Marco Pagni
- Heinz Stockinger

## *EBI, Hinxton*

- Henning Hermjakob
- Florian Reisinger

## *CNR-ITB, Bari*

- Andreas Gisel

## *CNIO, Madrid*

- Alfonso Valencia
- José María Fernández González
- David González Pisano
- Ana Rojas

## *CBS, Lyngby*

- Søren Brunak
- Kristoffer Rapacki

## *University of Dundee*

- Geoff Barton
- Mark McDowall
- Michelle S. Scott
- Tom P. Walsh

... and myself, Andrew Clegg

