

## LE SYSTÈME D'INFORMATION DU PROGRAMME "ARCHÉOLOGIE DU BASSIN PARISIEN". ENTRE SIG ET SGBD, VERS UN APPLICATIF OPEN SOURCE ADAPTÉ

### 1. INTRODUCTION

#### 1.1 *Objectif de la base de données*

La constitution de la base de données du programme est l'axe méthodologique fort de celui-ci (BRUN, SOULIER 2007). Loin de vouloir imposer un schéma universel de modélisation des données, l'objectif de la base "Archéologie du Bassin parisien" est de proposer un recensement des sites répertoriés dans les corpus d'études des aires géographique et chronologique du programme. L'intérêt et l'originalité de ce travail est de rendre compte des nombreuses recherches menées dans le cadre du programme. Il s'agit d'un projet ambitieux de grande envergure puisque de nombreux programmes collectifs (type ACR ou PCR) mais aussi des diplômes universitaires produisent des corpus de données susceptibles à terme, d'être intégrés à la base Bassin Parisien.

#### 1.2 *Vers une métabase de données*

Le programme Bassin Parisien couvre une aire géographique et chronologique très large, de la Préhistoire au Moyen-âge. Entre ces bornes chronologiques, la nature des sites, leur fonction, et leur degré de complexité sont très variables. De même, les bases de données destinées à être intégrées sont constituées selon des problématiques et des échelles très diverses (objet – site – réseau de sites – aire d'approvisionnement, etc.) qui sont spécifiques à chaque auteur et constituent autant de schémas de pensées différents. Vouloir mettre en cohérence des corpus conçus indépendamment les uns des autres oblige à ne prendre en compte que les éléments communs à l'ensemble, et non à reprendre *in extenso* tous les éléments de chaque corpus, ce qui équivaldrait à remettre en cause le modèle de la base dès l'ajout de nouvelles données.

Sur la base de ce constat, la présence systématique de trois critères a été relevée dans les différents corpus: 1) la localisation, 2) l'expression du temps ainsi qu'une 3) description fonctionnelle du site. La base de données étant conçue comme une base de sites archéologiques (AUBRY, FERJANI, QUILLIEC 2008), nous avons choisi comme entité élémentaire, ou entité archéologique, de cette base le triplet d'informations suivant 1) un lieu géographique occupé à 2) une époque donnée avec 3) une fonction définie. Chaque triplet d'information géographique, chronologique et fonctionnelle doit être unique dans la base.

**Archéologie du Bassin Parisien  
Schéma relationnel Base de Données**

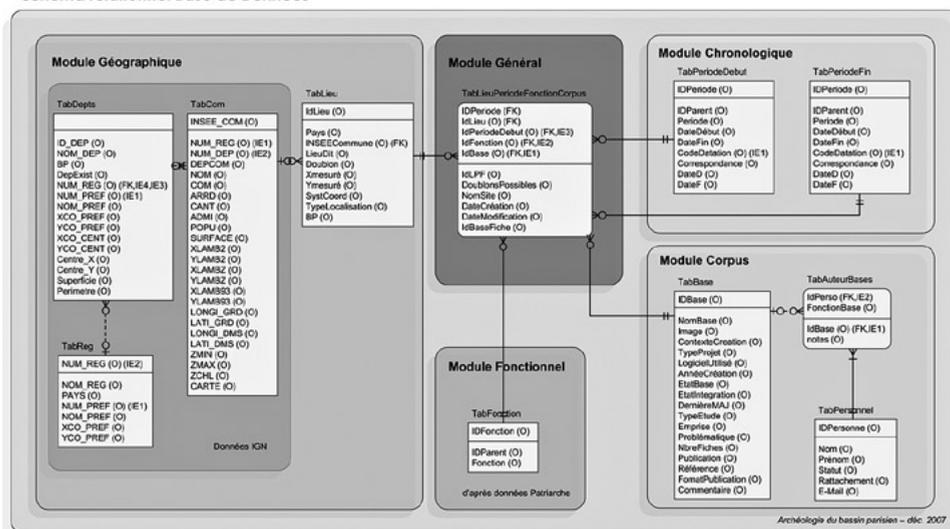


Fig. 1 – Schéma relationnel de la base du programme archéologie du Bassin parisien.

On peut l’interroger selon les questions élémentaires suivantes:

- Inventaire chronologique. Quels sont les sites pour une période donnée?
- Inventaire fonctionnel. Quels sont les sites pour une fonction donnée?
- Inventaire spatial. Quels sont les sites présents sur une zone géographique donnée?

Ces trois critères ont constitué la première ossature de la base. Toutefois, il est vite apparu que s’ils étaient nécessaires, ils n’étaient pas suffisants pour autant, car ils ne permettaient pas de gérer le phénomène de doublonnage. En effet, un même triplet d’informations peut se retrouver dans plusieurs corpus. Nous avons donc modifié la structure du modèle en y ajoutant une information supplémentaire à notre enregistrement élémentaire garantissant son unicité: le nom du corpus originel et donc son (ou ses) auteur(s). Cette solution est d’autant plus intéressante qu’elle replace les auteurs des corpus au centre du dispositif.

En modifiant légèrement sa structure, cette base de sites archéologiques est devenue, de fait, une base de bases de données, ou une base de corpus archéologiques, ou encore une métabase de données.

Aux interrogations précédentes, on peut maintenant ajouter la question élémentaire suivante:

- Inventaire des recherches. Quelles sont les études menées sur une zone géographique, une fonction, et/ou une période donnée?

La structure de la base est constituée de 5 modules distincts composés chacun d'une à plusieurs tables (Fig. 1). Les modules "Géographique", "Chronologique", "Fonctionnel" et "Corpus" se retrouvent liés au module "Général", pivot structurel de la base de données (cfr. AUBRY, FERJANI 2009 pour une description plus détaillée de ces modules). À noter que cette structure radiale autour d'un module central n'est pas sans rappeler le schéma de données OLAP<sup>1</sup> dit en étoile, i.e., une table de faits et une série de tables descriptives aussi nommées tables de dimension autorisant des analyses multidimensionnelles (CELKO 2006).

## 2. CAHIER DES CHARGES

Si l'on veut que cette base de données puisse se développer et être utilisée – et donc remplir son rôle d'outil collaboratif du programme – elle doit répondre à certaines exigences qui relèvent des méthodologies propres aux bases de données, de la technique de modélisation et de l'ergonomie du mode d'interrogation. Le système doit posséder les caractéristiques suivantes:

### 2.1 *Cahier des charges utilisateur*

#### Un accès personnalisé

Cette base étant un outil de travail ergonomique, elle doit proposer à l'utilisateur une mémoire ou un historique de ses interrogations. Pour ce faire, il est nécessaire de proposer un accès personnalisé avec login et mot de passe pour chaque utilisateur.

#### Une interface d'interrogation

L'utilisateur de la base ne doit pas obligatoirement connaître le fonctionnement du logiciel choisi pour l'utiliser avec succès. Pour cela, il est nécessaire de développer une interface d'interrogation intuitive (choix par menus déroulants, cases à cocher, etc.).

#### Un système multiplateforme

La majorité de la communauté des archéologues susceptibles de consulter cette base travaille sous les environnements Windows et MacOS. Il faut donc développer une solution multiplateforme, afin de ne pas faire du choix d'un environnement un obstacle pour une partie de la communauté.

