



**HAL**  
open science

## **ARSOL. La chaîne de gestion des données de fouilles du Laboratoire Archéologie et Territoires**

Henri Galinié, Philippe Husi, Xavier Rodier, Christian Theureau, Elisabeth  
Zadora-Rio

► **To cite this version:**

Henri Galinié, Philippe Husi, Xavier Rodier, Christian Theureau, Elisabeth Zadora-Rio. ARSOL. La chaîne de gestion des données de fouilles du Laboratoire Archéologie et Territoires. Les petits cahiers d'Anatole, 2005, 17, pp.36772 signes. halshs-00005416

**HAL Id: halshs-00005416**

**<https://shs.hal.science/halshs-00005416>**

Submitted on 9 May 2011

**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.

# **LES PETITS CAHIERS D'ANATOLE**

**n°17, mai 2005**

## **ARSOL**

La chaîne de gestion des données de fouilles  
du Laboratoire Archéologie et Territoires

## **CITERES**

### **LABORATOIRE ARCHEOLOGIE ET TERRITOIRES**

UMR 6173  
CNRS – Université de Tours  
3, place Anatole France, 37000 Tours  
[lat@univ-tours.fr](mailto:lat@univ-tours.fr)

<http://www.univ-tours.fr/lat>



## ARSOL

### La chaîne de gestion des données de fouilles du Laboratoire Archéologie et Territoires

*Arsol*

*The System in use at the Laboratoire Archéologie et Territoires  
for processing excavations data.*

**Henri GALINIE, Philippe HUSI, Xavier RODIER, Christian THEUREAU, Elisabeth ZADORA-RIO**  
UMR 6173 CITERES, Laboratoire Archéologie et Territoires

**Mots-clefs :** archéologie, enregistrement de terrain, stratigraphie, données, spatialisation

**Key-words :** *archaeology, site records, stratigraphy, data, spatialisation*

**Référence bibliographique :** GALINIE H., HUSI P., RODIER X., THEUREAU C., ZADORA-RIO E., ARSOL, La chaîne de gestion des données de fouilles du Laboratoire Archéologie et Territoires, Les petits cahiers d'Anatole, n° 17, 27/05/2005, 36772 signes, [http://www.univ-tours.fr/lat/pdf/F2\\_17.pdf](http://www.univ-tours.fr/lat/pdf/F2_17.pdf)

Le système utilisé par le Laboratoire Archéologie et Territoires, *ArSol* (pour Archives du sol), pour l'exploitation des données des fouilles conduites par ses membres comporte aujourd'hui trois modules principaux qui concernent :

- les données stratigraphiques et mobilières dans l'optique d'une analyse des sites fouillés (*Archives de Fouilles*) et leur spatialisation,
- les données issues de l'étude de la céramique à des fins de datation des sites et d'étude intrinsèque des produits céramiques *BaDoC* (Base de données céramique),
- les données issues de l'étude des squelettes humains à des fins d'analyse du fonctionnement des lieux funéraires et d'étude biologique des populations *BaDoA* (Base de données anthropologie).

Ces trois volets, présentés ici, seront suivis d'autres à l'avenir concernant d'autres catégories de matériel archéologique.

## Archives de Fouilles, BaDoC et BaDoA

Le système de gestion des données de fouilles a d'abord été élaboré en 1990 pour la fouille de Rigny (Indre-et-Loire) (Zadora-Rio 1994, Zadora-Rio, Galinié *et al.* 2001) sous le nom d'*Archives de Fouilles* pour l'enregistrement stratigraphique ; comme tel, il a ensuite été utilisé et adapté à d'autres chantiers de fouilles conduits par le LAT à Tours (Galinié, Husi, Rodier 2001) et à Chinon (Lorans à paraître), par l'AFAN à Tours et à Poitiers (Jouquand *et al.* 1999) et plus récemment par le Service archéologique d'Indre-et-Loire à Chinon, en 2004.

*Archives de fouilles* est fondé sur le système d'enregistrement stratigraphique manuel utilisé au préalable pour les fouilles de Tours (Galinié 1980, Galinié, Randoïn 1987). Construit comme un système ouvert, c'est-à-dire modulable et surtout non contraint par des thésaurus intégrés, il peut être utilisé pour toutes fouilles stratigraphiques.

Conçu sur le logiciel 4e Dimension<sup>1</sup>, il a été progressivement intégré à partir de 1998 à une chaîne de traitement de l'information graphique mise au point par Xavier Rodier et qui consiste à associer les relevés numériques de terrain aux données stratigraphiques dans un système d'information géographique. Ce système permet d'une part la gestion des données de terrain et leur exploitation post-fouille (analyse stratigraphique et spatiale) et d'autre part l'intégration des données utiles à des échelles plus petites dans d'autres systèmes d'information. Par exemple, les données des fouilles de Tours sont intégrées dans ToToPi, Système d'Information Géographique pour l'étude de la Topographie historique de Tours Pré-industriel (Rodier 2000, Galinié, Rodier 2001, 2002). ToToPi est également utilisé dans le cadre du programme d'ACR *Tours antique* entre l'INRAP et le LAT. A l'échelle de la fouille cela permet par exemple de consulter l'ensemble de l'information à partir des données planimétriques, de constituer automatiquement des plans selon les découpages chronologiques inclus dans la base stratigraphique, ou encore de cartographier des répartitions de mobilier.

La céramique est étudiée dans la base *BaDoC* réalisée par Philippe Husi. Ce module permet de traiter les données céramiques de toutes fouilles stratigraphiques, quelque soit le faciès typologique et la chronologie du site. Ce système a une vocation de gestion de la céramique mais surtout de recherche, puisque les principaux développements concernent, outre la typologie, la quantification en fonction des échelles d'analyse archéologique, ainsi que la datation. L'interface de cette base est suffisamment aboutie pour que des céramologues peu experts en informatique puissent assez rapidement s'en servir. Cet outil a été utilisé pour toutes les études céramologiques du Laboratoire Archéologie et Territoires, y compris dans le cadre de diplômes universitaires (maîtrises, thèses) (Galinié, Husi, Rodier 2001 ; Husi 2001 ; Husi *dir.* 2003 ; Husi, Bellanger 2003 ; Lorans à paraître). Ce système a

---

<sup>1</sup>4e Dimension est produit par la société 4D, <http://www.4d.fr/>

également été utilisé pour l'étude de la céramique de fouilles préventives de l'AFAN à Tours et à Poitiers (Jouquand *et al.* 1999).

Les données anthropologiques sont analysées dans la base *BaDoA* réalisée par Christian Theureau. L'analyse anthropologique vise d'abord à définir des éléments d'identité individuelle (sexe, âge) directement utiles à l'interprétation archéologique, par exemple lors d'analyses de répartitions spatiales. Dans son propre domaine, elle s'attache à décrypter les situations dont le squelette a conservé des traces, en particulier les variations diachroniques ou mésologiques qui, étendues à la population, sont confirmées en fonction de l'importance des effectifs examinés.

Le logiciel choisi, 4<sup>e</sup> dimension, présente les avantages suivants :

- il s'agit d'un SGBD robuste permettant de gérer des structures complexes et possédant un langage de développement ;
- il offre plusieurs niveaux d'accès selon les compétences de l'utilisateur : interface personnalisée (menus créés), utilisation avancée (utilisation directe), programmation/développement (structure) ;
- il permet de déployer une application selon une architecture client/serveur et multi plates-formes (Mac/PC) ;
- il permet de créer une application autonome ;
- il permet d'ouvrir une base de données sur internet.

## 1. Archives de Fouilles, l'enregistrement des données du terrain

La structure de la base de données est représentée par la fig. 1. Elle comprend treize tables (cf. fig. 1, n° 1 à 13). Les tables 1 à 9 comprennent l'enregistrement des données du terrain et de l'analyse du mobilier, des données anthropologiques extraites de *BaDoA* et le résumé des informations céramiques extraites de *BaDoC* ; les tables 10 à 13 contiennent l'enregistrement des archives de la fouille. A l'intérieur des cadres qui délimitent chacune des tables, les lignes correspondent aux titres des champs (c'est-à-dire aux titres des colonnes des tableaux de données).

