



RAPPORT DE STAGE DUT INFORMATIQUE

Amélioration de la base de données de gestion d'échantillons de l'Unité de Recherche de Zoologie Forestière d'Orléans

UDAHINYUKA Innocent

Année spéciale 2013/2014

Durée - date du stage : 10 Semaines - du 23 Juin au 29 Août 2014

Maîtres de Stage : Béatrice COURTIAL & Régis PHELUT

Tuteur de Stage : Emmanuel MELIN

REMERCIEMENTS

Je tiens à remercier plus particulièrement et à témoigner toute ma reconnaissance aux personnes suivantes :

Monsieur Alain ROQUES, Directeur de l'unité de Recherche de Zoologie Forestière (URZF), de m'avoir accordé ce stage au sein de l'unité, stage concluant ma formation de DUT informatique.

Madame Béatrice COURTIAL, ingénieure d'études dans l'unité URZF, et également Animatrice et Coordinatrice de l'Assurance Qualité pour l'INRA centre Val de Loire pour m'avoir accueilli, intégré au sein de l'INRA et m'avoir accordé toute sa confiance.

Monsieur Régis PHELUT, PRI (Personne Ressources Informatiques) de l'unité d'avoir accepté de partager son bureau, ainsi que pour les explications qu'il m'a fourni sur le fonctionnement de la base de données sur laquelle j'ai eu à travailler.

Monsieur Daniel SAUVARD, chargé de recherche dans l'unité, qui a mis à ma disposition une copie de la base pour que je puisse effectuer les tests dans les mêmes conditions que les biologistes.

Monsieur Jean-Léandre HATON, PRI à l'Unité de Recherche d'Amélioration, Génétique et Physiologie Forestières, pour m'avoir prêté des livres sur les langages de programmation PHP et JavaScript.

Je remercie aussi l'ensemble du personnel de l'unité pour leur accueil sympathique et leur coopération professionnelle tout au long de mon stage.

Je remercie également Monsieur Emmanuel MELIN, mon tuteur de stage, pour sa visite et ses nombreux conseils.

Et enfin, je remercie ma famille pour le soutien qu'ils m'ont apporté tout au long de mon stage et mes remerciements vont plus particulièrement à mon épouse et à mes enfants d'avoir accepté mon absence pendant ma formation.

Sommaire

REMERCIEMENTS	1
ABREVIATIONS	3
INTRODUCTION	4
1. PRESENTATION DE LA STRUCTURE D'ACCUEIL	5
1.1 L'Institut National de la Recherche Agronomique	5
1.2 Le site d'Orléans	6
1.2.1 Historique du site d'Orléans	6
1.2.2 Organisation du site	7
2 SUJET DE STAGE ET OBJECTIFS	10
2.1 Sujet de stage	10
2.2 Objectifs du stage	11
3. DEROULEMENT DU STAGE ET TECHNOLOGIES UTILISEES	12
3.1 Les activités dans l'unité et la base	12
3.2 L'existant	15
3.2.1 La base "SAMPLES"	15
3.2.2 L'interface graphique	18
3.2.3 Traitement d'un échantillon au niveau de la base de données	21
3.3 Les technologies utilisées	22
3.4 Le travail réalisé	24
3.4.1 Travail réalisé au niveau de la base	24
3.4.2 Travail réalisé au niveau de l'interface	25
CONCLUSION	31
BIBLIOGRAPHIE	32
WEBOGRAPHIE	32
Annexe	33

ABREVIATIONS

AGPF Amélioration, Génétique et Physiologie Forestières

AJAX Asynchronous JavaScript and XML

DOM Document Object Model

DUT Diplôme Universitaire de Technologie

EA Environnement et Agronomie

EFPA Ecologie des Forêts, Prairies et milieux Aquatiques

EPST Établissement Public à caractère Scientifique et Technologique

GBFOr Génétique et Biomasse Forestière d'Orléans

HTML Hypertext Markup Language

INRA Institut National de la Recherche Agronomique

JSON JavaScript Object Notation

PHP Hypertext Preprocessor

PRI Personne Ressources Informatiques

SDAR Services Déconcentrés d'Appui à la Recherche

SGBDR Système de Gestion de Base de Données Relationnelle

URZF Unité de Recherche de Zoologie Forestière

XAMPP X Apache MySQL Perl PHP

XML Extensible Markup Language

INTRODUCTION

La formation de DUT INFORMATIQUE (année spéciale) que j'ai suivie à Orléans se conclut par un stage d'une durée de 10 semaines en entreprise.

Cette partie de la formation a pour but de :

- se familiariser avec le monde professionnel,
- mettre en application les connaissances théoriques acquises pendant la période de formation,
- compléter les connaissances acquises,
- acquérir une première expérience en informatique.

C'est dans ce cadre que j'ai effectué mon stage au sein de l'Institut National de la Recherche Agronomique (INRA), et plus précisément dans l'Unité de Recherche de Zoologie Forestière (URZF) d'Orléans du 23 juin au 29 août 2014. J'ai été accueilli et encadré par M^{me} Béatrice COURTIAL, Ingénieure d'études en Biologie et M^r Régis PHELUT PRI de l'unité.

Les travaux de recherches réalisés à l'URZF portent sur des échantillons de natures diverses (insectes, feuilles, baies ...) et toutes les informations liées à ces échantillons doivent être conservées (localisation du prélèvement, date, identification...) dans une base de données pour pouvoir ré-exploiter les informations ultérieurement. Cette base de données doit donc toujours rester propre et fonctionnelle.

Mon stage a donc eu pour objectif l'amélioration de cette base de données de gestion d'échantillons.

Ce rapport résume les différentes étapes du déroulement de mon stage. Je vais donc commencer par la présentation de l'INRA et de l'URZF. Ensuite, je parlerai du sujet du stage et des objectifs qui m'ont été fixés, du travail réalisé et des technologies utilisées. Enfin, je terminerai par le bilan de mon stage dans la partie «Conclusion».

1. PRESENTATION DE LA STRUCTURE D'ACCUEIL

1.1 L'Institut National de la Recherche Agronomique

L'Institut National de la Recherche Agronomique (INRA) a pour mission de produire et diffuser des connaissances scientifiques et des innovations dans les domaines de l'alimentation, de l'agriculture et de l'environnement. L'INRA a été fondé en 1946 dans le contexte de la reconstruction nationale d'après-guerre et du projet de modernisation de l'agriculture française. Depuis 1984, l'institut a un statut d'Établissement Public à caractère Scientifique et Technologique (EPST) et il est placé sous la tutelle conjointe du ministère chargé de l'Agriculture et des ministères chargés de l'Education Nationale, de l'Enseignement Supérieur et de la Recherche. Par sa production scientifique, l'INRA est en troisième position mondiale en recherche agronomique.

