

PALEOPYR : un système d'information pour les données paléoenvironnementales nord-pyrénéennes

Emilie Lerigoleur, Florence Mazier, Laurent Jégou, Marie Perret, Didier Galop

▶ **To cite this version:**

Emilie Lerigoleur, Florence Mazier, Laurent Jégou, Marie Perret, Didier Galop. PALEOPYR : un système d'information pour les données paléoenvironnementales nord-pyrénéennes. Revue des Sciences et Technologies de l'Information - Série ISI : Ingénierie des Systèmes d'Information, 2015, 20 (3), pp.63-87. 10.3166/ISI.20.3.63-87 . hal-01659525

HAL Id: hal-01659525

https://univ-tlse2.hal.science/hal-01659525

Submitted on 8 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PALEOPYR : un système d'information pour les données paléoenvironnementales nord-pyrénéennes

Emilie Lerigoleur[1], Florence Mazier[1], Laurent Jégou[2], Marie Perret[1], Didier Galop[1]

1. GEODE UMR 5602 CNRS, Université Jean Jaurès,
5 allée Antonio Machado 31058 Toulouse Cedex 9, France
emilie.lerigoleur@univ-tlse2.fr

2. Université de Toulouse-2 Jean Jaurès Département Géographie
UMR Laboratoire Interdisciplinaire Solidarités, Sociétés, Territoires (LISST)
5 allée Antonio Machado 31058 Toulouse Cedex 9, France

RESUME. Les données paléoenvironnementales collectées sur le versant nord des Pyrénées font l'objet d'un archivage au sein d'un système d'information pour les données PALEOenvironnementales PYRénéennes (PALEOPYR) et visent à reconstruire les dynamiques spatio-temporelles de la végétation durant l'Holocène. Trente années de recherche ont permis de produire de grandes quantités de données (pollen, microfossiles non-polliniques, charbons, diatomées…). Ce travail illustre les grandes étapes méthodologiques de conception d'un système d'information et propose des recommandations d'optimisation du serveur PostgreSQL/PostGIS. La nouvelle base de données archive les données de 90 enregistrements sédimentaires ainsi que les métadonnées associées. La création d'interfaces graphiques et cartographiques sur Internet facilite la visualisation et l'exploitation de ces données. PALEOPYR représente ainsi une avancée méthodologique fondamentale pour la recherche paléoécologique à l'échelle des Pyrénées.

ABSTRACT. This paper i) reviews the data on Holocene paleoenvironmental sequences collected on the northern slopes of the Pyrenees, ii) attempts to store them in the PALEOenvironmental information system for PYRenees data (PALEOPYR) and iii) outlines a possible use for synthesis on spatio-temporal dynamics of the vegetation. Studies on paleoenvironmental stratigraphies over the last thirty years have yielded a very large amount of data (including pollen, non-pollen palynomorphs, charcoal, diatoms…). This paper aims at describing the major methodological steps of designing the database and provides recommendations for optimizing the PostgreSQL/PostGIS server. Data from 90 sequences and their relative metadata are stored in the database in relational tables. This database will allow a broad range of queries to be addressed in order to visualize and explore the data with both graphical user and web mapping interfaces. PALEOPYR will help the scientific community to study past landscape dynamics at the Pyrenean scale.

MOTS-CLES : données paléoenvironnementales, pollen, base de données, cartographie interactive, évolution de la végétation, montagnes, Pyrénées

KEYWORDS: paleoenvironmental data, pollen, database, web mapping, vegetation evolution, mountains, Pyrenees

DOI:10.3166/ISI.20.3.63-87

1. Introduction

Dans le champ des recherches paléoenvironnementales, les données sont très diversifiées. Elles sont issues de l'analyse d'indicateurs biotiques ou abiotiques fossiles, conservés dans des enregistrements sédimentaires terrestres (lacs, tourbières, paléochenaux, sols et sédiments archéologiques) ou marins. L'analyse des grains de pollen, des microfossiles non-polliniques (MNP : restes fongiques, algues…), de micro- et macro-charbons et d'autres indicateurs (diatomées, stomates, macrorestes végétaux, insectes…) s'ajoutent à la production de mesures sédimentologiques, biogéochimiques voire moléculaires en constituant le socle d'une démarche multi-indicateurs. Ainsi la confrontation de nombreux indicateurs permet d'affiner la connaissance et la reconstitution des environnements passés (paysages, climat). A ces données fossiles s'additionnent de grandes quantités d'informations actuelles ou subactuelles utilisées en tant que référentiels contemporains pour calibrer, par exemple, la relation pollen/végétation (Mazier *et al.*, 2006), construire des fonctions de transfert (Heiri *et al.*, 2001 ; Velle *et al.*, 2012), et ainsi améliorer les reconstitutions quantitatives de différents paramètres, qu'ils soient climatiques (Peyron *et al.*, 2000) ou en lien avec les activités humaines passées (Mazier *et al.*, 2009).

Si les études paléoenvironnementales avaient jusqu'ici une portée scientifique principalement locale à l'échelle du site d'analyse, l'enjeu actuel est de constituer et de gérer des jeux de données pour répondre à des questions scientifiques plus globales à l'échelle d'une région, d'un massif, d'un continent. Ainsi la gestion des données nécessite le développement de systèmes d'information pour organiser et structurer ces données, mais également pour les consulter et les visualiser afin d'en comprendre les dynamiques spatio-temporelles à large échelle. Depuis plus de 20 ans, plusieurs systèmes d'information ont été proposés dont la plupart possède un portail Web et un accès aux métadonnées et données. L'*European Pollen Database*[1] (EPD), créée à la fin des années 1990, permet l'archivage et la consultation des données polliniques fossiles provenant de 877 sites de la partie occidentale du continent eurasiatique (Fyfe *et al.*, 2009). Depuis 2007, le *Fossil Pollen Database Viewer*[2] permet la visualisation des données de l'EPD sous la forme de courbes, de

1 http://www.europeanpollendatabase.net
2 http://www.europeanpollendatabase.net/fpd-epd

cartes dynamiques et de tableaux. En complément de l'EPD, Davis *et al.* (2013) ont développé l'*European Modern Pollen Database* pour gérer plus de 5 000 échantillons polliniques modernes. Enfin, la *Neotoma Paleoecology Database*[3], en cours de développement, recensera l'ensemble des bases de données paléobotaniques, animales, régionales et nationales existantes couvrant les cinq derniers millions d'années, du Pliocène à l'Holocène (Grimm, 2008 ; Grimm *et al.*, 2013 ; Goring *et al.*, 2015). *Neotoma Explorer*[4] est l'explorateur cartographique en ligne permettant d'interroger cette base, de visualiser et d'exporter les jeux de données selon différents formats (*e.g.* cartes, graphes, exports). Un package sous le logiciel de statistiques R a été développé dans le but de faciliter les transferts de données de la base *Neotoma* vers l'environnement de R (Goring *et al.*, 2015).

Sur le versant nord-pyrénéen, la grande quantité d'informations produites par l'étude de plus de deux cents enregistrements sédimentaires au cours des trente dernières années constitue une précieuse source d'informations, disparate et hétérogène, d'accès difficile, sans gestion centralisée ni traçabilité des métadonnées. Ainsi la construction d'un système d'information robuste et sécurisé s'imposait. Ce système vise à centraliser, homogénéiser, sauvegarder et diffuser les données paléoenvironnementales et les métadonnées associées gérées par l'unité de recherche dans une base de données optimisée, performante et évolutive. Le défi est de mettre à disposition des producteurs de données un outil de travail facilitant l'accès aux informations et proposant une interface cartographique pour la visualisation et l'exploitation des données en fonction de leurs questions de recherche.

Une première base de données sous MS Access® a été créée au milieu des années 2000, sous l'impulsion d'un projet de recherche à l'initiation du système d'information des données PALEO-environnementales PYRénéennes (PALEOPYR) (Galop *et al.*, 2005 ; 2006 ; Le Flao, 2004 ; 2005). Néanmoins, en 2005, PALEOPYR ne recensait pas l'exhaustivité des données pyrénéennes et sa structure n'était pas suffisamment flexible pour intégrer la diversité des données et des indicateurs. De plus, la consultation des données (extraction de tableaux), leur visualisation et la génération de cartes dynamiques (suivi de l'évolution du couvert végétal sur une fenêtre de temps identifiée) par une interface fonctionnelle n'avaient pas été envisagées. Un effort pour informatiser les données manuscrites et harmoniser les jeux de données en vue de leur future intégration dans une base de données adaptée a été récemment réalisé (Lerigoleur, 2013). Ce travail portait également sur une analyse sémantique des données ainsi qu'un premier essai de modélisation conceptuelle et relationnelle des données (formalisation de la structure et de la signification des informations), offrant ainsi l'expérience et le recul nécessaires à la construction d'un nouveau système d'information et des jeux de données prêts à être intégrés.

3 http://www.neotomadb.org
4 http://apps.neotomadb.org/explorer/

Nous présentons ici le développement de la construction du système d'information PALEOPYR. Dans un premier temps, nous dresserons un état des lieux des données disponibles, puis nous préciserons la méthodologie d'analyse et de conception du système. Nous évoquerons et illustrerons ensuite la construction d'interfaces cartographiques interactives pour la visualisation et l'interrogation des données et des métadonnées. Nous conclurons enfin sur les avancées technologiques acquises et la poursuite des actions pour répondre aux futurs besoins techniques et scientifiques.

