



<http://www.natsca.org>

The Biology Curator

Title: ENHSIN: Remote Access to European Natural History Databases

Author(s): Scoble, M.

Source: Scoble, M. (2001). ENHSIN: Remote Access to European Natural History Databases. *The Biology Curator*, Issue 20, 25 - 27.

URL: <http://www.natsca.org/article/854>

NatSCA supports open access publication as part of its mission is to promote and support natural science collections. NatSCA uses the Creative Commons Attribution License (CCAL) <http://creativecommons.org/licenses/by/2.5/> for all works we publish. Under CCAL authors retain ownership of the copyright for their article, but authors allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in NatSCA publications, so long as the original authors and source are cited.

ENHSIN

Remote Access to European Natural History Databases

Malcolm Scoble – Natural History Museum

ENHSIN - European Natural History Specimen Information System is a network supported by the European Commission under its Improving Human Potential Programme of Framework V. The aim is to create an institutional and computerised network across the Community to deal with remote access to specimen databases, examine data standards and sources, look at IP issues (which restrict the free flow of information), assess user needs, and organise a management model for the system.

ENHSIN stands for the European Natural History Specimen Information Network funded under the Improving Human Potential programme of the EU Framework Programme V. Its purpose is to enable the development of a shared interoperable European network of specimen databases. Further information can be found on the web-site dedicated to the Network: <http://www.nhm.ac.uk/science/rco/enhsin>

Background

The Network is an initiative arising from CETAF, the Consortium of European Taxonomic Facilities - a body created to promote research in systematic biology and access to collections, information and expertise. The Consortium includes several major European natural history museums, herbaria and botanic gardens. With the growing enthusiasm for databasing material in collections across Europe, it was clear that enormous value would be added if a system could be developed to network relatively uncoordinated efforts across a diverse continent. There are two fundamental aims in ENHSIN. One is to provide a means of allowing remote access to specimen databases across the Internet. ENHSIN, however, is also a network of people and institutions. Professional interaction between systematists

has existed from the very dawn of the process of documenting and classifying organisms; but with an increase in the number of collections-based institutions and curators, a more formal means of facilitating contact is highly desirable. If better organization can be combined with the power of the Internet, enormous potential exists for providing access to information about biological collections. A successful outcome of the project will depend both on the extent to which collections data become digitised and the effectiveness of the system by which remote access across the Internet is achieved.

Natural history collections are a valuable part of the European infrastructure. They can be perceived as physical databases (West & Nielsen, 1992) holding information on specimens. Their value increases over time because the temporal details they contain provides us with at least some guide to the changing distributions of species. Although older specimens rarely have associated co-ordinate detail, access to what is admittedly patchy, qualitative data can help. Collections, therefore, are a source of information for environmental research, including that on the impact of changing patterns of land use on the distribution of organisms and their conservation. They also hold data relevant to health issues, such as the distribution, current and past, of vectors of disease. And finally, specimens, their acquisition and their collectors are part of the rich European history of natural history and they form part of our European cultural heritage.

The partners Seven European organizations comprise the ENHSIN partnership. They are: the Natural History Museum, London, which is the co-ordinating institution; the Royal Botanic Gardens, Kew; the Zoological Museum, University of Copenhagen; Museo Nacional de Ciencias Naturales, Madrid; Muséum national d'Histoire naturelle, Paris; Botanischer Garten und Botanisches Museum, Berlin-Dahlem; and the Zoölogisch Museum, Universiteit van Amsterdam. It is essential to appreciate that the Network is not a closed club. Its purpose is to create a system to which institutions will wish to communicate.

The use of collections

Although care is needed before attempting to extrapolate across Europe, figures from the Natural History Museum, London, give some idea of the magnitude of usage of natural history collections. About 9000 researchers visited the collections in the financial year 1999/2000 for a total of 18,000 days. Museum collections also inform public displays, and there were 1.9 million gallery visitors to the Natural History Museum over the course of the year. The fact that the museum's web-site received over one million hits each month from the public and researchers, gives a dramatic indication of the growing importance of the Internet to collections-based institutions.

Access and adding value

Access to collections has traditionally been gained almost exclusively by visiting institutions or by receiving specimens on loan by post. While personal visits will remain essential, gaining access to data via the Internet has the potential to broaden the users base. It will also enable direct users of collections to undertake background work prior to making expensive visits to institutions, so allowing valuable time to be spent working on the collections to maximum effect. The effectiveness of the Internet as a means of delivering access will depend largely on the number of collections that become digitised and the quality of software enabling users to interoperate across a variety of sites.