<sup>1</sup> OLAP: dans le domaine des bases de données, le traitement analytique en ligne (anglais *OnLine Analytical Processing*) est un type d'application informatique orienté vers l'analyse d'informations selon plusieurs axes ou dimensions indépendantes, dans le but de synthétiser l'information contenue dans des entrepôts de données (data warehouse). Les représentations multidimensionnelles sont aussi appelées Hypercube ou cubes OLAP.

## 2.2 Cahier des charges techniques d'échange

Une ouverture vers d'autres logiciels et notamment SIG

La base de données doit pouvoir se connecter vers un logiciel de traitement de données géomatiques pour proposer une représentation cartographique du résultat des interrogations.

Un système d'échange

L'utilisateur de la base doit pouvoir récupérer le résultat de ses requêtes dans un fichier de données tabulées tel qu'un fichier texte ou un fichier Microsoft Excel. Les données cartographiques doivent pouvoir être récupérées dans un fichier de type bitmap (.bmp, .jpg, .tif) et/ou vectoriel.

Un accès distant par une architecture client-serveur

La base de données s'enrichit régulièrement de nouvelles données. Diffusée sur plusieurs postes de travail, il faudrait mettre tous les postes à jour en même temps. La solution client-serveur, avec n clients qui se connectent à la machine serveur contenant les données, est plus pertinente car plus simple en termes de maintenance des données (suppression, ajout et modification).

Un système diffusable libre de droit<sup>2</sup>

La solution doit pouvoir être diffusable aisément. Il faut donc privilégier des solutions techniques libres de droit.

## 2.3 Choix techniques

Afin de répondre au mieux aux différents points abordés dans ce cahier des charges, la solution choisie se compose du logiciel Microsoft Access 2007. Ce logiciel est un SGBDR bureautique interrogeable par le langage SQL (CELKO 2005; CHURCHER 2008). Il a de plus, la particularité d'être fortement configurable afin de s'adapter aux besoins des projets le mettant en œuvre. C'est cette particularité que nous avons utilisée pour développer, par programmation en Visual Basic for Application, un applicatif possédant diverses interfaces d'interrogation de la base de données (HENNING *et al.* 2007; FEDDEMA 2007).

Cet applicatif est enrichi de fonctionnalités cartographiques insérées directement dans l'applicatif par l'adjonction du module libre de droit de fonctionnalités SIG MapWinGIS<sup>3</sup>, développé par l'Idaho State University (AMES *et al.*

<sup>2</sup> Quelques notions de logiciel libre ou encore de "Free and Open Source Software" à rappeler. La Free Software Foundation (GNU 2010) définit la notion de logiciel libre par 1) la liberté de l'utiliser, 2) la liberté de le modifier pour ses propres besoins, 3) la liberté de le redistribuer et enfin 4) la liberté de l'améliorer et de rendre ces améliorations publiques.

<sup>3</sup> <http://www.mapwindow.org/>.

2007b). Cette solution répond en partie au cahier des charges défini et représente une étape vers une solution plus pérenne. Toutefois, cette solution que l'on peut considérer comme un prototype de la solution définitive, va servir à éprouver la base et à en définir, par l'usage, les fonctionnalités et les attendus:

#### *2.4 Choix utilisateur*

Un accès personnalisé

Par programmation, un gestionnaire de comptes a été développé. Actuellement, un gestionnaire de compte avancé est en cours de développement, autorisant à terme l'enregistrement d'un historique des recherches.

Une interface d'interrogation

Là encore par programmation, il a été possible de développer un applicatif proposant de naviguer dans la base et de l'interroger par le biais d'interfaces ergonomiques et conviviales. Ces interfaces et les fonctionnalités qui s'y rattachent sont décrites plus loin.

Un système multiplateforme

Le couple Access et MapWinGIS ne fonctionne pour le moment que sur Windows. La solution choisie est donc une solution propre au monde Windows et non multiplateforme comme souhaité. À l'heure actuelle, une solution multiplateforme passerait par un développement en langage JAVA<sup>4</sup> ou par la conception d'un site Internet de type WebGIS. Les concepteurs de MapwinGIS développent actuellement une version web de leur produit ainsi qu'une version multiplateforme utilisant la technologie libre MONO<sup>5</sup>.

#### *2.5 Choix des techniques d'échange*

Une interface SIG intégré

L'emploi du module MapWinGIS remplace la connexion à un logiciel tel qu'ArcGIS (ESRI). Toutefois, celle-ci reste possible sur les postes possédant la licence d'exploitation de ce logiciel. En effet, Access a la possibilité de se connecter à d'autres logiciels par le protocole ODBC propre au monde Windows.

<sup>4</sup> JAVA: le langage Java est un langage de programmation orienté objet développé par Sun Microsystems depuis les années 90. Les logiciels écrits avec ce langage ont la particularité d'être très facilement portables sur plusieurs systèmes d'exploitation tels que UNIX, Microsoft Windows, MacOS ou Linux avec peu ou pas de modifications. C'est la plate-forme qui garantit la portabilité des applications développées en Java.

<sup>5</sup> MONO est le pendant actuel de la technologie JAVA. Cette technologie, développée par Miguel de Icaza (le créateur de l'interface GNOME) est basée sur un portage multiplateforme des développements dot.NET de Microsoft.

## Un système d'échange

Par programmation, il a été possible de développer des modules d'export de données au format souhaité. L'applicatif permet à l'utilisateur de récupérer le résultat de ses recherches sous différents formats:

- Le listing peut s'ouvrir automatiquement sous Microsoft Excel;
- La requête peut être sauvegardée dans Access sous format SQL. Cette option, réservée aux administrateurs de la base, permet de faire des tests dans le cadre du développement;
- La cartographie peut être exportée sous forme d'image au format bitmap (.bmp);
- Les données sélectionnées peuvent aussi être exportées au format SHAPE-FILE<sup>6</sup> (.shp) afin d'être intégrées dans un logiciel spécifique de cartographie avancée.

## Un système diffusable libre de droit

Actuellement, la base de données est pilotée par un SGBD commercial. Toutefois Microsoft a développé un logiciel gratuit "Microsoft Access Runtime" permettant de lire les applicatifs développés sous Access. Ce logiciel devrait être inclus dans une distribution à l'étude actuellement. La solution logicielle actuelle est donc diffusable gratuitement.

## Un accès distant par une architecture client-serveur

Pour l'instant, la solution choisie est locale, consultable dans l'unité, installée sur les ordinateurs du programme. La mise en place d'une architecture client-serveur nécessite un logiciel de gestion de base de données adapté et Microsoft Access n'est pas le logiciel idéal pour ce genre d'emploi. Toutefois, sa structure de SGBD gérant les données *via* le langage SQL confère une grande portabilité des données hébergées vers d'autres systèmes communiquant en SQL. Rappelons que le langage SQL est majoritairement employé par les SGBD qu'ils soient commerciaux comme Oracle ou libre de droit comme MySQL. Des premiers tests de portage de la base de données vers PostgreSQL<sup>7</sup> ont été réalisés avec succès.