2. Matériel et méthodes

2.1. Données paléoenvironnementales et métadonnées

Durant les trente années de recueil des données paléoenvironnementales sur le massif nord-pyrénéen, plus de deux cents enregistrements sédimentaires (lacs, tourbières, sols et mousses) fossiles et modernes ont été étudiés. Dans le cadre du système d'information PALEOPYR, les phases préalables de saisie des données manuscrites et de centralisation des jeux de données ont permis d'intégrer, pour le moment, 73 sites et 90 enregistrements sédimentaires (16 enregistrements lacustres et 74 tourbeux, Fig. 1).

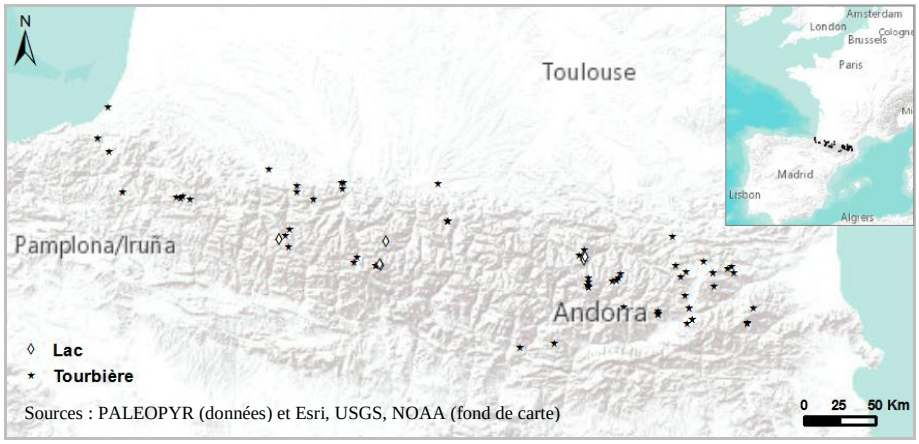


Figure 1 : Localisation des 90 enregistrements sédimentaires intégrés dans le système d'information PALEOPYR (74 issus de tourbières en noir, 16 de lacs en blanc). Le fond de carte est fourni par Esri®.

Pour bien gérer les données paléoenvironnementales, il faut en comprendre les étapes d'acquisition, de préparation, de stockage et d'analyse. Les enregistrements sédimentaires sont au cœur du système. Chacun d'eux est associé à des métadonnées (norme ISO Standard 15836 - *Dublin Core Metadata Element Set* Version 1.1[5] et métadonnées spécifiques aux données spatiales) renseignant notamment l'auteur, les coordonnées géographiques du site, le type et la taille du bassin de sédimentation, la liste des indicateurs disponibles (pollen, MNP, lithologie, radiocarbone…), les photographies et les publications associées.

Chaque enregistrement sédimentaire, défini par sa profondeur, est séquencé en lots d'échantillons d'épaisseur variable. Chaque échantillon peut faire l'objet d'analyses d'indicateurs différents et complémentaires :

– analyses sédimentologiques (lithologie, susceptibilité magnétique, colorimétrie, perte au feu) et géochimiques (organique, minéral, isotopique),

– analyses biologiques, polliniques ou anthracologiques : il s'agit de comptages (données brutes), de pourcentages, de concentrations et d'influx (données transformées) de grains de pollens, de MNP, de macrorestes, de micro- ou de macro-charbons, etc.,

– analyses des dates et de leurs incertitudes obtenues par dosage des radionucléides ou du radiocarbone ^{14}C.

Cela représente, pour les 90 enregistrements fossiles recensés, un total de 21 277 échantillons analysés, 2 243 données brutes multi-indicateurs, 87 401 données de comptage de grains de pollen, 68 684 données transformées, 249 références vers des documents annexes (*i.e.* photographies de sites et d'échantillons, publications). Bien que PALEOPYR ait vocation à archiver toutes les données multi-indicateurs, dans cet article nous illustrerons le cas des données polliniques et radiocarbones (Fig. 2).

L'analyse pollinique consiste à identifier morphologiquement (taxon pouvant être identifié au niveau de l'espèce, du genre ou de la famille) et à compter le nombre de grains de pollen contenus dans un échantillon d'une profondeur connue. L'identification s'appuie sur des clés de détermination, des atlas photographiques et des collections de référence. Pour chaque enregistrement sédimentaire, le diagramme pollinique représente graphiquement les taxons (exprimés en pourcentage, concentration ou influx) ordonnés par profondeur, ce diagramme représente la variation du taxon au cours du temps.

Ainsi pour donner une dimension chronologique aux données polliniques, des datations sont réalisées sur plusieurs échantillons d'un même enregistrement sédimentaire en utilisant différentes méthodes : i) le dosage du radiocarbone adapté aux 50 derniers millénaires et généralement réalisé sur sédiment total ou restes végétaux (âges radiocarbones *Before Present* – BP, l'âge de référence 0 étant 1950), ii) le dosage des radionucléides (isotope du plomb ^{210}Pb et du césium ^{137}Cs) permet de dater les échantillons pour les périodes plus récentes (150 dernières années). Ces âges sont ensuite calibrés et exprimés en cal BP en utilisant, dans le cadre de

5 Schéma de métadonnées générique (15 éléments) : http://dublincore.org/documents/dces/

PALEOPYR, le programme CLAM (Blaauw, 2010) et la courbe de calibration IntCal09 et IntCal13 (Reimer *et al.*, 2009 ; 2013). Un ou plusieurs modèles d'âges-profondeurs (courbes temps *vs* profondeur) sont générés permettant de relier les dates calibrées et de fournir l'âge de chaque échantillon analysé.

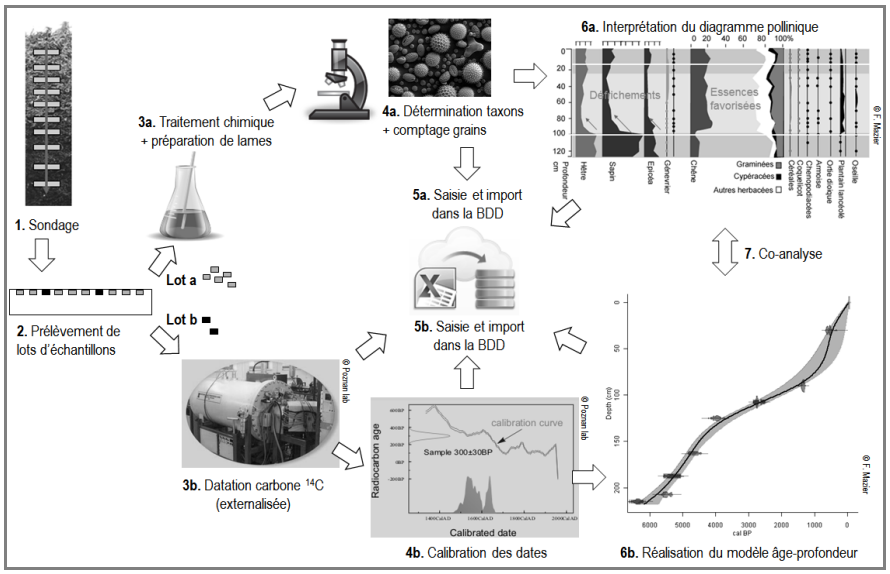


Figure 2 : Exemple de schéma illustrant : i) le prélèvement d'un enregistrement sédimentaire (étape 1), ii) la production de données polliniques (étapes 2a à 4a) et de datations par la méthode du radiocarbone ^{14}C (étapes 2b à 4b), iii) leur analyse conjointe permettant la génération d'un modèle d'âge-profondeur (étape 6b) puis d'un diagramme pollinique (étape 6a), iv) l'enregistrement des différents lots de données sous MS Excel® en vue de leur import futur dans la base de données (BDD)(étapes 5a et 5b).

Les données polliniques sont structurées selon le triplet *échantillon/taxon/nombre*, ce dernier renseignant des données brutes ou transformées. Par exemple, le triplet B122/Abies/34 correspond à 34 grains de pollen de sapin comptabilisés dans l'échantillon 122 de la carotte B. Cette structuration standardisée de l'information fut un élément majeur pour la modélisation conceptuelle et relationnelle future des données.

2.2. Méthode d'analyse et de conception du système d'information PALEOPYR

La démarche méthodologique (Fig. 3) adoptée se décline en trois grandes phases : l'analyse et la conception de système d'information, la mise en œuvre puis la mise en production du système.

La méthode reprend notamment les grands principes de la méthode d'analyse "Merise" la plus utilisée en France et dans une grande partie de l'Europe. Cette dernière se décline en quatre niveaux d'abstraction : conceptuel (quoi ? pourquoi ?), organisationnel (qui ? où ? quand ?), logique (comment ? abstrait) et physique (comment ? concret). Selon cette méthode, la conception d'un système d'information repose sur l'enchaînement des étapes suivantes appelé "cycle d'abstraction pour la conception des systèmes d'information" : système d'information manuel → expression des besoins → modèle conceptuel → modèle logique → modèle physique → système d'information automatisé (Libourel, 2010).