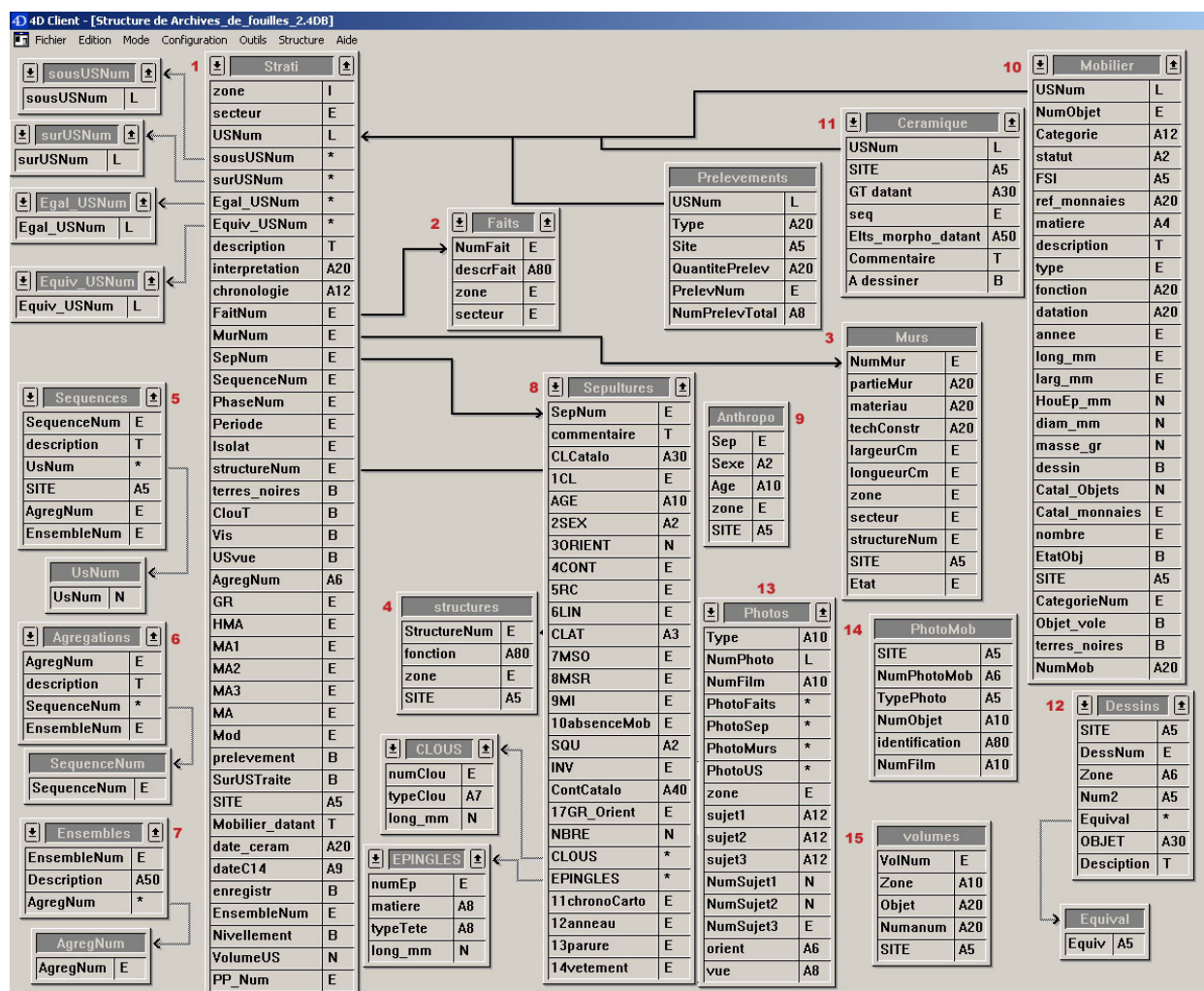


Figure 1

## **Strati** (fig. 1, n° 1)

Les données stratigraphiques sont enregistrées sur le terrain et l'enregistrement vérifié, et le cas échéant corrigé et complété, pendant l'étude post-fouille.

Chaque enregistrement de cette table correspond à une Unité Stratigraphique (ou US) repérée par son n°. Chaque fiche comprend la description de l'US, son interprétation et sa position dans la chronologie relative : par exemple US 3265 sous (=avant) US 3450 et sur (=après) US 3150. On indique aussi les relations d'égalité (même strate ayant reçu deux n° d'US dans deux secteurs différents de la fouille) ou d'équivalence (même strate dont la continuité physique a été interrompue et qui pour cette raison a reçu des n° d'US différents). Il est possible d'indiquer autant de relations qu'on le souhaite : tous les enregistrements "sous" et "sur" une US donnée sont regroupés dans des sous-tables attachées à l'enregistrement principal ; il en va de même pour les US "égalisées" ou "équivalentes" (fig. 1, n° 1a, 1b, 1c, 1d).

Les Unités Stratigraphiques, qui constituent les éléments de base de l'enregistrement, peuvent être regroupées dans des entités spatiales de quatre types, correspondant chacun à une table particulière de la base : les Faits, les Sépultures, les Murs et les Structures (fig.2) :

- **Faits** (fig. 1 n° 2) contient tous les trous de poteaux, fosses, tranchées etc (identifiés par leur numéro), leur description et, s'il y a lieu, la Structure à laquelle ils appartiennent.

- **Sépultures** (fig. 1 n° 8) contient toutes les sépultures (identifiées par leur n°) avec leur description (cf. *infra*).

- **Murs** (fig. 1, n° 3) contient tous les murs (identifiés par leur n°) avec leur description (matériau, technique de construction, dimensions) et le n° de la structure à laquelle ils appartiennent.

- **Structures** (fig. 1 n° 4) contient toutes les Structures (identifiées par leur n°) avec leur fonction.

Ces quatre tables sont liées à la table Strati. Cela signifie que chaque fois qu'on enregistre un nouveau n° de Fait, de Sépulture, de Mur ou de Structure dans Strati, la table liée correspondante s'ouvre automatiquement et présente un nouvel enregistrement à saisir. La saisie dans la table Strati et dans les quatre tables liées est faite au cours de la fouille, ce qui permet de mieux contrôler l'enregistrement sur le terrain.

