

UCM-MACB 2.0: A COMPLUTENSE UNIVERSITY VIRTUAL HERBARIUM PROJECT

A. Avalos, A. Barrera, J.M. Gabriel y Galán, T. Gallardo, A. Gómez, J.M. Hernández, R. Lahoz, P. López, N. Marcos, L. Martin, A. G. Moreno, G. Oliván, E. Pérez-Urria, B. Pintos

Complutense University (SPAIN)

avagar@bio.ucm.es, genista@bio.ucm.es, jmgabriel@bio.ucm.es, tgallar@bio.ucm.es, magom02@bio.ucm.es, jmh@bio.ucm.es, lahozraf@bio.ucm.es, maplopez@bio.ucm.es, nmsaman@bio.ucm.es, lmartin@bio.ucm.es, agmoreno@bio.ucm.es, golivan@bio.ucm.es, elenapuc@bio.ucm.es, bpintos@bio.ucm.es

Abstract

UCM-MACB 2.0 is the name of a project led by a multidisciplinary team composed by university professors, students and collaborators as well as national and foreign visiting professors and researchers under the European Area of Higher Education. This is an e-learning project oriented to training and teaching being available to the general public, students and professionals from different areas on the basis of the plant biology, plant popular uses and their applications. At present, the initial steps of the project have received the financial support of Vicerrectorado de Calidad y Desarrollo, Complutense University of Madrid (PIMCD 3/2010). The target of the project is the developing and configuration of a computer system implementing a virtual herbarium based on plant species from the MACB herbarium (Faculty of Biology, Complutense University of Madrid, Spain). The system also will include a web page and an IT application (as a client-servant application) to transfer herbarium information to the potential interested community. MACB herbarium sheets will be associated with a set of databases containing the following information: taxonomic information, geographic distribution, specimen photographs, links to external online resources such as educational, scientific and other databases containing freely available molecular data (genomics, proteomics, metabolomics), for instance, GenBank or GenomeNet, PDB, InterPro, Enzyme, Prosite, Pir, TrEMBL, EMBL-Bank, Swiss-prot or NCBI taxonomic data as Taxonomy Browser, GBIF, Species 2000, or Tree of Life, etc. It is important to note that auxiliary information (geographical, ecophysiological, genomics, etc) is one of the main distinguishing features of UCM-MACB 2.0. Furthermore, plant images are enriched with information in multimedia format (MP3 voice) to make available UCM-MACB 2.0 to visually impaired users. In this sense, we plan to cooperate with the Office for the Integration of Persons with Disabilities at the Complutense University of Madrid. Thus, one of the main contributions of this project is that we extend the classical idea of herbarium enriching the usual botanical information with multimedia elements, opening the definition of 'herbarium sheet' to new educational and scientific possibilities.

Keywords: Virtual herbarium, Virtual learning, Experiences in knowledge management.

1 INTRODUCTION

MACB Herbarium (Fig. 1) is part of museums and collections of the Complutense University of Madrid, being administrated by the Department of Plant Biology I, Faculty of Biological Sciences. The herbarium was borne in 1968 as a part of The Botanical Garden (Madrid), from which became independent in 1975. Since its independence is called "MACB Herbarium" as referred in the Index Herbariorum including a collection of 100,988 sheets. Most specimens come from the Iberian Peninsula, Balear Islands and North Africa and about 75% are vascular plants. Among the 100,988 sheets are 82 types (the original material used for the description of a new taxon). At present, the storage requirements of the herbarium specimens restrict the use to professionals and researchers visitors. In this sense, the project opens the possibility to offer this collection to general people, sharing the 'botanical knowledge' and comprehensive information to many other visitors via web. However, this possibility will not replace the researchers and professionals visits "in situ" to the herbarium.



Figure 1. Herbarium MACB facilities and samples of scanned sheets. A. Store plants cabinets. B. *Chichona officinalis* L., MACB 74724. C. *Digitalis purpurea* L., MACB 50527. D. *Pteridium aquilinum* (L.) Kuhn, MACB 43325.

The aim of UCM-MACB 2.0 is the design and implementation of a virtual herbarium (Fig. 2), which can be accessed either directly via Internet (<http://herbariomacb.bio.ucm.es>) or through the Virtual Campus of the Universidad Complutense de Madrid (<https://www.ucm.es/campusvirtual/>). In this project the botanical knowledge necessary for conservation, study and research on biodiversity is translated to an electronic format. The process requires several stages from information organization, thus the management of the data from the herbarium MACB, to the design and implementation of a software system which allows access to databases.