Dans cette partie nous préciserons les éléments méthodologiques clés conduisant à la création d'une base de données performante, apte à répondre aux besoins scientifiques et techniques.

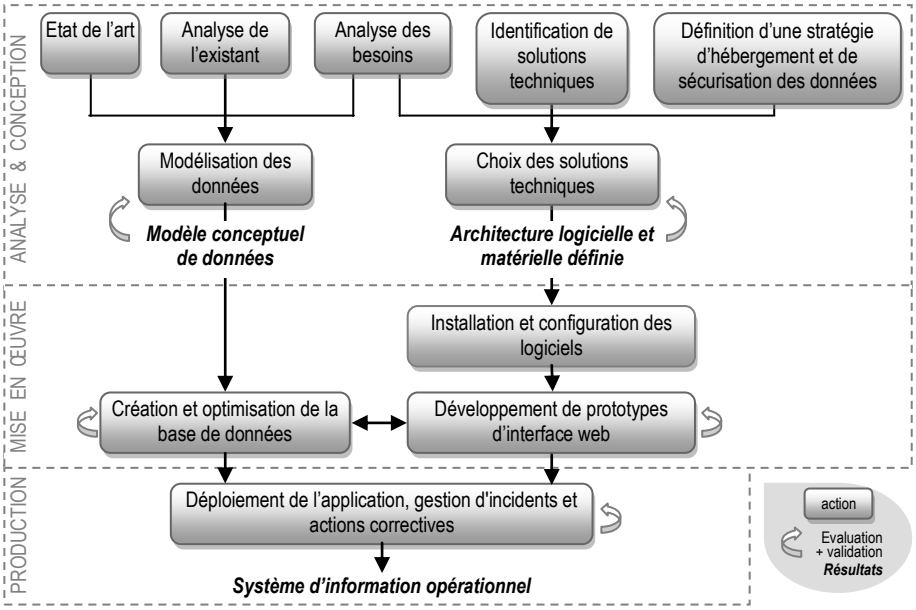


Figure 3. Démarche méthodologique pour la conception, la mise en œuvre et la mise en production du système d'information PALEOPYR.

2.2.1. Analyse de l'existant : données, flux et acteurs

Cette analyse a consisté à 1) dresser le bilan des échantillons et des données, 2) analyser la base de données embryonnaire créée en 2005 et les premiers essais de modélisation conceptuelle des données, 3) recenser les flux d'information à travers la description fine des processus d'acquisition, de préparation, de stockage et d'analyse des données, 4) clarifier le rôle des acteurs.

Afin de parvenir à inventorier finement l'existant, deux outils ont été utilisés pour le recensement des flux et des informations entre unités actives du système : le tableau des rubriques et le modèle organisationnel de flux.

L'exercice de constitution du tableau des rubriques a permis de compléter la liste des champs de l'ancienne base de données par de nouvelles rubriques.

La compréhension préalable des processus d'acquisition, de préparation, de stockage et d'analyse des données a conduit à la conception d'un modèle organisationnel de flux avec ses cinq unités actives internes et ses trois unités actives externes. La figure 4 illustre notamment les échanges d'informations ou d'échantillons entre le responsable de l'échantillonnage, le producteur de données, la personne en charge de la saisie des données dans la base PALEOPYR et l'administrateur du serveur de bases de données.

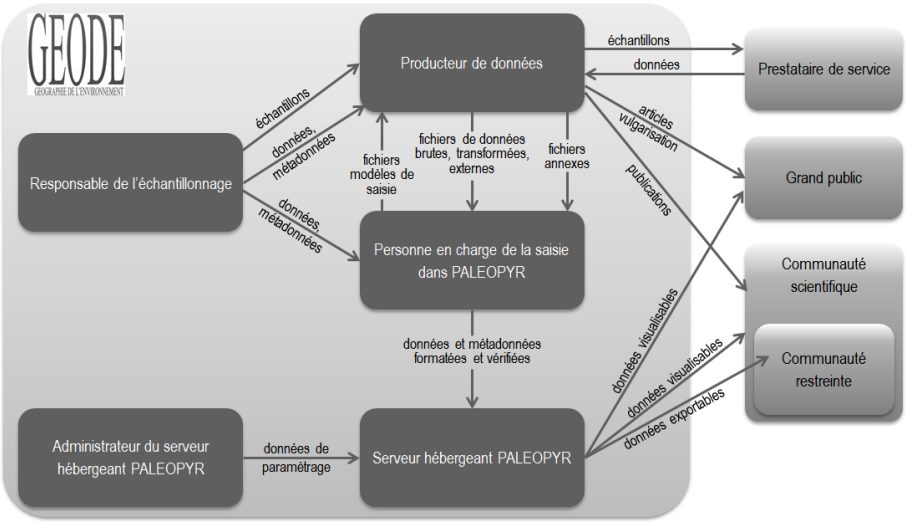


Figure 4 : Modèle organisationnel de flux pour le système d'information PALEOPYR. Cinq unités actives internes à l'unité GEODE s'échangent des données ou des échantillons et en échangent avec trois types d'unités actives externes.

Le producteur de données joue un rôle central puisqu'il produit des données brutes et transformées (*e.g.* pourcentages polliniques). Il peut commanditer des analyses complémentaires, considérées comme données externes (*e.g.* datations radiocarbone), à un prestataire de services. Enfin, il peut valoriser les résultats sous forme de publications scientifiques et de vulgarisations grand public.

La description des rôles des acteurs du système a permis de définir des profils d'utilisateurs du futur système d'information PALEOPYR, à savoir l'administrateur (tous les privilèges et droits de super-utilisateur), le producteur de données faisant partie de la communauté scientifique restreinte (droits d'ajouter, modifier, supprimer, exporter des données) et l'utilisateur grand public ou de la communauté scientifique (droit de consultation). Un système d'authentification a été mis en place afin de gérer les contraintes d'accès à PALEOPYR.

2.2.2. Modélisation conceptuelle des données

La mise en œuvre du système d'information PALEOPYR doit doter l'unité de recherche d'un Système de Gestion de Base de Données (SGBD). Celui-ci doit répondre aux besoins standards de maintenance de la base de données à l'aide de comptes-rendus automatisés (*e.g.* liste des tables, nombre d'enregistrements, etc.), fournir des fonctions pour interroger et calculer des données, mais également veiller à l'intégrité, la sécurité et l'indépendance des données (Sumathi et Esakkirajan, 2007). Il doit être performant, évolutif et permettre l'ajout de nouvelles variables sans en modifier la structure.

La modélisation a consisté en l'analyse sémantique de la signification de chaque descripteur du tableau des rubriques, puis à la schématisation des groupes de descripteurs (classes d'entités), en appliquant des règles strictes de cohérence. L'ajout des liens et des cardinalités (relations 0 à 1, 0 à *n*, etc.) entre les classes d'entités a permis d'aboutir à un Modèle Conceptuel de Données (MCD). Le modèle a été conçu pour permettre le moissonnage des métadonnées selon le protocole standard *Open Archives Initiative Protocol for Metadata Harvesting*[6] (OAI-PMH) par des catalogues de métadonnées.

2.2.3. Analyse des besoins et choix de solutions techniques adaptées

Sur la base du bilan de l'existant, identifiant de nouveaux besoins techniques, une grille des besoins a été établie notifiant les solutions techniques potentielles et les priorités d'actions.

Outre les besoins techniques liés au SGBD (*i.e.* centraliser, homogénéiser, sauvegarder, archiver, sécuriser, ajouter/modifier/supprimer/exporter, interroger les données), des besoins scientifiques ont été stipulés comme la compatibilité des données avec les outils d'analyse (*e.g.* module CLAM sous R pour la calibration des âges), la possibilité d'interroger les données pour répondre à diverses questions scientifiques, thématiques et cartographiques (*e.g.* Quand est apparu le taxon A le

6 http://www.openarchives.org/pmh/

long de la chaîne pyrénéenne ?, Quel est le pourcentage pollinique du groupe de taxons AB à la date T cal BP à l'échelle des Pyrénées ?, Quels sont les changements du couvert forestier en fonction de l'altitude à une date T cal BP ?).

Le choix du SGBD a été orienté vers PostgreSQL (logiciel open source, gratuit, documenté), qui offre en outre des potentialités d'analyse spatio-temporelle des données. Sa cartouche spatiale PostGIS, évolutive[7], apporte de puissantes fonctions pour le traitement des données spatiales, comme le ferait un logiciel SIG. PostgreSQL offre de nombreuses fonctionnalités comme les requêtes SQL (*Structured Query Language*) complexes pour interroger les données, la gestion de clés étrangères, de déclencheurs, de vues, etc.

Pour la création d'interfaces cartographiques, les langages et briques logicielles suivants ont été privilégiés : HTML pour les pages web de l'interface, des scripts PHP pour interroger la base de données ainsi que la bibliothèque de fonctions JavaScript GeoExt[8], cette dernière rassemble celles d'OpenLayers[9] et d'ExtJS[10] pour l'interface cartographique.

La sécurisation des données (intrusions extérieures et erreurs de saisie) a été identifiée comme un besoin majeur. La stratégie de l'utilisation d'un serveur de bases de données non ouvert sur l'extérieur, mais en relation avec un serveur Web distinct a ainsi été préférée.

