

La gestion informatisée des collections anatomiques au Muséum national d'Histoire naturelle

Alexis Martin *



Fig.1. Exemple de quelques spécimens : reptile en fluide, montage ostéologique d'oiseau et mammifères en boîte, avant reconditionnement définitif
© Alexis Martin

* Alexis Martin est technicien de collection au Muséum national d'Histoire naturelle, département Systématique et Évolution, USM 602 Taxonomie et Collections : collections Mammifères et Oiseaux, site Anatomie comparée
55 rue Buffon, 75005 Paris
téléphone + 33 1 40 79 56 05
alexis@mnhn.fr

Développée par l'auteur dans le cadre de la rénovation et l'inventaire des collections dites d'Anatomie comparée au Muséum national d'Histoire naturelle, la base de données *Anatomic* permet d'informatiser non seulement les collections des réserves, spécimens ostéologiques et objets formolés mais également un constat d'état détaillé de chaque spécimen. Les techniques de modélisation informatique décrites ici peuvent être déclinées dans d'autres contextes de conservation et pour d'autres types de collections.

L'informatisation des collections est une priorité qui s'impose aujourd'hui à de nombreux établissements patrimoniaux toutes thématiques confondues. Ce mode d'inventaire, ou plus exactement d'organisation du stockage de la documentation et de l'information, est repris dans la récente loi sur les Musées de France, signe que tous les établissements doivent ou devront s'y conformer un jour ou l'autre.

Le problème de l'informatisation se complique dès lors que les collections sont numériquement importantes : au temps très long que va nécessiter l'accomplissement de la tâche, il convient de bien calibrer en amont les besoins afin de ne pas commettre d'oubli ou d'erreur lors de la conception de l'outil informatique. Autrement dit, une bonne prévision permet de faire face aux différents cas particuliers, qui ne manquent pas de se rencontrer parmi les grands ensembles de collection.

Le Muséum national d'Histoire naturelle regroupe de très nombreuses collections de Vertébrés actuels, conservés selon différentes typologies de préparation. Parmi ces collections, celles dites anciennement d'Anatomie comparée regroupent l'essentiel des préparations ostéologiques de l'établissement, ainsi qu'une série importante de spécimens formolés, conservés entiers ou partiels. La plus grande partie des taxons de Vertébrés y est bien représentée par des animaux collectés dans le monde entier (Mammifères, Oiseaux, Reptiles, Amphibiens, Poissons) : l'ensemble compte environ 40 000 spécimens ostéologiques, plus ou moins complets, et environ 6 000 spécimens en fluide (Fig.1).

En avril 2002, dans le but de répondre aux nouvelles exigences conservatoires définies par l'établissement et surtout pour faciliter son exploitation, un chantier de rénovation de cette collection a été ouvert. L'inventaire informatisé de chacun des objets qui la compose constitue l'étape ultime de l'intervention en cours, à la suite des opérations de reconditionnement conservatoire et de rationalisation des conditions de stockage des spécimens. Pour mener à bien cet inventaire informatisé, la création d'une base de données dédiée et d'un outil d'interface spécifique s'est vite imposée. La contrainte principale était la nature particulière et distincte des objets : le développement de ces deux éléments avait pour but d'une part de permettre le stockage de toutes les données générées par l'examen de chacun des spécimens, et d'autre part d'intégrer les informations contenues dans les registres d'inventaire lors de sessions de récolement.

La finalité de ce travail est le développement de la base de données *Anatomic* et de l'application dédiée *Transanatomic*, qui intègre une série de modules pour l'alimentation et l'exploitation concrète des informations contenues dans la base. La principale particularité de la base *Anatomic* est de permettre l'informatisation d'un constat d'état détaillé pour chaque spécimen.

Outils et méthode

La base *Anatomic* a été développée sous le logiciel *Access* de manière expérimentale, à l'occasion de l'informatisation des spécimens ostéologiques initiée en 2002. Les champs ont été déterminés en fonction des besoins identifiés et en prévision d'une phase ultérieure d'exploitation des informations. Par ailleurs, des champs contenus dans une ancienne base monutable distincte, utilisée pour les inventaires antérieurs menés dès 1990 et qui abritent une série d'informations d'intégrité variable, ont été repris notamment pour permettre la migration et l'intégration de ces

anciennes données vers le nouveau modèle. Le choix des champs et de leur syntaxe a été, tout au long du développement, validé par Christine Lefèvre, chargée de conservation des collections anatomiques ostéologiques.