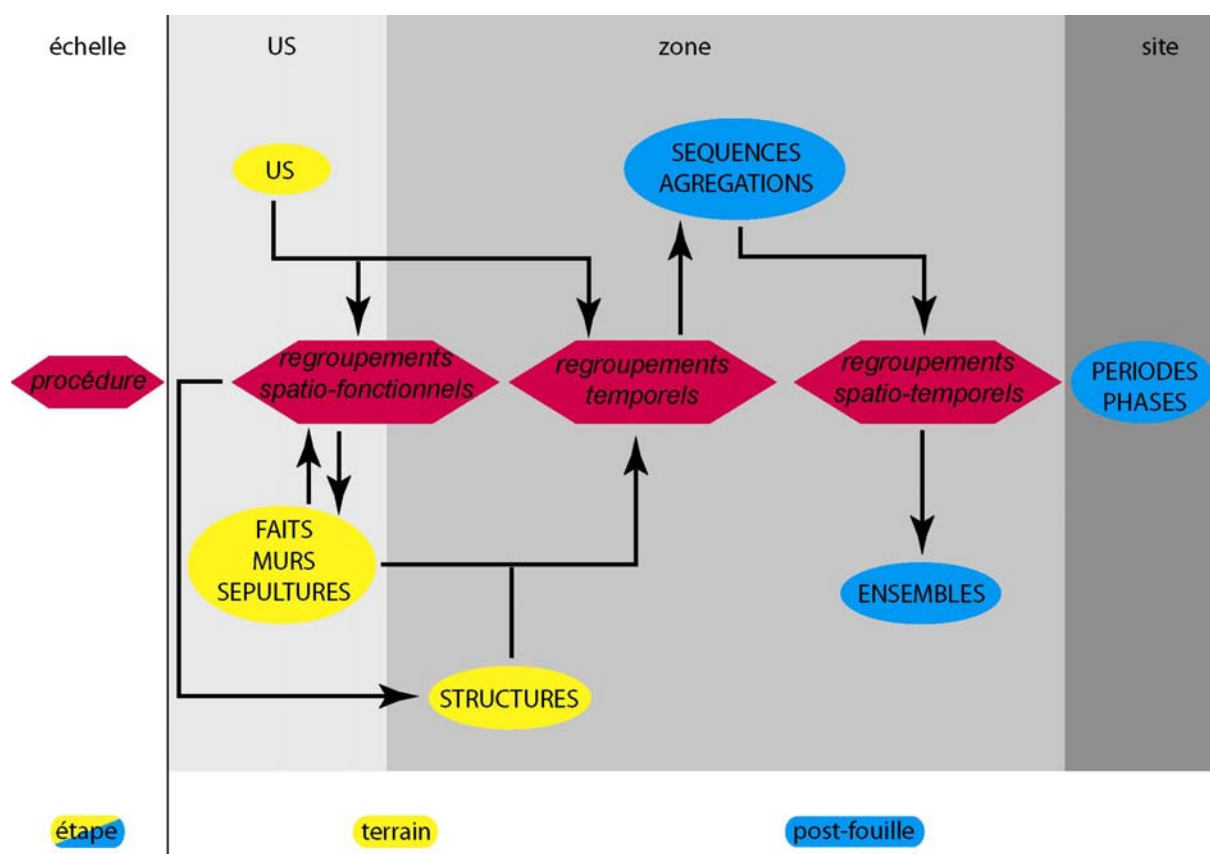


Figure 2

Les Unités Stratigraphiques sont ensuite regroupées, pour les besoins de l'analyse, dans des entités chronologiques : les Séquences, elles-mêmes regroupées en Agrégations, réunies à leur tour dans des entités de rang supérieur qui sont les Ensembles dont la valeur peut être à la fois chronologique et spatiale (fig.2).

- **Séquences** (fig. 1 n° 5) contient toutes les Séquences (identifiées par leur n°) avec leur définition.

- **Agrégations** (fig. 1 n° 6) contient toutes les Agrégations (identifiées par leur n°) avec leur définition.

- **Ensembles** (fig. 1 n° 7) contient tous les Ensembles (identifiés par leur n°) avec leur définition.

Pour chacune de ces entités, les identifiants des unités du niveau de regroupement inférieur sont saisis dans des sous-tables utilisées ensuite pour reporter sur les fiches stratigraphiques (fig. 1, n° 1) les n° de Séquence, Agrégation, Ensemble. Cela permet d'une part d'éditer des tableaux de mises en séquences et de consulter les éléments de datation associés à chacun de ces regroupements.



Les informations concernant la périodisation du site et ses subdivisions en phases peuvent être saisies selon le même mode opératoire ou simplement intégrer dans les fiches d'US à partir de sélections sur les n° d'Agrégations ou d'Ensembles.

Si les termes Séquence, Phase et Période appartiennent au jargon commun des archéologues, ce n'est pas le cas des termes Agrégation et Ensemble. Ces deux niveaux ont été introduits à Rigny pour répondre aux besoins de l'analyse du cimetière moderne et contemporain, afin de regrouper des éléments dont la contemporanéité était fondée sur des critères autres que stratigraphiques. Agrégation et Ensemble ont ensuite été substitués à Phase pour l'analyse des stratifications urbaines à Tours, parce que ces termes rendent mieux compte des réalités stratigraphiques tronquées et ne renvoient pas à des subdivisions chronologiques continues comme le font les Phases dans une même Période.

Ces différents niveaux de regroupement sont modulables en fonction des besoins spécifiques de chaque fouille.

### **Sépultures** (fig. 1, n° 8)

Cette table reprend les observations enregistrées sur les bordereaux de description des sépultures remplis sur le terrain : position du squelette, position des membres supérieurs et inférieurs, nature du contenant (inhumation en pleine terre, cercueil, coffrage...), présence ou absence de mobilier funéraire. Chaque enregistrement de sépulture contient deux sous-tables (fig. 1, n° 6a et 6b) dans lesquelles sont enregistrés, le cas échéant, les épingles de linceul et les clous de cercueil associés au squelette. Le contenu des champs relatifs à l'âge au décès et au sexe est importé de la table Anthro (cf. ci-dessous).

### **Anthro** (fig. 1, n° 9)

Les données anthropométriques sont enregistrées dans la base de données *BaDoA* (cf. ci-dessous). Sont importées dans la table Anthro les seules données anthropologiques nécessaires à l'interprétation de l'organisation spatiale des lieux d'inhumation (sexe et âge au décès). Une procédure (méthode) accessible par un menu permet ensuite de transférer ces informations dans la table Sépultures en recodant les variables.

### **Mobilier** (fig. 1, n° 10)

La table Mobilier comprend des enregistrements sur tous les "objets" trouvés en cours de fouille (mobilier métallique, verre, objets en os, monnaies etc.) à l'exception de la vaisselle en céramique qui est enregistrée dans une table distincte, des épingles de linceul et des clous de cercueil associés à des inhumations qui sont, eux, enregistrés dans la table Sépultures.

Les objets sont repérés par le n° de l'US dans laquelle ils ont été trouvés, suivi d'un n° d'ordre : l'objet 7632-3 est ainsi le troisième objet enregistré dans l'US 7632. Les

enregistrements comportent, outre une description et une identification de l'objet, des informations sur le contexte dans lequel il a été trouvé, et des indications chronologiques. On distingue la datation individuelle de l'objet et celle de l'US dont il provient, qui peut être largement postérieure (cas des objets redéposés).

### **Céramique** (fig. 1, n° 11)

La table céramique contient uniquement la vaisselle en terre cuite, à l'exclusion de tout autre objet en céramique. Les champs comportent les principales informations relatives à la morphologie, à la technique de fabrication, dans un but purement chronologique. Il s'agit d'un résumé de l'information provenant de la base *BaDoC*, module spécifiquement développé pour le traitement des données céramiques (cf. ci-dessous).

### **Dessins** (fig. 1, n° 12)

La table Dessins récapitule tous les dessins et plans réalisés ; elle comprend 4 séries ; série A : relevés de terrain ; B : documentation élaborée (plans d'assemblage) ; C : dessins prêts pour la publication ; D : documents graphiques extérieurs à la fouille (cadastre ancien, extrait de plan de ville, etc.).

### **Photos et PhotoMob** (fig. 1 n° 13 et 14)

La table Photos contient l'inventaire des photos de chantier. Les photos noir et blanc, diapos couleur et les photos numériques constituent trois séries séparées. Pour chaque prise de vue, sont enregistrés le n° du Fait, de la Sépulture, ou du Mur représenté. La table PhotoMob contient l'inventaire des photos d'objets.

Les photos numériques, ou les scans de photos argentiques, ne sont pas stockés dans la base elle-même mais dans des répertoires séparés. Cela présente deux avantages : alléger le volume (poids de fichier) de la base en n'y enregistrant que les chemins d'accès aux fichiers photos et permettre l'accès à ces mêmes photos depuis d'autres logiciels. Elles sont consultables soit par sélection directe dans les tables Photos ou PhotoMob, soit depuis les fiches des enregistrements de mobiliers ou de Faits, Murs, Sépultures ou Structures.

