

Flora Bellissima, un logiciel expert pour découvrir la botanique et identifier les plantes

Thierry Pernot, Daniel Mathieu

Abstract — Promouvoir et faciliter l'accès à la botanique en vue de préserver la biodiversité de manière indirecte, c'est l'objectif qui est donné à Flora Bellissima. Ce progiciel s'appuie sur une base de connaissances complétée de nombreuses fonctionnalités ainsi que sur un système expert de reconnaissance des végétaux nommé *Ophélie*. Destiné aussi bien aux novices et amateurs qu'aux experts, Flora Bellissima tente de démontrer qu'il est possible de concilier rigueur scientifique et vulgarisation en vue de rassembler un large public intéressé par la botanique.

Index Terms — Gestion de données botaniques, système expert, aide à la reconnaissance des plantes, pédagogie par le jeu.

1 INTRODUCTION

Comment sensibiliser le public à la découverte et à la protection de la biodiversité ? Une solution est de les aider découvrir la nature qui les entoure afin de mieux la connaître et la préserver. Pour que le public s'approprie cette connaissance, il lui faut des outils qui soient à la fois ludiques, faciles à utiliser et reposant sur des connaissances scientifiques éprouvées. Il faut également des outils capables de s'adapter à différents publics, du débutant au plus expérimenté, et qui prennent en compte l'évolution des connaissances que chaque individu va acquérir au cours de son apprentissage.

C'est dans cet objectif que Flora Bellissima [1], logiciel de gestion intégré dédié à la botanique a été réalisé. Organisé comme une sorte de mini ERP le logiciel est organisé autour d'une base de données centrale qui est constituée par un index scientifique des plantes (BDNFF [2]).

Autour de cette base de données sont organisés quatre modules permettant : i) une découverte pédagogique de la flore, ii) une aide à la détermination des plantes avec le système expert *Ophélie*, iii) un système de gestion des relevés botaniques et iv) un module de jeux.

Ce logiciel a été conçu pour être utilisé par toutes les personnes qui s'intéressent à la flore, quel que soit leur niveau de connaissances. Les novices qui souhaitent faire leurs premiers pas en botanique peuvent apprendre en s'amusant grâce au module de jeux, au glossaire imagé et aux photographies. Les amateurs qui souhaitent approfondir leurs connaissances peuvent mémoriser le résultat de leurs observations (textes et photographies) et se constituer ainsi leur propre base de données. Les experts peuvent consulter facilement et rapidement la nomenclature botanique avec tous ses synonymes et ses différents niveaux taxonomiques : espèces, sous-espèces, variétés et formes... Les associations naturalistes peuvent enregistrer et échanger rapidement leurs relevés botaniques. Chacun peut utiliser le système expert *Ophélie* pour déterminer une plante inconnue et lui attribuer un nom avant de l'enregistrer.

2 CONCEPTION DE L'OUTIL

2.1 Spécifications générales

L'analyse détaillée des différents objectifs nous a conduits à faire des choix que l'on retrouve dans les grands concepts à l'origine de Flora Bellissima :

- le logiciel est conçu comme un ERP, c'est-à-dire constitué de plusieurs applications qui partagent une unique base de données ;
- il repose sur un index de référence scientifique de la nomenclature botanique ;
- il est ouvert pour permettre l'ajout de textes et de photographies ;
- il est multi-niveaux pour être accessible à tous : novices, amateurs et experts ;
- le logiciel contient des données sur la flore de France pour 1 400 espèces ;

-
- *Thierry Pernot, the autor of the software, is with Yourproject Informatique 27 rue Saint-Georges 39360 Larrivoire France Yourproject@orange.fr www.yourprojectinfo.fr Il est biologiste de formation (Maîtrise de biologie des organismes et des populations à l'université de Besançon) et informaticien de profession (Chef de projet informatique en SSI puis en entreprise),*
 - *Daniel Mathieu is the president of Tela Botanica, the NGO which publishes and distributes software*

ERP= Enterprise Resource Planning : système d'information composé de plusieurs applications partageant une seule et même base de données. via un système automatisé.

- il propose un outil d'aide à la reconnaissance adapté aux différents publics avec le système expert *Ophélie* ;
- il présente plusieurs photographies par plante (9 800 photographies au total avec l'aspect général, l'inflorescence, la fleur, les feuilles etc.) pour confirmer la détermination ;
- il est disponible sur micro-ordinateurs PC et diffusé sur DVD.

