

Les rencontres Inra

La botanique numérique : une approche collaborative et intégrée pour l'agriculture et l'environnement

Rencontre organisée par l'Institut national
de la recherche agronomique
dans le cadre
du Salon international de l'agriculture

Lundi 23 février 2009

RÉSUMÉ DES INTERVENTIONS

Stand Inra • Hall 3 • allée C • n°13
Parc des expositions de Paris • Porte de Versailles

Photos : © Inra



INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE



Pl@ntNet : un réseau collaboratif et une plate-forme informatique dédiés à la compilation et au partage d'outils et de connaissance en Botanique

DANIEL BARTHÉLÉMY

• Inra - UMR Botanique
et bioinformatique
de l'architecture des plantes
AMAP Cirad-Inra-Cnrs-Ird-
Université Montpellier 2
daniel.barthelemy@cirad.fr

Pl@ntNet est une initiative originale proposée par un consortium international regroupant de nombreux organismes autour de trois équipes aux compétences complémentaires : l'UMR Amap (<http://amap.cirad.fr/>), l'équipe-projet IMEDIA de l'INRIA (<http://www-rocq.inria.fr/imedia/cbir-demo.html>) et le réseau de botanique francophone Tela Botanica (<http://www.tela-botanica.org/>). Ce projet ambitieux de 4 ans a été sélectionné par Agropolis Fondation (<http://www.agropolis-fondation.fr/>) comme son premier projet étandard. Prévu pour un démarrage au premier trimestre 2009, il a pour objectif principal la création d'une plateforme informatique destinée à faciliter l'acquisition et l'exploitation collaborative de données sur le monde végétal par trois grandes catégories d'acteurs : les scientifiques, les gestionnaires de milieux naturels ou anthropisés et les citoyens.

Tout au long de la mise en place de cette plate-forme, l'ergonomie et l'efficacité des logiciels développés seront évaluées par un panel d'utilisateurs variés, au travers du réseau Tela Botanica, ainsi que dans le cadre de projets pilotes appliqués, réalisés avec de nombreux partenaires internationaux (UICN - Union Internationale pour la Conservation de la Nature, Institut Français de Pondichéry, Université de Davis [USA], Kruger National Parc [République Sud-Africaine], les principaux organismes internationaux de recherche sur les céréales [IRRI : Institut International de Recherche du Riz ; WARDA : Centre du riz pour l'Afrique ; CIMMYT : Centre international pour l'amélioration du maïs et du blé ; CIAT : Centre international pour l'agriculture tropicale], etc.) afin d'assurer l'adéquation des applications aux besoins et attentes des utilisateurs finaux. Plusieurs thèses en recherche de contenu visuel et en écologie sont programmées ; leurs résultats seront directement intégrés dans la plate-forme mise en place.

L'ensemble de ces recherches devra permettre des avancées significatives en Agronomie, Ecologie, Bioinformatique, Botanique tropicale et méditerranéenne. Pl@ntNet s'inscrit ainsi dans le vaste mouvement international pour une meilleure gestion et valorisation des informations taxinomiques, en complément de projets tels que GBIF (Global Biodiversity Information Facility), EDIT (European Distributed Institute of Taxonomy), Encyclopedia of Life, ou KeyToNature, et utilisera donc les standards définis par le *Taxonomic Database Working Group* (TDWG).



Pl@ntNote : un logiciel libre de gestion locale de données naturalistes centrées sur la plante

PHILIPPE BIRNBAUM

• Cirad - UMR
Botanique
et bioinformatique
de l'architecture des plantes
AMAP Cirad-Inra-Cnrs-Ird-
Université Montpellier 2
philippe.birnbaum@cirad.fr

La connaissance des plantes est très dépendante de la quantité de données disponibles de sorte que la botanique est depuis toujours une science collaborative dont l'outil historique principal, l'Herbier, représente la plus importante source au monde. Les outils informatiques modernes, permettent d'agréger des données de différente nature dans des bases dont le volume permet des synthèses inaccessibles au niveau individuel. Cependant le travail de recueil des données reste individuel, réalisé par un botaniste (ou un petit groupe) qui observe une plante et note ses propriétés. Dans ce contexte, il a des besoins très spécifiques dépendant de son objet d'étude et qui ne peuvent pas être prédéfinis par une structure standardisée. Par ailleurs, il réalise toujours cette gestion au niveau local et en aucun cas il ne saisira ses observations exclusivement sur un serveur Web distant dont il ne peut pas contrôler la pérennité. Le partage des données au sein d'une communauté s'inscrit donc toujours comme une seconde étape de la démarche.

