

Walter G. Berendsohn & Anton Güntsch

Resource identification for a Biological Collection Information Service in Europe (BioCISE)

Abstract

Berendsohn, W. G. & Güntsch, A.: Resource identification for a Biological Collection Information Service in Europe (BioCISE). – *Bocconea* 13: 257-260. 2001. – ISSN 1120-4060.

“BioCISE Resource Identification” is a Concerted Action project funded by the European Union with members from 10 EU countries and Israel. The project has the aim of preparing the ground for the implementation of a general biological collection information service in Europe. For this purpose a survey is under way to identify available data sources as well as the needs and expectations of potential users of biological collection information. The results will serve as a base for a proposal for the implementation of a European Biological Collections Information Service.

Introduction

Biological collections - living and conserved materials - are useful not only for taxonomists. They constitute an immensely rich resource of both, physical material and vouchered information, e.g. on the occurrence, ecological role, behaviour, and capabilities of organisms. Collections thus have the potential to serve other branches of science as well as private enterprises by providing information or material.

In many fields Europe houses the largest collections world-wide, which have been maintained over hundreds of years with great financial and intellectual efforts (Scoble 1997). Somewhat following the lead of the United States, it has now widely been recognised also in Europe that electronic inventories are needed to make optimal use of collection information, and many such efforts are under way in individual institutions. Unfortunately there is no common electronic information system that leads a potential user to the specimen, samples, or data he or she is looking for. Such an interface - comparable to common catalogues of public and scientific libraries - would help to avoid time consuming investigations. Modern computer and networking technologies enable the implementation of a common access to already existing and planned local collection information systems (CIS). The BioCISE resource identification project tackles the problem of unavailability and incompatibility of collection information. A large scale survey examines biological collections information systems, available expertise in biodiversity informatics, and the potential users' needs and expectations of a collection information system.

Phases of the general project

BioCISE is a multidisciplinary Concerted Action project funded by the European Commission DG XII, Biotechnology Programme. It is the second phase of a possibly four step general project (Fig. 1).

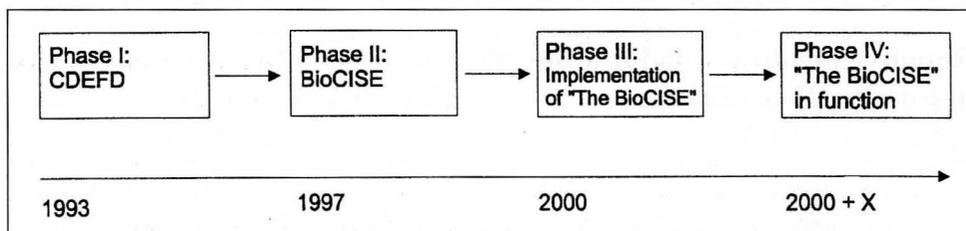


Fig 1: Phases of the general project

The CDEFD project ("A Common Datastructure for European Floristic Databases", 1993-1996) initially had the aim to provide a complete information model for the design of botanical collection databases. It soon became clear, however, that there is no fundamental difference between the information related to plant collections and those of other organism groups. Even non-biological objects (e.g. minerals) are often treated in a very similar way.

The CDEFD model (Berendsohn & al. 1996), which is based on the rules of modelling defined by Chen (1976) covers every aspect of collection information in the strict sense, i.e. excluding descriptive information (see Hagedorn, this volume) and taxonomic information (covered for plants by the IOPI model, Berendsohn 1997). A noteworthy aspect is the full incorporation of the derivation history of objects, e.g. analytical samples taken from a herbarium specimen which in turn has been collected from a plant accessioned in a botanical garden, which has been grown from seed collected in the field. This feature allows to cover living and conserved collections by the same base model. The comparison with the structure of implemented databases (see TDWG 1998, for a list of models; and Anagnostopoulos, this volume) has proven that the CDEFD project has created a complete reference model for biological collections in general.

Building on this foundation, BioCISE Resource Identification (the ongoing project) has set out to identify and analyse biological collection databases in the European Union and Israel. In parallel, the needs of potential users of a common collection information systems are determined. A catalogue of persons, companies, and institutions with know-how in biological information science will be compiled. Finally a project proposal for the implementation of "The BioCISE" will be drawn up. As in the case of CDEFD, the project is concentrating on collection objects, not on taxa, because the latter area is already widely investigated and it is the focus of several other projects. Of course, linkage to taxonomic name data will become significant when details of the access system (user interface) are planned.

The results of the survey will lead to a proposal for the implementation of "The BioCISE" as a demonstration prototype (phase III). Based on the experiences gained with the prototype, and after the important question of funding the further maintenance of the

system has been solved, a fully functional collection information system can be made accessible to the biological community (phase IV).

The BioCISE Resource Identification Project

The BioCISE project started in the end of 1997 by the installation of a four person project secretariat at the Botanical Garden and Botanical Museum Berlin-Dahlem. During an initial project meeting in Berlin, 20 participants from a wide variety of scientific fields came together to define the project's goals, to built up internal communication structures, and to form subgroups.