Les principales missions confiées à l'INRA sont de :

- Produire et diffuser les connaissances scientifiques
- Concevoir des innovations et des savoir-faire pour la société
- Eclairer, par son expertise, les décisions des acteurs publics et privés
- Développer la culture scientifique et technique et participer au débat science/société
- ❖ Former à la recherche et par la recherche

L'INRA exerce une politique de partenariat et de collaboration avec la communauté scientifique nationale (organismes de recherche et enseignement chercheur) et internationale, le monde agricole, les entreprises et les collectivités territoriales.

L'INRA compte au total (voir figure 1):

- ❖ Près de 9000 agents au 31 décembre 2012 dont 1921 chercheurs, 2804 ingénieurs, 4249 techniciens et administratifs.
- ❖ 13 départements de recherche,
- ❖ 17 centres de recherche régionaux et un siège à Paris.

Le budget qui est alloué à l'institut en 2013 se monte à 881 millions d'euros.

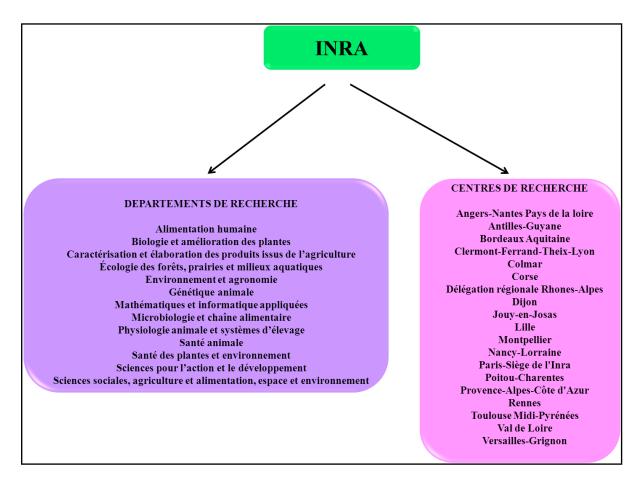


Figure 1 : Organisation de l'INRA

1.2 Le site d'Orléans

1.2.1 Historique du site d'Orléans

En 1964, la première mission de prospection pour l'implantation des recherches forestières en terres orléanaises a lieu.

En 1976, les bâtiments de l'amélioration des arbres forestiers sont terminés.

En 1977, le centre est inauguré.

En 1980, le domaine expérimental de Bourges est rattaché au centre d'Orléans.

En 1983, la Station 'Sciences du Sol' est inaugurée.

En 1986, les locaux de Zoologie Forestière sont inaugurés.

En 1999, le bâtiment de Zoologie Forestière brûle dans un incendie.

En 2001, l'agrandissement du bâtiment de l'unité 'Sciences du Sol' et du simulateur de pluie est inauguré.

En 2004, le nouveau bâtiment de Zoologie Forestière est inauguré.

A partir du 1^{er} janvier 2013, les Centres d'Orléans et de Tours ont fusionné pour donner le Centre de Recherche **INRA Val de Loire**.

1.2.2 Organisation du site

Situé sur les communes d'Ardon et de Saint-Cyr-en-Val, le site de Recherche d'Orléans s'étend sur 63 ha, et compte près de 8500 m² de bureaux et laboratoires. Il est composé de 6 unités et est présidé par M^{me} Catherine Beaumont (cf figure 2).

Les recherches développées sur le site concernent 3 domaines :

- ❖ la forêt,
- les insectes.
- ❖ les sols avec une orientation marquée en matière d'environnement.

Les 6 unités présentes sur le site sont :

• 3 unités de recherche :

- Unité de Recherche **Amélioration, Génétique et Physiologie Forestières** (AGPF) : valorisation des ressources génétiques forestières en vue d'une production durable de bois d'œuvre et de biomasse, tout en prenant en compte l'impact écologique des populations domestiquées sur l'écosystème et un contexte climatique changeant.
- Unité de Recherche de Zoologie Forestière (URZF) : caractérisation aux plans génétique, physiologique et comportemental des populations d'insectes forestiers envahissants ou en expansion sous l'effet du changement global, analyse les mécanismes gouvernant la réussite des invasions biologiques d'insectes, mesure des risques écologiques et économiques et déduction des mesures de gestion.
- Unité de Recherche **Science du Sol** : étude des interactions entre les propriétés physiques des sols et leurs fonctionnements hydriques et biogéochimiques.

•1 unité de service :

 Unité de Service InfoSol: constitution et gestion d'un système d'information à vocation nationale sur les sols, par rapport à leur distribution spatiale, leurs propriétés et l'évolution de leurs qualités.

• 1 unité expérimentale :

- Unité Expérimentale **Génétique et Biomasse Forestière d'Orléans** (GBFOr) : description et conservation de la diversité génétique des espèces forestières.

• 1 unité d'appui à la recherche :

- Services Déconcentrés d'Appui à la Recherche (SDAR) : gestion collective et organisation des fonctions transversales de support et de soutien, dans une relation de partenariat avec l'ensemble des autres unités.

Les 2 départements de recherche représentés sur le site d'Orléans sont :

- Ecologie des Forêts, Prairies et milieux Aquatiques (EFPA)
- Environnement et Agronomie (EA)

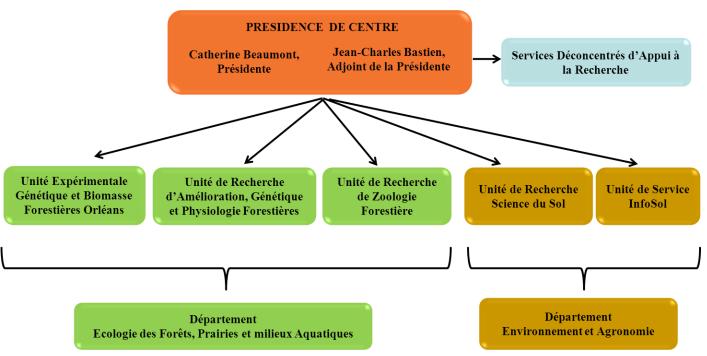


Figure 2 : Organisation de l'INRA d'Orléans

1.3 L'Unité de Recherche de Zoologie Forestière

L'Unité de Recherche de Zoologie Forestière (URZF) a pour vocation de comprendre les mécanismes qui régissent les populations d'insectes en expansion sous l'effet des activités humaines et des changements environnementaux et d'étudier les invasions biologiques.

