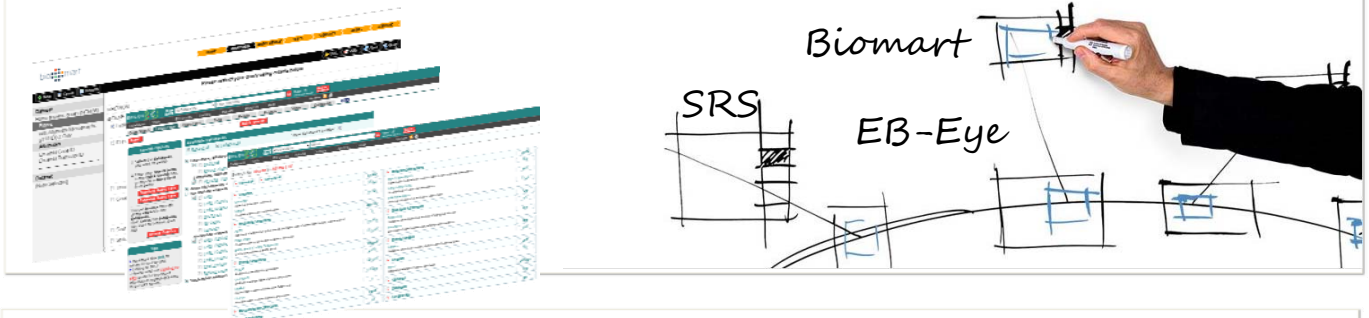


Unified access to bioinformatic resources: SRS, Biomart and EB-Eye

Programmatic access to online data and tools using Web services has an important role in bioinformatics. In this poster we present a web service built over WS-I and REST technologies that provides an integration layer for three commonly used search and retrieval systems for biological information: EB-Eye, Biomart and SRS, and present and a common interface to them.



Simple and unified interface

info()

Get information about the servers, databases, available fields, etc

```
embl
swissprot
uniprot
uniref
uniparc
pdb
pride
intact
Medline
...
```

search()

Search all databases for your favorite gene or protein, your biological process or pathway of interest



related()

Explore related entries in other databases navigating existing and inferred cross references



query()

Retrieve required info (full entry, sequence, specified fields) for a given query (list of ids, filters, etc)

Accession	Accession	Accession	Accession	Accession	Accession
AL048422	AL048422	AL048422	AL048422	AL048422	AL048422
AL048422	AL048422	AL048422	AL048422	AL048422	AL048422
AL048422	AL048422	AL048422	AL048422	AL048422	AL048422
AL048422	AL048422	AL048422	AL048422	AL048422	AL048422
AL048422	AL048422	AL048422	AL048422	AL048422	AL048422
AL048422	AL048422	AL048422	AL048422	AL048422	AL048422
AL048422	AL048422	AL048422	AL048422	AL048422	AL048422
AL048422	AL048422	AL048422	AL048422	AL048422	AL048422
AL048422	AL048422	AL048422	AL048422	AL048422	AL048422
AL048422	AL048422	AL048422	AL048422	AL048422	AL048422



The tools we use

build by maven

Maven is a software project management and comprehension tool. Based on the concept of a project object model (POM), Maven can manage a project's build, reporting and documentation from a central piece of information.

maven.apache.org

programmed using jax-ws/cfx

Apache CXF is an open source services framework. CXF helps you build and develop services using frontend programming APIs, like JAX-WS. These services can speak a variety of protocols such as SOAP, XML/HTTP, RESTful HTTP

cfx.apache.org

deployed in tomcat

Apache Tomcat is a Servlet container implementing the Java Servlet and the JavaServer Pages (JSP) specifications and provides a "pure Java" HTTP web server environment for Java code to run.

tomcat.apache.org

developed with eclipse

Eclipse is an open source community whose projects are focused on building an open development platform comprised of extensible frameworks, tools and runtimes for building, deploying and managing software across the lifecycle.

www.eclipse.org



SOAP and REST interfaces



Future work

We intend to implement **new interfaces** to other data repositories such as Uniprot, Entrez, WikiProfessional and integrate **semantics** in the systems so more complex queries can be performed.

We will also add a new class of methods **extend()** that allows the generation of new inferred links by calling external tools such as BLAST for sequences or Whatizit for literature data



References and further reading

Labarga A., Valentin F., Andersson M. and Lopez R. (2007) Web Services at the European Bioinformatics Institute *Nucleic Acids Research Web Services Issue*, 2007.

Stockinger H. et al. (2008) Experience Using Web Services for Biological Sequence Analysis *Briefings in Bioinformatics*, 2008.