A subgroup consists of one or more project participants with a strong interest in a certain biological or technical subject area (Fig. 2). It provides the secretariat with information from its specific field and recommends action to be taken. The major tasks of most of the subgroups are to define user's requirements for a collection information service, to help the secretariat in identifying collection information database resources, and to foment liaison with existing or projected systems in specific areas. The latter point will be significant for the project proposal, because BioCISE does not intend to duplicate efforts in areas where other organisations have already created federations of collection databases (see for example the CABRI project, Rechaussat 1997). To further restrict the scope of the project the area of human biology was excluded. The survey of database and expertise resources is carried out by means of questionnaires. Separate institutional and database questionnaires were

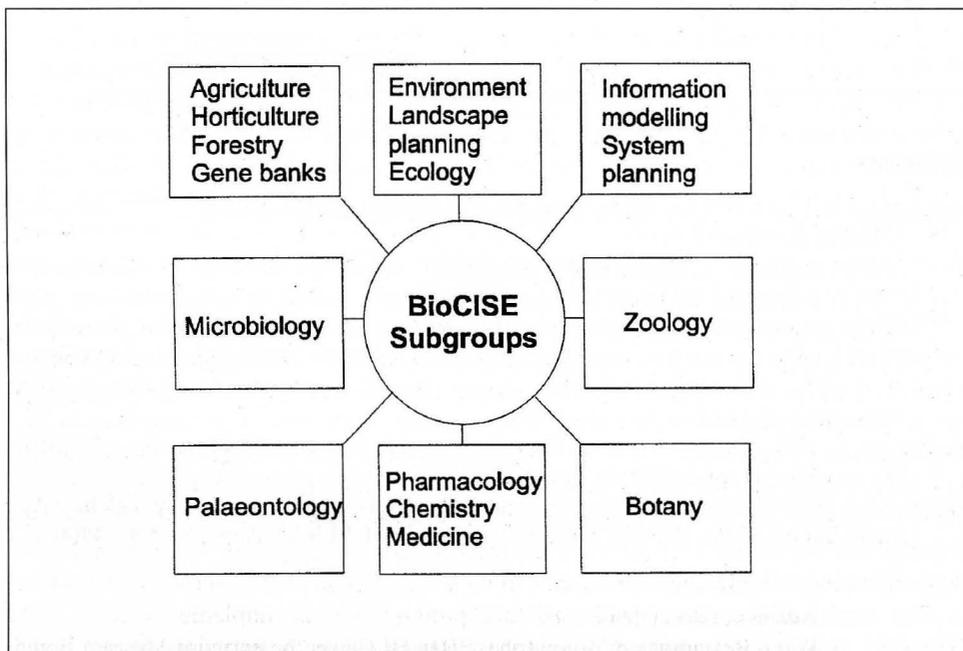


Fig 2: BioCISE subgroups

devised (Table 1), which are currently available in English, French, and German. They are sent to institutions by mail, but can also be downloaded from the BioCISE WWW site or requested from the secretariat in Berlin. The data from returned questionnaires are entered into a project documentation database and - provider's explicit consent given - they are published on the World Wide Web. As soon as sufficient data have been obtained, statistics will be derived and published. In this way, the current project will also be of assistance to those who are currently in the process of designing a collection database.

Table 1. Main topics covered by the BioCISE survey

Institution/Laboratory Questionnaire	Database Questionnaire
<ul style="list-style-type: none"> • Name, address, contact person • Co-operation with other labs in the field of collection databasing • Collection information systems experts' names and specialities • Presentation of the institution 	<ul style="list-style-type: none"> • Collection managed by the database system • Data acquisition, metadata standards used • Software name, development, and documentation • Software features, missing features, and on-line accessibility

References

- Berendsohn, W. G. 1997: A taxonomic information model for botanical databases: the IOPI model. — *Taxon* **46**: 283-309.
- , Anagnostopoulos, A., Hagedorn, G., Jakupovic, J., Nimis, P. L. & Valdés, B. (in press): The CDEFD Information Model for Biological Collections — in: Los, W. (ed.), Proceedings of the European Science Foundation Workshop "Dissemination Biodiversity Information", Amsterdam, 25.-27.3.1996. Preprint: <http://www.bgbm.fu-berlin.de/CDEFD/CollectionModel/cdefd.htm>.
- Chen, P. P. 1976: The Entity-Relationship Model - Toward a unified view of data. — *ACM Transactions on Database Systems* **1**: 9-36.
- Réchaussat, L. 1997: Common Access to Biotechnological Resources and Information — CABRI. <http://www.cabri.org/description.html>.
- Scoble, M. J. 1997: Natural history museums and the biodiversity crisis: the case for a global taxonomic facility. — Pp. 55-59 in: *Museum International* (UNESCO, Paris), No. 196. **49(4)**.

Address of the authors:

W. G. Berendsohn & A. Güntsch, Botanical Garden & Botanical Museum Berlin-Dahlem, Free University Berlin, Königin-Luise-Str. 6-8, D-14191 Berlin, Germany