La figure 5 illustre l'architecture matérielle et logicielle mise en œuvre, avec d'une part le cœur du système avec les serveurs PostgreSQL et Web (sous système Linux Debian) et d'autre part les interactions entre le serveur Web, la machine cliente et celle du producteur de données et les serveurs de *webservices* (selon les protocoles *Web Map Service* WMS ou *Web Feature Service* WFS de l'OGC[11]).

La vérification de la qualité des données saisies se déroule à différents niveaux : i) dès la saisie sous MS Excel, via la relecture par les personnes impliquées dans la production des données puis par la personne en charge de l'importation des données dans la base de données, ii) par le SGBD au moment de l'importation des données via la gestion des contraintes appliquées à chaque champ de la base de données. Dans ce dernier cas il s'agit d'une vérification de format et du maintien de l'intégrité référentielle des données (*i.e.* s'assurer des correspondances d'identifiants entre les tables).

Identifiée comme tâche prioritaire à réaliser dans la grille des besoins, une procédure automatisée de sauvegarde des données de la base PALEOPYR a été définie. Le planificateur de tâche périodique CRON a été privilégié pour réaliser la copie de la base (*backup*) vers le répertoire cible du serveur de sauvegarde (Pouliot,

7 http://docs.postgresql.fr/9.1/preface.html#intro-whatis
8 http://www.geoext.org/
9 http://openlayers.org/
10 http://dev.sencha.com/deploy/ext-3.4.0/examples/
11 http://www.opengeospatial.org/

2010). La même technologie a été utilisée pour la sauvegarde du répertoire du site Web.

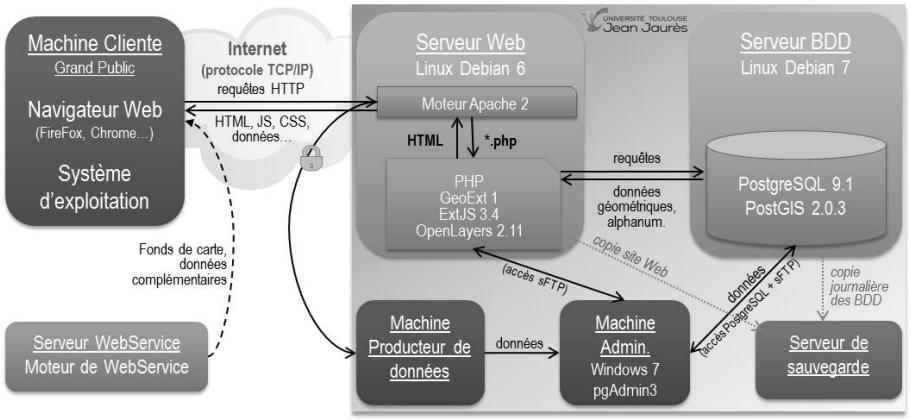


Figure 5 : Architecture technique retenue pour le système d'information PALEOPYR, utilisant les bibliothèques de fonctions JavaScript GeoExt, ExtJS et OpenLayers et des webservices.

2.2.4. Optimisations du système d'information

Dans le cas d'une utilisation du SGBD PostgreSQL et de son extension spatiale PostGIS, les recommandations de Cornet (2011) et de Lubaczewski (2007) ont été mises en œuvre afin d'optimiser et de sécuriser la base de données PALEOPYR.

Dans cette partie, deux spécificités relatives à la traçabilité des modifications d'une part et à la gestion des informations géographiques d'autre part sont détaillées.

2.2.4.1. Tracer les propriétaires des données et les dates de modification

La qualité et la fiabilité d'une base de données dépendent des éléments de traçabilité. L'extension *spi* de PostgreSQL rajoute des fonctions usuelles en adéquation avec les besoins de suivi des modifications, comme l'ajout du propriétaire de la donnée (fonction *insert_username*) et de la date de modification (fonction *moddatetime*). Après l'activation de ces fonctions, deux champs 'username' et 'moddate' sont créés dans chaque table et remplis automatiquement grâce à un déclencheur *trigger* lors de l'événement "avant mise à jour ou ajout de données" (voir code Fig. 6).

```
# Pour la fonction insert_username[12]
ALTER TABLE paleo.entity ADD COLUMN username text NOT NULL;
CREATE TRIGGER entity_insert_username
    BEFORE INSERT OR UPDATE ON paleo.entity
        FOR EACH ROW
            EXECUTE PROCEDURE insert_username (username);

# Pour la fonction moddatetime[13]
ALTER TABLE paleo.entity
    ADD COLUMN moddate timestamp DEFAULT CURRENT_TIMESTAMP NOT NULL;
CREATE TRIGGER entity_moddatetime
    BEFORE UPDATE ON paleo.entity
        FOR EACH ROW
            EXECUTE PROCEDURE moddatetime (moddate);
```

Figure 6 : Commandes SQL pour la création des déclencheurs pour l'ajout du propriétaire de la donnée (insert_username) et de la date de modification (moddatetime). Dans ces exemples, il s'agit de la table entity du schéma paleo.

2.2.4.2. Gérer correctement les informations géographiques dans la base

L'extension PostGIS gère les informations de type point, ligne, polyligne, ou polygone. Dans notre cas, il s'agit de points correspondants aux lieux de prélèvement des enregistrements sédimentaires. Dans la base, deux tables gérant des informations géométriques ont été créées, une pour les sites (SITELOC) et une pour les enregistrements sédimentaires (ENTITY) (cf. MCD Fig. 8). Elles sont pourvues d'un champ 'geom' de type *geometry* et de champs gérant les longitudes et latitudes en degrés décimaux. Pour que le champ 'geom' soit calculé à partir des valeurs de longitude et latitude renseignées, il a fallu exécuter la requête suivante (concernant ici la table ENTITY) et combinant plusieurs fonctions spatiales ST de PostGIS : *UPDATE paleo.entity SET geom = ST_Transform(ST_setSRID (ST_Point(elondd,elatdd),4326),3857);*. Cette requête utilise tout d'abord la fonction *ST_Point* qui permet, à partir des longitudes et latitudes, d'obtenir une description de la géométrie dans un format binaire compressé et de le stocker dans le champ 'geom' de type *geometry*. Puis la fonction *ST_setSRID* affecte un système de coordonnées géoréférencées, ici en l'occurrence le code EPSG 4326 qui correspond au WGS84 et à l'absence de projection. Enfin, la fonction *ST_Transform* permet de transformer le type de système de projection, passant ici du code 4326 au code 3857.

La projection *Spherical Mercator*, communément appelé *Google Mercator* (code EPSG 3857) a été choisie car elle permet d'afficher la répartition spatiale des données paléoenvironnementales à l'échelle pyrénéenne sur une carte interactive Web dont le fond de carte peut être l'un des fonds de l'API Google Maps®.

12 https://github.com/adunstan/postgresql-dev/blob/master/contrib/spi/insert_username.example
13 https://github.com/adunstan/postgresql-dev/blob/master/contrib/spi/moddatetime.example

Pour automatiser le calcul du champ 'geom' à chaque mise à jour ou insertion de nouvelles coordonnées spatiales, la fonction exécutant la requête précédemment énoncée doit être associée à un déclencheur via une commande SQL spécifique (Fig. 7).

Enfin, pour accélérer l'exécution des requêtes faisant appel au champ 'geom', il a été utile de créer un index spatial pour chaque champ 'geom' via une commande SQL (*e.g. CREATE INDEX entity_geom_gist ON paleo.entity USING gist(geom);*).

```
CREATE OR REPLACE FUNCTION public.entity_calcul_geom()
    RETURNS trigger AS
    $BODY$
    BEGIN
        UPDATE paleo.entity SET
            geom=ST_Transform(ST_setSRID(ST_Point(elondd,elatdd),4326),3857)
            WHERE (geom is null AND elatdd is not null AND elondd is not null);
        RETURN NULL;
    END;
    $BODY$
    LANGUAGE plpgsql VOLATILE
    COST 100;
ALTER FUNCTION public.entity_calcul_geom() OWNER TO paleo_admin;
CREATE TRIGGER trig_entity_calcul_geom
    AFTER INSERT OR UPDATE ON paleo.entity
    FOR EACH ROW
        EXECUTE PROCEDURE public.entity_calcul_geom();
```

Figure 7 : Commande SQL pour la création d'une fonction de calcul automatique du champ 'geom' à partir des latitudes (elatdd) et longitudes (elondd) de la table entity du schéma paleo.

2.3. Exploitation des données et diffusion des métadonnées

2.3.1. Création de requêtes

Plusieurs requêtes ont été développées pour : 1) extraire et visualiser les métadonnées via une interface cartographique interactive dédiée, 2) construire des tableaux de données synthétiques simples (*e.g.* liste des enregistrements sédimentaires, liste des taxons) ou de données calculées (*e.g.* tableau affichant les données polliniques en face de données d'âges cal BP pour un enregistrement sédimentaire donné) et 3) calculer le pourcentage pollinique d'un groupe de taxons ciblé pour une période chronologique choisie.