The ENHSIN partners aim to create an operational system for what is intended to become a pan-European network. Over the duration of the project, which is supported by the Commission for three years, a variety of issues will be addressed. These include user needs, data standards and sources, intellectual property issues, and network management. Central to the project is achieving interoperability across specimen databases. The partners are aware of existing interoperability software, notably 'Species Analyst', which has been developed at the University of Kansas (<http://habanero.nhm.ukans.edu/TSA/>) and which provides simultaneous access to multiple biological collection databases from a web browser. An experimental interface providing common access to distributed specimen data has also been developed within the partnership to facilitate a pilot network for the ENHSIN project (for de-

tails see <http://www.bgbm.fu-berlin.de/BioDivInf/projects/ENHSIN/XMLClient.htm/>). Already, four specimen databases have been identified and made accessible (lichens, fruits and seeds, Homoptera insects, fishes). Further data sources suitable for linking in the pilot are being sought. Implementation of the pilot network will provide the opportunity to evaluate its effectiveness both in terms of technical development and by allowing priorities identified by users to be addressed.

Associated tasks

To gain a better understanding of European user needs, in both scientific and other sectors, a questionnaire was constructed and mailed to 2287 institutions or individuals. The results are in the process of being analysed. The questionnaire has also been placed on the ENHSIN web-site and, to encourage responses, is available in five European languages. This task emphasizes the fundamental importance that the partners and the European Commission place on addressing users in the development of the infrastructure.

Since specimen databasing is at a relatively early stage in its development, it is possible to gain a reasonable degree of agreement over standards and protocols for data access and quality and for the exchange of specimen information. To lay the foundation for data exchange, it is important to agree the core information to be searched and shared within the infrastructure collaboration. Since Europe is an area of great cultural diversity, this is a demanding task.

A potential barrier to Internet access to data exists in restrictions imposed by the protection of intellectual property. This field is complex and IP issues are being identified and addressed. Complete open access to geographical information will inevitably and rightly be restricted where threatened and protected species are involved.

A management model is being constructed both to track and guide the course of the project and also to sustain the network for the future. Certainly forming policy and frameworks to implement and bind a pan-European infrastructure will be critical if, as is intended, the network is to be sustained and expanded after completion of the ENHSIN initiative.

Complementary initiatives ENHSIN, clearly, should not be seen in isolation. A particularly notable initiative, which is intended to provide access to the wider content of European natural history collections is the BioCISE (the Biological Collection Information Service in Europe (<http://www.bgbm.fu-berlin.de/biocise/>). BioCISE deals with collections metadata above the level of the specimen. The system enables questions to be asked such as "in which European collections can I find specimens of a particular taxon or from a particular geographic region?". ENHSIN, by contrast, is creating a system to enable users to gain access to unit data - information pertaining to actual specimens such as geographical co-ordinates or observations. Related to these two initiatives is 'Species 2000' (<http://www.sp2000.org>), the purpose of which is to enable interoperability across global species databases. If links can be established to other databases, species names are a major means of access to collection metadata (at or above the level of the specimen). There exist many other initiatives from the global to the local, but those mentioned here have a particularly close association.

Hopes, reality and the future

Although there exists a wave of enthusiasm for creating specimen databases, we are far from the goal of access to full, standardized, digitised data of high quality for the vast number of natural history specimens housed in European collections. By developing systems such as BioCISE and ENHSIN, however, a means of gaining remote access to data in collections is, at least, rendered possible. Furthermore, it is hoped that these networks will provide a focus and serve as an encouragement for the digitisation of specimen metadata and encourage ways of improving data standards and systems of access.

Natural history collections are housed largely in what Lorcan Dempsey termed "memory institutions" (museums, libraries and archives) (<http://www.ariadne.ac.uk/issue22/dempsey>). Such institutions hold a wealth of information on the distribution of organisms through time across vast geographical areas: collections held in many European institutions span the globe in their representation. Although much of the data within these collections is

(inevitably) uneven and qualitative, we have nothing to equal it. Modern samples lack the time dimension, are often restricted to one or a few species, and, typically, are focused on narrow geographical areas. Bioinformatics has changed profoundly the science of genomics. Informatics may not have had quite the same impact within biodiversity, but it is making great strides and shows every sign of developing much further.

Acknowledgements

I thank my colleagues in the ENHSIN partnership and in the Research and Consulting Office of the Natural History Museum, London, for the work on which this article is substantially based.

Reference

West, J.G. & Nielsen, E.S. 1992. Management and accessibility of biological collections. *Australian Biologist* 5: 68-75.

The Virtual Store Natural History Collections at Stoke Museum

Keith Bloor Senior Museum Officer, The
Potteries Museum & Art Gallery

Introduction

The award of designation was made because of the strengths of all collections, the Natural History collections are of local and regional importance and is the only major collection of its kind in the County. The designation application identified a number of weaknesses in the section:

1. "most groups of invertebrates are poorly represented except mollusca."
2. "storage space is now greater than 80% utilised"
3. "little time is available to adequately research and document collections at item level, which in turn has decreased oppor-