En pratique, les recherches sont de :

- caractériser aux plans génétique, physiologique et comportemental les populations d'insectes forestiers envahissants ou en expansion sous l'effet du changement climatique,
- analyser les mécanismes gouvernant la réussite des invasions biologiques d'insectes,
- mesurer les risques écologiques et économiques et en déduire des mesures de gestion.

L'unité dispose de 17 personnes titulaires (cf figure 3) dont 8 chercheurs INRA, 1 maître de conférences de l'Université d'Orléans, 4 ingénieurs et 5 techniciens. Depuis 2006, l'URZF a accueilli 6 doctorants, 7 étudiants en Master II des Universités d'Orléans, Rennes et Valencia (Espagne), 12 en Master I d'Orléans et de Metz et 11 BTS en Biotechnologie. Elle a régulièrement accueilli des post-doctorants étrangers (Chine, Russie, Mexique).

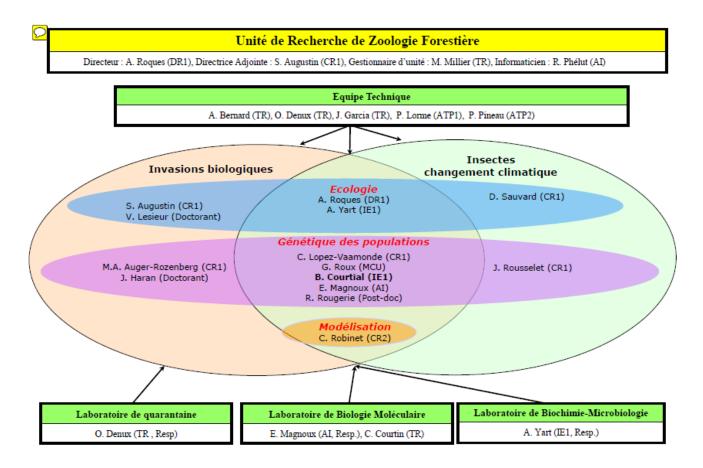


Figure 3 : Organigramme de l'Unité de Recherche de Zoologie Forestière.

2 SUJET DE STAGE ET OBJECTIFS

2.1 Sujet de stage

Le personnel de l'unité URZF reçoit par colis, ou récolte par lui-même, des échantillons de divers types (Insectes, Graines, Feuilles,...) sur lesquels sont réalisés des analyses pour répondre à des questions scientifiques. L'unité a impérativement besoin d'une base de données bien structurée pour pouvoir enregistrer, consulter et modifier les informations sur les échantillons en sa possession.

Cette base de données informatisée leur permet de gagner du temps dans leurs missions de tous les jours. Par exemple, une recherche d'échantillon avec la base est plus rapide avec l'interface qu'une recherche avec des fiches papier.

La base de données « SAMPLES », avec son interface graphique, ont été développées en 2008 pour répondre aux besoins de l'unité. En effet, il était crucial de pouvoir attribuer à chaque échantillon collecté un identifiant unique de type code-barres, consultable automatiquement (lecteur optique), permettant une traçabilité immédiate et fiable de son origine et de celle de ses sous-échantillons (pattes, séquences,...). Cette base de données contient ainsi les différentes informations sur les échantillons : date et lieu de prélèvement, lieu de stockage, généalogie Pour qu'elle réponde à son objectif initial, il est nécessaire d'être sûr de sa fiabilité.

Au fil et à mesure de l'utilisation de la base, les utilisateurs se sont rendu compte de l'existence ou de l'apparition d'anomalies de fonctionnement ou "bugs". L'apparition des bugs peut être le résultat de changement de version du navigateur, de données mal saisies, de la non-maîtrise de la part des utilisateurs de certaines fonctions, etc... De plus, la base de données contenant de plus en plus d'entrées (plus de 20 000 échantillons actuellement), certaines requêtes trop lourdes à gérer n'aboutissent pas ou prennent du temps.

Enfin, aujourd'hui, les besoins au sein de l'unité ont évolué avec le temps et d'autres fonctionnalités sur la base sont apparues nécessaires telles que la consultation sur une localisation géographique ciblée (département français, province étrangère...) ou l'affichage des échantillons sur une carte géographique.

Le projet consistait donc à :

- corriger les "bugs",
- fournir de nouvelles fonctionnalités à l'interface graphique.

Mon stage s'est déroulé au sein de l'Unité de Recherche de Zoologie Forestière sous la direction de M^{me} Béatrice COURTIAL (ingénieure d'études) et de M^r Régis PHELUT (PRI de l'Unité).

2.2 Objectifs du stage

Les objectifs à atteindre consistaient, à partir d'une application déjà existante, à modifier son interface, mais également à nettoyer la base de données associée. Après le développement, l'interface doit permettre aux utilisateurs de :

- réaliser des requêtes avec de nouveaux critères (par exemple par département ou par province dans le cas de pays étrangers)
- **donner la possibilité de créer un nouvel expéditeur**
- ❖ visualiser les échantillons issus d'une requête sur une carte géographique
- ❖ pouvoir éliminer définitivement un échantillon entré par erreur (en mode administrateur)
- permettre de restorer des échantillons détruits par erreur
- ***** modifier une analyse
- éliminer la possibilité de lier une radio comme résultat d'analyse (redondant avec l'option d'ajout de fichier déjà existante)
- * pouvoir déplacer un ensemble d'échantillons d'un lieu de stockage à un autre en mode utilisateur (fonction déjà disponible en mode administrateur)
- * nettoyer les listes déroulantes proposées aux utilisateurs
- * modifier l'ordre d'affichage des champs pour respecter la logique taxonomique
- afficher un message de retour (confirmation) suite à une opération sur la base
- * avoir une consultation des résultats d'analyse simplifiée
- * modifier les champs affichés suite à une requête
- afficher graphiquement la généalogie des échantillons

- * maintenir l'ordre de saisie à l'issue d'une requête
- ***** afficher correctement les compteurs d'entités restantes
- afficher le nombre d'échantillon par lieu de stockage en mode utilisateur (fonction déjà disponible en mode administrateur)
- * maintenir la casse des données lors de l'exportation sous Excel suite à une requête

3. DEROULEMENT DU STAGE ET TECHNOLOGIES UTILISEES

3.1 Les activités dans l'unité et la base

Le personnel de l'unité va récolter sur le terrain, ou reçoit par colis, les échantillons. Quand un échantillon est récolté, les opérateurs remplissent une feuille de terrain (cf figure 4) sur laquelle ils notent toutes les informations concernant l'échantillon (code-barres, type, localisation, date ...).