2.2 Le module de découverte de la flore

Ce module aide à découvrir la flore de la France au travers de nombreuses photographies et permet de se constituer une base de données personnelle avec ses propres observations.

Il propose plus de 30 fonctionnalités qu'il n'est pas possible de détailler ici. En résumé, ce module permet de consulter la flore suivant divers points de vue : par type de plantes (arbre, arbustes, fougères, etc.), par leurs utilisations (médicinales, cultivées, toxiques), par famille botanique, par genre, par noms latins ou vernaculaires. Il permet également de consulter et de saisir des informations selon différents thèmes : description, propriétés médicinales, protection, répartition géographique [3]. À noter qu'il est possible d'ajouter ses propres photographies sur chacun des taxons. Ses trois niveaux d'accès, son glossaire imagé et son ergonomie procurent à ce module une forte valeur pédagogique.

L'idée importante qui ressort de ce module est le regroupement et la centralisation des informations afin d'en faciliter l'accès.

2.3 Le module de gestion des stations et des relevés botaniques

Ce module permet de définir des stations avec leurs caractères géographiques et écologiques, puis de leur associer des relevés botaniques. L'outil est conçu pour saisir de manière sûre et rapide les noms latins complets des plantes. En plus de ses fonctions de classement des stations et des relevés, ce module propose les options suivantes :

- recherche une plante parmi tous les relevés ;
- analyse de l'évolution du cortège floristique d'une station dans le temps ;
- duplication des relevés ;
- importer / exporter des stations et des relevés ;
- imprimer des relevés et les exporter sous Word, Excel, PDF etc.

2.4 Le module d'aide à la reconnaissance

Ce module est composé de deux grandes fonctions : un système de comparaison entre plusieurs plantes et un système expert d'identification.

Le système de comparaison permet de comparer entre 2 et 5 plantes en affichant tous les caractères qui les différencient. Cela permet de

trouver rapidement les caractères importants pour distinguer les plantes entre elles, sans avoir besoin de consulter l'ensemble de la description de chacune d'elles.

Le système expert, dénommé *Ophélie*, est un outil original d'aide à la reconnaissance des plantes utilisant la base de données de Flora Bellissima. 1 400 espèces sont décrites actuellement, avec environ 30 caractères par plante, soit 45 000 caractères décrits pour l'ensemble de la base.

Pour que le système soit efficace et performant il fallait prendre en compte deux contraintes importantes : i) tous les utilisateurs du système ne sont pas des experts et peuvent commettre des erreurs, ii) la morphologie des plantes varie d'un échantillon à l'autre. Aussi, pour fonctionner correctement le système doit être suffisamment souple pour accepter des erreurs et des variations de descriptions.

De ce fait, le système ne peut pas travailler de manière binaire avec les caractères - vrai ou faux -. Il doit travailler sur l'ensemble des caractères décrits simultanément, et non pas pris les uns à la suite des autres, comme dans une clé dichotomique.

Chaque valeur saisie permet d'affecter un « poids » à toutes les valeurs possibles du caractère, pour cela nous avons besoin de connaître « les distances » séparant les différentes valeurs d'un même caractère. C'est le rôle des matrices de proximité qui dans *Ophélie* représente plus de 11 000 informations. Pour pouvoir saisir une telle quantité d'informations, des outils informatiques spécifiques ont été développés. Ces outils ont permis d'accélérer la saisie tout en la contrôlant. *Ophélie* gère des caractères de type matrice, intervalle, valeurs quantitatives, etc.

Les « poids » qui sont affectés à chacune des valeurs choisies le sont grâce à des algorithmes qui dépendent du type de caractère : par exemple, si j'affecte la valeur *rose* à la couleur d'une fleur, la valeur *rose* dans *Ophélie* aura le « poids » le plus élevé, la valeur *rouge* un « poids » plus faible, la valeur *jaune*, un « poids » encore plus faible...

Afin de limiter le nombre de questions auxquelles l'utilisateur doit répondre, *Ophélie* dispose d'un algorithme qui optimise le chemin en fonction des réponses précédentes. Le système détermine ainsi pour chaque critère encore non utilisé, lequel permettra d'apporter le plus d'informations au système. C'est l'ordinateur qui gère la complexité, pas l'utilisateur !

Un autre problème important a dû être résolu, il s'agit de celui du très grand nombre de caractères descriptifs à définir (environ 700) pour les 1 400 espèces. Ce nombre est si grand qu'il est difficilement concevable de le renseigner « à la main » pour toutes les espèces.