Anatomic est hébergée sur un serveur distant et fonctionne de manière partagée sous *Oracle*. Elle est administrée à l'aide des outils *SQL Plus*, *Access* via *ODBC*, et *Jacim*. *Jacim* est un outil d'interface polyvalente écrit en Java et développé au Muséum, qui a pour vocation l'alimentation et l'exploitation des bases de données sous *Oracle* des différentes grandes collections de l'établissement.

Transanatomic a été réalisé en *Visual Basic* sous *Access* et voit son développement se poursuivre conjointement aux opérations de rénovation en cours.

La base *Anatomic* est constituée d'une suite de 9 tables reliées entre elles selon le principe des bases de données relationnelles (Fig. 2). L'organisation méthodique du stockage de l'information et l'utilisation de liens numériques, simples ou complexes, a pour but l'exploitabilité et la fiabilité maximales des données saisies. Des principes de modélisation, couramment utilisés dans les bases de données scientifiques et dans les bases de collections d'Histoire naturelle du Muséum national, ont été ici repris et intégrés en fonction de leur pertinence quant à la résolution des problématiques rencontrées.

Anatomic permet en l'état la saisie de l'ensemble des informations issues des étiquettes et des notes accolées aux spécimens, ainsi que celles qui ont été reportées dans les différents catalogues papier. Les catégories d'information communes à tout objet de la collection sont contenues dans la table « Objets », de même que les références documentaires ou le lien vers une possible numérisation. Les différents embranchements reliés à cette table ont pour vocation de sectoriser des informations spécifiques : déterminations systématiques et éléments de taxinomie de chaque spécimen, localisation de sa provenance géographique, identification des intervenants dans la vie de chaque objet et constat d'état.

Le principe de modélisation retenu pour la caractérisation des intervenants, permet d'enregistrer et de caractériser pour chaque objet, et sans limitation quantitative, tous les intervenants ayant conduit à sa présence en collection. Les collecteur(s), donateur(s), mais aussi les préparateur(s) et conservateur(s) peuvent y être enregistrés par l'opérateur de saisie, et leur intervention datée. Le principe de la modélisation de cet embranchement s'inspire largement de la technique

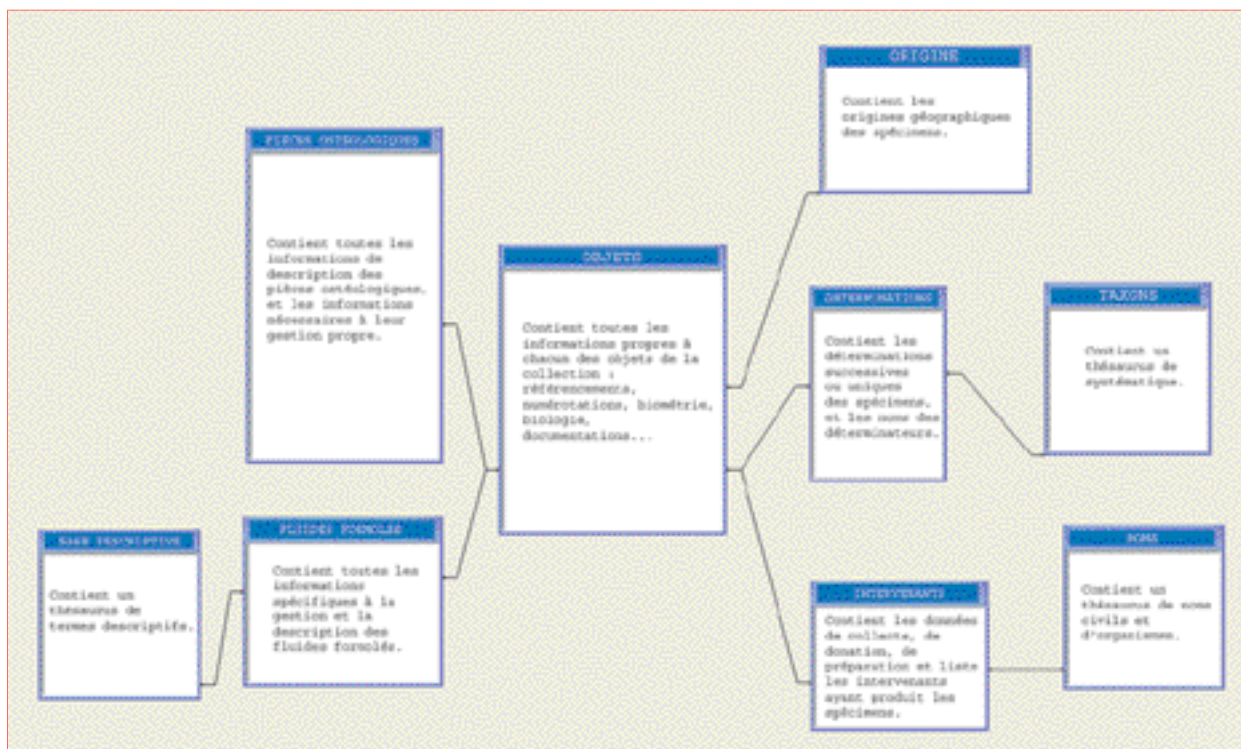


Fig. 2. Représentation synthétique du modèle relationnel utilisé pour *Anatomic* : liste des tables et description de leur contenu