FEUILLE DE TERRAIN					
Scientifiq	Opérate Cont ue demande	ate: / /20 (jj/mm/aaaa) Code Barre échantillon rat: rat: sur: rot: Piègeage lumineux			
LOCALISATION	N°GPS :				
ESSENCE	Pin d'Alep Pin à croch Pin maritin Pin noir Pin pignon Pin radiata Pin sylves	Genévrier Cédre Mélèze tets Genévrier Commun Cyprès de l'Atlas Marronnier ne Genévrier de Phénicie Cyprès d'Italie Rosa Genévrier Sabine Douglas Sapin (parasol) Genévrier Thurifère			
z	Insecte	Nom : Stade : Ponte □ Larve □			
ECHANTILLON	Nid	Nom: Nombre prélevé : Nombre de nids / arbre¹ : Taille des nids² :			
H	Autre	Cône Feuille Baie, Fruit Graine Autre Nom: Nombre prélevé: Position: sur l'arbre sur le sol			
ENVI- RON- NEMENT	Situation	Milieu³: Positionnement⁴:			
COMMENTAIRES					
1 2 2 2 3 4 5 5 5 5 5 5 5 5 5					

Figure 4 : Feuille de terrain utilisée par les opérateurs lors d'un prélèvement

Ensuite, des analyses sont effectuées pour obtenir des résultats de recherche. Ces analyses s'effectuent soit directement sur l'échantillon, soit après traitement d'un ou plusieurs échantillons. Les traitements possibles sont de différents types (cf figures 5 et 6) :

- Prélèvement : une partie de l'échantillon donne un ou plusieurs échantillons fils
- Transformation : un échantillon est modifié pour devenir un nouvel échantillon
- Division : un échantillon est séparé et donne plusieurs nouveaux échantillons
- Regroupement : plusieurs échantillons sont associés pour donner un nouvel échantillon

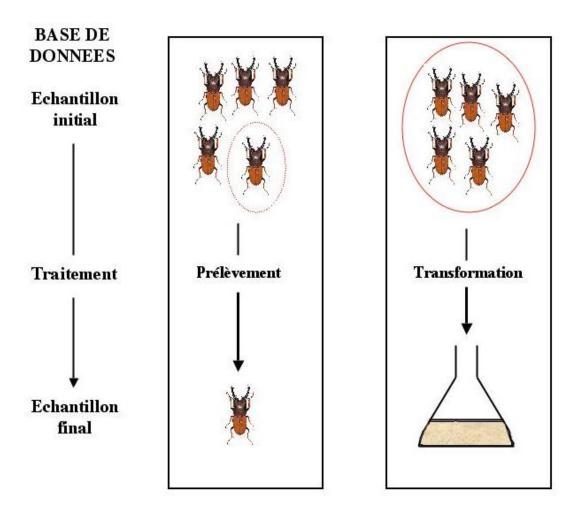


Figure 5 : Prélèvement et transformation d'un échantillon

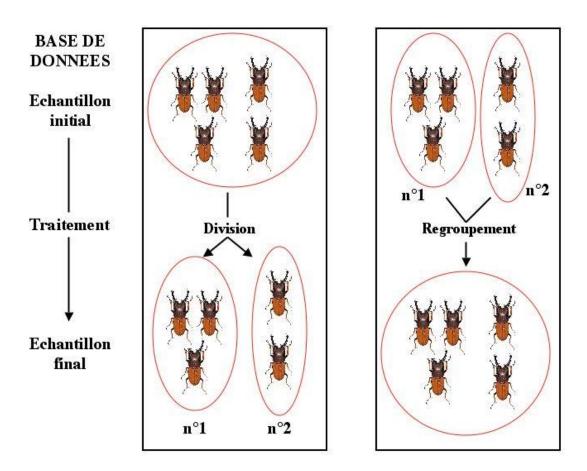


Figure 6 : Division et regroupement d'échantillons

Une fois les travaux de recherche achevés sur un échantillon, il est conservé dans un lieu de stockage approprié (cf figure 7) ou bien il est détruit (analyse destructive, fin de projet,..).



Figure 7 : Exemples de modes de conservation des échantillons

3.2 L'existant

3.2.1 La base "SAMPLES"

La base de données "SAMPLES" a été développée en 2008. Elle a été conçue et mise en place par Joshua AMAVI, étudiant en Master "Sciences et Technologies" Mention "Sciences et technologies de l'information et de la communication" de l'université d'Orléans.

En 2011, Elom AMAVI, stagiaire en Licence 3 "Informatique" Mention "Sciences et Technologies pour l'Information et la Communication" à l'Université d'Orléans, a travaillé sur la base pour corriger les premières anomalies de fonctionnement et pour répondre aux nouveaux besoin des utilisateurs.

La base de données a été créée sous PostgreSQL qui est un SGBDR (Système de Gestion de Base de Données Relationnelles). Il s'agit d'un système de gestion de base de données non commercial disposant de nombreuses fonctionnalités. L'efficacité d'écriture, l'accès aux données par les seuls index et les autres opérations de bas niveau permettent à PostgreSQL de gérer des volumes conséquents.

Aujourd'hui, cette base de données est composée de 32 tables (cf figure 8) qui, par convention, ont des noms et des attributs en anglais. Les tables composant la base de données "SAMPLES" sont les suivantes :

sample : table principale concernant les caractéristiques des échantillons

analysis: liste de toutes les analyses réalisées sur les échantillons

analysis_operator : table qui relie les tables analysis et operator en cas de réalisation d'une analyse

analysis_result: table qui contient tous les résultats des analyses

barcode: table contenant tous les codes-barres utilisés dans la base de données

dictionary : dictionnaire de données qui aide à la saisie des échantillons

destruction: table renseignant l'état (destruction ou non) d'un échantillon

file: table contenant tous les fichiers qui sont liés aux échantillons

operator : liste de tous les utilisateurs de la base de données

origin_qualfiers : liste des niveaux d'identification taxonomique utilisés dans la base de données

origin_shortcut : liste des raccourcis utilisés pour l'identification taxonomique des échantillons

exploités le plus couramment dans l'unité

place : liste recensant les différents lieux de stockage de l'unité

radio: table contenant les radiographies effectuées sur les échantillons

reception : table renseignant sur le prélèvement d'un échantillon (localisation, date, opérateur...)

relationship : table gérant la hiérarchie des échantillons et les traitements (manipulations) subis par un échantillon

result_precision : liste des intitulés précis d'analyses réalisées

result_type : liste de tous les types généraux d'analyses effectuées dans l'unité (séquençage, radiographie...)