Finally, the project complements images and other materials with a voice system in MP3 format in order to make UCM-MACB 2.0 accessible to disabled users

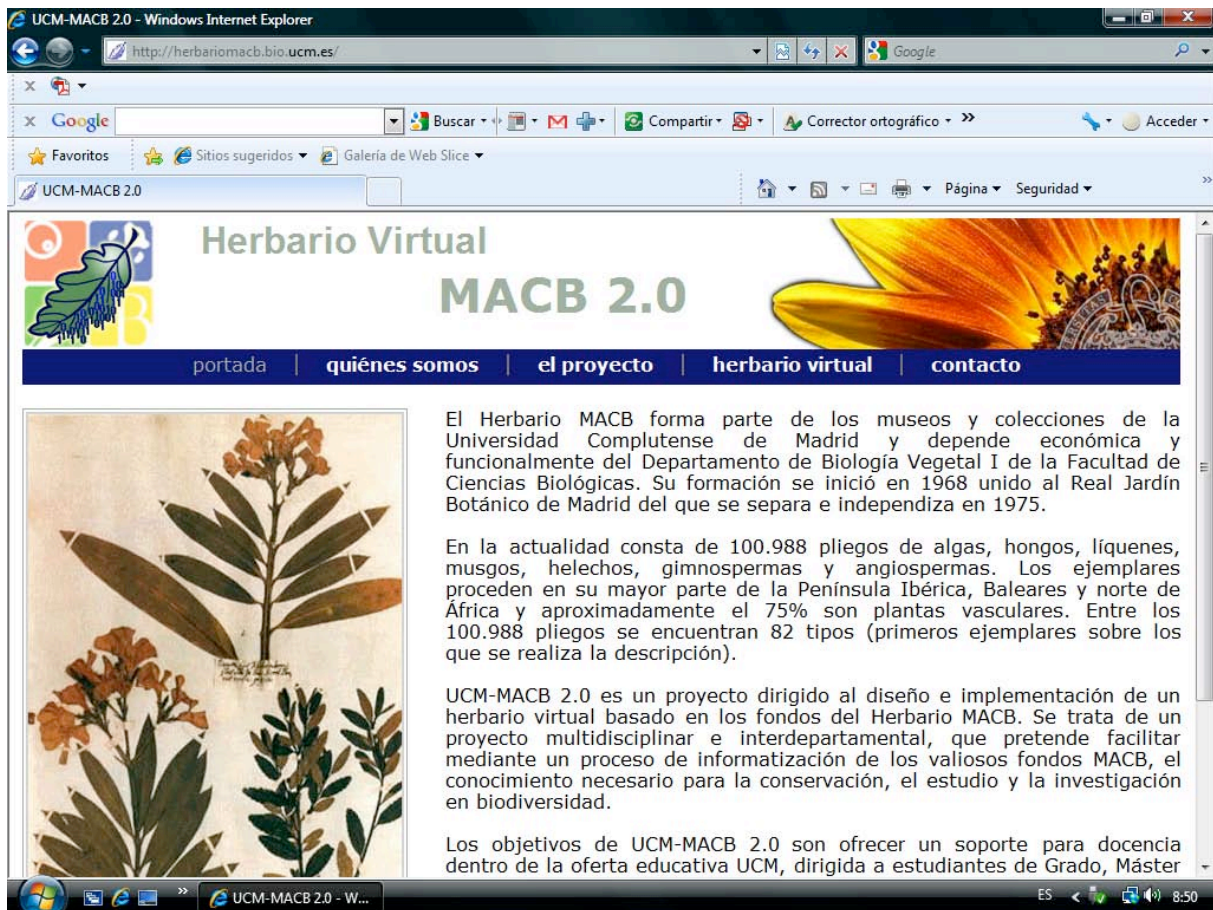


Figure 2. Home page of UCM-MACB 2.0 Virtual herbarium

2 METHODOLOGY

The main objective of the project is the design and implementation of a virtual herbarium, named UCM-MACB 2.0 (as a counterpart of MACB herbarium). This is achieved as follows:

1. Design and development of 'electronic sheets' to record the data contained in the herbarium MACB. We start with 82 types, thus building a database of the specimens stored and preserved. At present the herbarium MACB contains a database of about 100.988 sheets to be used for the purposes of this project.
2. Construction of high definition (.TIFF) herbarium sheets image bank.
3. Natural photographs enriched with MP3 voice will be
4. Geographical distribution data input and storage.
5. Ecophysiological data input and storage.
6. Building an information network of databases.