Ces requêtes ont été optimisées, conformément aux recommandations de Cornet (2011), à savoir : i) utiliser des sous-requêtes pour éviter les multiples conditions et jointures et décomposer le problème en sous-problèmes plus simples, ii) placer en premier les conditions les plus rapides dans le WHERE de la requête SELECT, car,

dans certains cas, l'analyseur de requêtes pourrait ne pas trouver le meilleur ordre d'exécution des conditions.

2.3.2. Création d'interfaces interactives pour la diffusion des métadonnées et l'exploitation des données

Pour chaque interface identifiée, une maquette a été dessinée pour positionner les différents éléments structurants et la liste des données ou métadonnées à afficher.

Dans le cas d'une interface cartographique interactive, l'API GeoExt bénéficie des fonctionnalités complémentaires de l'API cartographique OpenLayers et de l'API d'interfaçage ExtJS (Jégou, 2011 ; 2012 ; Quesnot, 2011 ; Quesnot T. et Jégou L., 2012). Par exemple, les tableaux d'information sont réalisés avec ExtJS alors que l'accès à une infobulle par réaction au clic sur les objets ponctuels est rendu possible grâce à OpenLayers. Pour ce qui concerne les fonds de carte, OpenLayers se connecte aux sources de données spatiales distantes via les *webservices* WMS puisqu'il s'agit en l'occurrence ici de cartes rasters tuilées. La légende des couches dérive de la fonction spécifique de GeoExt *GeoExt.LegendPanel*, dont les règles de style graphiques ont été préalablement définies via la fonction *OpenLayers.Rule*. Par ailleurs, les graphes sont des objets de la bibliothèque de fonctions ExtJS, utilisant les fonctions *Ext.chart.PieChart* et *Ext.chart.ColumnChart* pour les diagrammes en secteurs et en bâtons respectivement.

Plus précisément, trois étapes permettent l'affichage dynamique des données ou métadonnées sur une carte interactive : 1) la transformation des données au format texte GeoJSON, 2) l'exécution d'une requête asynchrone (selon le principe AJAX : *Asynchronous JavaScript And XML*) dans un script PHP pour que les données soient retournées au navigateur, 3) le navigateur stocke les données et les coordonnées des points dans un *FeatureStore* (de la bibliothèque GeoExt) pour la génération de la couche vectorielle et des données attributaires correspondantes. Ces dernières peuvent être affichées dans une infobulle grâce à la fonction d'OpenLayers *createPopup(nomFeatureStore).*

Pour la visualisation interactive des tableaux de synthèse des données créés en dehors de l'interface cartographique, la bibliothèque de fonctions Javascript JQuery[14] a été utilisée afin de présenter des tableaux esthétiques, ergonomiques et téléchargeables aux formats CSV, XLS ou PDF.

2.3.3. Création d'une interface cartographique dynamique illustrant l'évolution de la végétation forestière nord-pyrénéenne de la fin de l'Antiquité au haut Moyen Âge

Des outils fonctionnels ont été développés pour visualiser les dynamiques du couvert végétal au cours du temps. Pour illustrer notre propos, nous avons choisi de tester l'opérationnalité de la base PALEOPYR sur un exemple simulé (*i.e.* jeu de données fictif de pourcentages polliniques) visant à reconstruire les dynamiques

14 http://jquery.com/

forestières du milieu nord pyrénéen de la fin de l'Antiquité au haut Moyen Âge (de -100 avant JC à l'an 1000).

L'interface devait calculer et représenter ponctuellement pour chaque enregistrement sédimentaire la somme des pourcentages polliniques des taxons forestiers par fenêtre temporelle de 100 ans, soit onze fenêtres. Il s'agissait de : i) vérifier la faisabilité du calcul à partir de requêtes d'interrogation de la base de données, ii) tester le prototype en simulant l'évolution cartographique du couvert forestier via l'animation des scènes générées, et via des graphes ou des tableaux.

3. Résultats et discussion

3.1. Modèle conceptuel de données optimisé et base de données opérationnelle

Le MCD issu de la modélisation comprend 18 classes d'entités et de nombreuses relations entre elles. La figure 8 en illustre une version simplifiée. La classe d'entité ENTITY (*i.e.* enregistrements sédimentaires) est au cœur du système, elle est reliée à ses lots d'échantillons (SAMPLE_SET) définis par des segments ou niveaux de profondeur (SAMPLE). Les métadonnées sont réparties dans différentes classes d'entités (ENTITY, SITELOC, ANNEX, WORKER, STUDY, CHRON, CALIB, etc.) et sont extraites et formatées selon la norme *Dublin Core* via une vue, permettant ainsi le moissonnage futur par des catalogues de métadonnées.

Le modèle a permis de générer une synthèse des données et indicateurs disponibles pour chaque enregistrement sédimentaire, et d'ajouter de nouveaux types d'information (en plus de ceux listés dans INFOTYPE) et de nouveaux paramètres (en plus de ceux listés dans VARIABLE), rendant ainsi la base évolutive. Ainsi, une relation de plusieurs à plusieurs (*n* à *n*) contenant les valeurs des triplets *échantillon/variable/valeur* fut créée entre une classe d'entité stockant les échantillons (SAMPLE) d'une part et une classe d'entité stockant la liste des variables d'autre part (VARIABLE). De même, une relation *n* à *n* pour les données de comptage de grains de pollen permet de gérer le triplet *échantillon/taxon/nombre* via le triplet SAMPLE/VAR/COUNT dans le MCD.

Cette structuration se différencie de celle proposée par la base EPD ou de la *Neotoma Paleoecology database* qui contiennent généralement une table pour chaque indicateur (*e.g.* une table pour les dates ^{14}C, une table pour les données lithologiques, etc.), les rendant difficilement évolutives en cas de nouveaux indicateurs.

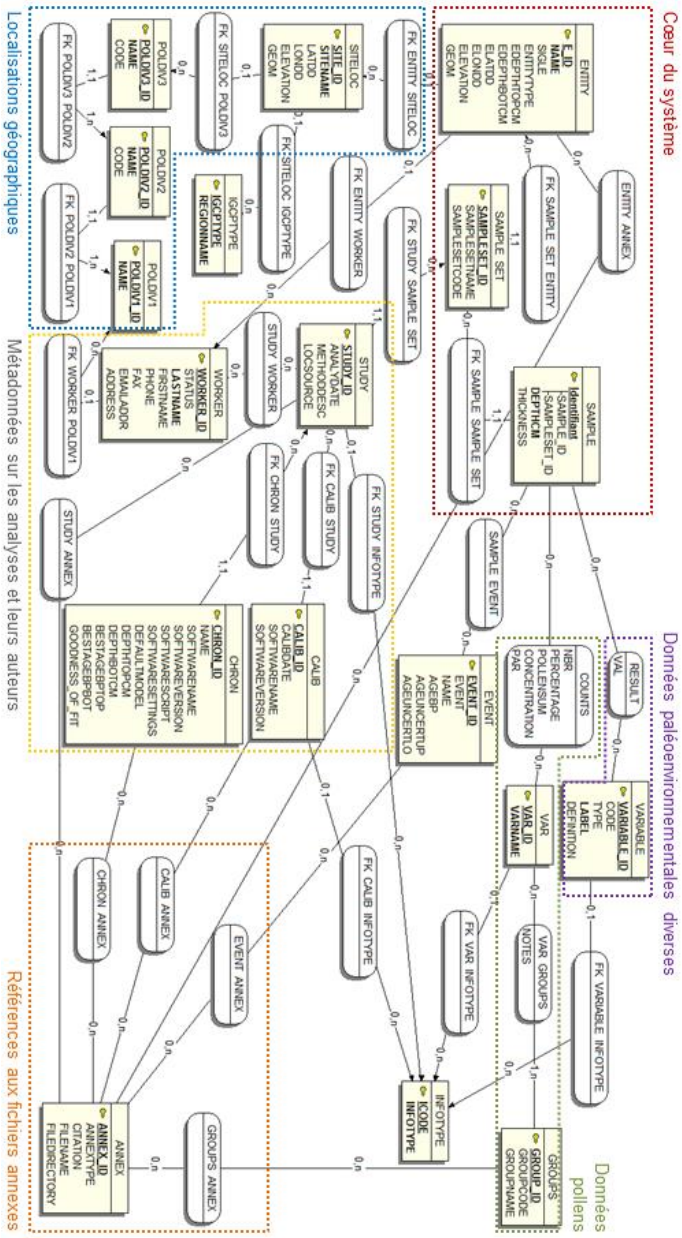


Figure 8 : Modèle conceptuel de données de la base de données PALEOPYR (logiciel Win'Design®). Chaque classe d'entité est représentée par un rectangle, chaque relation par une ellipse. Les cardinalités (0,1 ; 0,n ;…) sont indiquées sur chaque lien sémantique. Certains descripteurs de métadonnées des classes d'entité ENTITY, SITELOC et ANNEX ne sont pas affichés.

Le résultat de cette modélisation conceptuelle des données conduit à une simplification de la structure de la base de données allégée en nombre de tables et de relations sous-jacentes, soit 29 tables pour PALEOPYR au lieu de 45 pour l'EPD (cf. manuel EPD[15]) ou de 67 pour la *Neotoma Paleoecology database*[16]. Néanmoins, dans la perspective d'alimenter la base de données EPD, la langue, la nomenclature et le format de certaines données ont été autant que possible conservés (cf. manuel EPD[15]), rendant ainsi compatible le cœur du système de PALEOPYR avec l'EPD (*e.g.* tables ENTITY, SITELOC, POLDIV).