Ces spécificités sont au centre du cahier des charges de l'application Pl@ntNote qui se définit comme un gestionnaire local de données. Ce logiciel repose sur une organisation dont l'unité de base est l'individu-plante, à laquelle sont associées des propriétés de différente nature. Une propriété est par exemple le nom (scientifique, vernaculaire,...) de la plante, sa localité, mais également une photo, une récolte d'herbier, une description, un prélèvement ADN, un diagnostic architectural, une mesure biométrique, etc. Un individu-plante peut donc posséder plusieurs propriétés et chaque propriété peut à son tour avoir des propriétés sous-jacentes. Il s'établit une hiérarchie descendante dont le nœud de départ est l'individu-plante. L'utilisateur peut générer de nouvelles propriétés, modifier celles existantes puis les partager avec une communauté d'utilisateurs. Adoptant une structure commune, la communauté des botanistes peut, en second lieu, échanger et publier le contenu vers un serveur Web.

Cette organisation est gérée au sein d'une interface simple, utilisant les standards courants (multi-fenêtres), utilisation de la souris, glisser/déplacer, raccourcis clavier, menus déroulants...). Elle permet de faire des recherches multicritères sans connaissance préalable des standards SQL (Structured Query Language), d'exporter les résultats dans divers formats, d'interroger des bases de données en ligne, de visualiser les images sous la forme de planche contacts, etc.

Pl@ntNote propose des méthodes qui s'adaptent au contexte, sans a priori sur les besoins des utilisateurs. Cette flexibilité constitue la force du logiciel qui attire des profils d'utilisateurs très variés. Actuellement, il est utilisé pour gérer des données de nature différente telles que la plus grande collection des herbiers de l'Afrique sahélienne (Herbier ALF-Cirad avec près de 70 000 échantillons), une collection vivante des Orchidées du Laos (700 spécimens), la photothèque des arbres du Mali (Cirad, 6200 illustrations), les inventaires des arbres de Guyane (IRD, 75 000 individus), le suivi de croissance des plantes de sous-bois (IRD-Cayenne). D'autres partenaires sont d'ores et déjà sur le point d'utiliser Pl@ntNote notamment l'herbier de l'Institut Français de Pondichery en Inde, le suivi des mauvaises herbes de l'île de la Réunion (Cirad) ou le projet de paléobotanique « Terrestrialisation » soumis à l'ANR-2009 (Cnrs).



Tela Botanica, le réseau collaboratif des botanistes francophones

DANIEL MATHIEU¹,
RÉNALD BOULNOIS²

• 1. Tela Botanica
dmathieu@tela-botanica.org

• 2. Biotope

Depuis bientôt 10 ans, le réseau Tela Botanica développe des projets autour de la botanique en collaboration directe avec des chercheurs, des professionnels et le grand public. Il participe ainsi à la valorisation des savoirs sur la biodiversité végétale, et facilite la coopération des acteurs au niveau national et international.

La botanique, une science citoyenne et participative

Avec Tela Botanica, la botanique se construit et se diffuse grâce à la participation de tous. Ce sont des milliers d'inscrits au réseau, issus de pays différents et de compétences variées, qui contribuent bénévolement au développement de cette activité. Tela Botanica offre ainsi la possibilité à chacun de s'impliquer dans la réalisation de projets. Les données, ainsi collectées, sont consultables et réutilisables de façon libre et gratuite.

Exemple d'un projet participatif : la répartition des plantes par département

Contrairement à ce que l'on peut trouver sur les sites Internet de nombreux pays, il n'existe pas, en France métropolitaine, de cartes de répartition de l'ensemble de notre flore consultable en ligne. L'objectif de ce projet est donc de produire une cartographie départementale des plantes de France métropolitaine à partir du dépouillement de la littérature, et sur la base d'un critère simple, celui de la présence / absence dans le département. Atteindre cet objectif a nécessité un travail important de collecte d'informations mettant à contribution l'ensemble des botanistes du réseau. Le projet est animé par un coordinateur bénévole qui a la charge d'harmoniser l'ensemble des listes qu'il reçoit des membres du réseau. Pour organiser et faciliter cette collecte, il utilise des outils collaboratifs mis à disposition par Tela Botanica. Plus de 150 personnes, organisées par départements, participent à ce projet et 95 départements disposent aujourd'hui d'une liste de références. Des mises à jour régulières sont effectuées, avec ajout de nouveaux départements, actualisations et corrections de ceux déjà présents en fonction des remarques et des investigations de terrain.