- The UCM-MACB 2.0 node (Fig. 3) will be connected with sites, web pages and other databases containing free available molecular and taxonomic data such as GenomeNet, GenBank, PDB, InterPro, Enzyme, Prosite, Pir, TrEMBL, EMBL-Bank, Swiss-prot, NCBI taxonomic data as Taxonomy Browser, GBIF, Species 2000, Tree of Life and others.

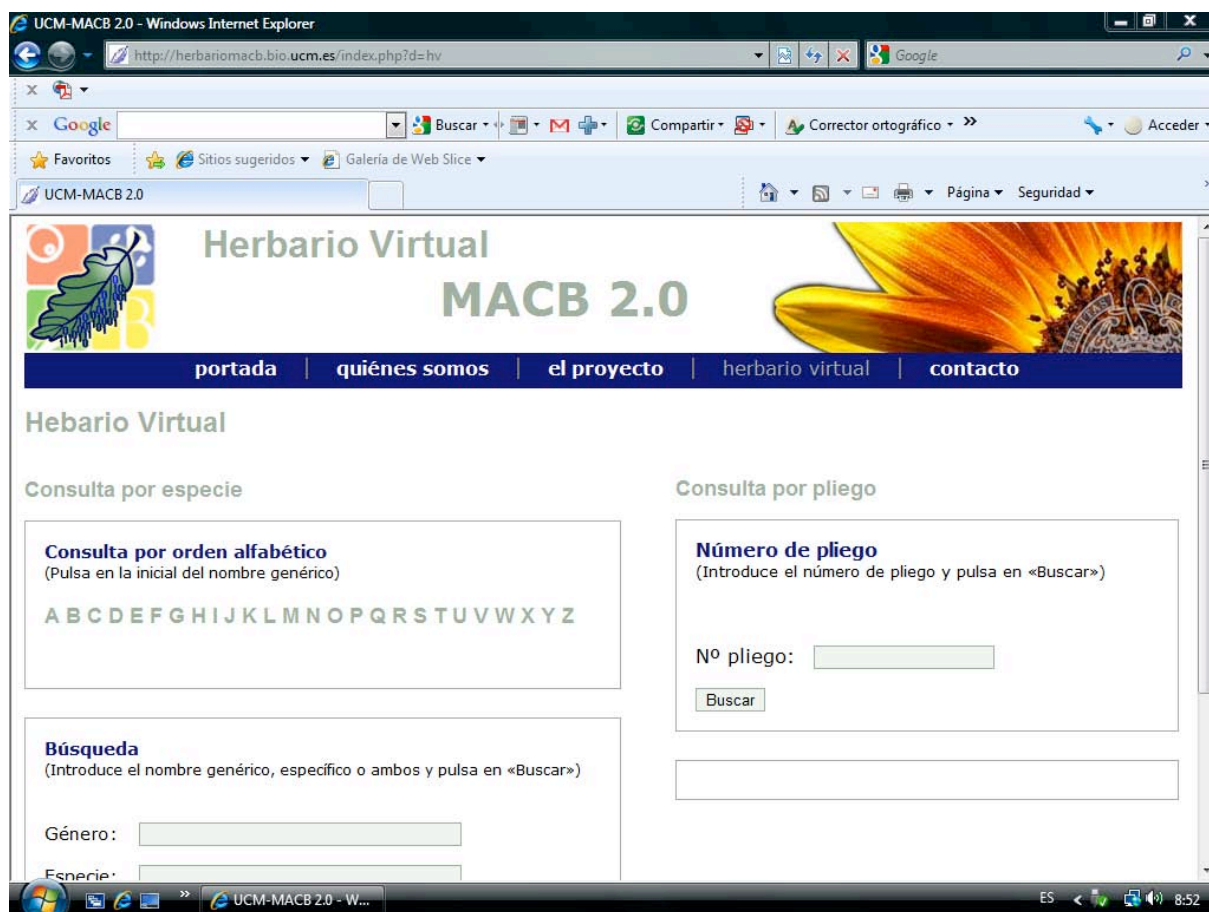


Figure 3. Search page of UCM-MACB 2.0 Virtual Herbarium.

3 RESULTS

3.1 Geographical information acquisition

The auxiliary geographical information is one of the distinguishing features of UCM-MACB 2.0 and is related to the geographic distribution data of specimens. The virtual herbarium web page will have several links to this geographical information. The aims are two: first, to get images by Google Earth of the geographical places where the specimens were collected; and, second, to facilitate to the users a file with the geographical information.