communément employée pour l'informatisation des déterminations systématiques successives des spécimens de collection d'Histoire naturelle. *Transanatomic* est l'outil dédié à cette base, et sert d'interface entre l'utilisateur et les données. Il intègre une série de modules permettant à l'opérateur de saisie d'enregistrer de l'information et de l'exploiter de manière transparente, c'est-à-dire sans maîtrise et gestion manuelle des liens entre les tables. Les différents modules de *Transanatomic*, développés selon les besoins courants du chantier, permettent chacun de traiter une tâche spécifique via une fenêtre unique : outils de recherche et de mise en forme, outils de saisie, écran de correction et de récolement, module d'étiquetage automatique, impression d'un mémo de synthèse sur une séquence d'objets...

L'informatisation du constat d'état : objectifs et contraintes

L'informatisation est menée sur une collection constituée et exploitée depuis plus de deux siècles. Elle porte par conséquent sur des spécimens en état de conservation variable ou sur des lots d'objets parfois incomplets d'un point de vue anatomique (prélèvements, pertes partielles, dégradations diverses...). Or, afin que la collection garde à la fois toute sa valeur, d'exploitation scientifique entre autres, et que son informatisation

tienne compte de l'usure de certains objets, le choix d'intégrer dans la base *Anatomic* les informations relatives au constat d'état est prépondérant. Ces informations sont récoltées par les différents opérateurs de saisie lors de l'examen des spécimens, que ces derniers soient ostéologiques ou conservés en fluide.

Le premier objectif de l'informatisation du constat d'état via *Anatomic* est de faciliter et d'améliorer la qualité des réponses faites aux utilisateurs quant à l'exploitabilité de la collection (notion de requête). Grâce au descriptif informatisé de chaque objet, à la possibilité de les trier et de les rechercher dans la base selon ces différentes catégories d'information, *Anatomic* doit permettre un gain de productivité non négligeable dans la recherche et la localisation des spécimens utilisables pour chaque exploitation particulière. Les spécimens peuvent en effet être consultés ou prêtés à des fins scientifiques, muséographiques ou artistiques, à des personnels internes de l'établissement ou bien externes. La demande de consultation des spécimens se fait rarement sur la base d'une demande ciblée sur des objets précis, mais sur une série d'objets répondant à certains critères (taxon, origine de collecte, type de préparation...). La possibilité pour le conservateur ou le régisseur d'obtenir une liste pertinente à l'aide d'un outil simple est donc visée. Le second objectif est de permettre, sur la base des informations récoltées par les « inventaristes », de

dresser des vues synthétiques de la collection. Ces vues synthétiques correspondent à un cliché instantané de la collection pour un taxon donné, comme le ferait le résultat d'une évaluation fine. Cette synthèse peut aussi permettre au conservateur de prévoir les travaux à réaliser, qu'il s'agisse d'actions urgentes comme le marquage ou d'orienter l'accroissement ou les compléments à apporter à la collection. Enfin, ces vues synthétiques portant également sur l'état physique des objets ou sur la description de leur contexte conservatoire, elles peuvent constituer un outil d'aide à la décision et à la programmation d'opérations particulières de restauration ou de rénovation.

Pour permettre l'informatisation du constat d'état, une modélisation particulière a dû être développée. Un constat d'état détaillé intègre nécessairement une grande série d'informations qu'il est nécessaire de structurer pour en permettre l'exploitation en aval. La contrainte principale, dans le cas de la collection dite d'Anatomie comparée, tient dans deux éléments : la nature distincte des spécimens ostéologiques et des spécimens formolés, et le contexte conservatoire propre à chacun de ces deux ensembles d'objets. Une modélisation différenciée selon la nature de préparation des objets a ainsi été mise en place.

Enfin, le développement de l'outil *Transanatomic* répond également à deux contraintes principales, outre l'adaptation au modèle : la rapidité et la facilité de saisie des informations. L'objectif de cet outil est de permettre une réduction du temps de saisie, pour permettre à l'« inventariste » de se consacrer plus facilement à un examen complet des objets.