Il est maintenant possible d'obtenir pour chaque espèce la liste des départements où elle est présente ainsi qu'une représentation cartographique sur une carte de France à l'échelle départementale. Ces données sont consultables et réutilisables de façon libre et gratuite sur le site internet de Tela Botanica sous licence libre Creative Commons.

Présentation de l'utilisation par un professionnel

BIOTOPE est une société reconnue notamment pour ses éditions et ses expertises environnementales. Cette société compte à ce jour 13 agences en France et à l'étranger. Pour élaborer ces études, elle utilise certains résultats produits par le réseau Tela-Botanica :

- utilisation de l'outil de gestion de la synonymie des taxons, notamment dans le cadre des aspects réglementaires des études d'impact ;
- utilisation du référentiel à jour de la taxonomie des espèces végétales françaises pour produire des états initiaux à jour sur le plan de la nomenclature ;
- utilisation de la répartition départementale des espèces pour alimenter le dire d'expert dans les dossiers nécessitant un zoom sur une espèce précise.

Par ailleurs, dans le cadre du montage en cours de sa base de données interne, BIOTOPE réalise un travail de mise en correspondance des textes réglementaires de protection des espèces au niveau régional et départemental et a besoin de la nomenclature à jour produite et mise à disposition par le réseau Tela Botanica.



IDAO : Un outil graphique d'identification des plantes.

PIERRE BONNET

• Inra – UMR Botanique et bioinformatique de l'architecture des plantes
AMAP Cirad-Inra-Cnrs-Ird-
Université Montpellier 2
pierre.bonnet@cirad.fr

L'identification précise des espèces végétales est primordiale pour de nombreuses disciplines, notamment pour l'élaboration de moyens de lutte performants contre les adventices des cultures. Or cette identification est généralement difficile pour des non-spécialistes, tels que des agronomes, des techniciens ou des étudiants. Les difficultés rencontrées avec les outils classiques, comme les flores ou les manuels, sont généralement liées aux méthodes d'identifications structurées sous forme de clés dichotomiques, et à la terminologie employée qui reste réservée surtout aux spécialistes. De plus, ces méthodes sont généralement basées sur l'emploi de caractères des organes reproducteurs (fleurs, fruits), qui ne s'expriment qu'une courte partie de l'année chez les plantes, alors qu'il est parfois nécessaire de les identifier avant leur stade de développement adulte, notamment pour toutes les espèces invasives.

Afin de remédier à ces contraintes nous développons depuis plusieurs années une nouvelle approche pour l'identification des plantes, à travers le logiciel libre et le procédé IDAO (Identification Assistée par Ordinateur). Celui-ci se caractérise par l'utilisation d'un portrait-robot graphique, qui permet aux utilisateurs de construire le portrait d'une plante à partir de caractères morphologiques aussi bien de l'appareil végétatif que reproducteur. Les caractères peuvent être choisis librement, sans contrainte d'un ordre de sélection, et modifiés à n'importe quel moment du processus d'identification. Les espèces répertoriées dans ce logiciel sont alors organisées en fonction de leur ressemblance avec le portrait réalisé par l'utilisateur. Celui-ci peut visualiser des illustrations de chacune d'elle, ainsi que des descriptions morphologiques, dont les termes scientifiques sont définis et également illustrés. Cette méthode d'identification tolère un certain nombre d'erreurs d'observations et prend en compte le polymorphisme de chaque taxon, ce qui en facilite d'autant plus l'utilisation. Les erreurs d'observation peuvent être mises en évidence, contribuant ainsi à la formation des utilisateurs. Une fois l'espèce identifiée, une présentation de sa biologie, de son écologie et des moyens de lutte les plus efficaces est directement accessible dans sa fiche descriptive.