Les multiples requêtes fonctionnelles créées pour répondre aux questions techniques ou scientifiques ont permis de valider le MCD. A titre d'exemple, la requête la plus complexe obtenue fait appel à 17 sous-requêtes emboîtées dans le but de calculer le pourcentage pollinique de taxons ciblés pour une période chronologique choisie et pour chacun des enregistrements sédimentaires. La base PALEOPYR est donc opérationnelle et répond aux objectifs d'archivage et d'interrogation des données paléoenvironnementales et des métadonnées associées.

3.2. Interface cartographique publique de diffusion des métadonnées

Lors du dessin de la maquette, la carte est privilégiée en termes d'espace écran. Elle permet de visualiser la distribution des lieux de prélèvement des enregistrements sédimentaires. Une infobulle pour chaque point permet la visualisation des métadonnées réparties dans quatre onglets intitulés *Record characteristics*, *Paleoenvironmental data*, *Bibliographic references*, *Pictures*.

Le prototype obtenu a été mis en ligne : http://paleopyr.univ-tlse2.fr (Fig. 9). Il présente, comme un logiciel de SIG, des outils classiques de navigation, de zoom et l'échelle. Le fond de carte habille esthétiquement la carte active et permet le repérage géographique des informations. Plusieurs fonds de carte ont été choisis parmi ceux identifiés dans la grille des besoins : les différents fonds de l'API Google Maps® V.3 et le fond OpenStreetMap.

Un menu "en accordéon" permet de fournir ergonomiquement des données complémentaires qui aident à la compréhension de la couche spatiale des données paléoenvironnementales. Il détaille la légende, un graphe en secteurs présentant le nombre de points par type de site (*e.g.* tourbière, lac), un diagramme en bâtons pour donner le nombre de sondages par type de données, les auteurs, puis le tableau des enregistrements sédimentaires.

Le développement technique de l'interface de consultation des métadonnées s'est avéré difficile (répartition des champs de métadonnées dans de nombreuses tables et restrictions d'accès à certaines données). La diversité des types de données (*e.g.* données géographiques, non géographiques, images, champ textuel de grande longueur comme les références bibliographiques) ne facilitait pas la réalisation du

15 http://www.europeanpollendatabase.net/data/downloads/
16 http://api.neotomadb.org/doc/resources/dbtables

programme sous JavaScript. De plus, les bibliothèques de fonctions d'OpenLayers et d'ExtJS étaient relativement peu documentées et illustrées. Parmi les limites des applications développées sous GeoExt, l'incompatibilité avec le navigateur Internet Explorer® est le principal défaut. Enfin, par rapport à la maquette initiale, le menu de recherche des métadonnées reste à créer afin de faciliter l'exploration des résultats. Malgré ces verrous techniques, l'interface est fonctionnelle via les autres navigateurs et le programme JavaScript accessible à tous les internautes peut servir d'exemple illustrant de nombreuses fonctionnalités d'OpenLayers et d'ExtJS.

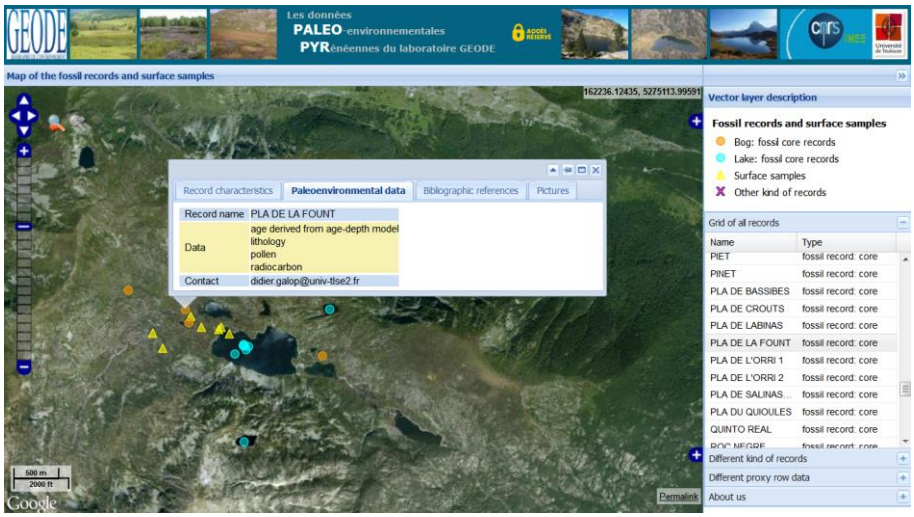


Figure 9 : Aperçu de l'interface cartographique interactive de visualisation des métadonnées http://paleopyr.univ-tlse2.fr, sa carte interactive (zoom sur un site d'étude, Bassiès), sa légende et sa grille d'enregistrements sédimentaires. Le deuxième onglet de l'infobulle liste les indicateurs pour lesquels l'enregistrement sédimentaire cible présente des jeux de données.

3.3. Interface privée de visualisation, interrogation et exportation des données

Un premier prototype d'interface interactive privée a été développé dans le but de donner rapidement un accès aux producteurs de données (Fig. 10). Cette première version, permet de lister différents éléments de la base de données (*e.g.* enregistrements sédimentaires, taxons, variables) dans des tableaux téléchargeables. Elle permet également de rechercher toutes les informations relatives à l'enregistrement sédimentaire de son choix afin d'obtenir les treize tableaux de synthèse correspondants.

Le prototype actuel offre la possibilité d'explorer dynamiquement et de télécharger les données. Il s'agit désormais de permettre la saisie ou la modification de ces dernières en ligne. Ainsi l'interface devra évoluer vers un système intégrant les fonctionnalités d'un véritable outil appelé "CRUD" (pour *Create*, *Read*, *Update*, *Delete*). Deux possibilités ont été identifiées pour le développement du CRUD : la programmation "manuelle" ou l'utilisation d'un générateur automatique de CRUD tel que celui du *framework Symfony*[17] ou *Grocery Crud*[18], un générateur de CRUD en langage PHP qui est une extension du *framework CodeIgniter*[19].

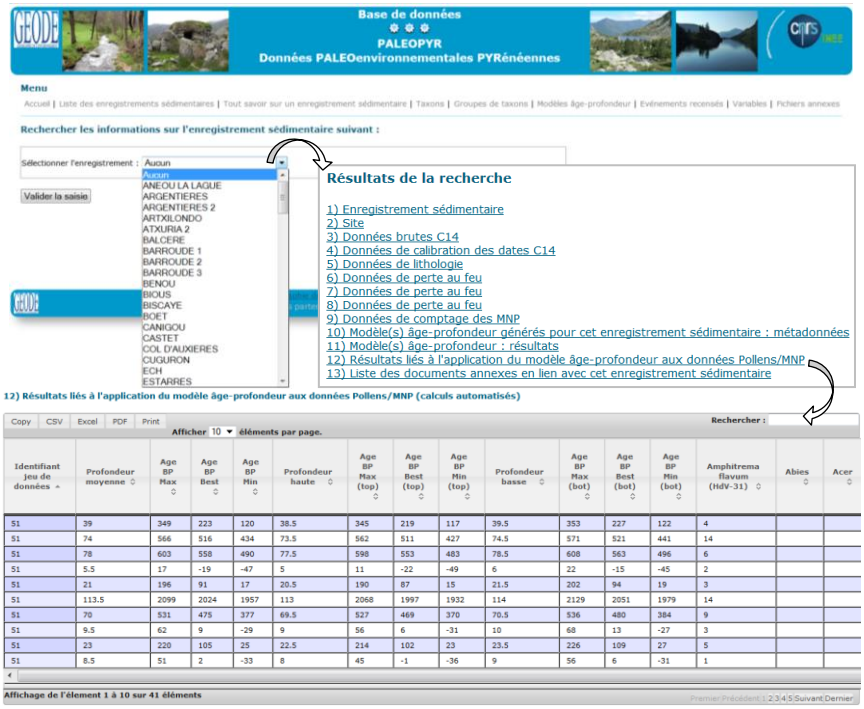


Figure 10 : Aperçu de l'interface de visualisation, d'interrogation et d'exportation des données paléoenvironnementales du système d'information PALEOPYR. La sélection d'un enregistrement sédimentaire à partir d'une liste déroulante permet d'obtenir un menu présentant les treize intitulés des tableaux de synthèses. Chaque tableau propose un affichage de dix éléments par page par défaut, avec des fonctions de tri, de navigation d'une page à une autre et de téléchargement.

17 http://symfony.com/
18 http://www.grocerycrud.com/
19 http://codeigniter.fr/user_guide/index.html

3.4. Résultats du test de simulation de l'évolution de la végétation forestière nord-pyrénéenne de la fin de l'Antiquité au haut Moyen Âge via une interface cartographique dédiée

La maquette du prototype d'interface a tout d'abord été produite, positionnant en son centre une carte similaire à celle de l'interface grand public de métadonnées. La nouveauté réside dans la création d'un bloc dédié à l'exploration des résultats via un système de boutons.