Les collections en fluide

Pour les spécimens en fluide (principalement en solution aqueuse de formol), j'ai choisi une méthode d'organisation verticale des informations, inspirée de la technique utilisée pour l'informatisation des déterminations systématiques ; ici, les informations ne sont pas contenues dans une ligne unique relative à un objet, mais sur plusieurs lignes, chacune porteuse d'une information spécifique.

Cette méthode répond à la contrainte posée par le rapport quantitatif des types d'information spécifiques aux objets formolés. Peu d'informations sont en effet nécessaires à la caractérisation anatomique de chaque objet, tandis que la description de leur contexte conservatoire en requiert une grande quantité (par exemple, nature des contenants, type de scellement des couvercles, état du fluide...).

Le constat d'état est explicité dans une table de croisement appelée « Fluides formolés », intégrant pour chaque ligne de donnée la référence à un objet et celle d'une information descriptive contenue dans le thésaurus que constitue la table appelée « Base descriptive » (Fig. 3). Dans la table « Base descriptive », les données de description sont hiérarchisées afin de permettre les requêtes thématiques. Ces données sont regroupées en deux blocs principaux : données de description anatomique et données de description du contexte conservatoire. Ces informations ont été déterminées suivant les besoins liés à l'organisation des activités de conservation et de rénovation.

Les données de description anatomique sont constituées d'une liste hiérarchisée d'ensembles anatomiques (organes complets, sections partielles ou animaux entiers). Les données de description de contexte recouvrent quant à elles l'ensemble des informations liées au milieu de fixation et au conditionnement : qualité, niveau et saturation du milieu, dimensions et nature des conditionnements... Elles ont été déterminées en collaboration avec Marc Herbin, chargé de conservation des collections anatomiques en fluide. Le niveau de détail maximum de l'information descriptive se trouve dans le champ DETAIL de la table « Base descriptive », et peut être caractérisé par une donnée variable dans la table de croisement ; c'est le cas par exemple de l'information relative à la localisation du lieu de stockage, pour laquelle le référencement de toutes les variables n'était pas pertinent en l'état actuel des locaux. La saisie par objet de chaque information descriptive se fera à terme grâce à un module de saisie en cours de développement au sein de *Transanatomic*, en collaboration avec le chargé de conservation.

Outre sa performance en ce qui concerne l'informatisation des collections en fluide, cette technique présente deux intérêts dans le cadre de certaines problématiques de gestion de collection. Elle permet de spécifier au cours des séances de travail de nouvelles informations descriptives qui n'auraient pu être prévues à la rédaction de son protocole, et cela sans intervention sur le modèle lui-même.

Par la simple saisie d'une ligne dans le thésaurus que constitue la table « Base descriptive », le conservateur ou le technicien peut également, suivant les nécessités ou l'évolution de ses besoins gestionnaires, spécifier et caractériser une nouvelle information descriptive à adjoindre aux objets. Associée à cette première caractéristique, l'absence de limite dans la quantité d'informations qu'il est possible d'adjoindre à chaque objet permet par ailleurs d'utiliser cette méthode pour

