Nimis P. L., Vignes Lebbe R. (eds.)
Tools for Identifying Biodiversity: Progress and Problems – pp. 121-125.
ISBN 978-88-8303-295-0. EUT, 2010.

Flora Bellissima, an expert software to discover botany and identify plants

Thierry Pernot, Daniel Mathieu

Abstract — To promote and facilitate access to botany for protecting biodiversity indirectly is the objective of Flora Bellissima. This sofware package is based on an integrated and complete database with numerous fonctionalities, as well as on an expert system of recognition of plants called "*Ophélie*". Aimed at beginners, amateurs and experts, Flora Bellissima attempts to show that it is possible to reconcile scientific rigor and popularization, in order to bring together everyone interested in botany.

Index Terms — botanical data management, expert system, plant recognition help, education through play.

1 Introduction

How to raise awareness in the public of what biodiversity is and of its protection? A solution is to help them to discover the nature around them, so they can know and protect it better. To acquire this knowledge, people need tools which are fun, easy, and based on scientific knowledge. The tools must be also adaptable to different audiences, from novices to experts, and they must take into account the progression of each individual's learning.

It is in this perspective that Flora Bellissima [1], an integrated management sofware dedicated to botany, has been developed. With a organisation similar to a " mini ERP" (Enterprise Resource Planning), this software is based on a central database composed of a scientific index of plants (BDNFF [2]). Four modules are organised around this database. They allow : (i) an educational discovery of the flora, (ii) assistance in plant determination thanks to the expert system « *Ophélie* », (iii) a management system for botanical data collected during field trips, and, (iv) a game module.

T. Pernot, the autor of the software, is with Yourproject Informatique 27 rue Saint-Georges 39360 Larrivoire France Yourproject@orange.fr - www.yourprojectinfo.fr. He is a biologist by training (Master of biology of organisms and populations at the University of Besançon) and computer scientist (project manager in corporate and SSI).
Daniel Mathieu is the president of Tela Botanica, 163 rue A. Broussonet, 34090 Montpellier, France, the NGO which publishes and distributes the sofware. dmathieu@tela-botanica.org www.tela-botanica.org.

This software has been developed for anybody interested in the flora, whatever their knowledge level. Novices who wish to make their first steps in botany can learn in an entertaining way thanks to the game module, the image glossary and the photographs. Amateurs who wish to improve their knowledge can store the result of their observations (texts and photographs) and build up their own database. Experts can consult easily and quickly the botanical nomenclature with all its synonyms and its taxonomic levels : species, sub-species, varieties and forms. The naturalist organisations can quickly save and exchange their field botany notes. Everybody can use the expert system *Ophélie* to identify a plant and give it a name before recording it.

2 Tool conception

2.1 General specifications

The detailed analysis of the different objectives led us to make the choices found in the main concepts of Flora Bellissima:

- the software is made like an ERP, i.e. composed of several applications sharing the same and unique database ;
- It is based on a scientific nomenclatural referential index ;
- It is open to allow the addition of text and photographs ;
- It is multi-level to be accessible to anyone : novices, amateurs and experts ;
- It contains data on the French flora for 1 400 species ;
- It proposes a help tool to determine plants, adapted to all audiences with the expert system *Ophélie* ;
- It has several photographs for each plant (9 800 photograhs in total with general appearance, inflorescence, flower, leaves etc.) to validate the determination ;
- It is available on PC and is distributed on DVD.

2.2 The flora discovery module

This module helps the discovery of the flora of France through numerous photographs and allows users to build up their own database based on their own observations. It also includes more than 30 functionalities which are not detailed here.

In summary, this module allows the consultation of a flora depending on different points of view : plant type (tree, shrub, fern, etc.), plant use (medicinal, cultivated, toxic), botanical family, genus, Latin or vernacular name. It also allows the consultation and capture of information on different themes : description, medicinal properties, protection, and geographical distribution [3]. Note that it is possible to add and link your own photograhs to each taxon. The three access levels, the picture glossary and the ergonomy give a strong educational value to this module.

The important idea emerging from this module is the gathering and centralisation of information in order to facilitate the access to it.

2.3 The « plant station » note management module

This module allows “plant stations” to be defined with their geographical and ecological characters, then to connect them to “botanical field trip notes”. This tool has been designed for capturing efficiently the complete Latin names of plants. Besides its functionalities of classification of stations and botanical field trip notes, this module proposes the following options :

- plant search among all field trip notes ;
- analysis of the evolution of the plant population of a station through time ;
- copying field trip notes ;
- import/ export of stations and field trip notes ;
- printing field trip notes and exporting them as a Word, Excel, PDF etc file.

2.4 The determination aid module

This module is made up of two main functions : a system of comparison between plants and an expert system of determination.

