BioWareDB: the biomedical software and database search engine

M. W. Matthiessen

Department of Medical Gastroenterology (54O3), Herlev University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark

Received on March 13, 2003; revised on May 9, 2003; accepted on May 29, 2003

ABSTRACT

Summary: A wealth of bioinformatics tools and databases has been created over the last decade and most are freely available to the general public. However, these valuable resources live a shadow existence compared to experimental results and methods that are widely published in journals and relatively easily found through publication databases such as PubMed. For the general scientist as well as bioinformaticists, these tools can deliver great value to the design and analysis of biological and medical experiments, but there is no inventory presenting an up-to-date and easily searchable index of all these resources. To remedy this, the BioWareDB search engine has been created. BioWareDB is an extensive and current catalog of software and databases of relevance to researchers in the fields of biology and medicine, and presently consists of 2800 validated entries.

Availability: BioWareDB is freely available over the Internet at http://www.biowaredb.org/

Contact: biocomp@madswichmann.dk

Finding articles on a specific biomedical topic can be easily accomplished in PubMed (Canese et al., 2002, http://www.ncbi.nlm.nih.gov/books/), but it is far from trivial to find a software tool or a relevant database for a specific bioinformatics task through either PubMed or a general Internet search tool. Although, a few specialized journals (Bioinformatics, BMC bioinformatics and Nucleic Acids Research) and online repositories (Bioinformatics.Org, http://bioinformatics.org/ and Sourceforge.Net, http://sourceforge.net/) have dedicated themselves to promote the use and development of bioinformatics tools and databases, most are used primarily by bioinformaticians and developers. Across the Internet a large number of web sites contain lists of recommended links to software, web tools and databases, and these sites are ideal for providing inspiration and an opinion on the quality of tools and databases. However, a common problem with these lists is that they fail to keep the material updated, maintaining the validity of the Internet addresses, and are often difficult to navigate (such as the exhaustive and valuable Database Issue of Nucleic Acids Research, http://nar.oupjournals.org/).

Some journals publish Application Notes and repositories are available for scientific open-source software development (e.g. Sourceforge.Net), but although advances in biological software development over the last decades have led to a large amount of high-quality applications developed both by the commercial and the open-source software community it is far from easy to identify and obtain a software tool for a specific task, particularly for investigators with little experience in the field of bioinformatics.

A few databases of biocomputing tools are available but these either contain specialized content (i.e. The Structural Biology Software Database, http://www.ks.uiuc.edu/Development/biosoftdb/) or are outdated and contain many invalid links.

DATABASE DESCRIPTION

To aid both the bioinformatics community and to provide the general scientist with easy access to biomedical software and databases, BioWareDB has been created. The database is an attempt at gathering the largest possible number of entries relevant to biological and medical researchers in an accessible format. BioWareDB is updated every month and entries are checked to reflect the status of their Internet links as either working or invalid at the time of testing. Even if a link is found to be non-functional, verification in the following month will establish if this is only a passing problem. Additionally, over time this generates statistical data on the quality of links, which will be included in the database.

Data on biomedical tools and databases is obtained from online repositories of software and from journals covering bioinformatics resources (links to the source data can be found on the BioWareDB web site). The creation of BioWareDB does not involve interactive reviewing or analysis, instead records are derived from the collected source material and contain as a minimum a title, a description and the original source. If a valid web site can be extracted, it is automatically checked and the result of the verification is displayed as either OK or Failed. Where available, supported platforms and license information are included. The automated creation and validation of entries counters the problem of relying on
submissions to the database and subsequent time-consuming reviewing of submissions, yet it dispenses with the equally potent problem of the database filling with invalid entries.

BioWareDB offers two search methods: (1) A fulltext search that supports boolean terms (AND, OR and NOT) which mimics natural language and is therefore straightforward to use. (2) An advanced search that can be limited to specific fields (title, web site, data source, etc.). BioWareDB is based on a MySQL database, that is accessed through a PHP web interface. Collection and manipulation of source data is carried out with Perl scripts.

CONCLUSION

With more than 2800 records of biomedical software and databases, BioWareDB offers an extensive and easily searchable index of validated entries to current and available biomedical computing resources.

REFERENCES