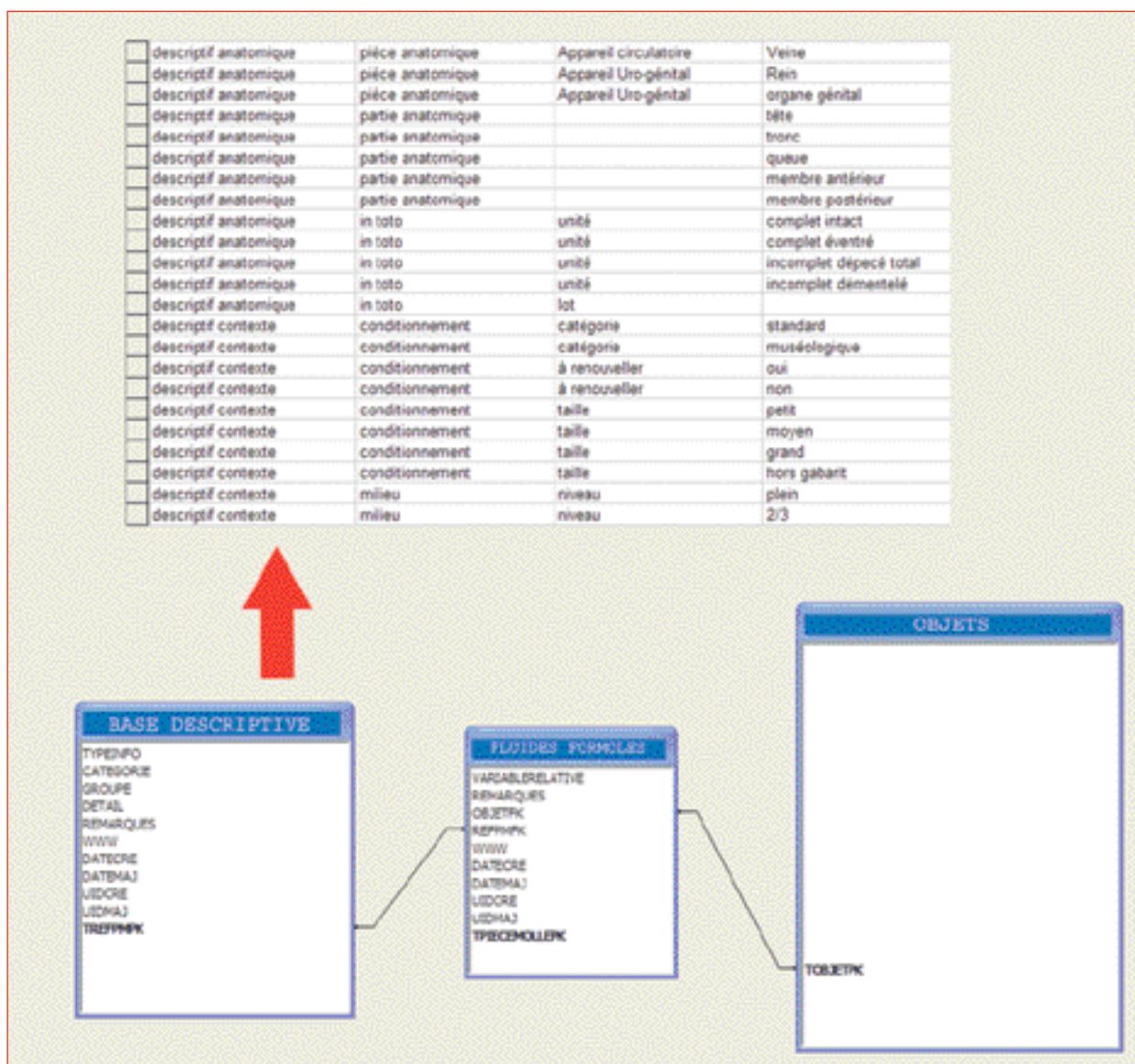


Fig. 3. Modélisation pour la description des fluides formolés : détail des tables spécifiques, et exemple séquentiel de quelques données descriptives thésaurisées

informatiser une collection hétérogène quant à la typologie de préparation des spécimens qui la composent.

Les collections ostéologiques

Pour les spécimens ostéologiques, le rapport quantitatif entre les informations de description anatomique et celles du contexte de conservation est l'inverse de celui des spécimens en fluide. De plus, le fait que les collections ostéologiques aient bénéficié historiquement et très globalement parlant de plus de soins que les collections en fluide, les données à intégrer ainsi que les informations à récolter pour l'organisation de la gestion en aval sont mieux connues. Pour l'informatisation des spécimens ostéologiques, j'ai donc choisi une modélisation moins

souple, mais plus directement adaptée à la gestion de ce type d'objet (Fig. 4).

Chaque spécimen est ici considéré comme un lot d'os différenciés, dont le nombre et la qualification dépendent à la fois du taxon d'appartenance (l'anatomie ostéologique dépend pour un individu, de sa place dans le règne animal) et de l'état du spécimen en tant qu'objet de collection. Le descriptif global est intégré dans une seule table « Pièces ostéologiques » dont les champs ne sont pas hiérarchisés entre eux. Une série restreinte de champs permet la saisie d'informations diverses liées à la gestion des objets ostéologiques ainsi qu'au contexte de chacun : rangement, référence du conditionnement utilisé et champs de contrôle pour