Google Earth supports managing three-dimensional geospatial data through Keyhole Markup Language (KML), generating a kml file. Keyhole Markup Language (KML) is an XML notation for expressing geographic annotation and visualization within Internet-based, two-dimensional maps and three-dimensional Earth browsers. KMZ is the default format for KML because it is a compressed version of the file and it stands for KML-zipped. One of the more powerful features of KMZ is that it allows any images you use - say custom icons, or images in your descriptions - to be zipped up within the KMZ file. That way you can share these details without having to reference the files through some link to the Internet. Network links are one of the most powerful features of Google Earth enabling a KML file to reference data dynamically on a server somewhere out on the Internet.

At UCM-MACB 2.0 database is also stored the information of geographical coordinates of the specimens and, therefore, it is possible to read these coordinates, to modify a KML file and to generate the adequate Google Earth image corresponding to the place where the specimen was

collected. UCM-MACB 2.0 will also supply the user the possibility of unloading the KML file to save it and use it later (Fig. 4).

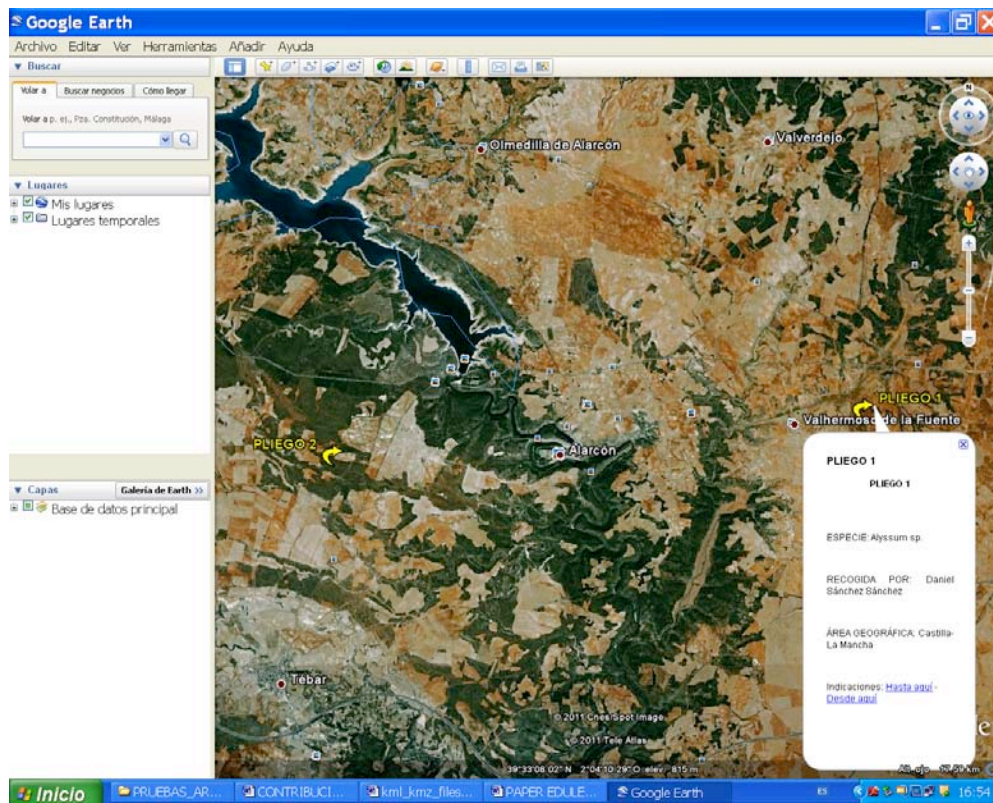


Figure 4. Geographical information related to the geographic distribution data of specimens

3.2 Building an information network

The project is first on the hardware and software configuration system that will house UCM-MAC 2.0 virtual herbarium. In this first phase requirements are derived from two basic needs: the generation of high resolution images and the storage of these images with other data. All this information must be available for applications and services offered by the virtual herbarium. The continuous generation and integration of data and images requires a system redundant backup to ensure the integrity of the system.

About data acquisition, we proceed to the design of standard card format that will be implemented at computer as a frame. Each sheet present in the MACB herbarium will be associated with the card containing taxonomic, geographical and ecophysiological data, and linked to different free access molecular and taxonomic database (Fig. 5).

The auxiliary information (geographical, ecophysiological genomics data) will be one of the distinguishing elements of the UCM-MAC 2.0 respect to other virtual herbarium.