### **Volumes** (fig. 1 n° 15)

La table Volumes comprend un index de tous les classeurs, cahiers, etc. qui contiennent les enregistrements manuels primaires (enregistrement des données de terrain et bordereaux de saisie du mobilier et des *Archives de Fouilles*).

## **L'aide à la réalisation du diagramme stratigraphique**

Outre les opérations habituelles des bases de données informatisées (recherches, tris, constitution d'index, réalisation de graphiques et d'histogrammes...), la base *Archives de Fouilles* effectue une tâche spécifique d'aide à la réalisation du diagramme stratigraphique<sup>2</sup>. La réalisation du diagramme demande un travail préalable qui consiste à représenter, sous forme de tableau, les relations d'antériorité, de postériorité, d'égalité ou d'équivalence qu'une US entretient avec toutes celles avec lesquelles elle est en contact immédiat. Cette tâche est indispensable parce qu'elle permet de regrouper les informations dispersées entre différentes fiches d'enregistrement de terrain et parce qu'elle permet de déceler d'éventuelles erreurs d'enregistrement (dans le cas, par exemple, où l'US 3220 aurait été enregistrée à la fois sous et sur l'US 3125). Ce travail très long et fastidieux, qui demande une attention soutenue et de nombreuses vérifications lorsqu'il est exécuté manuellement, est effectué désormais par un programme de la base informatisée. Celui-ci complète l'enregistrement stratigraphique (par exemple, si la fiche de l'US 3410 indique que cette dernière se trouve sous l'US 3402, le programme ajoutera sur la fiche de l'US 3402 qu'elle se trouve sur l'US 3410). Il vérifie également la cohérence de l'enregistrement et affiche sous forme de liste les erreurs éventuelles à corriger.

## **L'interface**

L'utilisation de la base de donnée en mode dit « menus créés » autorise son utilisation sans avoir besoin de connaître le fonctionnement de 4<sup>e</sup> Dimension. L'interface proposée permet d'une part de saisir, interroger et consulter l'ensemble des données et d'autre part de gérer les données : aide à la constitution du diagramme stratigraphique, intégration des données anthropologiques et céramologiques, import/export (pour des catalogues de mobilier par exemple).

En outre, dans ce mode d'utilisation, la base peut fonctionner sous un *runtime*, c'est-à-dire sans l'obligation de posséder 4<sup>e</sup> Dimension.

## **La spatialisation des données**

Dans *ArSol*, la spatialisation des données attributaires enregistrées dans les modules *Archives de Fouilles*, *BaDoC* et *BaDoA* repose sur une chaîne opératoire d'acquisition et de traitement de l'information spatiale dans un Système d'information géographique. Ce SIG est

---

<sup>2</sup>Le diagramme stratigraphique représente les relations chronologiques entre les Unités Stratigraphiques : chaque US est placée "sous" (=avant) toutes celles qui lui sont postérieures et "sur"(=après) toutes celles qui lui sont antérieures.

construit sur une liaison entre les logiciels 4<sup>e</sup> Dimension™ et ArcGIS™ par une interface ODBC (Open DataBase Connectivity).

### **Les étapes de la chaîne opératoire**

Le relevé de terrain au 1/20e sur papier ou calque a été remplacé par une photographie numérique géoréférencée. Cela consiste à prendre une photo numérique la plus verticale possible de l'élément (US, Fait, Mur, Sépulture, Structure) à relever en ayant placé autour, au préalable, 4 amers sur un même plan horizontal. Les coordonnées x, y, z et id (identifiant de l'élément relevé) des 4 amers sont relevées au tachéomètre. Un cinquième, ou plus si nécessaire, au centre de la structure permet d'enregistrer une cote d'altitude dans ou sur l'élément.

Toutes les étapes suivantes s'effectuent dans ArcView™. Le transfert des fichiers de coordonnées des amers permet de créer des points à l'aide desquels les photos sont géoréférencées puis rectifiées.

Les photos sont ensuite imprimées de manière à pouvoir y apposer toutes les notes nécessaires à l'interprétation sur le terrain. Le niveau de détail des photos et le fait qu'elles soient géoréférencées permettent de s'affranchir des relevés pierre à pierre par exemple. Toutefois, ceux-ci peuvent être produits, à partir des photos elles-mêmes.

L'étape suivante consiste à vectoriser le contour des éléments relevés pour constituer un fichier de polygones identifiés par le numéro de Fait, de Mur... Cet identifiant permet de construire une jointure avec la table de données attributaires du module *Archives de fouille*. De la même manière, les relations entre les différentes table de donnée sous 4<sup>e</sup> Dimension™ sont reconstruites.

Cette organisation des données offre trois types d'utilisation :

la consultation de l'ensemble des archives du sol à partir des objets spatiaux et des possibilités de requêtes spatiales (localisation, proximité, distance...)

la production de plans thématiques fondés sur l'ensemble des données enregistrées dans *Arsol* (plans de Structure, plans par phase, répartition de mobilier...)

l'accès à des outils de calcul et d'analyse fondés sur les propriétés spatiales des objets eux-mêmes (localisation, périmètre, surfaces, topologie, distance...)

## **Les conséquences**

Ce changement de pratique permet d'une part de gagner du temps sur le terrain pour les relevés et pendant l'étude pour la production des plans et d'autre part autorise des traitements de données difficilement envisageables dans le temps imparti en général au travail post fouille. La mise en place de cette chaîne opératoire entraîne de profondes modifications des habitudes de travail qui ne se discutent pas en terme d'avantage ou d'inconvénient mais dont il faut analyser et accepter les implications. Sur le terrain elles interviennent dans deux registres : l'organisation du travail et les moyens mis en œuvre ; le déroulement du processus d'interprétation des données. Pendant la fouille, cette chaîne opératoire mobilise une personne à plein temps disposant de l'équipement nécessaire : un tachéomètre, un appareil photo numérique, un PC avec une licence ArcView™ et une imprimante couleur.

La phase d'observation et d'interprétation indispensable qui s'effectuait au moment du relevé est transférée à une prise de note sur l'impression de la photo géoréférencée très rapidement.

## **2. BaDoC : un système de traitement des données céramiques**

Cette base de données peut servir à la gestion de grandes quantités de mobilier céramique, mais elle a surtout été développée pour la recherche dans ce domaine. Outre la structure de la base, il existe une série de procédures (méthodes – programmation) utiles à un traitement précis, notamment quantitatif du mobilier.

Ce système est ouvert parce qu'il est adaptable à tout type de mobilier provenant de fouilles stratifiées (il n'est pas lié à un thesaurus prédéfini) ; parce qu'il permet d'intégrer de l'information extérieure (par exemple : résultats de modélisation statistique pour la datation) et d'exporter d'autres informations (par exemple datation des contextes archéologiques ou des US dans la base *Archives de fouilles*).

La structure de la base de données est représentée par la *Figure 3*. Elle comprend seize tables (cf. fig. 3, n° 1 à 16). Les tables 1 à 4 comprennent les données typologiques ; la table 5 les données quantifiées en nombre de restes (NR) et en poids par Groupe Technique ; les tables 6 à 11 servent de clefs de recherche sur plusieurs tables et pour certains champs (forme, rebord...) ; la table 12 permet de stocker temporairement les données quantifiées en fonction d'une requête ; les tables 13 à 15 servent à la datation ; enfin les tables 16 et 17 récupèrent certaines informations utiles de la base *Archives de Fouilles*. Les liens relationnels

entre les tables sont en rouge lorsqu'ils relient les tables typologiques et en bleu lorsqu'ils servent à gérer l'information provenant d' *Archives de Fouilles*. Il faut préciser qu'une grande partie des fonctionnalités du système passe par des procédures (programmation) interfacées (boutons) donc non visibles sur le graphique, mais servant également à mettre en oeuvre les données provenant de différentes tables.