<sup>6</sup> Le format SHAPEFILE développé par ESRI est l'un des formats les plus utilisés en géomatique. Ce format permet d'associer un fichier d'objets vectoriels ou formes (les objets affichés à l'écran) avec un fichier d'informations de nature textuelle et/ou numérique. Il permet de plus, d'associer des informations concernant la projection cartographique des données ainsi que les métadonnées de celles-ci. Le SHAPEFILE est une spécification ouverte à tous d'ESRI (ESRI 1998).

<sup>7</sup> PostgreSQL est un système de gestion de base de données relationnelle et objet (SGBDRO). C'est un outil libre disponible selon les termes d'une licence de type BSD. Ce système est concurrent d'autres systèmes de gestion de base de données, qu'ils soient libres (comme MySQL et Firebird), ou propriétaires (comme Oracle, Sybase, DB2 et Microsoft SQL Server). Comme les projets libres Apache et Linux, PostgreSQL n'est pas contrôlé par une seule entreprise, mais est fondé sur une communauté mondiale de développeurs et d'entreprises.

### 3. FONCTIONNALITÉS DE L'APPLICATIF

#### 3.1 *Accès personnalisé*

Lors du lancement de l'appliquatif, une fenêtre de connexion apparaît. Cette fenêtre liste les membres enregistrés du programme (Fig. 2). Lors de l'enregistrement d'un nouveau membre, celui-ci reçoit un identifiant de connexion ainsi qu'un code d'accès ou mot de passe. Trois types de connexion ont été prévus: 1) Administrateur, 2) Utilisateur et 3) Anonyme. Le compte Anonyme permet d'accéder aux fonctionnalités d'interrogation ainsi qu'aux fonctions d'export. Le compte Utilisateur permet d'enregistrer l'historique d'interrogation. Le compte Administrateur donne accès à des informations techniques de gestion de la base et permet d'agir sur les données (Modification, Ajout, Suppression).

Dès l'apparition de cette interface, l'utilisateur est invité à s'identifier (nom et mot de passe) afin de se connecter à la base. Une série de boutons apparaît proposant de visualiser des informations concernant les auteurs des corpus intégrés, les corpus eux-mêmes ainsi que les données.

#### 3.2 *Interrogation des données*

Le bouton "Données" de la fenêtre de navigation (Fig. 2) donne accès à la fenêtre principale (Fig. 3)<sup>8</sup>. La sélection des données se fait selon les quatre modules du schéma de données (Fig. 1), c'est-à-dire les modules "Géographique", "Chronologique", "Fonctionnel" et "Corpus". La zone centrale du formulaire représente la liste des sites sélectionnés. Cette zone se recalcule automatiquement dès modification d'un des critères de sélection.

Bien entendu, les quatre critères de sélection (Géographie/Chronologie/Fonction/Corpus) peuvent être mobilisés simultanément. Dans ce but, ils fonctionnent de façon interdépendante.

Les choix portés dans les modules d'interrogation permettent de filtrer les données. Les résultats sont représentés dynamiquement dans la zone de liste de la Fig. 3. Il est à noter que dans la liste résultante, figure pour chaque enregistrement le nom du corpus d'origine ainsi que le numéro de la fiche dans celui-ci. Le consultant pourra par la suite contacter l'auteur pour avoir accès à la donnée dans son intégralité. Enfin, lors des procédures d'export, c'est cette liste qui est envoyée vers Microsoft Excel et/ou qui constitue les données attributaires du fichier SHAPEFILE de sortie.

<sup>8</sup> Note des auteurs: dans cet article les formulaires concernant les auteurs et les corpus ne sont pas abordés car le développement n'a pas posé de problème particulier. Les auteurs ont préféré concentrer leur discours sur les interfaces de consultation et d'interrogation.

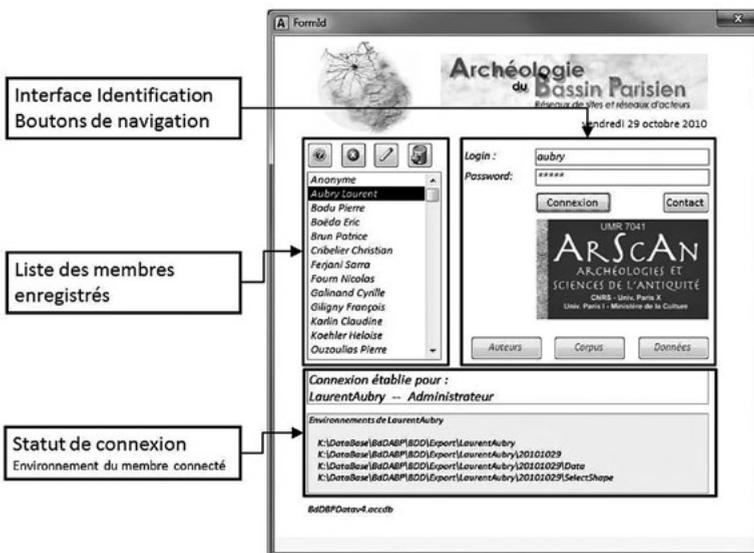


Fig. 2 – Fenêtre de gestion des comptes utilisateurs et d'accès à la base.

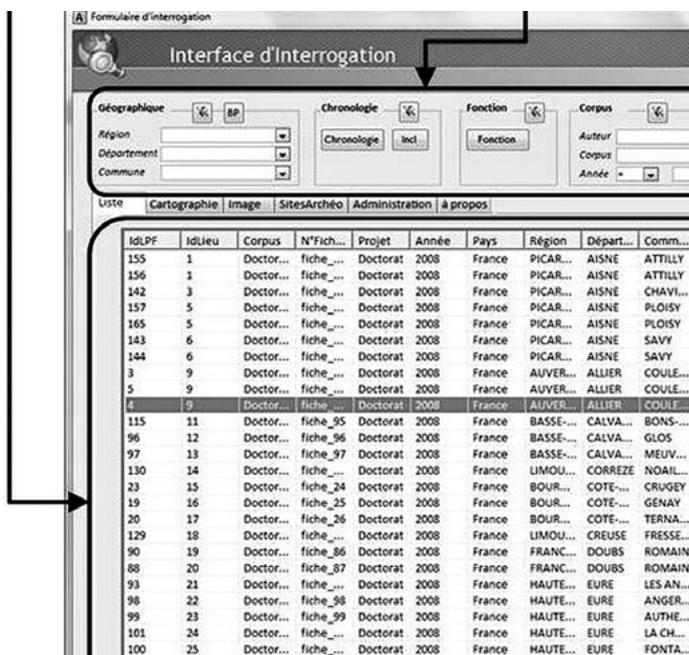


Fig. 3 – Fenêtre de l'interface d'interrogation.

### *3.3 Le module d'interrogation géographique: un référentiel administratif*

Les tables du module Géographique sont extraites des données GEOFLA de l'IGN exprimées dans le système de projection Lambert II étendu. Le module d'interrogation Géographique se compose de trois menus déroulants liés entre eux et correspondant aux choix "Région", "Département", "Commune". Dès sélection d'une région, seuls les départements et les communes de la région choisie où des enregistrements existent sont accessibles (Fig. 4).

### *3.4 Le module d'interrogation chronologique: un écueil méthodologique*

Les corpus à intégrer dans cette base de données n'ont aucune raison d'avoir la même échelle chronologique même pour des périodes similaires, notamment dans le cas de particularismes régionaux. Ainsi par exemple, un enregistrement d'un site du paléolithique pourra être défini par une culture matérielle et/ou le phénomène climatologique qui s'y rapporte. De même, l'appartenance des données à des périodes préhistoriques ou historiques entraîne une grande variation de précision et de subdivision. Enfin, chaque auteur va choisir le degré de granularité de son information en fonction de ses objectifs de recherche.