result_type_precision : table faisant le lien entre le type général d'analyse et l'intitulé précis de l'analyse réalisée

sample_additionnal: table qui fait le lien entre la table sample et la table sample_qualifiers

sample_file : table qui relie un échantillon à tous les fichiers auxquels il est rattachés

sample_initial: table qui lie un échantillon à son(ses) parent(s) initial(initiaux) lorsqu'il y a eu un (des) traitement(s)

sample_origin : table contenant l'identification taxonomique des échantillons

sample_place : table indiquant les conditions de stockage d'un échantillon (localisation, date, opérateur)

sample_qualifiers : liste des informations complémentaires associées à un échantillon (hôte végétal, sexe...)

sample_team : table qui fait le lien entre un échantillon et un projet de recherche

sample_type : table qui fait le lien entre un échantillon et le type de matériel constituant l'échantillon

sender: liste de tous les expéditeurs d'échantillons (personnes extérieures à l'unité)

stype: liste de types d'échantillons exploités au sein de l'unité (larve, adulte, ponte, baie...)

team : liste des différents projets de recherche de l'unité

terrain_paper : table contenant toutes les feuilles de terrain remplies lors des prélèvements d'échantillons par les opérateurs

treatment : table renseignant sur les différents traitements qu'a subis un échantillon

treatment_operator : table reliant un opérateur à la réalisation d'un traitement

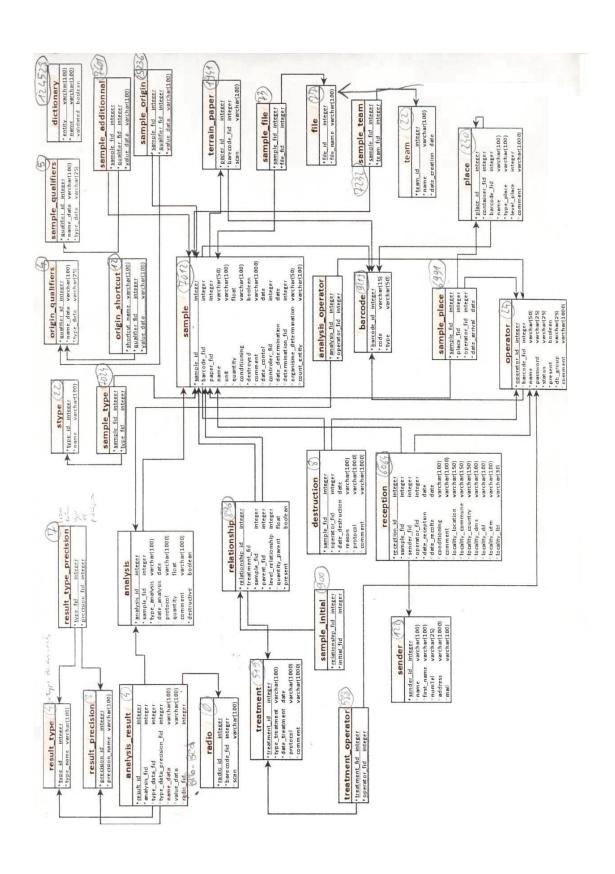


Figure 8 : Architecture de la base de données

3.2.2 L'interface graphique

Les utilisateurs se connectent à la base "SAMPLES" via une interface WEB divisée en deux parties. Les différents niveaux d'accès aux données dépendent des droits (rôles) de la personne connectée. Il existe des fonctionnalités qui sont à la portée de toute personne connectée et d'autres qui ne sont disponibles que pour les administrateurs. On trouve donc :

- Une interface 'Opérateur' : réservée aux utilisateurs possédant un login et un mot de passe qui permet la gestion courante des échantillons
- Une interface 'Administrateur' : réservée au super-utilisateur (ou administrateur), elle permet de créer de nouvelles données (lieux de stockage, opérateurs, projets de recherche, champs qualifiant les échantillons) et de recharger des données mises à jour dans les dictionnaires.

Pour s'authentifier, les utilisateurs utilisent cette page :

GESTION DES ECHANTILLONS						
Accueil	Feuille de terrain	Echantillon	Lieux de stockage	Traitement	Analyse	
Quoi faire Rôle des M		Connexion à Identifiant *: Mot de passe *				
			© Copyright 20	009 INRA Design	n by: stylesh	out Valid XHTML CSS

Figure 9: Interface d'authentification

L'interface est codée en HTML (Hypertext Markup Language), en PHP (Hypertext Preprocessor) et en JavaScript. Chaque page est structurée en 5 parties (cf figure 10 et 11) :

- L'entête où se situe le nom de la base et quelques images
- Le **menu horizontal** qui est constitué de différents onglets qui diffèrent en fonction des deux parties :
- Partie d'administrateur avec 6 onglets :
 - * <u>Opérateurs</u> : permet de créer, de consulter, de modifier et de supprimer un opérateur
 - * <u>Projet</u> : permet d'ajouter, de modifier, de consulter et de supprimer un projet
 - * <u>Lieux de stockage</u> : permet de gérer les lieux de stockage
 - * <u>Identification</u>: permet d'ajouter ou de supprimer des niveaux d'identification taxonomique, de créer et gérer les raccourcis d'identification
 - * <u>Infos supplémentaires</u> : permet de créer et de supprimer des champs de caractérisation d'échantillon
 - * <u>Dictionnaire</u> : réservé à la mise à jour des éléments des listes déroulantes mises à la disposition des opérateurs
- Partie opérateur avec 6 onglets :
 - * Accueil : Présentation des fonctionnalités de la base de données
 - * <u>Feuille de terrain</u>: permet la création, la modification, la consultation et la suppression d'une feuille de terrain
 - * <u>Echantillon</u>: permet la consultation, la création, la modification le changement de lieu de stockage des échantillons. Dans cette partie, on trouve également la liaison d'un fichier à un échantillon, l'importation de plusieurs échantillons à partir d'un fichier Excel et la consultation de la liste des expéditeurs
 - * <u>Lieux de stockage</u> : permet de voir la liste de tous les lieux de stockage et aussi d'obtenir la liste de tous les échantillons qui se trouvent dans un lieu de stockage donné
 - * <u>Traitement</u> : permet de générer un traitement sur un (ou plusieurs) échantillon(s)
 - * <u>Analyse</u>: permet de compléter et de consulter les résultats d'analyse sur un (ou plusieurs) échantillon(s)
 - Le **menu vertical** : à ce niveau sont listées toutes les actions possibles dans chaque onglet du menu horizontal. Exemple : consulter, modifier, supprimer, créer ...
 - Le **corps de la page** : emplacement où se trouvent les différents formulaires et les pages de sortie des requêtes. Cette partie est donc amenée à changer d'une page à l'autre
 - Le **pied de la page** : emplacement qui se trouve en bas de toutes les pages (il ne change pas d'une page à l'autre).