Cette base est conçue pour intégrer le matériel de différents sites ; elle gère donc les problèmes de numérotation, afin que chaque enregistrement possède un identifiant unique, par exemple pour un même n° d'US mais de sites différents. Son interface a été thématifiée de manière à suivre la chaîne opératoire d'un traitement de masse de données céramiques (une barre de menu de saisie et consultation générale des données ; une barre de préparation des données qui sert également de contrôle de la saisie ; une barre de quantification et de datation.

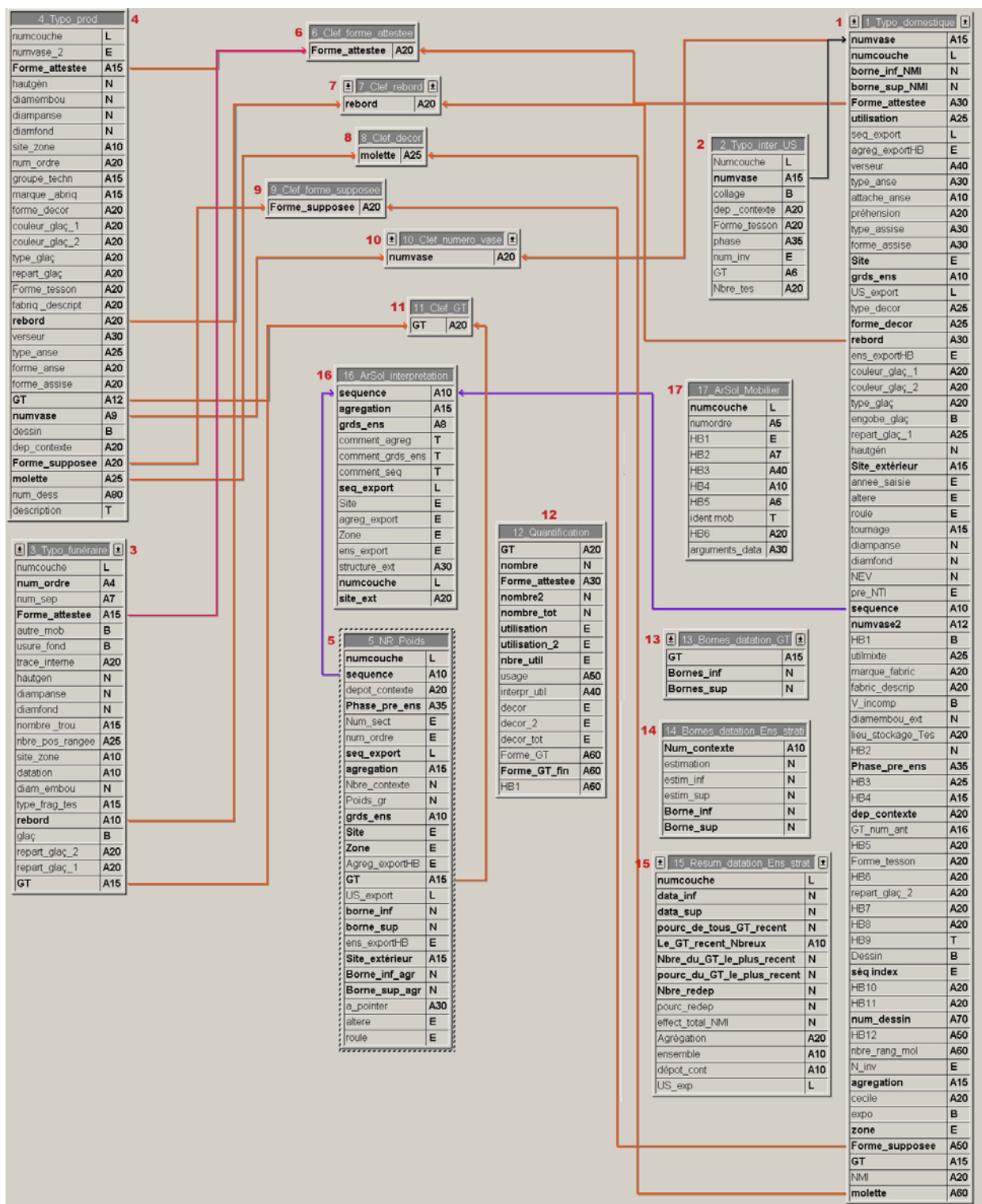


Figure 3

### Tables 1, 2, 3, 4 : typologie

Nous avons fait le choix de séparer dans des tables différentes le matériel provenant des contextes domestiques, funéraires ou de production pour des raisons de commodité. Les tables typologiques contiennent les critères morphologiques et technologiques discriminants de chaque tesson attestant une forme, du vase complet au fragment de bord, de fond,

d'anse... ; chaque enregistrement contient les informations d'un seul élément typologique. Ainsi, un enregistrement contient : la numérotation du vase, de l'Ensemble, de la Séquence ou Agrégation, de l'US, à laquelle nous considérons que le récipient appartient dans sa globalité ; puis les champs décrivant la morphologie du vase : deux champs consacrés à la forme générale du vase, se référant à un répertoire de forme, l'un renseigné lorsque la forme est attestée, l'autre lorsqu'elle est supposée ; puis une série de champs permettant de donner le détail morphologique du vase ou du tesson (forme du rebord, du fond, de l'anse, du verseur...) ; puis le détail technologique avec le Groupe Technique et le type de couverture et sa couleur (glaçure, engobe...) ; puis les décors et leur descriptions, les traces d'utilisation (trace d'exposition au feu...) ; enfin quelques champs réservés aux dimensions du récipient.

C'est dans les tables 1, 3 et 4 que les informations utiles à la quantification du NMI (Nombre Minimum d'Individus), du NTI (Nombre Typologique d'Individus) et du NEV (Nombre d'Equivalent Vases) sont enregistrées. Nous reviendrons ultérieurement sur la quantification et la datation.

Les tables 1 et 2 sont liées pour la gestion des récipients dont les tessons proviennent de différentes US ; le lien sur le n° de vase, permet de décliner toutes les US et le nombre de tessons composant un vase par US, dans la table 2.

Pour la table 3 (vases funéraires médiévaux et modernes), d'autres champs viennent préciser le nombre, la position et le type de trous dans le vase.

Il est possible d'afficher les dessins des récipients ou des rebords (bouton dans l'interface). Le choix a été fait de stocker les dessins et les photos à l'extérieur de la base et de passer par des chemins d'accès à d'autres logiciels, afin de ne pas alourdir le système.

### **Table 5 : NR et poids**

Cette table permet de gérer ces deux techniques de quantification qui se font avant recollage, par US et par groupe technique (production). Un enregistrement contient le nombre et le poids des tessons pour une US et un Groupe Technique. Tous le matériel est ainsi inventorié dans cette table, éléments de forme comme tessons de panse ; ces derniers ne font pas l'objet d'un enregistrement individuel (sans intérêt), puisqu'ils n'apparaissent pas dans les tables typologiques (sauf dans certains cas pour le calcul du NMI qui n'est pas détaillé ici).

### **Tables 6, 7, 8, 9, 10, 11 : clefs de recherches croisées**

Ces tables, qui sont liées aux précédentes (tables 1, 3, 4, 5), permettent de faire des requêtes croisées sur un des critères énoncés (forme attestée ou supposée, rebord, décor, n°



de vase, Groupe Technique). L'intérêt est d'afficher par exemple tous les récipients d'une forme ou les rebords... quels que soient leurs contextes de découverte (funéraire, production, domestique).