Afin de pallier les problèmes structurels liés à la diversité de la chronologie et de saisir l'information définie par chaque auteur, il a été choisi de donner à la table de référence chronologique la structure d'une table hiérarchique, i.e., une table qui a la propriété de s'appeler elle-même dans des interrogations en utilisant les propriétés de la récursivité<sup>9</sup>. Cette structure de table étant liée à elle-même, il est possible d'avoir autant de niveau de précision que nécessaire. Ainsi, tout enregistrement d'une période renverra à une période "parente" plus générale. À titre d'exemple, un enregistrement "La Tène B" renverra à "La Tène ancienne", qui renverra à "La Tène", puis à "Âge du fer", puis "Protohistoire". Avec cette structure hiérarchique, lors d'une interrogation, tous les sites "enfants" – qui dépendent de la période sélectionnée – seront pris en compte.

La structure particulière de cette table ne peut être représentée sous forme de menu déroulant, il a donc fallu le développement spécifique d'un formulaire de sélection. Ce formulaire est accessible par le bouton "Chronologie" de l'interface d'interrogation (Fig. 3). L'affichage des résultats se fait au travers d'un objet de visualisation adapté à l'aspect hiérarchique, le Treeview, qui représente l'arborescence de la structure. Pour information, cet objet de navigation graphique a été développé pour les systèmes d'exploitation afin de faciliter la navigation dans l'arborescence des répertoires et des fichiers de données (Fig. 5).

<sup>9</sup> Récursivité: en informatique et en logique, un algorithme qui contient un appel à lui-même est dit récursif.

**Formulaire d'interrogation**

**Interface d'Interrogation**

**Géographique**  BP  
 Région: BOURGOGNE  
 Département: NIEVRE  
 Commune: CHEVENON

**Chronologie**  Incl  
 Chronologie: -3000000 3000

**Fonction**  Fonction

**Corpus**   
 Auteur:   
 Corpus:   
 Année:   
 Année:

Liste: **Ca** **esArchéo** **Administration** **à propos**

IDLPF	N°Fich...	Projet	Année	Pays	Région	Départ...	Comm...	LieuDit	XL2e		
21	fiche_27	Doctorat	2008	France	BOUR...	NIEVRE	CHEVR...	Abri d...	6912		
4411	fiche_...	Doctorat	2004	France	BOUR...	NIEVRE	CHEVR...	L'Atelier	6681		
4412	fiche_...	Doctorat	2004	France	BOUR...	NIEVRE	SAINT-...	Les Tai...	6780		
777	fiche_32	Master...	2004	France	BOUR...	NIEVRE	CLAME...	Sembe...	6892		
781	fiche_33	Master...	2004	France	BOUR...	NIEVRE	LAVAU...	Fou de...	7272		
818	869	Maîtris...	fiche_34	Master...	2004	France	BOUR...	NIEVRE	MENES...	La Cha...	6692
771	870	Maîtris...	fiche_36	Master...	2004	France	BOUR...	NIEVRE	VAREN...	La Plai...	6616
1071	871	Maîtris...	fiche_35	Master...	2004	France	BOUR...	NIEVRE	VAREN...	La Gra...	6616
1510	1336	Master...	fiche_20	Master...	2007	France	BOUR...	NIEVRE	ENTRA...	---	6692
2008	1752	BiensP...	fiche_...	Master...	2003	France	BOUR...	NIEVRE	POUG...	Champ...	6581
239	186	PCR_P...	fiche_85	PCR	2005	France	BOUR...	NIEVRE	SAINT-...	---	6472
238	187	PCR_P...	fiche_84	PCR	2005	France	BOUR...	NIEVRE	SUILLY...	Les Re...	6551
202	233	PCR_P...	fiche_28	PCR	2005	France	BOUR...	NIEVRE	CHEVR...	---	6912
208	256	PCR_P...	fiche_31	PCR	2005	France	BOUR...	NIEVRE	OISY	---	6836

Fig. 4 – Fenêtre du formulaire d’interrogation. Détail des menus de sélection géographique.

**FormSelectChrono**

**Sélection Chronologique**

- Neolithique
  - Neolithique ancien
  - Neolithique moyen
    - Neolithique moyen I
      - Chambon
      - Ceiry
      - Post-Rössen
      - Rössen tardif
    - Neolithique moyen II
      - Chasséen
      - Groupe de Noyen
      - Michelsbeem

Neolithique moyen Début -4700 - Fin -3400

nb fiches : 227

Période	C.	Début	Fin
Neolithique moyen I	B.P.	4700	4200
Neolithique moyen II	B.P.	4200	3400

Début Fin Annuler Ok

Fig. 5 – Fenêtre de sélection chronologique.

Le choix de la période s'effectue à l'écran. La requête récursive appliquée à la table hiérarchique parcourt l'arborescence et sélectionne tous les sous-niveaux chronologiques quel que soit le nombre de niveaux hiérarchiques. Par exemple, dans la Fig. 5, le "Néolithique Moyen" est la période choisie. Les sous-périodes telles que "Néolithique Moyen I", "Néolithique Moyen II", "Chasséen", "Michelsberg", "Cerny", "Chambon", etc. sont automatiquement choisies. Les enregistrements de ces périodes seront donc sélectionnés.

Soulignons que ces notions de récursivités qui apportent une souplesse de modélisation des phénomènes hiérarchiques, n'existent pas de manière native dans les SGBDR bureautiques. Un développement algorithmique particulier a été effectué pour cet applicatif.

### *3.5 Le module d'interrogation fonctionnelle*

La table des fonctions s'inspire du thésaurus de la base de données Patriarche. Elle respecte donc sa structure hiérarchique. Le module d'interrogation Fonctionnelle, accessible par le bouton "Fonction" de l'interface d'interrogation (Fig. 3), a donc été construit sur le même modèle que le module de sélection chronologique (Fig. 6).

L'exemple de la Fig. 6 montre que l'occurrence "Réseau routier" a été choisi. Les occurrences, "Borne routière", "Relais routier", "Borne miliaire" et "Voie"... sont donc automatiquement intégrées dans les critères de sélection.

### *3.6 Le module d'interrogation des corpus et des auteurs*

Le module d'interrogation "Corpus" permet d'afficher toutes les données d'un corpus défini ou bien de filtrer les données par auteur. Là encore, les menus "Auteur" et "Corpus" sont liés afin d'afficher les corpus d'un auteur dès que son nom est sélectionné. Rappelons que dans la liste résultante, figure pour chaque enregistrement le nom du corpus d'origine ainsi que le numéro de la fiche dont il est issu.

### *3.7 De la base de données à la carte...*

L'expression du filtrage des données sous une forme tabulaire n'est que l'une des représentations possibles des données. L'intégration au sein de l'applicatif du module MapWinGIS permet de proposer la représentation cartographique de celles-ci (Fig. 7). Lors du développement, nous avons choisi d'associer à la représentation cartographique des outils permettant là encore d'optimiser l'ergonomie et le confort de consultation. Nous avons programmé des outils d'import de données géoréférencées (geotif, .bil, .shp), de *Zoom Avant* et *Zoom Arrière*, de *Mise à l'échelle automatique*, de *Déplacement*, ainsi qu'un gestionnaire de couches de données et ses outils spécifiques de



déplacement et suppression des couches. Une cartographie de taille réduite représente l'emprise géographique de la carte principale. Enfin, un petit éditeur de symbologie permet de changer la couleur, la taille et la transparence de la représentation des couches de données. L'objectif est de pouvoir créer, simplement et rapidement, des documents de travail.

