

DES BASES DE DONNÉES À LA PUBLICATION ÉLECTRONIQUE: UNE INTÉGRATION DES DONNÉES ET DES OUTILS DE RECHERCHE¹

1. INTRODUCTION

L'application de l'informatique à l'archéologie a commencé, dans les années 60, par la recherche documentaire, héritière des premiers travaux de mécanisation qui avaient été menés dans les années 50. Progressivement, ce mouvement d'informatisation a gagné toutes les étapes du travail de l'archéologue. Après plusieurs décennies, on peut affirmer que cette démarche a conduit à des avancées indéniables, même si les progrès ne se sont pas toujours produits là où on les attendait en priorité. Il ne peut être question ici de refaire l'histoire des systèmes d'information en archéologie, ni de faire de la prospective mais plutôt de la «veille» en analysant certains aspects de la situation actuelle. On sait que le développement d'un système d'information – c'est-à-dire sa conception, sa mise au point et son usage – se fonde sur des logiques d'ordre divers, parmi lesquelles on peut reconnaître les logiques techniques (les technologies qu'il met en oeuvre), les logiques d'utilisation (les besoins qu'il satisfait) et les logiques des pratiques sociales, logiques propres aux organisations mais aussi aux individus qui composent la communauté scientifique.

2. LES LOGIQUES TECHNIQUES

Si on envisage d'abord les **logiques techniques informatiques**, on constate que les systèmes d'information bénéficient de progrès constants, correspondant à divers aspects du traitement du signal numérique, et que leur mise en oeuvre *en synergie* offre des conditions techniques jamais atteintes jusqu'ici. Sans établir une liste exhaustive, on peut citer les progrès:

¹ Cette réflexion s'inscrit dans la continuité des travaux menés depuis plusieurs décennies au Centre de Recherche «Archéologie et Systèmes d'information» fondé, au début des années 70, par René Ginouvès à l'Université de Paris X-Nanterre. C'est la raison pour laquelle les exemples et les références cités sont choisis principalement dans l'environnement proche de la Maison qui porte désormais le nom de ce savant disparu en 1994. De plus, cette Maison abrite aussi le Centre de Recherche qui fut fondé par Jean-Claude Gardin, comme celui d'André Leroi-Gourhan: ce triple héritage méthodologique devrait nous rendre particulièrement vigilants dans la qualité de notre traitement de l'information en Archéologie! Mais cette réflexion concerne bien évidemment l'ensemble de la communauté scientifique, en France comme ailleurs, et bien des exemples peuvent être choisis dans d'autres pays, dont certains sont aujourd'hui plus avancés que la France dans ce domaine.

- de la puissance de traitement des matériels informatiques (à prix quasi constant),
- des technologies de saisie, de restitution et de stockage (là encore à prix quasi constant),
- des réseaux (le câblage à haut débit en cours de réalisation à grande échelle et l'interconnexion mondiale des réseaux),
- des systèmes d'exploitation, des logiciels réseaux, des logiciels applicatifs et des interfaces d'utilisation,
- des standards et des normes relatifs aux documents numériques et à leur traitement.

Les deux derniers points présentent un intérêt particulier pour notre discipline. En effet, c'est la convivialité des interfaces – qui reste à améliorer – qui nous permet une réelle appropriation de ces techniques et leur insertion dans notre pratique quotidienne: une communauté d'origine souvent littéraire, dans laquelle le niveau de culture informatique reste relativement faible – et qui a tant à apprendre dans son propre champ scientifique pour devenir efficace – ne peut, ni ne veut consacrer trop de temps à l'apprentissage de technologies en constante évolution, et dont la nouveauté même ne constitue pas un facteur d'intérêt particulier (BÉGUIN 1996). Par ailleurs, le travail archéologique ne peut être ni ponctuel – il se situe dans la durée –, ni solitaire – nous travaillons nécessairement en équipes –, ce sont donc l'adoption et la pérennité de standards et de normes qui garantissent la transmission, la restitution et l'archivage de nos données une fois numérisées.

Les **technologies documentaires**, quant à elles, sont les héritières de techniques déjà anciennes et éprouvées. Il n'est pas inutile de rappeler que la représentation du contenu des documents – nécessaire lorsque les documents sont trop nombreux pour que la mémoire humaine puisse s'y retrouver sans aide – a commencé à être pratiquée dans la Bibliothèque d'Alexandrie au IIIe siècle avant J.-C., sous l'impulsion de Callimaque, à la fois «bibliothécaire» et savant, c'est-à-dire utilisateur des informations contenues dans les documents qui y étaient rassemblés et conservés. Depuis l'Antiquité, les techniques de l'indexation manuelle ont été perfectionnées, ainsi que les outils – matériels comme intellectuels – qui permettent de gérer ces indexations (catalogues, fichiers, mais aussi lexiques, thésaurus...). Le fichier mécanographique et le code d'analyse des outils de l'âge du bronze, mis au point par J. Deshayes et J.-Cl. Gardin en 1955, étaient à la fois les héritiers de ces techniques anciennes et les précurseurs des «nouvelles» technologies documentaires d'aujourd'hui. Les banques – devenues bases – de données archéologiques leur doivent beaucoup, ainsi que les systèmes d'analyse mis au point par plusieurs équipes à partir des années 60.