The comparison system allows the comparison of between 2 to 5 plants displaying all the characters which can differentiate them. This quickly gives the important characters to differentiate plants from each other. The system, called *Ophélie*, is an original tool which helps in the determination of plants using the Flora Bellissima database. One thousand four hundred species are now described, with about 30 characters for each plant, i.e. 45 000 characters described for the whole database.

To make sure that the system is efficient, it is important to take into account two main constraints: i) not all users are experts, and they can make mistakes, ii) plant morphology differs from one plant sample to another. Also, to work properly, the system must be able to accept mistakes and variations in the descriptions.

The system can work only on a binary way with the characters: true or false. It must work on all the characters described simultaneously, but not one after the other, as it is the case with a dichotomous key.

Each entered value allows a “weighting factor” to be allocated to all possible values of a character. To do this, we need to know “the distances” separating the different values of the same character. This is done by the similarity matrices which represent over 11 000 datapoints in Ophélie. In order to be able to enter such a quantity of datapoints, specific computing tools have been developed. These tools enable data capturing to be speeded up while checking it at the same time. Ophélie can manage different types of characters such as matrix, interval, quantitative values, etc.

The “weighting factors” are allocated to each of the chosen values thanks to algorithms which depend on the character type: for instance, if we input the value “pink” to the colour of a flower, this value will have the highest weighting factor, the value “red” would have a lower weighting factor, and the value “yellow” an even lower value...

In order to limit the number of questions to which the user must answer, Ophélie has an algorithm which optimizes the path to the final answer (plant name) by taking into account the previous answers. The system can determine which criterion - until then unused in the process - can provide the best information. Anything complex is handled by the computer, not by the user!

Dealing with the large number of descriptive characters to fill in - approximately 700 for 1 400 species, is another issue which had to be worked out. It was obviously not conceivable to enter all this information manually for each species!

In Ophélie, the solution was to introduce hierarchical levels of description which allow the factorization of descriptions. Three hierarchical levels are used: (i) “general” to differentiate families, (ii) “family” to differentiate genera and (iii) “genus” to differentiate species. This arrangement has led to a great reduction in the number of descriptive characters to type in, from 700 to approximately 30.

Note that the system works in a global way and that consequently, the “general” level can be enough for the determination of some species! However, the management of level changes is a delicate step. To do so, two solutions are used:

- delay the moment we activate the level change, but not excessively in order to maintain a reasonable number of questions;
- checking descriptive characters for level change.

Once the determination is finished, the system *Ophélie* only proposes a plant name if the difference between this plant and other plants is big enough, and if the degree of similarity with the description of the given plant has reached a certain level. This allows the reduction of the risk of getting an incorrect determination. The quick display of all photographs concerning all plants corresponding closely to the description is also a good control to validate the result.

The *Ophélie* system is therefore based on the principle of separation of species, *i.e.* individuals of the same species are generally more similar than individuals from related species.

With this structure, the expert system *Ophélie* could work with a number of species superior to 1 400, and must be able to take into account the 6 000 species of the French flora. The performance of this system is solely dependent on the quality of the information of the database and the sharpness of the parameter setting.

In conclusion, *Ophélie* solves the problem of absence of an answer which occurs when using a determination key and which stops the determination. Also, with *Ophélie*, a bad choice for a few answers is not a major problem for the final determination, because it is the mean of all the answers which is important.

2.5 The game module

This module proposes two games with three levels (Novice/ amateur/ expert). Their aim is to make users learn how to recognize plants in an entertaining way through the detailed observation of photographs..

3 Conclusions

Conceived in order to take action on the problems of biodiversity in a sustainable and indirect way, Flora Bellissima is an attempt to bring together different categories of botanists : novices, amateurs and experts in proposing an educational and entertaining software based on strong scientific knowlegde.

Flora Bellissima represents 4 000 hours of work including both the software design and the setting up of the knowledge base. Its extension to the whole French flora will require the setting up of a collaborative working group in order to capture characters concerning all taxa. The organization of such a group will be managed by the Tela Botanica association, which groups together all French-speaking botanists, and which has the qualifications and abilities required to achieve such a complex task. Tela Botanica also plans, in the long term, the online consultation of Flora Bellissima and could propose its application to other floras in the world within the framework of the research project Pl@ntNet [4] as a counterpart of existing systems.

Acknowledgement

We wish to thank the Tela Botanica team for their help in the realisation of this project as well as for its diffusion. We also want to thank Paul Fabre for translating this presentation in English.

References

[1] Flora Bellissima is a registered trademark of the Yourproject Informatique company.
[2] BDNFF: “Base de Données Nomenclaturale de la Flore de France” conducted by Michel Kerguélen † and Benoit Bock, in the Tela Botanica network.
[3] Maps of geographic distribution of plants are provided by Tela Botanica.
[4] Pl@ntNet is an Interactive plant identification and collaborative information system supported bay Agropolis Internatioonal Montpellier, France, built around three core teams that possess complementary skills : AMAP, IMEDIA and Tela Botanica.