GESTION DES ECHANTILLONS L'entête					
Accueil Feuille de terrain Echantillon	Lieux de stockage Traitement Analyse				
	Menu horizontal				
Quoi faire ? © Rôle des Menus	Connexion à l'interface				
Menu Vertical	Identifiant * : • Mot de passe * : Envoyer				
© Copyright 2009 INRA Design by: styleshout Valid XHTML CSS					

Figure 10 : Structure d'une page de la partie opérateur

Opérateurs	Projet	Lieux de stockage	Identification	Infos supplémentaires	Dictionnaire	Déconnexion

Menu horizontal

Figure 11 : Menu horizontal de la partie administrateur

3.2.3 Traitement d'un échantillon au niveau de la base de données

La base « SAMPLES » doit permettre aux utilisateurs de créer un échantillon qui possède un numéro d'identification unique (son code-barres), d'enregistrer toutes les manipulations effectuées et d'assurer la traçabilité. La base « SAMPLES » permet après création de l'échantillon de :

- ❖ Consulter (cf figure 12),
- * modifier,
- détruire,
- stocker,
- consulter les expéditeurs de colis,
- * attacher des fichiers,
- * réaliser des traitements,
- lier des résultats d'analyse.



Figure 12 : Affichage pour la consultation d'un échantillon

3.3 Les technologies utilisées

Les différents outils mis à ma disposition pour mener à bien mon projet sont de différents ordres :

• Moyens techniques

Matériel physique:

Ordinateur:

TOSHIBA Satellite INTEL® CORETM CPU 1.90 GHz, 4 Go de RAM Système d'exploitation : Windows 7 professionnel

Serveur de développement :

La base de données « Samples » est hébergée sur le serveur « DUBA » dont les caractéristiques sont les suivantes :

- Marque : DELL

- Modèle : PowerEdge 2950

- Processeur : Quad-CoreIntel® Xeon® E5420 2.5GHz

- Capacité : 2 disques Hot-Plug 750 Go, 4 disques SAS 146 Go et 8 RAM

- Système d'exploitation : Debian Linux

Langages de programmation

L'interface de la base « Samples » est codée en :

- HTML (*HypertextMarkupLanguage*) : format de données conçu pour représenter les pages Web. C'est un langage de balisage. HTML permet également de structurer sémantiquement et de mettre en forme le contenu des pages.
- PHP version 5.2.0 (*HypertextPreprocessor*): langage de programmation côté serveur, utilisé principalement pour produire des pages web dynamiques.
- JavaScript : langage de programmation de scripts côté client. Il permet de dynamiser une page web et de faciliter son ergonomie.
- AJAX (*Asynchronous JavaScript And XML*): ce terme évoque l'utilisation conjointe d'un ensemble de technologies libres couramment utilisées sur le Web (cf figure 13):

HTML (ou XHTML) pour la structure sémantique des informations

CSS (Cascading Style Sheets) pour la présentation des informations

DOM (*Document Object Model*) et JavaScript pour afficher et interagir dynamiquement avec l'information présentée

objetXMLHttpRequest pour échanger et manipuler les données de manière asynchrone avec le serveur Web

XML (eXtensibleMarkupLanguage) qui est un langage de balisage générique.

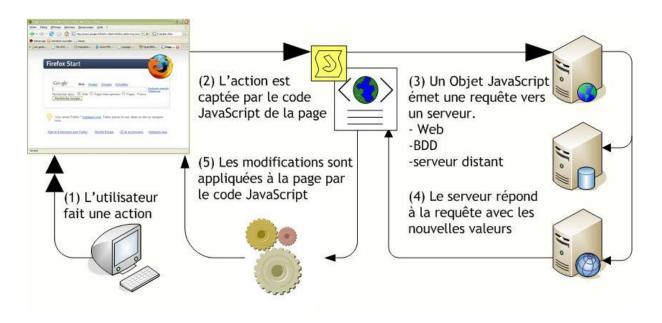


Figure 13 : Schéma de fonctionnement de l'Ajax

Base de données :

PostGreSQL version 8.3 : SGBDR (Système de Gestion de Base de Données Relationnelles) utilisé pour la base de données "SAMPLES".

Logiciels et utilitaires :

pgAdmin III version 1.10.5 : interface graphique pour administrer la base données sous PostGreSQL.

Navigateur Mozilla Firefox : navigateur Web libre qui offre la possibilité d'ajouter des extensions et des thèmes graphiques.

XAMPP: (**X** Apache MySQL Perl PHP): ensemble de logiciels permettant de mettre en place facilement un serveur web. Il s'agit d'une distribution gratuite offrant une bonne souplesse d'utilisation, réputée pour son installation simple et rapide.

NotePad++ version 6.6.6 : éditeur de code source, il a été utilisé pour éditer les fichiers CSS, HTML, JavaScript et PHP.

3.4 Le travail réalisé

3.4.1 Travail réalisé au niveau de la base

Le travail demandé était de corriger des anomalies et de nettoyer les informations erronées entrées par erreurs par les utilisateurs. Il s'agissait par exemple de fautes d'orthographe (*Scrophulia* au lieu de *Scrofularia* pour le genre végétal hôte de l'échantillon). Pour corriger ce genre d'anomalies, j'ai procédé à une mise à jour des tables concernées par la commande SQL 'update table'.

Il m'a été aussi demandé d'éliminer certaines données entrées dans la base qui n'avaient pas d'utilité dans les listes déroulantes et qui avaient été introduites par erreur. Il s'agissait par exemple de valeurs numériques pour le champ Hôte animal. La commande SQL *delete* a été utilisée pour réaliser cette opération.

Ces modifications ont été réalisées via pgAdmin III car l'interface graphique actuelle ne permet pas de réaliser ce genre d'opérations.

❖Eliminer définitivement un échantillon entré par erreur (en mode administrateur)

Un échantillon peut-être détruit physiquement, son enregistrement dans la base quant à lui ne pouvait pas être détruit totalement même pour des échantillons entrés par erreur. On m'a donc demandé de pouvoir supprimer définitivement un enregistrement entré par erreur en mode administrateur (voir figure 14).



Figure 14: Elimination définitive d'un échantillon

Pour éliminer un échantillon dans la base, j'ai commencé par sélectionner son code dans toutes les tables où il est présent. Après sélection, je procédais à sa suppression dans toutes les tables en finissant par sa suppression dans la table « barcode » (voir Annexe).

* Restorer des échantillons détruits par erreur

Il y a aussi eu des cas où des échantillons ont été déclaré détruits par erreur par les opérateurs dans la base de données. Ils apparaissaient comme détruits alors que physiquement ils existaient et que les travaux de recherche pouvaient continuer sur ces échantillons. Il m'a donc été demandé de pouvoir corriger cette erreur de saisie.