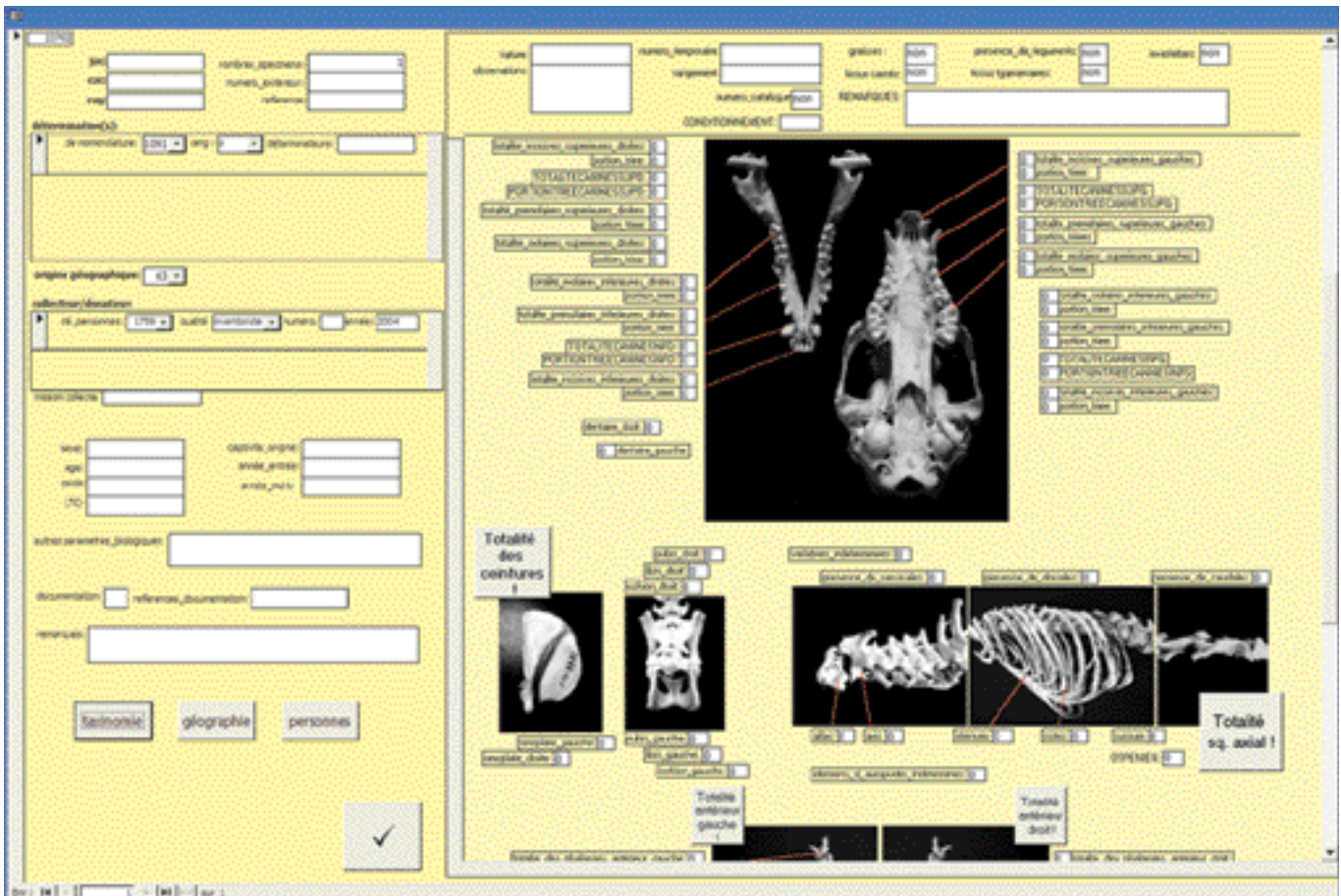


Fig. 5. Exemple d'un module de saisie réalisé pour *Transanatomic* et qui permet, grâce à un standard, l'inventaire des Carnivores. Pour chaque fiche à saisir, la planche anatomique à droite défile et permet la saisie en détail du codage de tout le squelette. L'incrément de la valeur « 1 » pour la présence d'un os individualisé se fait par le clic de sa cellule liée ; la saisie plus rapide pour tout un ensemble anatomique complet (crâne, ceinture, membre...) se fait par le clic de boîtes génériques attenantes aux illustrations. La sélection des champs à coder se fait automatiquement en tâche de fond lors du lancement du module.

L'informatisation de plusieurs milliers de spécimens, chacun constitué de plusieurs centaines d'éléments squelettiques, imposait la mise au point d'une technique de saisie rapide et efficace, fondée sur la modélisation des données et l'ergonomie de l'outil de saisie. Ensuite, pour la recherche de l'information anatomique, une organisation structurée et rationnelle permet de réaliser des requêtes tant sur des ensembles que sur des éléments isolés, en adaptant le rang de précision aux besoins de l'utilisateur.

Par ailleurs, le codage numérique permet de faciliter la manipulation des données pour des exploitations plus complexes que la simple consultation des informations brutes. *Transanatomic* intègre par exemple un module d'étiquetage automatique des conditionnements, pouvant générer pour une sélection d'objets un affichage graphique et conditionnel de l'état ostéologique par spécimen, fondé sur le calcul et la manipulation des

valeurs numériques (Fig. 6). La mise en forme et le choix du modèle d'étiquette sont fixés automatiquement par l'outil selon la référence du conditionnement employé, lors du lancement du calcul.