Jimenez R., Quinn A., Labarga A., O'Neill K., Garcia K., Hermjakob H. Dasty2, an Ajax protein DAS client *Bioinformatics*, 2008



Alberto Labarga

Evangelos Pafilis



Visual exploration of biological data

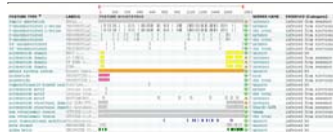
Current access to biological databases is commonly done through the web. Usually, the different search interfaces present results as lists of entities that correspond to a search query, with links to both the entry and entries in related databases. In our work we are exploring visual approaches to navigate subsets of the searched data. Exploiting the network-structure of data resources, we can search, retrieve and visualize multiple data related to a given query.

Visit www.scientifik.info for more details



We explored the web for interactive user interfaces that can be applied to biological data and selected key visualizations related to biology from visualcomplexity.com. Three open source visualization toolkits allow us to reproduce and extend these concepts into biological data realms they were not originally applied to.

Dasty, the open source DAS, can be extended to present new types of sequence related data, such as alternative splicing detected in microarray experiments, BLAST searches and multiple alignments.



Simple graph visualization using **tree or force-directed layouts** have been frequently used to present related data, author networks, protein-protein interactions, ontologies, etc



Integrating keywords and documents in a seamless interface similar to **6pli.com** provides a nice concept exploration of related literature.



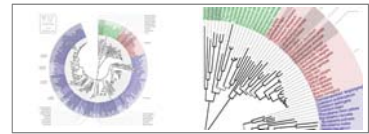
Treemaps allow access to data details without leaving the overview of the data. Size and color provide users with a rapid mechanism to evaluate data, and zooming allows access to details about genes of interest.



A **radial, space-filling layout** of different relationships (is-a, talks-about, etc) can be generated with interactive techniques of zoom, filter, and details-on-demand for the task of document visualization.



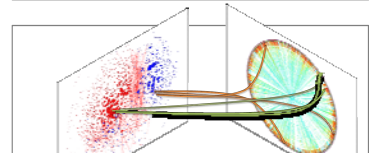
The radial tree used in the **Interactive Tree of Life** can also be extended to support database searches, biological enrichment analysis, etc.



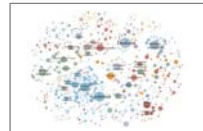
A dynamic search interface inspired in **blenderbox.com** allows us to present related results in different databases with up to three layers of complexity maintaining all relationships.



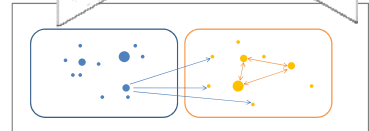
We work on the concept of **VisLink**, a method by which visualizations and the relationships between them can be interactively explored.



The **diseasome** paper presented an original way of relating diseases and genes that can be extended to any kind of gene set enrichment visualization in gene expression analysis



Network Visualization by Semantic Substrates enables users to specify regions to place nodes, and then control over link visibility.



prefuse

Prefuse is a set of software tools for creating rich interactive data visualizations. Prefuse supports a rich set of features for data modeling, visualization, and interaction. It provides **optimized data structures for tables, graphs, and trees**, a host of **layout** and visual encoding techniques, and support for animation, dynamic queries, integrated search, and database connectivity.

prefuse.org

piccolo

Piccolo2D is a toolkit that supports the development of 2D structured graphics programs, in particular, **zoomable user interfaces**, a new kind of interface that presents a huge canvas of information on a traditional computer display by letting the user smoothly zoom in, to get more detailed information, and zoom out for an overview. It maintains a hierarchical structure of objects and cameras, allowing the application developer to orient, group and manipulate objects in meaningful ways.

piccolo2d.org

processing

Processing is an open source **programming language** and environment to program images, **animation and interactions**.

processing.org



Future work

We intend to make our results publicly available, including source code for the visualizations, and most relevant results published in visualcomplexity.com.

While we are currently focused on functional genomics data, but we are always looking for new challenges and ideas for visualizations, so if you have a new concept you are happy to share of looking for new ways of presenting your data, please contact us at visual@scientifik.info



References and further reading

Visual Complexity: www.visualcomplexity.com

- The human disease network (the human diseasome)**
Goh K-I, Cusick ME, Valle D, Childs B, Vidal M, Barabási A-L
Proc Natl Acad Sci USA (2007)
- Interactive Tree Of Life (ITOL)**
Letunic I, Bork P.
Bioinformatics(2007)
- VisLink: Revealing relationships amongst visualizations.**
C. Collins et al.
IEEE Tran. on Visualization and Computer Graphics (2007)
- Network Visualization by Semantic Substrates**
B. Shneiderman et al.
IEEE Tran. on Visualization and Computer Graphics (2006)

Alberto Labarga

Alexander García-Castro