Cet outil a été adapté aux adventices de l'Afrique de l'Ouest, de la plaine Indo-Gangétique, de la Péninsule Indochinoise, et des îles du S-O de l'Océan Indien. Son interface graphique et conviviale le rend particulièrement adapté à une utilisation par des non-spécialistes, c'est pourquoi il a également été utilisé sur d'autres flores que celles des adventices, notamment des flores de plantes ligneuses ou des groupes d'espèces protégées comme les Orchidées.

Ce logiciel est utilisable aussi bien sur Internet qu'hors ligne, sur ordinateur personnel, que sur ordinateur ultramobile de terrain, et gère des versions multilingues des descriptions taxonomiques pour favoriser sa diffusion dans les pays concernés.

IDAO est actuellement en pleine évolution, notamment pour élargir la gamme des utilisateurs, et permettre son application à de plus large flores, et son exploitation pour un plus grand nombre de caractères d'identification (écorce, anatomie, etc.).



IKONA : un moteur d'indexation, de navigation et de recherche d'information par le contenu visuel - Application aux bases de données images en botanique

**NOZHA BOUJEMAA,
RAFFI ENFICIAUD**

• INRIA, Paris-Rocquencourt
Équipe Projet IMEDIA
Images et Multimédia :
indexation, navigation
et recherche
Nozha.Boujemaa@inria.fr
Raffi.Enficiaud@inria.fr

Ikona est un moteur de recherche d'images par le contenu. Il est développé depuis plusieurs années dans le cadre des recherches du projet IMEDIA de l'INRIA. Son objectif est de pouvoir retrouver des images à partir de signatures visuelles calculées par des descripteurs d'images. Ikona a fait l'objet de nombreuses applications sur différents types d'images, comme des images satellites, des visages humains, des collections de photos professionnelles ou encore des objets en 3D. Des travaux de recherche, actuellement en cours de finalisation, utilisent Ikona pour la reconnaissance d'espèces végétales à partir d'images de leurs organes constitutif (feuille, organe sexuel...). Il est ainsi possible de retrouver en très peu de manipulations l'espèce associée à une image de feuille par exemple.

L'application se présente selon une architecture client-serveur. Le client est une interface graphique mettant en évidence les similitudes entre les images. Le serveur analyse les images pour en extraire des descripteurs visuels, et effectue des comparaisons avec les requêtes formulées par le client. Il gère également les bases de données d'image ainsi que les annotations, ce qui permet aux utilisateurs finaux d'avoir accès à de grandes bases et à des informations supplémentaires sur les images comme, dans le cas particulier des plantes, les groupes taxonomiques et/ou les stades phénologiques des prises de vue. Le serveur fonctionne sur des machines spécialisées. Il existe par contre des clients sur plusieurs plateformes, et il existe également une passerelle Web accessible à un grand nombre d'utilisateurs : <http://www-rocq.inria.fr/imedia/cbir-demo.html>

Ikona a été utilisé pour la reconnaissance des orchidées asiatiques. Fortement menacées par la destruction de leurs habitats et par leur collecte non-raisonnée, elles présentent des enjeux de conservation importants, d'où la nécessité du développement d'outils d'identification rapides et accessibles par des non-spécialistes. Or, cette famille d'orchidées est complexe et l'identification est reconnue comme étant difficile. Par ailleurs, il n'existe en Asie que peu de ressources multimédia fiable sur ces plantes.

Dans cette étape d'analyse, les prises de vue se sont concentrées sur les feuilles et selon plusieurs protocoles. Les images, analysées par Ikona, ont ensuite été annotées textuellement selon un formulaire.

L'utilisation d'Ikona dans ce cadre d'étude s'est montrée très pertinente. Elle a donné lieu à plusieurs publications dans des revues scientifiques spécialisées, ce qui valide l'approche et le rapprochement des deux communautés scientifiques. L'analyse a mis également en relief des améliorations possibles, liés aux traitements internes à Ikona, au niveau des descripteurs visuels et des méthodes de recherche. Des cas d'utilisation nouveaux se sont manifestés, comme l'utilisation d'Ikona dans un environnement complet dédié à l'identification des plantes. Ces problématiques de recherche et d'ingénierie font l'objet du projet Pl@ntNet, qui commence au début de l'année 2009.