Un générateur automatique semblable à celui utilisé pour l'étiquetage permet aussi la mise en forme d'un mémo prêt à imprimer : il comptabilise pour le taxon sélectionné le nombre total de spécimens, le nombre de spécimens nécessitant des réparations, le nombre par type de réparation (réduction, dégraissage...), le nombre de spécimens par nature ostéologique, la liste des espèces contenues et la liste des localisations géographiques connues ; par ailleurs, la saisie de la référence du conditionnement permet à l'outil le calcul du volume physique de la série d'objets relative. La synthèse est faite automatiquement sur les données saisies par les opérateurs lors de l'informatisation.

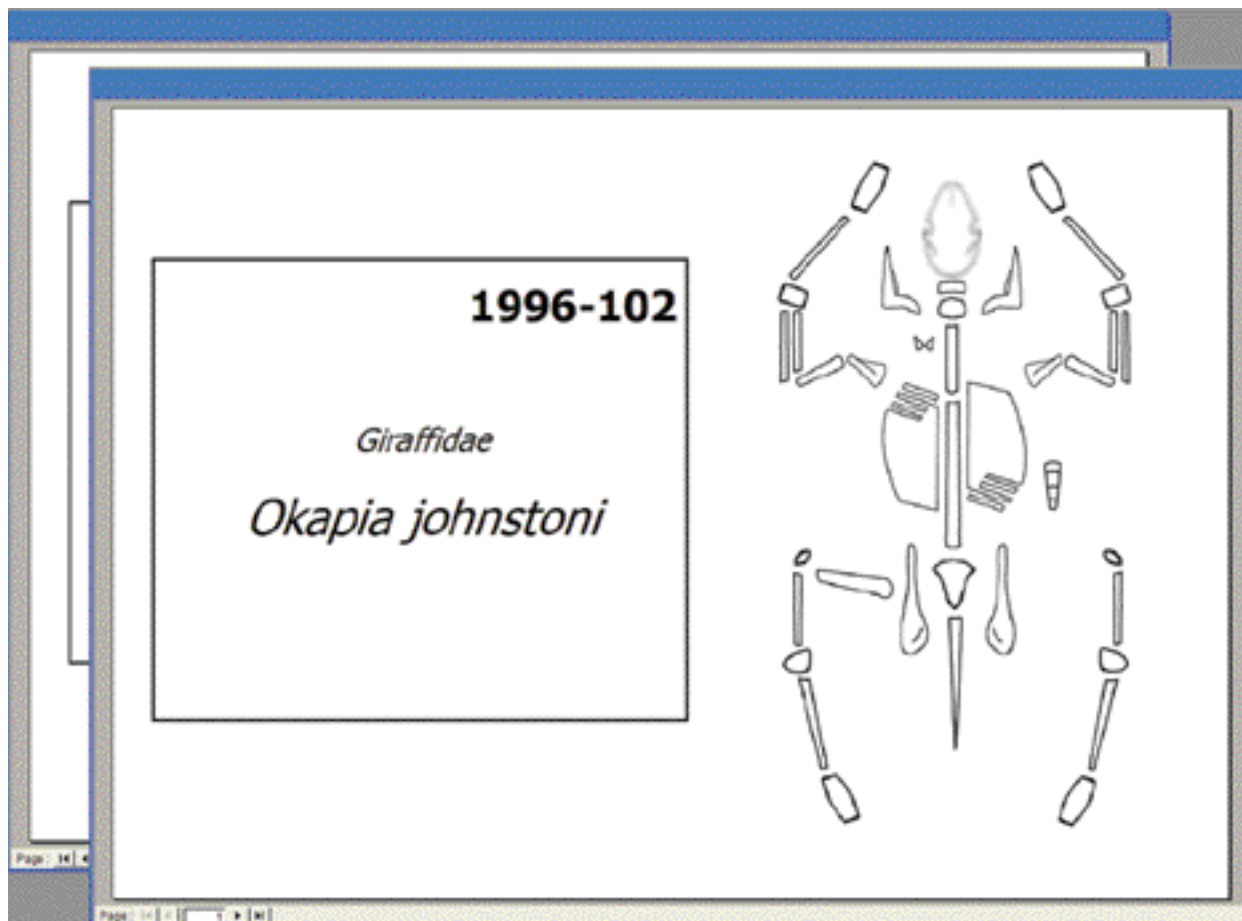


Fig. 6. Générateur d'étiquettes pour le conditionnement des spécimens de grande taille. La taille des conditionnements utilisés pour les Artiodactyles, permet sur une surface A4 de détailler l'état de chaque spécimen : pour chacun, le générateur établit et met en forme automatiquement un éclaté de squelette schématique, affichant les parties anatomiques présentes, en vue de faciliter le repérage des objets dans leur réserve de stockage. Sur l'étiquette de l'Okapi ici figurée, on voit qu'il manque le fémur et que la partie supérieure du crâne (calvarium), floue, est incomplète.

