EnviroInfo 2004 (Geneva)
Sh@ring – EnviroInfo 2004
 ISBN: 28 29 30 275-3

Biological Information System (BIS): Sh@ring Data and Experiences

Peter T. Strobl[1] and Albin Blaschka[2]

Abstract

The technologies underlying the BIS project (Strobl/Blaschka/Heiselmayer 2003) enable a scientific information system paying special attention to open standards and platform independence. We believe we should not strive for homogenizing the scientific database landscape using only one single data exchange path, but focus on optimizing interoperability of diverse systems.
The current BIS implementation supports taxonomical, syntaxonomical and distribution data, reflecting research projects hosted at our department. The system does not cover all user requirements, but only those which are really needed, since "return of investment" topics played a central role in the design. BIS permits access with various client platforms. Replacing proprietary software with open source software to implement our XML and OOP based approach we anticipated a general trend in the public sector. BIS' technologies are freely available for low-funded research.

1. BIS – Biological Information System

The BIS project (Strobl/Blaschka/Heiselmayer 2003) is the materialization of experiences gained with its predecessor database system. The problem analysis required to build an information system with paying special attention to using open standards and remaining platform independent. As an information system, BIS focuses not only on data storage like database management systems, but additionaly offers support for data analysis, data maintenance and data presentation.

2. Data exchange paths

Up to date we have recognized three different strategies for data interchange:

[1]Universität Salzburg, Department of Organismic Biology, Hellbrunnerstr. 34, A-5020 Salzburg, Austria, email: peter.strobl@sbg.ac.at
[2]Same department, email: albin.blaschka@sbg.ac.at

1. Data-centric exchange standardization like seen by the ABCD initiative (URL in bibliography) focuses on the physical and logical data structures, trying to take most (if not all) cases of representation issues into account. The data structures are very sophisticated but difficult to comprehend for willing standard adopters.
2. Document-centric data representation mimics the simple idea of putting information into textual shape, like HTML web pages. Automated extraction and interpretation of data from such sources is a complicated issue.
3. Interface-centric architectures originate from the world of object-oriented programming. This approach enables hiding system internals from the user.

BIS mainly uses the last option to fulfill user requirements. But each category has its pros and cons (which cannot be discussed here). So we believe it is not desirable to strive for homogenizing the scientific database landscape using only one exchange path. In contrast, we try to optimize interoperability of diverse systems.

3. System complexity

BIS utilizes modern technologies from the XML domain and from the field of object oriented programming. Using such open industry standards led to replacing proprietary software with open source software almost three years ago.

Figure 1
Demonstration of a typical query scenario for a species list in field work. BIS is open for several data exchange paths, including support for common cell-phones.

This way we anticipated a general trend in the public sector, for example shown by EU's eEurope 2005 action plan (Commission of the European Communities 2002).

Currently BIS supports taxonomical, syntaxonomical and distribution data for research projects hosted at our department. The system should not cover all user requirements, but only those which are really needed (in order to keep system maintainance costs low). Instead it is possible to add or replace modules, when needed.

When designing the system, “return of investment” topics played a central role because scarcely used components of underlying technologies require a huge amount of development resources, without producing a proper “revenue”. This was a key argument that led to the decision not to bind to popular W3C standards like WSDL and SOAP, since we did not expect sufficient advantages in incorporating their complexity. The remote procedure calls we need for our interface centric system design could be achieved with about a tenth of system complexity (measured by the number of pages of the specifications involved). Still we are able to implement all requested work flow scenarios and address even “exotic” platforms like cell phones (figure 1).

4. Open access

With it's versatile system architecture BIS is an open system, permitting access with various client platforms. These include several operating systems, browser programs, or even PDA’s or cell phones, which will gradually become more and more important in the future (for instance in field work scenarios).

We freely offer BIS' technologies to low-funded research in order to enrich the scientific information community. We do this as a small contribution for freeing scientific work from financial obstacles. Thus, the BIS development team invites you to make use of our experience and contact us.

Bibliography

ABCD, Task Group on Access to Biological Collection Data: http://www.bgbm.org/TDWG/CODATA/Schema/

Commission of the European Communities (2002): eEurope 2005: An information society for all, http:// europa.eu.int / information_society / eeurope / 2005 / all_about / action_plan / index_en.htm

Strobl, P., Blaschka, A., Heiselmayer, P. (2003): BISxml 1.0: An Open Framework for Management and Interchange of Ecological Data, in: Gnauck, A., Heinrich, R. (eds.) (2003): The information society and enlargement of the European Union / 17th International Conference Informatics for Environmental Protection, Cottbus 2003 / Part 2: Applications, workshops, posters. Metropolis-Verlag, Marburg, pp. 609-612.