La conception d'une telle interface a permis d'atteindre l'objectif de visualisation de onze "scènes" correspondantes aux onze siècles sur la période historique ciblée. Chaque scène, activée par un bouton, permet de visualiser la somme des pourcentages polliniques des taxons arboréens sur l'ensemble des sites disponibles pour chaque siècle. La figure 11 illustre l'évolution du couvert forestier pour trois siècles.

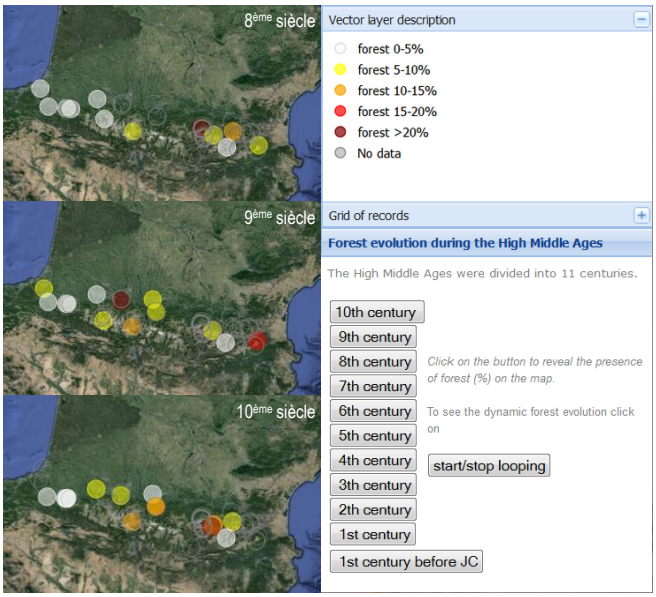


Figure 11 : Aperçus de l'interface cartographique dédiée à la visualisation des changements de la couverture forestière de la fin de l'Antiquité au haut Moyen Âge. À gauche, les trois scènes correspondant à trois siècles (700 à 1000 après JC) sur la base de données fictives pour trois taxons forestiers. À droite, la légende et le bloc de boutons permettant soit de sélectionner le siècle soit de lancer une animation sur l'ensemble de la période avec un changement de scène toutes les deux secondes.

Pour chaque point de la carte, une infobulle permet de visualiser l'évolution du couvert forestier à une résolution temporelle centennale ainsi que le modèle âge-profondeur retenu pour la génération du graphe (Fig. 12).

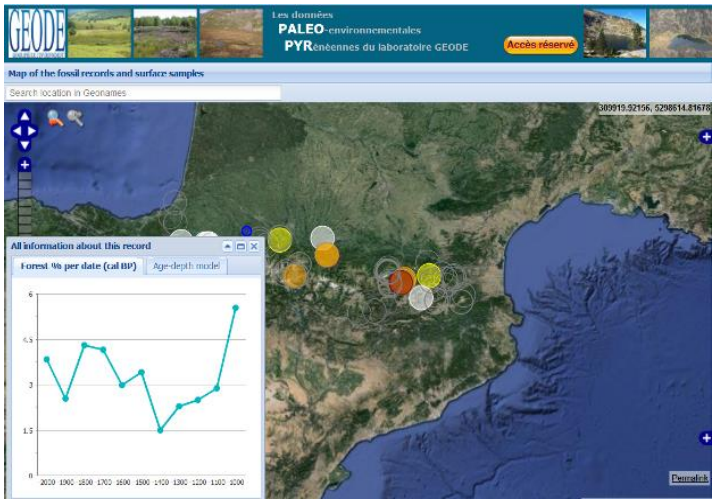


Figure 12 : Aperçu de l'interface cartographique permettant de fournir via une infobulle un graphe représentant la somme des pourcentages polliniques des taxons arboréens par siècle (ici données fictives). Les proportions polliniques totales par siècle sont calculées en temps réel en fonction du modèle d'âge-profondeur de référence pour chaque enregistrement sédimentaire. Il est possible de visualiser simultanément les graphes de plusieurs enregistrements afin de faciliter les comparaisons.

La réalisation de cette interface cartographique marque une première étape dans la visualisation de la dynamique de la végétation au cours du temps. Cette interface, comme un logiciel de SIG, offre la possibilité de superposer plusieurs couches d'informations, par exemple un modèle numérique de terrain, des cartes géologiques et d'occupation du sol. Elle permet également de fournir des résultats cartographiques proches de ceux mentionnés par le cahier des charges. En effet, le type de carte obtenu s'apparente à celui publié par Galop *et al.* (2013) qui représente une série chronologique des zones d'apparition des activités agropastorales sur un fond de carte des Pyrénées. Certaines modifications (ergonomie, ajout de fonctionnalités) la rendront encore plus attractive et facile d'utilisation. Par exemple l'utilisateur, en fonction de ses objectifs de recherche, pourra sélectionner le(s) taxon(s) et la fenêtre temporelle de son choix via un système de menus déroulants puis visualiser les résultats sur la carte et obtenir des graphes ou cartes disponibles au téléchargement.

La mise en relation directe des données de PALEOPYR avec un SIG est possible et reste à explorer. En effet, le fait d'avoir des tables ou des vues avec un champ de type '*geometry*' dans PALEOPYR offre la possibilité de générer des couches vectorielles dans les logiciels SIG et d'accéder aux données de la base de données via les tables attributaires correspondantes. Une interpolation des données ponctuelles de proportions polliniques pourrait permettre alors la visualisation de l'évolution spatiale continue de taxon(s) à l'échelle globale des Pyrénées, comme Branch et Marini (2013) l'ont démontré dans les Apennins pour plusieurs essences forestières via l'utilisation du logiciel de SIG ArcGis Esri®. Toutefois il sera nécessaire de définir une méthode d'interpolation fiable qui intègre la complexité des facteurs entrant en jeu dans le dépôt des pollens (Mazier *et al.*, 2008). C'est pourquoi une prise de recul sur la problématique générale posant la question de la spatialisation des données d'évolution de la végétation reste nécessaire.

4. Conclusion et perspectives

S'intéresser à la reconstitution spatio-temporelle des environnements passés nécessite au préalable de gérer et de centraliser les milliers de données quantitatives disponibles dans un système unique, optimisé, ainsi que de donner une dimension spatiale à ces données pour en permettre une pleine exploitation et, *in fine*, la visualisation dynamique des évolutions environnementales.

La sauvegarde des données paléoenvironnementales et de leurs métadonnées associées acquises depuis trente ans sur le massif nord-pyrénéen a nécessité le développement d'une base de données relationnelle, optimisée et sécurisée. Le nouveau modèle conceptuel de données PALEOPYR propose une simplification du contenu de la base sans pour autant en altérer la qualité, le niveau de précision des données à gérer ou l'évolutivité du système. L'ajout de nouveaux indicateurs ou paramètres s'en trouve ainsi facilité. Ce système va poursuivre son évolution technique par l'intégration d'un système de CRUD permettant aux producteurs de données de saisir, mettre à jour et supprimer des données plus directement.

Les interfaces cartographiques développées offrent pour la première fois la possibilité technique de visualiser les métadonnées sur une carte interactive, fournissant ainsi en libre accès la localisation et la nature des jeux de données disponibles sur le versant nord-pyrénéen. L'interface cartographique privée propose également une nouvelle façon de visualiser simultanément sur une carte les résultats des modèles d'âges-profondeurs sous la forme de graphes pour tous les enregistrements sédimentaires. Elle permet, en outre, une animation cartographique de l'évolution des proportions polliniques de l'ensemble des enregistrements sédimentaires présentant des données polliniques en face de données d'âges cal BP. La possibilité de superposer la couche de points avec d'autres couches informatives peut aider à l'interprétation des données paléoenvironnementales. L'évolution conjointe des technologies *open source* permettra à terme de lever les verrous techniques restants et offrira de nouvelles perspectives d'amélioration de l'ergonomie de ces applications clientes avec l'apport de nouvelles fonctionnalités.

PALEOPYR alimente des projets de recherche en cours. Cet outil représente une avancée méthodologique fondamentale pour la recherche paléoécologique pyrénéenne, offrant de nouvelles perspectives scientifiques de visualisation et d'analyse par le croisement des différents types d'information qu'elle contient, notamment via le couplage SIG-SGBD, et de proposer *in fine* des scénarios d'évolution de la végétation durant le Tardiglaciaire et l'Holocène à l'échelle des Pyrénées. Les interfaces cartographiques pourront à terme intéresser toute la communauté scientifique et administrative concernée par les dynamiques historiques de l'environnement (archéologues, historiens, les Services Régionaux de l'Archéologie, Office National des Forêts, Institut National de la Recherche Agronomique, organismes de recherche étrangers).

Remerciements

Les auteurs remercient les autres contributeurs des données archivées dans PALEOPYR (Guy Jalut, Damien Rius, Anaëlle Simonneau) et Anne Le Flao pour l'ensemble de ses travaux préliminaires de structuration de l'information. Cette étude a bénéficié des soutiens financiers du CNRS et du Ministère de la culture dans le cadre de l'ACR "Rythmes et causalités des dynamiques de l'anthropisation en milieu montagnard. L'exemple de la construction des territoires pyrénéens de la fin du mésolithique à l'aube de notre ère" (dir. D. Galop), du projet MODE RESPYR ANR JC 2010 (dir. T. Houet) et du LABEX DRIIHM (OHM du Haut Vicdessos) (dir. D. Galop.).