Enfin, de nombreuses autres exploitations peuvent être envisagées grâce à la manipulation de ces données. Le calcul d'indices de qualité basés sur la manipulation des données numériques, pourrait constituer par exemple un outil d'aide à la décision pour les chargés de conservation, notamment lors d'une campagne de restauration. Par ailleurs, la représentation graphique de l'état des spécimens pourrait être déclinée vers un interfaçage Web, si le choix était fait de la mise à disposition de ces informations via Internet.

Cette technique de modélisation peut être réutilisée dans d'autres contextes de collection. On peut ainsi l'envisager pour des collections d'objets homogènes, tant dans leur nature physique que dans leurs exigences de conservation.

Perspectives de développement

La base de données *Anatomic* et son outil dédié *Transanatomic* sont aujourd'hui opérationnels et peuvent accompagner le travail d'inventaire des collections anciennement dites d'Anatomie comparée. Si ces deux outils sont fonctionnels, de nombreuses pistes permettent cependant d'envisager leur amélioration dans le futur.

Premièrement, l'utilisation de référentiels normalisés devra permettre à terme d'améliorer la qualité de l'information contenue dans les thésaurus de systématique et de géographie. Leur développement ne sera pas assuré dans le cadre de l'informatisation de cette collection, mais leur intégration est prévue dès lors que des collaborations larges auront permis leur production. En ce qui concerne la modélisation d'*Anatomic*,

des améliorations pourront être apportées sur les branches intégrant les informations de constat d'état. Concernant les données descriptives des objets en fluide, la confrontation de la technique avec un inventaire qui reste à mener permettra de préciser l'organisation et la définition des données thésaurisées. En ce qui concerne les spécimens ostéologiques, une modélisation utilisant des techniques de croisement pourrait peut-être permettre l'allègement de la structure et la réduction du temps-machine actuellement nécessaire pour la manipulation des données de codage.

Deuxièmement, *Transanatomic* intégrera, selon la poursuite de l'inventaire, des modules de traitement spécifiques aux problématiques rencontrées. Il mettra à disposition des intervenants des outils adaptés à leur utilisation de la base, tout en permettant la manipulation facilitée des données anatomiques. Par ailleurs, l'expérimentation dans d'autres contextes de collection des techniques développées ici, pourrait permettre de les améliorer et peut être de diffuser plus largement le principe de l'exploitation informatique des constats d'états des objets de collection.

Enfin, si *Anatomic* reste pour le moment un outil à vocation interne, permettant de satisfaire sur notre site certains besoins liés à la gestion de la collection, on pourra à terme envisager d'en étendre l'exploitation. Au-delà des objets eux-mêmes, la base de données pourra permettre l'informatisation des prêts de spécimens et de la consultation de la collection. Par ailleurs, toujours en lien avec les tâches liées à la mise à disposition du matériel, l'exploitation des données que contient la base pourra également être rendue possible via Internet grâce à un interfaçage Web.

Je tiens à remercier particulièrement Jacques Cuisin pour toute l'aide apportée et pour son soutien. Je remercie également Christine Lefèvre, Marc Herbin et Stéphanie Bréhard d'avoir bien voulu accompagner ce présent travail.

Bibliographie

A Preliminary Worldwide Survey of Systematics Collections Holdings conducted for The Global Biodiversity Information Facility, *The Library American Museum of Natural History*, 2003, pp. 1-26.

L'informatisation des collections d'Histoire naturelle, *la Lettre de l'OCIM*, n°6, 1989, p. 17.

Bordereaux minimums pour les collections de Sciences naturelles, *la Lettre de l'OCIM*, n°6, 1989, pp. 18-19.

Davis, P. La documentation des collections de Sciences naturelles en Grande-Bretagne, *la Lettre de l'OCIM*, n°38, 1995, pp. 14-18.

Lancelot G. et Bailly N., Rapport sur la réunion nationale BioCASE, MNHN, 20 mars 2003, pp.1-13.

Ueberschär, B. and Teltow, M. FishBase goes fishing Bulletin, *EC Fisheries Cooperation Bulletin*, vol. 12 n°2-3, 1999, pp. 38-39.

Pauly, D. La science et FishBase, *EC Fisheries Cooperation Bulletin*, vol. 10 n°2, 1997, p. 6.