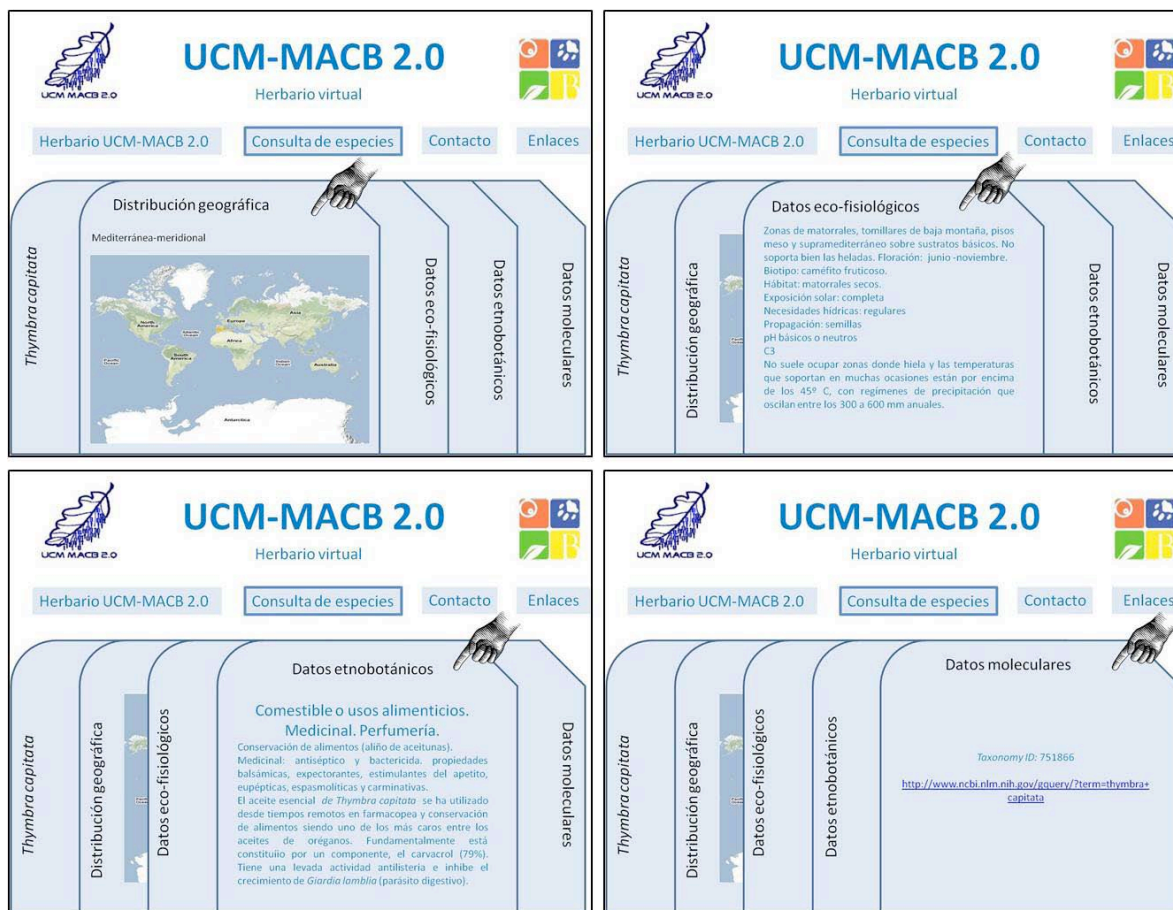


Figure 5. Standard card format

3.3 Public Service UCM-MACB 2.0

Considering the training and teaching, also the e-learning and life-long learning, this project will put the herbarium MACB and virtual herbarium UCM-MACB 2.0 available worldwide according with the open educational offer of the Complutense University in the context of the International Excellence Campus, through portals such as Universia, UCM Virtual Campus and through agreements established with universities and colleges around the world. In this sense it is important to consider the next project “The Ibero-American Knowledge Area”.

From the point of view of educational innovation and improvement of teaching quality, many students from different scientific fields would benefit the UCM-MACB 2.0 project as a source of information. The reason for this large projection on studies with very different goals is the involvement of basic uses and applications of plant biology, on different areas of study and research fields.

From the point of view of research, UCM-MACB 2.0 would support and assistance for research work in various fields. On the other hand, this project would enhance the public service performed by the MACB herbarium, through the loan, exchange or consultation of specimens in situ. Moreover, the design and implementation of the virtual herbarium (vHerbarium) is an objective in itself resulting novel software and applications in this area.

REFERENCES

- [1] <http://herbariomacb.bio.ucm.es/>
- [2] KML Tutorial: http://code.google.com/intl/es/apis/kml/documentation/kml_tut.html