3.4.2 Travail réalisé au niveau de l'interface

Au niveau de l'interface graphique, les utilisateurs souhaitaient de nouvelles fonctionnalités. La liste des points à traiter a déjà été énumérée au point 2.2 (Objectifs du stage).

* Réaliser des requêtes avec de nouveaux critères

Les utilisateurs de la base pouvaient interroger celle-ci, mais certaines requêtes prenaient beaucoup de temps pour aboutir. De plus, ils n'avaient pas toujours le niveau d'interrogation qui leur convenait. Les interrogations étaient soit trop globales, soit trop ciblées. Ils avaient donc besoin de requêtes leurs fournissant des informations au niveau intermédiaire. Ainsi, ils souhaitaient pouvoir effectuer des recherches ciblées sur un département français et sur une province pour les autres pays, et pas seulement des recherches par pays ou par commune.

J'ai implémenté de nouveaux champs de saisie pour optimiser la consultation des échantillons. J'ai ajouté sur l'interface un champ de saisie Département et Province pour les pays autres que la France.

Pour faciliter l'ergonomie de saisie, la consultation via le champ Département ou Province est proposée aux utilisateurs en fonction de leur choix de pays (cf figure 15). J'ai utilisé une fonction JavaScript permettant aux utilisateurs consultant les données pour la France de voir un département, une commune et le lieu- dit. Pour les autres pays, les utilisateurs ont cette fois-ci accès à la province et à la commune.

Pays:	FRANCE	Pays:	CHINE
Département :		_	1
Commune :		Commune :	choisissez la commune ▼
Lieu dit :	choisissez le lieu dit ▼	Province :	choisissez la province ▼

Figure 15 : Affichage obtenu suite au choix du pays

Dans la structure de la base de données, il n'y avait pas de champ spécifique correspondant à un département mais on disposait comme information de la concaténation entre le nom d'une commune et le numéro de son département. Pour pouvoir obtenir toutes les communes appartenant à un département, il a fallu faire une requête permettant de sélectionner toutes les communes ayant en commun le numéro saisi.

Par exemple si l'utilisateur saisit '45' dans la zone de saisie du département, les communes Orléans - 45, Olivet - 45, ainsi que toutes les communes du Département 45 pourront être sélectionnées.

Pour récupérer ces communes, j'ai utilisé la requête SQL suivante :

SELECT commune FROM ... WHERE locality_commune LIKE \$%numero_departement.

Le paramètre LIKE de cette requête SQL permet de récupérer toutes les lignes qui contiennent le numéro associé à la commune. L'utilisation du symbole % signifie que seules les communes se terminant par le numéro saisi sont concernées.

Certains caractères spéciaux ont une signification particulière dans les requêtes SQL sous PostGreSQL. C'est le cas de l'apostrophe (simple quote) qui met fin la ligne d'instruction comme on le voit dans l'exemple suivant :

SELECT * FROM data.reception WHERE locality_commune = 'SAINT-MARTIN-D'ABBAT';

Lors de l'exécution, cette requête sera interprétée comme suit :

SELECT * FROM data.reception WHERE locality commune = 'SAINT-MARTIN-D'.

La partie 'ABBAT qui se trouve après l'apostrophe ne sera pas considérée parce que la première apostrophe a mis fin à l'instruction et l'erreur suivante est retournée :

ERREUR: erreur de syntaxe sur ou près de « ABBAT »

LINE 2: ...eception WHERE locality_commune = 'SAINT-MARTIN-D'ABBAT - 45...

Pour remédier à ce problème, il a fallu utiliser la fonction PHP : str_replace. Cette fonction remplace toutes les occurrences dans une chaîne donnée par une valeur voulue. En effet, pour que cette requête puisse aboutir, il suffit d'échapper les apostrophes se trouvant en son milieu. Un antislash (\) a été utilisé pour échapper à une apostrophe.

\$commune = str replace("'", "\'", \$commune tmp);

En effectuant mes tests après l'utilisation de cette fonction, je me suis aperçu que certaines requêtes n'aboutissaient pas avec des erreurs sur les communes, comme par exemple pour le département 12 où j'ai eu comme message en retour :

ERREUR !ERREUR: erreur de syntaxe sur ou près de « OLT » LINE 1: ...=sample_id AND locality_commune='SAINT-GENIEZ-D\\'OLT - 12' ... ^

Apres vérification sur toutes les communes de France, je me suis aperçu que certaines communes étaient enregistrées avec un antislash(\). Pour corriger cette erreur, j'ai fait appel à la commande update de SQL.

❖ Donner la possibilité de créer un nouvel expéditeur

Pour importer plusieurs échantillons en même temps, les opérateurs ont la possibilité d'utiliser un fichier Excel. Par contre, si des échantillons avaient été envoyés par un expéditeur qui n'existait pas encore dans la base, le système d'importation ne fonctionnait pas. De plus, via l'interface Web, il n'était pas possible de créer un expéditeur nouveau sans création d'échantillon. Pour pallier à ce problème, l'administrateur avait créé des expéditeurs sans nom directement dans la table « sender » par l'intermédiaire de pgAdmin III. Ensuite, les opérateurs pouvaient attribuer des noms en fonction de nouveaux expéditeurs et ainsi importer les fichiers Excel.

Dans la partie « Expéditeur colis », j'ai ajouté un formulaire de saisie pour un nouvel expéditeur (cf figure 16), qui évite la saisie directe dans la base de données par l'intermédiaire de pgAdmin III. Une fonction JavaScript a été utilisée pour vérifier si le champ *nom* est rempli avec plus d'une lettre. Cette fonction empêche aux utilisateurs d'entrer dans la base des expéditeurs sans nom, ou des espaces vides dans la table « *sender* » en appuyant par erreur sur le bouton « Enregistrer ».

Nouvel Expéditeur :	
Nom * :	
Prénom :	
Téléphone :	
Email:	
Adresse:	
Enregistrer Annuler	

Figure 16: Enregistrement d'un nouvel expéditeur

❖ Déplacer un ensemble d'échantillons d'un lieu de stockage à un autre en mode utilisateur

Les utilisateurs de l'interface voulaient pouvoir transférer en bloc l'ensemble des échantillons d'un lieu de stockage à un autre en mode utilisateur. Cette possibilité existait déjà en mode administrateur. Pour cela, j'ai ajouté dans le menu vertical « Déplacer contenu d'un lieu à un autre » dans l'onglet « Lieux de stockage » et j'ai réutilisé le code du menu administrateur car l'opération à effectuer était exactement la même.