### *3.8 ... et de la carte à la base de données*

Classiquement, lorsqu'une base de données est connectée à un logiciel SIG bureautique, l'aspect cartographique n'intervient que pour cartographier le résultat de la requête, i.e., elle n'intervient qu'en fin de processus de définition des critères de sélection. La cartographie n'est alors qu'un support visuel ne participant pas directement à l'interrogation de la base de données.

L'implémentation du module de fonctions cartographiques directement au sein du SGBD autorise l'utilisation de la carte comme moyen d'interrogation de la base de données, au même titre que les modules sémantiques vus précédemment. Les modalités d'interrogation cartographiques sont compatibles avec les modules de sélection et peuvent donc être panachées à l'envie.

### *3.9 Saisie de formes géométriques*

Le module cartographique possède des fonctionnalités de digitalisation développées par les auteurs qui permet la saisie de polygones. Le dessin effectué à l'écran est transcrit en SQL avant d'être injecté dans la base de données. Les enregistrements, situés dans l'emprise des polygones tracés par l'utilisateur et/ou des formes géométriques prédéfinies et paramétrées par l'utilisateur, sont sélectionnés (Fig. 8).

### *3.10 Utilisation du format SHAPEFILE*

L'applicatif peut charger des fichiers de données de divers types et notamment du format SHAPEFILE. Ces données permettent d'enrichir visuellement la cartographie mais aussi, quand il s'agit de données surfaciques, apportent des possibilités avancées de sélection.

### *3.11 Utilisation du format SHAPEFILE: la saisie et l'extraction de formes*

Lorsque l'on souhaite filtrer les données selon des zones complexes et/ou précises, la saisie à l'écran peut se révéler vite malaisée. On peut alors utiliser un fichier de données – ou des formes le constituant – comme filtre. Cette fonctionnalité a fait l'objet d'un développement spécifique de sélection – et d'extraction – de formes élémentaires d'un fichier de référence. Les enregistrements compris dans la sélection cartographique seront alors sélectionnés (Fig. 9).

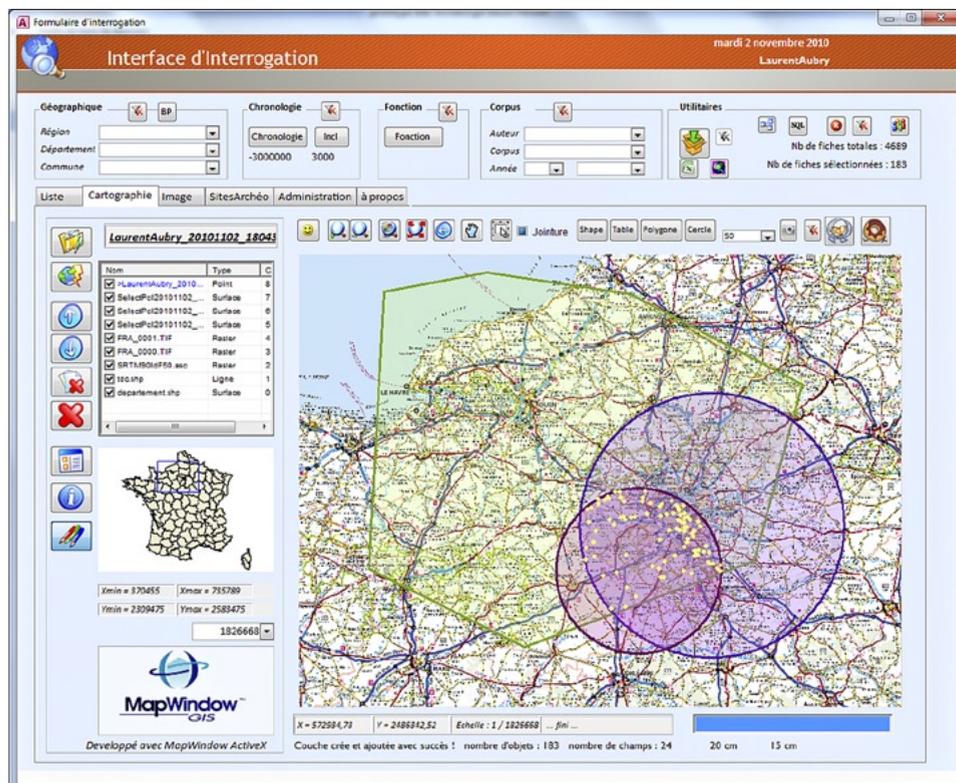


Fig. 8 – Fenêtre de l’interface d’interrogation. Sélection géographique par formes géométriques.

### 3.12 Utilisation du format SHAPEFILE: exploitation de la table attributaire

Dans le cas précédent, les données ont été sélectionnées en se basant sur la géométrie des formes choisies du fichier de référence. Il peut être plus judicieux, lorsque le fichier choisi est complexe, de sélectionner les enregistrements non pas par la géométrie des formes mais par les valeurs attributaires de celles-ci. Lorsque cette option est choisie, on définit le fichier ou la couche de données de référence, le champ attributaire sur lequel va porter l’interrogation et la valeur de ce champ (Fig. 10). Cette valeur est soit définie par un menu constitué des occurrences du champ, soit par une recherche libre. Le système extrait les formes choisies et les importe dans un fichier SHAPEFILE qui servira de filtre cartographique des données (Fig. 11).

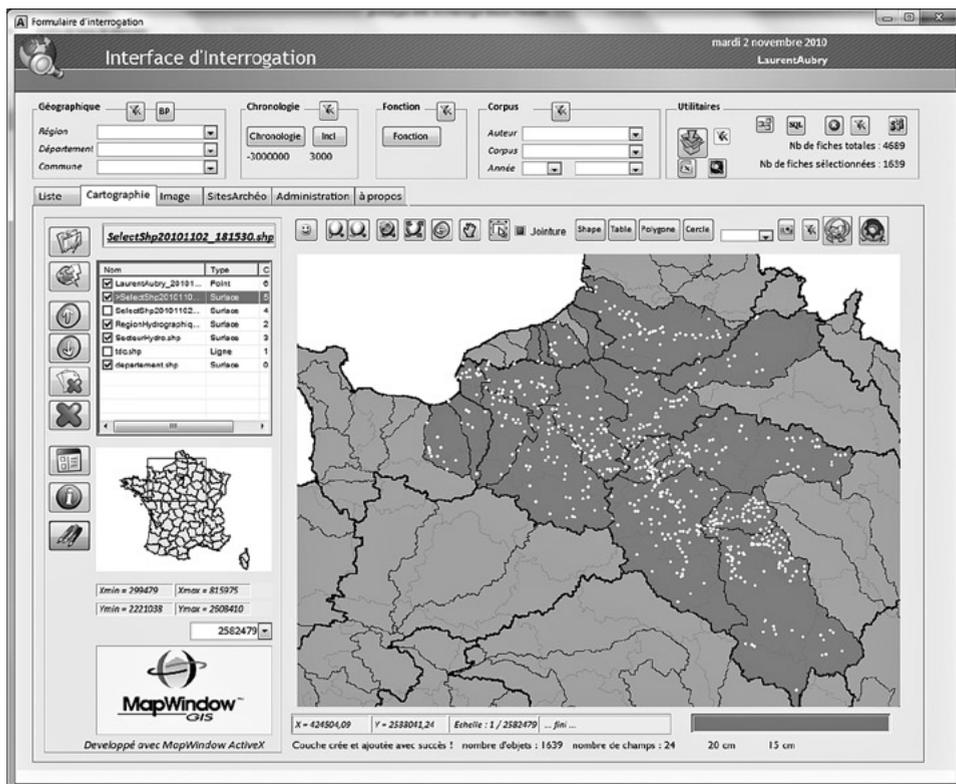


Fig. 9 – Sélection des sites présents dans les formes des secteurs hydrologiques sélectionnés à l'écran.