La solution retenue dans *Ophélie* a été d'introduire des niveaux de description hiérarchiques qui permettent de factoriser les descriptions.

Trois niveaux hiérarchiques sont utilisés : un niveau « général » permettant de différencier les familles, un niveau « famille » permettant de différencier les genres et un niveau « genre » permettant de différencier les espèces. On pourrait imaginer un niveau supplémentaire pour descendre à la sous-espèce.

Notons que le système fonctionne de manière globale et que pour certaines plantes, les critères de niveau « général » peuvent suffire à déterminer l'espèce ! Cependant, la gestion des changements de niveau est un point délicat. Pour le gérer deux solutions sont utilisées :

- retarder le moment où l'on déclenche le changement de niveau, mais pas trop pour ne pas augmenter inutilement le nombre de questions ;
- utiliser des caractères descriptifs de contrôle du changement de niveau.

En fin de détermination, le système *Ophélie* ne propose un nom de plante que si l'écart séparant cette plante des autres est suffisant et si le pourcentage de correspondance avec la description atteint un certain niveau qui est paramétrable. Ceci permet de réduire le risque de déterminations erronées. La visualisation rapide de l'ensemble des photographies concernant les plantes les plus proches permet également de renforcer le contrôle du résultat.

Le système *Ophélie* repose ainsi sur le principe de séparation des espèces qui veut que chaque individu d'une même espèce est plus proche globalement des autres individus de la même espèce, que de ceux des espèces voisines.

Ainsi structuré, le système expert *Ophélie* pourrait fonctionner avec un nombre d'espèces supérieur à 1 400, et doit permettre de prendre en compte les 6 000 espèces de la flore de France. Les performances du système sont uniquement liées à la qualité des informations de la base de données et au niveau de finesse du paramétrage saisi.

En conclusion, *Ophélie* résout le problème rencontré lors de l'utilisation de clefs de détermination où l'absence de réponse à une question stoppe la détermination, et quelques mauvais choix dans les réponses ne remettent pas en cause la détermination finale, car c'est la moyenne de l'ensemble des réponses qui compte.

2.5 Le module de jeux

Ce module propose deux jeux dont la difficulté varie suivant le mode d'accès choisi (Novice/Amateur/Expert). Leur but est d'apprendre à reconnaître les plantes de manière ludique via l'observation détaillée de photographies.

6 CONCLUSIONS

Conçu dans le but d'agir sur les problèmes de biodiversité de manière pérenne et indirecte, Flora Bellissima constitue une tentative de rapprochement entre différentes catégories de botanistes : novices, amateurs et experts en proposant un logiciel pédagogique et ludique fondé sur des connaissances scientifiques solides.

Flora Bellissima a demandé plus 4 000 heures de travail comprenant la conception du logiciel et la constitution de la base de connaissances. Son extension à l'ensemble de la flore de France nécessitera la mise en place d'un groupe de travail collaboratif pour la saisie des caractères relatifs à l'ensemble des taxons. L'organisation d'un tel groupe de travail serait prise en charge par l'association Tela Botanica qui regroupe les botanistes francophones et dispose des compétences nécessaires à la réalisation d'une tâche aussi complexe. Tela Botanica envisage également à plus long terme la consultation en ligne sur Internet de Flora Bellissima et pourrait proposer son application à d'autres ensembles floristiques que celui de la France dans le cadre du projet de recherche PI@ntNet [4] en complément de systèmes déjà existants.

ACKNOWLEDGEMENT

Nous remercions toute l'équipe de Tela Botanica pour l'aide qu'elle a apportée à la réalisation de ce projet ainsi qu'à sa diffusion. Nous remercions également Paul... pour la traduction en anglais de cette présentation.

[1] Flora Bellissima est une marque déposée de l'EURL Yourproject Informatique

[2] Base de Données Nomenclaturale de la Flore de France réalisée par Benoit Bock sur la base de l'Index Synonymique de la Flore de France (ISFF) de Michel Kerguélen, dans le cadre du réseau Tela Botanica.

[3] Les cartes de répartition géographiques proviennent du réseau Tela Botanica.

[4] PI@ntNet is an Interactive plant identification and collaborative information system supported by Agropolis International Montpellier, France, built around three core teams that possess complementary skills : AMAP, IMEDIA and Tela Botanica.