Bibliographie

Blaauw M. (2010). Methods and code for 'classical' age-modelling of radiocarbon sequences. *Quaternary Geochronology*, vol. 5, n° 5, p. 512-518.

Branch N.-P., Marini N.-A.-F., (2013). Mid-Late Holocene environmental change and human activities in the northern Apennines, Italy. *Quaternary International*. Available online 19 August 2013. *In Press*.

Cornet J. (2011). *PostgreSQL/PostGis: bonnes pratiques et astuces*. http://www.portailsig.org/content/postgresql-postgis-bonnes-pratiques-et-astuces

Davis B.-A.-S. *et al.* (80 authors), (2013). The European Modern Pollen Database (EMPD) project. *Veget. Hist. Archaebot*. vol. 22, p 521-530.

Fyfe R.-M., de Beaulieu J.-L., Binney H., Bradshaw R.-H.-W., Brewer S., Le Flao A., Finsinger W., Gaillard M.-J., Giesecke T., Gil-Romera G., Grimm E.-C., Huntley B., Kunes P., Kühl N., Leydet M., Lotter A.-F., Tarasov P.-E., Tonkov S. (2009). The European Pollen Database: past efforts and current activities. *Veget. Hist. Archaebot*. vol. 18, p. 417-424.

Galop D. (dir), Carozza L., Valdeyron N., (2005). *Rythmes et causalités des dynamiques de l'anthropisation en milieu montagnard : l'exemple de la construction des territoires pyrénéens de la fin du Mésolithique à l'aube de notre ère.* Action Collective de Recherche 2003 CNRS, Ministère de la recherche, INRAP, Ministère de la culture. Rapport d'étape : 98p.

Galop D. (dir), Carozza L., Valdeyron N., (2006). *Rythmes et causalités des dynamiques de l'anthropisation en milieu montagnard : l'exemple de la construction des territoires pyrénéens de la fin du Mésolithique à l'aube de notre ère.* Action Collective de Recherche 2003 CNRS, Ministère de la recherche, INRAP, Ministère de la culture. Rapport d'étape: 114p.

Galop D., Rius D., Cugny C., Mazier F., (2013). A History of Long-Term Human-Environment Interactions in the French Pyrenees Inferred from the Pollen Data. *Continuity and Change in Cultural Adaptation to Mountain Environments*, Ludomir R. Lozny Editor, p. 19-29.

Goring S., Dawson A., Simpson G.-L., Ram K., Graham R.-W., Grimm E.-C, Williams J.-W., (2015). *Neotoma: A Programmatic Interface to the Neotoma Paleoecological Database.* DOI: 10.5334/oq.ab http://www.openquaternary.com/article/10.5334/oq.ab/print/

Grimm E.-C. (2008). *Neotoma: An Ecosystem Database for the Pliocene, Pleistocene, and Holocene*. http://www.neotomadb.org/uploads/NeotomaManual.pdf

Grimm E.-C., Bradshaw R.-H.-W., Brewer S., Flantua S., Giesecke T., Lézine A.-M., Takahara H. and Williams J.-W., (2013). Databases and Their Application. In: Elias S.A. (ed.) *The Encyclopedia of Quaternary Science*, vol. 3, p. 831-838. Amsterdam: Elsevier.

Grobe H., Diepenbroek M., Dittert N., Reinke M., Sieger R., (2006). Archiving and distributing earth-science data with the PANGAEA information system. *Proceedings of the IX International Symposium of Antarctic Earth Sciences (Fütterer, D. et al., eds)*, p. 403-406, Berlin.

Heiri O., Brooks S.-J., Birks H.-J.-B, Lotter A.-E., (2001). A 274-lake calibration data-set and inference model for chironomid-based summer air temperature reconstruction in Europe. *Quaternary Science Reviews*, vol. 30, p. 3445-3456.

Jégou L. (2011). *Carte en symboles proportionnels avec GeoExt, OpenLayers et PostGIS* http://www.portailsig.org/content/carte-en-symboles-proportionnels-avec-geoext-openlayers-et-postgis

Jégou L. (2012). *Illustration de cours de Master 2 SIGMA Université Jean Jaurès Toulouse 2 sur les objets GeoExt*. http://www.geotests.net/test/sigma/

Le Flao A. (2004). "*Paleopyr*". *Premiers jalons pour une base de données paléoécologiques pyrénéennes sous Access*. Maîtrise de Géographie, Université Toulouse 2 Le Mirail.

Le Flao A. (2005). *Apports des reconstitutions cartographiques à l'histoire de l'environnement. Essai de spatialisation des données palynologiques pyrénéennes.* DEA Géographie. Université Toulouse 2 Le Mirail.

Lerigoleur E., (2013). *Base de données PALEOPYR pour les données PALEO-environnementales PYRénéennes: de la conception de la base à la production d'interfaces cartographiques*. Mémoire de Master 2 SIGMA, Université Toulouse 2 Le Mirail.

Libourel T. (2010). *Système d'Information et Méthode MERISE. Méthode d'Etude et de Réalisation Informatique pour les Systèmes d'Entreprise*. Support de Thèse. http:#www2.lirmm.fr/IC#Supports/FMIN114-ConduiteProjets/Supports_ThereseLibourel/MERISE_2010.pdf

Lubaczewski H. (2007). *Sécuriser votre base PostgreSQL*. http://blog.postgresql.fr/index.php?post/drupal/397

Lupovici C. (1998). *Le Digital Object Identifier : Le système du DOI.* http://bbf.enssib.fr/consulter/bbf-1998-03-0049-007

Mazier F., Galop D., Brun C., and Buttler A., (2006). Modern pollen assemblages from grazed vegetation in the Western Pyrenees Mountains (France). A numerical tool for more precise reconstruction of past cultural landscapes. *The Holocene*. vol. 16, n° 1, p 91-103.

Mazier F., Brostöm A., Gaillard M.-J., Sugita S., Vittoz P., and Buttler A., (2008). Pollen productivity estimates and Relevant Source Area for major taxa in a pasture woodland (Jura mountains, Switzerland). *Vegetation History and Archaeobotany*. vol. 17, p 479-495.

Mazier F., Galop D., Gaillard M.-J., Rendu C., Cugny C., Amaïa L., Peyron O., Buttler A., (2009). Multidisciplinary approach to reconstruct pastoral activities. An example from the Pyrenean Mountains (Pays Basque). *The Holocene*, vol. 19, n° 2, p. 171-188.

Peyron O., Jolly D., Bonnefille R., Vincens A., Guiot J., (2000). Climate of east Africa 6000 C-14 yr BP as inferred from pollen data. *Quaternary Research*. vol. 54, n° 1, p 90-101.

Pouliot E. (2010). *A simple yet effective PostgreSQL backup script.* http://www.defitek.com/blog/2010/01/06/a-simple-yet-effective-postgresql-backup-script/

Quesnot T. (2011). *Développement d'un prototype de cartographie thématique interactive en ligne*. Mémoire de Master 2 SIGMA, Université Toulouse 2 Le Mirail.

Quesnot T., Jégou L., (2012). Atlas cartographique de la CAF 31. http://www.sig-caf31.fr/

Reimer P. J., Baillie M.-G.-L., Bard E., Bayliss A., Beck J.-W., Blackwell P.-G., Bronk Ramsey C., Buck C.-E., Burr G.-S., Edwards R.-L., Friedrich M., Grootes P.-M., Guilderson T.-P., Hajdas I., Heaton T.-J., Hogg A.-G., Hughen K.-A., Kaiser K.-F., Kromer B., McCormac F.-G., Manning S.-W., Reimer R.-W., Richards D.-A., Southon J.-R., Talamo S., Turney C.-S.-M., van der Plicht J., Weyhenmeyer C.-E., (2009). IntCal09 and Marine09 radiocarbon age calibration curves, 0-50,000 years cal BP. *Radiocarbon*, vol. 51, n° 4, p. 1111-1150.

Reimer P. J., Bard E., Bayliss A., Beck J.-W., Blackwell P.-G., Bronk Ramsey C., Buck C.-E., Cheng H., Edwards R.-L., Friedrich M., Grootes P.-M., Guilderson T.-P., Haflidason H., Hajdas I., Hatté C., Heaton T.-J., Hoffmann D.-L., Hogg A.-G., Hughen K.-A., Kaiser K.-F., Kromer B., Manning S.-W., Niu M., Reimer R.-W., Richards D.-A., Scott E.-M., Southon J.-R., Staff R.-A., Turney C.-S.-M., van der Plicht J. (2013). IntCal13 and Marine13 radiocarbon age calibration curves, 0-50,000 years cal BP. *Radiocarbon*, vol. 55, n° 4, p. 1869-1887.

Sumathi S., Esakkirajan S., (2007). Database Management System. *Fundamentals of Relational Database Management Systems*. Studies in Computational Intelligence, vol. 47, p. 3-5.

Velle G., Telford R.-J., Heiri O., Kurek J., Birks H.-J.-B., (2012). Testing intra-site transfer functions: an example using chironomids and water depth. *Journal of Paleolimnology*. vol 48, n° 3, p. 545-558.