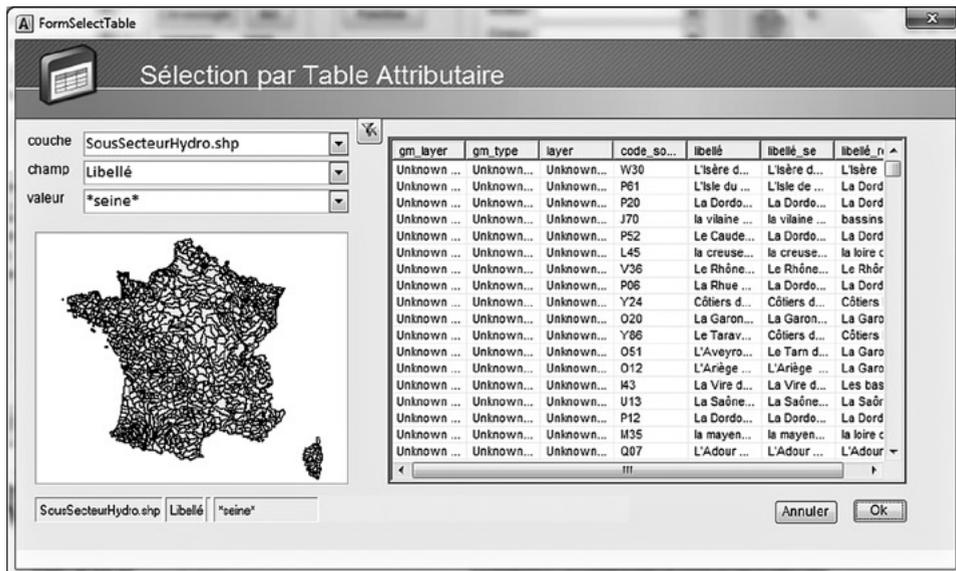


Fig. 10 – Descriptif du module de sélection des données attributaires.

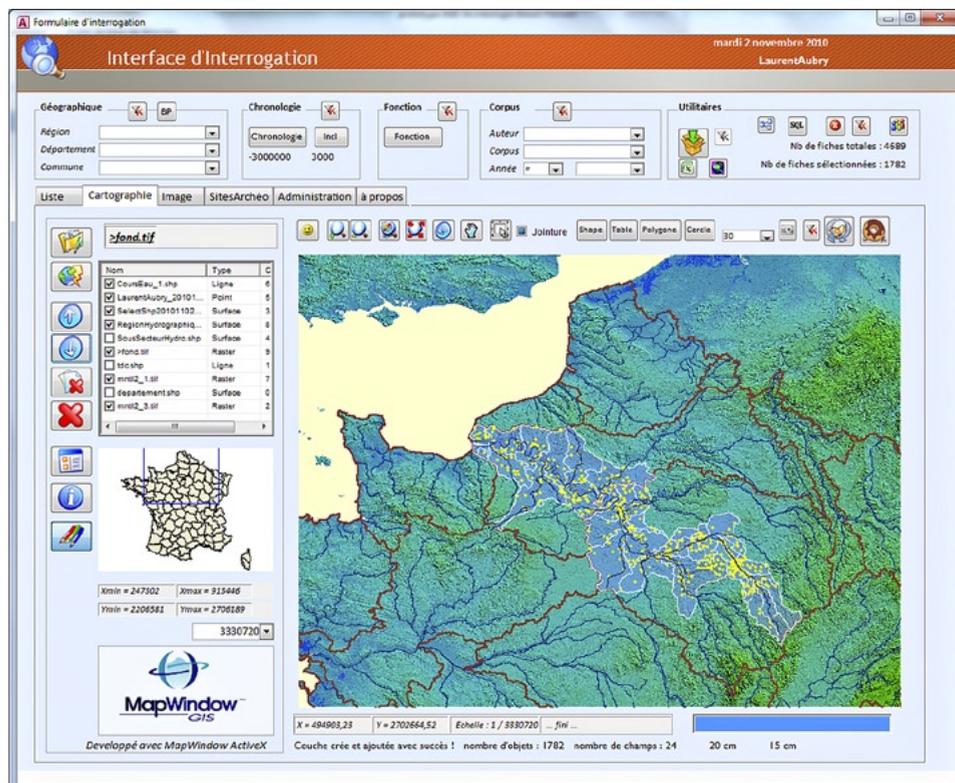


Fig. 11 – Sélection des sites présents dans les formes des secteurs hydrologiques ayant le terme «seine» présents dans le champ «libellé» de la table attributaire.

### 3.13 Utilisation du format SHAPEFILE: un modèle de données évolutif – la jointure spatiale

Lors de l'ajout d'un fichier au format SHAPEFILE, l'aspect graphique est représenté, et s'il est de type surface, le fichier DBF<sup>10</sup> associé est automatiquement importé dans la base de données au même titre qu'une table externe et est connecté au module Général par un lien décrivant, pour chaque site, la relation d'inclusion dans une ou plusieurs formes du fichier. Grâce à son architecture en étoile, le schéma de la base de données est donc capable de se modifier et d'évoluer en prenant en compte les données géographiques ajoutées

<sup>10</sup> L'information attributaire des formes contenues dans un fichier SHAPEFILE est portée par un fichier au format DBF, format de la base de données dBase.

Le système d'information du programme "Archéologie du Bassin parisien"

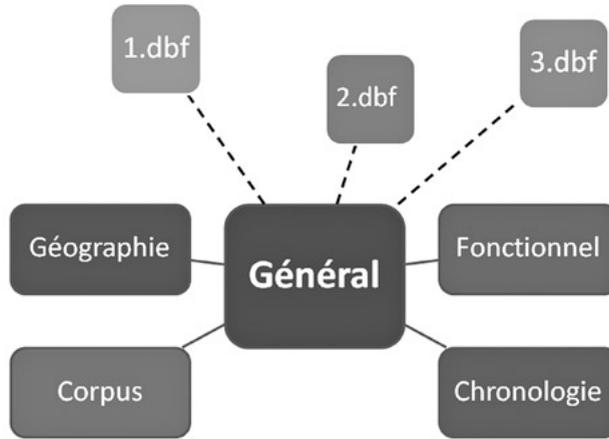


Fig. 12 – Schéma de jointure des fichiers DBF au modèle de données.