### **Table 12 : Quantification**

Cette table sert uniquement à stocker temporairement les données quantifiées, résultat d'une requête faite dans les tables typologiques (NMI, NTI, NEV) ou dans la table 5 NR\_Poids. Généralement, la requête se fait sur un contexte archéologique (fosse, niveau d'occupation...), peu importe l'échelle choisie, avec comme résultat la quantification par Groupe Technique et par contexte. Cette table ne possède aucun lien avec d'autres car tout passe par les procédures (méthodes), activées par des boutons, donc faciles d'accès. L'originalité ici est que la liste des groupes techniques quantifiés comprend la totalité des groupes contenus dans la base et se décline toujours de la même manière. Le résultat est une matrice, c'est-à-dire un tableau contenant également les données manquantes (o). C'est indispensable lorsqu'on veut sortir de l'application et faire un traitement statistique des données. En effet, la comparaison des matrices, lorsqu'elles sont fabriquées uniquement à partir des données présentes pour chaque contexte, devient très difficile (pas les mêmes groupes techniques dans les mêmes lignes ou colonnes, donc reconstruction très fastidieuse du tableau).

La table 12 affiche les résultats dans l'application (4D) et à l'échelle choisie (exemple : quantification en NR du matériel de la fosse F1). D'autres développements (méthodes), qui font également l'originalité de ce système, intègrent un tableur proche d'excel (plug-in 4D view). Il est alors possible de quantifier le matériel à n'importe quelle échelle, quelle que soit la technique de quantification choisie. Par exemple, nous pouvons faire une requête sur la fosse F1 à Fn et demander au tableur de quantifier le matériel en NMI par US (ou par Séquence ou par Ensemble...). Il construit alors une matrice (dans 4D view) qui contient les données quantifiées à l'échelle choisie. Si nous travaillons à l'échelle de l'US, il précise l'Ensemble et la Séquence à laquelle appartient chaque US. Ainsi, en quelques secondes, il est possible de construire une matrice qui pourrait demander plusieurs heures, voire plusieurs jours de travail, sans courir le risque d'erreur engendré par une trop grande manipulation des données.

### **Tables 13, 14, 15 : datation**

Ces trois tables intègrent les bornes chronologiques, inférieures et supérieures, d'un contexte stratigraphique, d'une production céramique ou d'un type de récipient (cette dernière datation n'est pas encore intégrée à la base). Nous ne préjugeons pas ici de la

manière dont ces bornes chronologiques sont construites. Pour plus de précisions sur la démarche suivie à Tours pour ces questions de datation, notamment par l'élaboration d'un modèle statistique, il suffit de se référer à un précédent article (Husi, Bellanger 2003). En d'autres termes, la fosse F1 est datée entre le milieu et la fin du 15<sup>e</sup> s. : Nous renseignons donc la table 14 avec trois champs, l'un contenant l'identifiant stratigraphique, le second la date inférieure, formulée en chiffre (ici 1450) le dernier, la date supérieure ici (1500) ; ces dates doivent ensuite être réinterprétées, l'usage des nombres étant indispensable à la machine. La même démarche est suivie pour la table 13 avec une date inférieure et supérieure pour une production céramique. On ne saisit donc qu'un enregistrement par contexte et par Groupe Technique. L'intérêt ici est qu'il existe des procédures de mise à jour des données dans les autres tables où cette information est indispensable (tables 1, 3, 4 typologiques et table 5 NR\_Poids). Ainsi, si un Groupe Technique (GT 12) est daté entre 1300 et 1400, la mise à jour (bouton), permet de répercuter et/ou de changer, quand nous le désirons et automatiquement, l'information (borne inférieure et borne supérieure) dans tous les enregistrements qui font référence à ce groupe GT 12 (table typologiques...). Nous avons choisi ici cette option plutôt que le lien entre tables, parce que ces nouvelles informations contenues dans les tables typologiques et NR\_Poids sont utilisées à d'autres fins. En effet, une aide à la datation (procédure ; bouton) permet, pour une échelle d'analyse choisie (requête), de préciser la plage chronologique durant laquelle toutes les productions sont contemporaines pour un contexte sélectionné en fonction des champs « bornes chronologiques » (tables typologiques ou NR\_poids au choix) ; puis en fonction du résultat, la part et le nombre du matériel redéposé ou intrusif. Ces informations, qui peuvent ensuite être exportées dans *Archives de Fouilles*, sont contenues dans la table 15 (Resum\_datation).

### **Tables 16 et 17**

Ces deux tables permettent d'importer de la base *Archives de Fouilles* les informations utiles au céramologue pour son étude, c'est-à-dire les principales interprétations stratigraphiques (table 16) et les objets datés comme les monnaies ou le verre... (table 17) ; les relations avec les autres tables (liens mais aussi procédures) permettent de visualiser ces informations ailleurs : par exemple dans le formulaire général de la table 1 (typologie) qui synthétise un grand nombre d'informations (pour un vase sélectionné, outre les informations typologiques et de datation, s'affiche également les objets datés de l'US où se trouve le vase et l'interprétation archéologique du contexte dont il provient).

## **Les atouts du système**

### **La quantification**

La possibilité de quantifier le matériel à n'importe quelle échelle d'analyse en NMI, NTI, NEV, NR et Poids, avec comme résultat une matrice de données formatée de la même manière et qui tient compte des données manquantes, est un outil indispensable.

A la différence de beaucoup d'autres systèmes, le Nombre Minimum d'Individus n'est pas figé lorsque l'archéologue étudie son matériel. Ne sont saisies dans *BaDoC* que les informations permettant ultérieurement de créer le NMI. La différence est significative, car dans le cas contraire, la quantification ne tient pas réellement compte des changements d'échelles d'analyse. Dans *BaDoC*, le NMI est recalculé à chaque nouvelle requête. En d'autres termes, la plupart des systèmes incluent un champ NMI contenant un nombre correspondant au NMI observé au moment *t* de l'étude, c'est-à-dire lorsque le matériel est sur la table ; la fosse F1 (ou l'US...) contient 10 NMI appartenant au GT 12. L'information est alors figée, elle ne peut être que cumulée en fonction de la requête (10 NMI dans F1 et *n* NMI dans F*n*). Notre approche est très différente : nous renseignons un champ « forme tesson » dans la table 1 (typologie) qui permet ensuite, par une procédure, de calculer à la demande et suivant l'échelle d'analyse le NMI (F1 : 10 lèvres, 4 fonds, 2 vases entiers = 13 NMI). Ainsi, pour F1, le calcul global du NMI sur l'ensemble de la fosse sera différent de celui réalisé pour chaque US composant cette même fosse. Cette différence n'est pas sans incidence sur les résultats puisqu'elle donne une réelle amplitude aux données, en fonction de l'échelle d'analyse choisie.

### **La datation**

Les développements faits dans ce sens sont plus une aide à la décision qu'une datation. Libre à chacun d'utiliser ces possibilités.

### **La saisie minimum et l'intégrité des données**

Afin de minimiser le temps de saisie, nous avons pris le parti de développer des procédures de mise à jour, de saisies automatiques (concaténation...), de duplications d'enregistrements... dès qu'il était possible de le faire. Ainsi, la saisie primaire, au moment de l'étude, est limitée puisqu'elle ne concerne en réalité que les tables typologiques (1 et 2 ou 3 ou 4), la table 5 NR\_Poids ; les tables 6 à 11 sont remplies automatiquement puisqu'elles ne servent qu'au bon fonctionnement du système relationnel ; les tables 12 et 15 ne servent qu'à récupérer des données secondaires et de manière temporaire sans intervention de

l'utilisateur ; les tables 13 et 14 ne comportant que peu de champs et d'enregistrements, ne sont saisies qu'ultérieurement ; les tables 16 et 17 stockent des informations déjà saisies et importées de la base *Archives de Fouilles*. Ce choix permet non seulement d'alléger la part de saisie mais aussi de conserver l'intégrité des données.

### **Un système interfacé**

L'interface proposée suit la logique d'une étude céramique : saisie primaire des données, préparation des données pour la quantification (procédures) ; quantification et datation (procédures). La navigation entre menus déroulants et boutons permet de suivre et d'appliquer la démarche, en sachant que les procédures qui activent toute la base sont toujours précédées d'un commentaire ou d'une mise en garde. Il est donc impossible d'affecter de manière irrémédiable la base par l'application incontrôlée de procédures.