❖ Eliminer la possibilité de lier une radio comme résultat d'analyse (redondant avec l'option d'ajout de fichier déjà existante)

Dans le menu de l'onglet Analyse, les utilisateurs avaient 2 fonctions redondantes : ajout (et modification) d'un fichier radio à un résultat d'analyse et possibilité de lier un fichier à un résultat. Pour éliminer les options inutiles de l'interface graphique, j'ai mis toute la partie concernée en commentaire dans le code source. En effet, je n'ai pas voulu effacer les scripts car dans le futur, les utilisateurs auront peut-être à nouveau besoin de ces options.

❖ Afficher un message de confirmation suite à une opération sur la base

Après certaines opérations (création, modification, enregistrement d'un nouvel opérateur, etc...), les utilisateurs se sont aperçus qu'il n'y avait pas toujours affichage d'un message de confirmation comme cela était prévu initialement dans la base. Ils souhaitaient obtenir cet affichage pour vérifier si l'opération avait réellement été réalisée ou si elle avait échoué.

J'ai analysé le code pour les pages en question et j'ai bien constaté la présence des messages de confirmation. En regardant de plus près, je me suis aperçu que le programmeur avait oublié de fermer la balise « span ». J'ai fait la correction et, actuellement, après chaque opération, les messages de confirmation s'affichent correctement (cf figure 17).



Figure 17 : Exemple de message de confirmation

Diverses tâches réalisées sur l'interface

Sur l'interface existante, il y avait de nombreuses incohérences à l'affichage des résultats comme les compteurs de nombre d'entité qui n'étaient pas corrects, les noms des espèces qui s'affichaient avec une majuscule au début après exportation, des exportations de listes qui ne respectaient pas l'ordre de saisie par les opérateurs, l'ordre d'affichage de champs ne respectant pas les règles taxonomiques (genre, espèce végétal).

Pour tous ces cas, j'ai effectué des modifications dans les pages codées pour remédier à ces problèmes.

CONCLUSION

Le stage effectué au sein de l'INRA a été riche d'enseignements pour moi sur différents aspects. Tout d'abord, il m'a permis de comprendre le mode de fonctionnement d'un laboratoire de recherche scientifique. De plus, le fait de travailler avec les biologistes m'a ouvert les yeux sur l'utilité de l'informatique au service des autres. On répond aux besoins informatiques des utilisateurs afin qu'ils puissent eux aussi mener leurs travaux de recherche dans de bonnes conditions. Le travail réalisé est donc indispensable au bon fonctionnement du laboratoire.

En ce qui concerne mon stage, la première difficulté que j'ai rencontrée concernait le code utilisé qui était dans un langage totalement inconnu pour moi (Ajax). La seconde difficulté était que je repartais d'un projet déjà existant. Cela m'a donc pris beaucoup de temps et d'investissement pour comprendre la logique du premier développeur puis, ensuite, faire des modifications sans toutefois impacter le fonctionnement existant. Heureusement, le premier développeur avait mis un grand nombre de commentaires dans son code ce qui m'a facilité la compréhension du mode de fonctionnement de la base de données.

Ce stage m'a également permis de développer mon sens de l'autonomie, de la réflexion et surtout m'a fait comprendre l'utilité des commentaires dans un code informatique. J'ai aussi pu améliorer mes connaissances en création d'application Web, et notamment en ce qui concerne le respect strict des standards du Web et l'utilisation des technologies comme Ajax, JavaScript et PHP.

En plus de mon projet de stage, grâce à M^r Régis Phélut et M^r Jean-Léandre Haton j'ai eu des explications sur l'architecture et la gestion du réseau informatique de l'INRA. Cela m'a permis de voir en pratique ce que j'avais appris lors des cours du DUT sur les réseaux. Ainsi, j'ai ainsi pu visiter aussi la salle de serveurs.

Au niveau personnel, ce stage m'a fait découvrir le milieu de la fonction publique. Jusqu'à présent, j'avais toujours travaillé dans le secteur privé et avec une énorme pression. J'ai été extrêmement étonné par le mode de fonctionnement dans le laboratoire au début de mon stage. Mais, au fur et à mesure, j'ai découvert une autre façon de travailler où la liberté qui m'était laissée, m'a poussé à me surpasser pour atteindre les objectifs de mon stage.

BIBLIOGRAPHIE

JavaScript pour les nuls - Emily A. Vander Veer.- ed. Pierre Chauvot - 2005

PHP 5.4 Développez un site web dynamique et interactif – Heurtel O.- ed. Eni Editions - 2012

PHP 5 avancé - DaspetE. et Geyer C.P. - ed. Eyrolles - 2004

PostGreSQL par la pratique - WORSELEY J. C. et DRAKE J. D. - traduit par Jacoboni E. - ed. O'Reilly - 2002

WEBOGRAPHIE

Sites généraux :

fr.wikipedia.org/wiki/ php.net/manual

Présentation de l'institut et de l'unité de Recherche :

www.inra.fr/ www.orleans.inra.fr/ intranet.orleans.inra.fr/organisation_scientifique

Tutoriels sur HTML, CSS, PHP:

www.siteduzero.com/ php.developpez.com/

Localisation géographique :

mapsengine.google.com

Annexe

```
while($row = pg_fetch_array($result_code)){
          $barcode_fid = $row['barcode_id'] ;
         $select from place =0 pg query($dbh, "SELECT barcode fid from place where barcode fid = '$barcode fid'");
         if(pg_num_rows($select_from_place) > 0 )
            $delete from place =0 pg query($dbh, "delete from place where barcode fid='$barcode fid'");
         $select_from_sample =@ pg_query($dbh, "SELECT barcode_fid from sample where barcode_fid = '$barcode_fid'");
         if(pg_num_rows($select_from_sample) > 0 )
           $delete_from_sample =0 pg_query($dbh, "delete from sample where barcode_fid='$barcode_fid'");
         $select_from_operator =0 pg_query($dbh, "SELECT barcode_fid from operator where barcode_fid = '$barcode_fid'");
         if(pg num rows($select from operator) > 0 )
           $delete_from_operator =0 pg_query($dbh, "delete from operator where barcode_fid='$barcode_fid'");
        $select_from_terrain_paper =0 pg_query($dbh, "SELECT barcode_fid from terrain_paper where barcode_fid = '$barcode_fid'
         if(pg_num_rows($select_from_terrain_paper) > 0 )
            $delete_from_terrain_paper =0 pg_query($dbh, "delete from terrain_paper where barcode_fid='$barcode fid'");
         $delete_from_barcode =0 pg_query($dbh, "delete from barcode where code=$codebare_echt");
         if($delete_from_barcode)
            echo "Le code bare". $codebare echt. " a ete supprimé ";
         else
            echo "Erreur lors de la suppression du codebare";
pg_close($dbh);
```

Annexe1. Extrait du code d'élimination d'un échantillon