Formulaire d'interrogation

Interface d'interrogation

mercredi 3 novembre 2010  
LaurentAubry

Géographique [BP] Chronologie [Incl.] Fonction [Fonction] Corpus [ ]

Région [ ] Département [ ] Commune [ ]

Chronologie [ ] Incl. [ ]

Fonction [ ]

Corpus [ ]

Auteur [ ]

Corpus [ ]

Année [ ]

UTILITAIRES

Nb de fiches totales : 4689

Nb de fiches sélectionnées : 1787

Liste Cartographie Image SitesArchéo Administration à propos

surface	zmin	zmax	zchl	Fonction	Début	Fin	DateD...	DateFin	code...	libellé	libellé_se
758	97	134	125	Indéte...	Weich...	Weich...	115000	75000	G72	Les bassins côtiers compris entre l'emb...	Le Valmont ...
758	97	134	125	Indéte...	Paléoli...	Paléoli...	130000	30000	G72	Les bassins côtiers compris entre l'emb...	Le Valmont ...
627	0	114	105	Dépôt	Bronze...	Bronze...	-1500	-1350	H31	L'Epte de sa source au confluent de la S...	La Seine du C...
208	5	102	94	Dépôt	Bronze...	Bronze...	-930	-800	H31	L'Epte de sa source au confluent de la S...	La Seine du C...
1031	10	119	96	Dépôt	Bronze...	Bronze...	-1600	-1500	G72	Les bassins côtiers compris entre l'emb...	Le Valmont ...
901	67	166	141	Dépôt	Bronze...	Bronze...	-1600	-1350	H31	L'Epte de sa source au confluent de la S...	La Seine du C...
1543	94	143	115	Dépôt	Bronze...	Bronze...	-1500	-1350	H31	L'Epte de sa source au confluent de la S...	La Seine du C...
1192	77	156	135	Dépôt	Bronze...	Bronze...	-1500	-1350	H31	L'Epte de sa source au confluent de la S...	La Seine du C...
946	10	139	20	Dépôt	Bronze...	Bronze...	-1500	-1350	H31	L'Epte de sa source au confluent de la S...	La Seine du C...
1667	47	142	59	Dépôt	Bronze...	Bronze...	-1500	-1350	H31	L'Epte de sa source au confluent de la S...	La Seine du C...
460	31	134	39	Dépôt	Bronze...	Bronze...	-1600	-1500	H31	L'Epte de sa source au confluent de la S...	La Seine du C...
1667	47	142	59	Dépôt	Bronze...	Bronze...	-1500	-1350	H31	L'Epte de sa source au confluent de la S...	La Seine du C...
1759	118	211	155	Dépôt	Bronze...	Bronze...	-930	-800	H31	L'Epte de sa source au confluent de la S...	La Seine du C...
1673	71	207	191	Dépôt	Bronze...	Bronze...	-1500	-1350	H31	L'Epte de sa source au confluent de la S...	La Seine du C...
800	2	82	6	Dépôt	Bronze...	Bronze...	-1500	-1350	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
800	2	82	6	Dépôt	Bronze...	Bronze...	-1600	-1350	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
2138	2	152	15	Dépôt	Bronze...	Bronze...	-1500	-1350	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
2138	2	152	15	Dépôt	Bronze...	Bronze...	-1500	-1350	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
2138	2	152	15	Dépôt	Bronze...	Bronze...	-1500	-1350	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
2138	2	152	15	Dépôt	Bronze...	Bronze...	-930	-800	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
955	2	131	5	Dépôt	Bronze...	Bronze...	-1500	-1350	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
794	44	171	98	Dépôt	Bronze...	Bronze...	-1500	-1350	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
2080	39	169	152	Indéte...	Stade l...	Stade l...	115000	25000	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
1019	1	129	125	Dépôt	Bronze...	Bronze...	-1500	-1350	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
507	65	162	155	Indéte...	Paléoli...	Paléoli...	300000	30000	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
1761	2	138	115	Dépôt	Bronze...	Bronze...	-1500	-1350	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
251	80	157	144	Indéte...	Paléoli...	Paléoli...	300000	30000	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
636	4	75	20	Indéte...	Paléoli...	Paléoli...	300000	30000	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
794	44	171	98	Indéte...	Paléoli...	Paléoli...	300000	30000	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
1761	2	138	115	Dépôt	Bronze...	Bronze...	-1150	-930	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...
1671	55	172	157	Indéte...	Stade l...	Stade l...	115000	75000	H50	La Seine du confluent de l'Eure (exclu) ...	La Seine du C...

Fig. 13 – Interface d'interrogation. Les informations ajoutées par jointure spatiale sont affichées sur fond orange.

lors de l'utilisation de la base de données (Fig. 12), i.e., chaque fichier DBF représente une dimension supplémentaire d'analyse.

Cette jointure effectuée sur la spatialité des données permet d'ajouter aux enregistrements de la base les informations attributaires de couches de référence (Fig. 13). Ces données "augmentées" et "augmentables" à volonté sont récupérables par l'utilisateur selon les modalités d'export vues précédemment.

#### 4. RÉFLEXIONS SUR L'OUTIL ET SON DEVENIR

##### 4.1 Type de l'applicatif

En 2006, Dan Ames, initiateur et porteur du projet MapWinGIS, rappelle les trois principaux types de développement possible d'outil logiciel SIG: 1) des modules ou plug-ins venant enrichir un logiciel SIG existant, 2) des sites web autorisant la visualisation et enfin 3) le développement d'applications autonomes adaptées aux besoins. Une synthèse critique récente de ces applications autonomes ou "StandAlone" a été proposée par STEINIGER et BOCHER (2009). Ces applications sont souvent des versions allégées de logiciels propriétaires présentant des bibliothèques de fonctions propres à des disciplines scientifiques (hydrologie, géographie physique, etc.). L'applicatif décrit ici ne s'inscrit que partiellement dans les premier et dernier cas proposés par D. AMES, car il correspond à un développement original adapté aux besoins d'un programme de recherche, qui ajoute des fonctionnalités à un logiciel – SGBD et non SIG – existant. Présentant des fonctionnalités liées, propres aux mondes des SGBD et de SIG, cet applicatif n'est ni l'un ni l'autre *stricto sensu*, mais une sorte de "chimère" logicielle.

Il s'inscrit toutefois dans l'esprit de l'OGC<sup>11</sup> qui tend à promouvoir les bases de données spatiales, bases de données qui respectent les spécifications de l'OGC du SQL autorisant les requêtes spatiales et qui intègrent les types d'objets géographiques comme n'importe quel type de format de données (SHERMAN 2008). Le SGBDR spatial le plus abouti est actuellement PostGIS<sup>12</sup> développé par Refraction Research. Le nombre important de communications lui étant consacré, lors de la dernière conférence Free Open Source Software for Geography (FOSS4G 2010)<sup>13</sup> organisée par OSGEO<sup>14</sup>, montre la vivacité de la communauté des développeurs et des utilisateurs. La nouvelle version de PostGIS doit intégrer les données raster (ARÉVALO 2010; RACINE 2010) et s'orienterait vers la 3D (COURTIN 2010).

<sup>11</sup> OpenGIS Consortium définition <http://www.opengeospatial.org/>.

<sup>12</sup> <http://www.refractorions.net/>.

<sup>13</sup> <http://2010.foss4g.org/>.