### **Un système ouvert où l'utilisateur reste maître du jeu**

Enfin, il faut revenir sur deux aspects importants. D'une part, le fait que ce système est utilisable sur tous les chantiers mettant en oeuvre des données stratifiées, quelle que soit la période ou le faciès céramique. L'utilisateur peut très bien adapter le système en modifiant le contenu des champs, ce qui n'affecte ni la structure de la base, ni les procédures de calcul. D'autre part, aucune programmation n'intervient dans les décisions d'ordre archéologique : seuls les calculs sont automatisés. En d'autres termes, l'utilisateur ne devient pas un « presse bouton » qui subit les résultats sans les comprendre.

En contrepartie, ce système est difficilement utilisable par un novice dans le domaine céramologique. L'utilisateur doit prendre des décisions d'ordre archéologique à certaines étapes de la démarche, sans quoi le système est bloqué.

## **3. BADOA : Traitement informatisé des données anthropologiques**

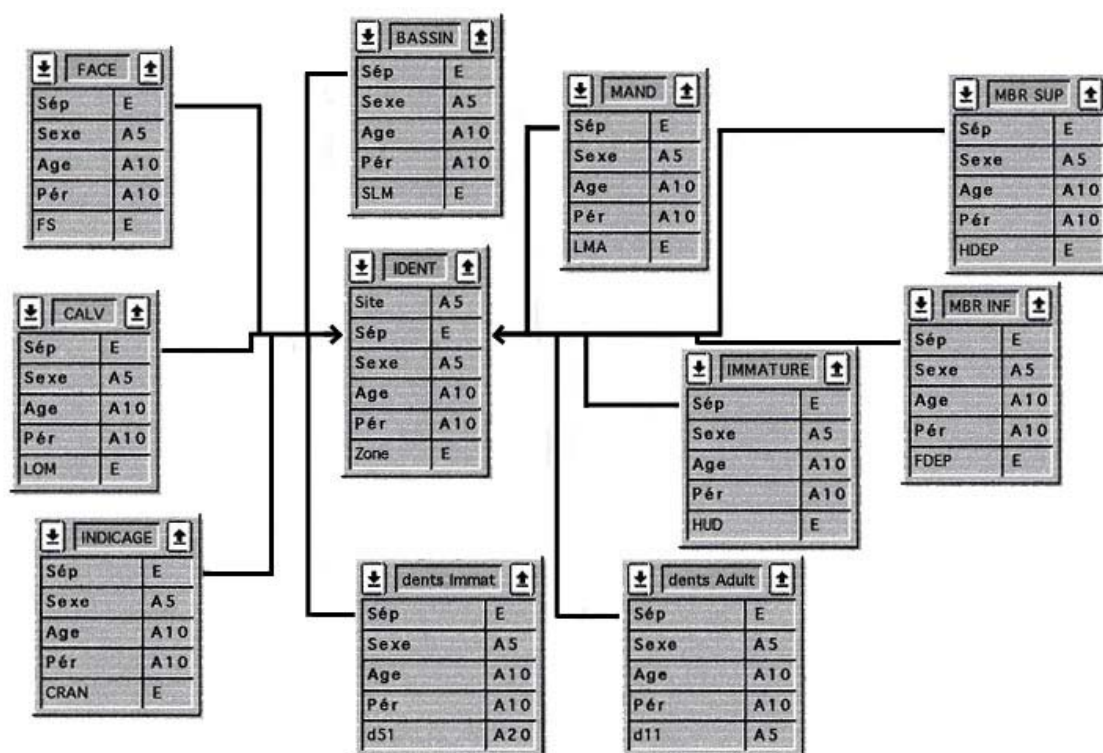
### **Clauses techniques**

La définition morphologique du squelette de l'adulte s'appuie sur trois sortes de caractères qui correspondent aux trois types de données classiquement utilisées en statistique : métriques (continues), descriptives (discontinues) et discrètes (alternatives ou booléennes). L'architecture de la base de données est conçue pour les données métriques, les autres étant presque totalement absentes des analyses.

La quantité des données collectées peut considérablement varier d'un individu à l'autre.

Les vestiges osseux des immatures sont entièrement conditionnés par la croissance. Ils ne sont à rapprocher de ceux des adultes que dans les décomptes des individus et ils en sont écartés pour tout le reste. Les informations qu'ils apportent se réduisent essentiellement aux longueurs des grands os et à l'état dentaire, la pathologie étant un sujet à part.

La complétude des vestiges étant très imparfaite et inconstante, le nombre des observations faites pour chaque sépulture varie grandement, surtout lorsque le squelette entier est pris comme référence permanente. Le squelette a donc été divisé en segments de sorte que chacun d'eux ne reçoit que les individus qui lui apportent des données. Une quantité appréciable de plages vides est alors écartée du système informatisé. Après cela, des données manquent encore. Or, dans l'ordre logique de son déroulement, l'analyse statistique fait d'abord appel à des calculs univariés puis à des multivariés. Ces deux phases de calcul sont nettement distinctes et s'appuient sur des paramètres de distribution différents. Un comblement de données doit donc être effectué entre les deux car les formules multivariées prennent en compte toutes les relations entre les valeurs individuelles des variables et, dans ces conditions, elles exigent une absence totale de lacunes. Par prudence et besoin de clarté dans ces manipulations, le travail a été réparti sur deux bases indépendantes : l'une pour les données brutes à conserver dans leur état d'origine et l'autre pour les données comblées.



Structure de la base anthropologique des données brutes. Les tables sont liées à la table "Identité" par la variable "Sép" (n° de sépulture).

## **La base de données brutes**

Sa structure comporte dix tables usuelles dont six correspondent à la segmentation du squelette. Une onzième concerne l'identification individuelle. Elle subordonne, en quelque sorte, les autres auxquelles elle est liée par le champ "n° de sépulture" configuré en variable obligatoire et unique. Outre les variables qui leur sont propres, les tables comportent, en rappel à côté du n° de sépulture, certains éléments de l'identité qui sont à la base des principaux tris utilisés.

### **Les formulaires**

Chaque table dispose d'un formulaire pour l'enregistrement et d'un autre pour l'organisation des données en liste, voire de deux formulaires en liste, pour présenter une version de valeurs appariées lorsqu'il y a des données bilatérales. La saisie des données a pour préalable obligatoire l'enregistrement du n° de sépulture sur le formulaire de saisie de la table d'identification. Ce dernier comporte des boutons pour continuer l'enregistrement sur n'importe quelle table en fonction des besoins. À la table choisie, le n° de sépulture est toujours en tête d'enregistrement puisqu'il est la référence du lien et il provoque l'importation automatique des éléments d'identification désirés. Il est toutefois possible de débiter la saisie à partir de n'importe quelle table, mais il se produit alors un retour automatique à la table d'identification quand le n° de sépulture ne figure pas encore dans cette dernière. Les formulaires de saisie copient les dispositions de la fiche papier de l'examen. Un simple coup d'œil permet de vérifier vite et sans peine la concordance entre les deux. Une procédure interne au formulaire d'enregistrement calcule automatiquement la version appariée des valeurs bi-latérales au moment de l'enregistrement.

Chacune des quatre tables appelées à ne fournir qu'un dénombrement (identification, dents des adultes, dents des immatures et indicateurs de l'âge des adultes) possède un formulaire spécifique pour présenter l'état désiré.

Les formulaires particuliers des sept autres tables (calvarium, face, mandibule, bassin, membre supérieur, membre inférieur et immatures) calculent les paramètres de distribution (effectif, valeurs minimale et maximale, somme des valeurs, somme des carrés des valeurs, moyenne et écart-type) de chaque variable en version appariée. Ces paramètres constituent la base des analyses univariées dont l'utilisation des nombreuses formes -surtout des tests- varie fortement selon les cas et est effectuée, à la demande, sur un tableur séparé (Excel).