<sup>14</sup> <http://www.osgeo.org/>.

## 4.2 Développement de l'applicatif

Le pari d'intégrer dans le développement de l'applicatif des technologies SIG a été dicté par 1) la nature des objets manipulés, i.e., des sites archéologiques nécessitant l'élaboration de cartographies thématiques 2) et aussi par la volonté d'inscrire ce projet dans le contexte actuel de l'émergence des bases de données spatiales, et d'associer la cartographie au processus d'interrogation des données.

Le développement du prototype a permis de mieux appréhender les contraintes tant techniques que méthodologiques et de mieux comprendre les enjeux liés à telle ou telle technologie. Il a aussi permis aux auteurs de s'intégrer dans la communauté des développeurs et d'utilisateurs de MapwinGIS, la solution choisie, de recueillir son soutien informatique mais aussi – et surtout – de participer aux débats concernant les choix de développement *via* des forums de discussion dédiés et aussi lors de la première rencontre internationale (AUBRY 2010). Cet échange est l'apanage des logiciels libres.

Le développement de l'applicatif s'appuiera sur le moteur de bases de données spatiales PostGIS, hébergé sur un serveur dédié, proposant une connexion distante accessible *via* un applicatif développé en C++ construit autour de la technologie MapwinGIS et/ou *via* un WebGIS développé en php.

Ces développements futurs dépendent de contraintes techniques mais aussi des choix et des évolutions du programme de recherche. Toutefois, en l'état, l'applicatif est opérationnel et les fonctionnalités – qui ont nécessité des développements spécifiques – sont utilisées dans d'autres programmes de recherche.

LAURENT AUBRY

UMR 7041 Archéologies et Sciences de l'Antiquité  
Équipe Protohistoire européenne

SARRA FERJANI

Université de Paris 1 Panthéon-Sorbonne  
UMR 7041 Archéologies et Sciences de l'Antiquité  
Équipe Protohistoire européenne

## *Remerciements*

Les auteurs tiennent à remercier les modérateurs, gestionnaires et intervenants des sites Internet de Développez.com (<http://www.developpez.com/>) et de MapWinGIS (<http://www.mapwindow.org/>), véritables mines d'informations et grandes sources d'inspiration. Ils remercient tout particulièrement Dan Ames et Paul Meems pour leurs conseils précieux.

## BIBLIOGRAPHIE

AMES DANIEL P. 2006, *Getting started with the MapWinGIS ActiveX Control* (<http://www.mapwindow.org/>) [consulté 26 Apr 2011].

- AMES DANIEL P. 2007a, *Map WinGIS Reference Manual: A function guide for the free Map Window GIS ActiveX map component* (<http://www.mapwindow.org/>) [consulté 26 Apr 2011].
- AMES DANIEL P., MICHAELIS C., DUNSFORD T. 2007b, *Introducing the Map Window GIS project*, «OSGeo Journal», 2, 2007.
- ARÉVALO J. 2010, *PostGIS WKT Raster. An Open Source alternative to Oracle GeoRaster* (FOSS4G 2010, 6-9 September 2010), Barcelona.
- AUBRY L. 2010, *A joined approach of geomatics and database management in archaeological research. The program "Archaeology in the Paris Basin": a case study*, in *1st International Conference of Map Window Users and Developers, Orlando USA (31 Mars-2 Avril 2010)*.
- AUBRY L., FERJANI S., QUILLIEC B. 2008, *Etat de la base de données du Programme*, in P. BRUN, P. SOULIER, *Rapport du programme "Archéologies du Bassin parisien", pour les années 2006-2007*, UMR 7041 Archéologies et sciences de l'Antiquité, Nanterre, Maison René Ginouvès.
- AUBRY L., FERJANI S. 2009, *La base de données "Archéologie du Bassin parisien". Vers un système d'information aux fonctionnalités avancées*, in P. BRUN, P. SOULIER, *Rapport du programme "Archéologies du Bassin parisien, pour l'année 2008"*, UMR 7041 Archéologies et sciences de l'Antiquité, Nanterre, Maison René Ginouvès.
- BRUN P., SOULIER P. 2007, *Archéologie du Bassin parisien. Rapport pour les années 2004-2005*, UMR 7041 Archéologies et sciences de l'Antiquité, Nanterre, Maison René Ginouvès.
- CELKO J. 2005, *Joe Celko's. SQL for Smarties. Advanced SQL programming*, Elsevier, Morgan Kaufmann Publishers.
- CELKO J. 2006, *Joe Celko's. Analytics and OLAP*, Elsevier, Morgan Kaufmann Publishers.
- CHURCHER C. 2008, *Beginning SQL Queries. From Novice to Professional*, New York, Apress.
- COURTIN O. 2010, *PostGIS meet the third Dimension*, in FOSS4G 2010, (Barcelona, 6-9 September 2010) ([http://2010.foss4g.org/presentations\\_show.php?id=3556](http://2010.foss4g.org/presentations_show.php?id=3556)) [consulté 26 Apr 2011].
- ESRI 1998, *ESRI shapefile Technical Description. An ESRI White paper* (<http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>) [consulté 26 Apr 2011].
- FEDDEMA H. 2007, *Access 2007 VBA Bible*, Indianapolis, Wiley Publishing.
- GNU 2010, *The Free software definition* (<http://www.gnu.org/philosophy/free-sw.html>) [consulté 26 Apr 2011].
- HENNING T., COOPER R., GRIFFITH G., STEIN A. 2007, *Access 2007 VBA Programmer's Reference*, Indianapolis, Wiley Publishing.
- JANSEN M., AVEN N. 2010, *What you type is what you see*, in FOSS4G 2010 (Barcelona, 6-9 September 2010) ([http://2010.foss4g.org/presentations\\_show.php?id=3630](http://2010.foss4g.org/presentations_show.php?id=3630)) [consulté 26 Apr 2011].
- RACINE P. 2010, *Introducing PostGIS WKT raster. Seamless raster/vector operations in a spatial database*, in FOSS4G 2010 (Barcelona, 6-9 September 2010) ([http://2010.foss4g.org/presentations\\_show.php?id=3221](http://2010.foss4g.org/presentations_show.php?id=3221)) [consulté 26 Apr 2011].
- SHERMAN G.E 2008, *Desktop GIS, Mapping the Planet with Open Source Tools*, Pragmatic bookshelf.
- STEINIGER S., BOCHER E. 2009, *An overview on Current Free and Open Source desktop GIS Developments*, «International Journal of Geographical Information Science», 23-10, 1345-1370.
- STEINIGER S., HAY G. 2009, *Free and opensource geographic information tools for landscape ecology*, «Ecological informatics», 4, 183-195.

ABSTRACT

One of the goals of the research program "Archaeology of the Paris Basin" has been to establish an Archaeological Information System, i.e., a meta-database of the archaeological sites studied in the geographic area of the program, from the Paleolithic to modern days, in order to produce thematic maps using appropriate software. This paper aims to present this meta-database, but also to consider the technical constraints and the methodological choices involved in its implementation. The development of the query interface of this database represented an opportunity to reconsider the initial geomatic approach. This application is based on Open Source technology, and is both appropriate and adapted to the objectives of the program. It proposes an original synthesis of DBMS and GIS functionalities. Its development is in line with the rereading of the discipline paradigms suggested – or imposed – by the rise of the geomatic Open Source.