### **Les méthodes**

Hormis le n° de sépulture qui est immuable, les données de l'identification ont la possibilité d'être saisies ou modifiées durant la progression du travail ou après un premier

enregistrement parfois réduit au seul n° de sépulture. Leurs modifications doivent donc être transmises dans les autres tables où elles ont été importées. L'opération est effectuée par la procédure (lancée à la demande) d'une méthode de mise à jour qui se réfère au contenu de la table d'identification.

À l'ouverture de chaque formulaire, une méthode adjointe règle automatiquement les détails variables de la présentation (site, définitions de la sélection utilisée, effectif en jeu, date...) ou de l'enregistrement (appel, sur lien, des éléments de l'identité et calcul des valeurs appariées). La saisie et les sorties en liste ne nécessitent pas d'autre commande alors que les dénombrements des états ainsi que les paramètres de distribution exigent des méthodes spécifiques. Elles sont, cette fois, séparées de leur formulaire d'accueil qui est appelé à la fin de leur programme.

## **La base de données comblées**

Cette seconde base est essentiellement destinée à produire les paramètres de distribution des analyses multivariées qui sont pratiquées seulement sur les variables de certains segments du squelette adulte. Elle ne nécessite pas de table d'identification et ne contient donc que six tables sans lien (calvarium, face, mandibule, bassin, membre supérieur et membre inférieur). Chaque table n'est plus accompagnée que par trois formulaires pour valeurs appariées. Celui de la présentation en liste et celui du calcul des paramètres de distribution de chaque variable sont identiques à leurs homologues créés pour les données brutes. Le troisième présente les paramètres (effectif, sommes des valeurs, sommes des valeurs au carré, sommes des carrés des écarts et sommes des produits des écarts) composant les éléments fondamentaux des formules statistiques multivariées.

Les méthodes sont semblables, voire identiques, à celles de la base des données brutes et leur nombre est réduit en rapport de celui des formulaires. Pour chaque table, on en retrouve effectivement trois liées aux formulaires dont elles gèrent les variables de présentation et deux indépendantes pour les paramètres univariés et multivariés. Les appels des formulaires et les commentaires à apporter aux méthodes sont les mêmes que ceux qui ont été décrits pour la base de données brutes.

Les tables reçoivent, par importation et en versions appariées, les valeurs de leurs homologues de la base des données brutes. Les importations de données et leurs modifications à venir sont faites en liste. La question des données manquantes doit être résolue avant tout début d'analyse. C'est une opération délicate à effectuer car elle ne doit pas dénaturer les caractéristiques initiales de l'ensemble. Aucun moyen automatique n'est capable de la mener à bien correctement. Elle s'effectue au cas par cas, en variant les

procédés (éliminations d'individus ou de variables, utilisations de moyennes ou de régressions) et en procédant à de fréquents contrôles par comparaisons avec les paramètres des données brutes. Au moment de sélectionner les valeurs à exporter, il est déjà possible de faire l'impasse sur d'éventuelles variables trop redondantes.

L'analyse multivariée ouvre sur de multiples comparaisons qui prennent des formes différentes (fonctions discriminantes pour déterminer le sexe, distances généralisées pour vérifier les similitudes) et dont les équations sont résolues par des calculs matriciels après importation des données dans un tableur.



## Références

### **DESACHY** à paraître

Desachy B. – Du temps ordonné au temps quantifié : application d'outils mathématiques au modèle d'analyse stratigraphique d'Edward Harris, CTHS, 129<sup>e</sup> congrès, Besançon 1-22 avril 2004, à paraître.

### **GALINIE** 1980

Galinié H., De la stratigraphie à la chronologie in Schnapp A. (dir.), L'Archéologie aujourd'hui, Paris, 1980 : 63-86.

### **GALINIE, RANDOIN** 1987

Galinié H., Randoin B., Enregistrement et exploitation des données des fouilles de Tours, Actes de la table ronde : *Enregistrement des données de fouilles urbaines*, Centre National d'Archéologie Urbaine, Tours 1987, 61-73.

### **GALINIE, HUSI, RODIER** 2001

Galinié H., Husi P., Rodier X. et al., L'archéologie de Tours : le site 3. Recherche sur Tours 9, version provisoire, publication électronique sur CD Rom, mars 2001.

### **GALINIE, RODIER** 2001

Galinié H., Rodier X. - Les modifications du trait de la rive gauche de la Loire dans l'ouest de Tours, in *Géoarchéologie de la Loire moyenne et de ses marges*, N. Carcaud et M. Garcin dir. : 29-36. [http://www.brgm.fr/projet\\_loire/publicat.htm](http://www.brgm.fr/projet_loire/publicat.htm), 2001.

### **GALINIE, RODIER** 2002

Galinié H., Rodier X. - ToToPI, un outil d'analyse urbaine, Les petits cahiers d'Anatole n° 11,09/12/2002, 21.065 signes, [http://www.univ-tours.fr/lat/pdf/F2\\_11.pdf](http://www.univ-tours.fr/lat/pdf/F2_11.pdf)

### **HUSI** 2001

Husi P. – La céramique des premières phases d'occupation (6<sup>e</sup>/7<sup>e</sup> – 11<sup>e</sup> s.). In : Zadora-Rio E. et Galinié H. et al. 2001.

### **HUSI** dir. 2003

Husi P. dir. – La céramique médiévale et moderne du Centre-Ouest de la France (11<sup>e</sup> – 17<sup>e</sup> siècle), chrono-typologie de la céramique et approvisionnement de la vallée de la Loire moyenne, supplément à la *Revue Archéologique du Centre de la France*, n° 20, FERAC, Tours, 2003, 110 p. et un cédérom.

### **HUSI, BELLANGER** 2003

Husi P., Bellanger L. - Estimer la datation, la durée et les discontinuités de l'occupation des sites archéologiques à l'aide de la modélisation des faciès céramiques, Les petits cahiers d'Anatole, [http://www.univ-tours.fr/lat/pdf/F2\\_13.pdf](http://www.univ-tours.fr/lat/pdf/F2_13.pdf)

### **JOUQUAND et al.** 1999

Jouquand A.-M., Champagne F., Rodier X., Husi P., Witmann A. – La fouilles des "abords de la cathédrale" de Tours (Indre-et-Loire) : Antiquité-haut Moyen Age, *Revue Archéologique du Centre de la France*, n°38, 1999 : 7-98.

### **LORANS** à paraître

Lorans E. (dir.) – Saint-Mexme de Chinon (5<sup>e</sup>-19<sup>e</sup> s.), CTHS, Paris, à paraître 2005.

### **RODIER** 2000

Rodier X. - Le système d'information géographique TOTOPI : Topographie de TOurs Pré-

Industriel, Les petits cahiers d'Anatole, 4, 22/12/2000, 28600 signes, 5 fig.,  
[http://www.univ-tours.fr/lat/pdf/F2\\_4.pdf](http://www.univ-tours.fr/lat/pdf/F2_4.pdf)

**THEUREAU 1998**

Theureau (C.) - *La population archéologique de Tours (IVe-XVIIe s.) : étude anthropologique*, Recherches sur Tours 7, Tours 1998.

**ZADORA-RIO 1994**

Zadora-Rio E. – Le système de gestion des données de fouilles de Rigny, Le médiéviste et l'ordinateur, n°29, Printemps 1994 : 31-35.

**ZADORA-RIO, GALINIE 2001**

Zadora-Rio E., Galinié H. et al., La fouille du site de Rigny, 7<sup>e</sup>-19<sup>e</sup> s. (commune de Rigny-Ussé, Indre-et-Loire) : l'habitat, les églises, le cimetière. 3<sup>e</sup> et dernier rapport préliminaire (1995-1999) Revue Archéologique du Centre 40, 2001 : 167-242.