Les recherches linguistiques alliées aux problématiques de la recherche documentaire ont conduit progressivement à mettre au point les processus de l'indexation automatique du texte intégral puis, plus récemment, le traitement du langage naturel grâce à des analyses morphologiques, syntaxiques puis sémantiques. Les outils les plus performants aujourd'hui savent allier les techniques de l'indexation manuelle – pour les types de données rebelles à l'indexation automatique, la notion de datation, par exemple, exprimée de façons si diverses selon les contextes – et celles de l'indexation automatique qui, au prix de moindres performances en matière de silence et de bruit, permettent à la fois de traiter de très vastes corpus de textes et de les indexer en fonction non du point de vue retenu a priori par l'analyste mais, à la demande, selon les centres d'intérêt de l'utilisateur. La GED, qui assure la gestion documentaire de documents numérisés, a enrichi les fonctionnalités offertes par ces bases de données.

Un autre progrès important tient, d'une part, aux possibilités du multimédia, d'autre part, à la diversification des modes d'accès aux informations: à l'accès indexé traditionnel, le plus efficace pour un utilisateur averti ou pour un volume important de données, se sont ajoutés l'accès hiérarchique et l'accès hypertexte. Dans l'accès hiérarchique, on propose à l'utilisateur une série de choix dont chacun conditionne le choix suivant: sans connaissance préalable, il accède donc aux informations selon un schéma préétabli qui le guide à chaque étape. L'accès hypertexte offre les mêmes avantages pour un utilisateur novice mais l'accès aux informations y est plus souple, l'utilisateur pouvant à chaque moment choisir le chemin qu'il va suivre – et la catégorie d'information qu'il souhaite – parmi ceux que l'auteur du système d'information a prévus.

Ces différents modes d'accès aux informations ont d'abord correspondu à des logiciels différents, puis de nouvelles versions ont permis de les combiner, que ce soit dans des applications en ligne du type client-serveur, des applications distribuées sur des supports optiques ou, enfin, sur Internet. Sur ce nouveau média de communication, en effet, on a d'abord construit des outils comme WAIS pour l'accès indexé, GOPHER pour l'accès hiérarchique, et enfin World Wide Web (WWW ou Web) pour l'accès hypertexte-hypermédia. On connaît le succès remporté par ce dernier outil qui permet de mettre quantités d'informations multimédia en ligne avec un accès à la fois souple et facile pour des utilisateurs ainsi affranchis de trop longs apprentissages techniques. Le volume des informations ainsi disponibles est devenu tel que l'accès hypertexte est désormais insuffisant et, sur le Web, on combine les divers modes d'accès, en y ajoutant en particulier l'accès indexé, grâce à des moteurs de recherche dédiés à un site, ou bien généralistes.

Ces moteurs de recherche travaillent en indexation automatique, mais aussi à partir d'une indexation «manuelle» des données qui apporte les avan-

tages complémentaires précédemment signalés. On le voit, au cours des dix dernières années, les progrès des techniques informatiques et des techniques documentaires sont allés de pair pour offrir des plates-formes techniques propres à la réalisation de meilleurs produits d'information. En outre, cette évolution des systèmes d'information correspond à un nouvel état d'esprit des documentalistes: «médiateurs de l'information spécialisée», ils s'occupent moins de l'offre que de la demande et ils prennent davantage en compte les besoins des utilisateurs, tant dans le choix des informations enregistrées que dans leur forme – notamment grâce au multimédia – et dans leurs modes d'accès.

3. LOGIQUES DES PRATIQUES PROFESSIONNELLES DES ARCHÉOLOGUES: LES BESOINS DE TRAITEMENT DE L'INFORMATION DES ARCHÉOLOGUES AUX DIFFÉRENTES ÉTAPES DE LEUR TRAVAIL

Comme toute science, l'Archéologie *est* traitement d'information. En effet, celui-ci intervient à toutes les étapes du travail – qu'il s'agisse de la prospection, de la fouille, de l'élaboration des corpus d'étude, de leur interprétation, de l'intégration de ces résultats dans le corpus des connaissances acquises puis de leur publication (GUIMIER-SORBETS 1996). Dans un schéma de formalisation du processus – non linéaire – du travail archéologique, on peut distinguer trois étapes: la collecte des données, leur traitement en vue d'une (nouvelle) structuration conduisant à leur (nouvelle) interprétation – phase de recherche proprement dite –, et la diffusion des résultats. A chacune de ces étapes, on travaille sur des représentations de divers types (description en langage naturel, description formalisée, dessin, photographie, tableau chiffré...) constituant chacune un document – numérique ou non – qu'il faut archiver pour un traitement ultérieur: représentation et archivage précèdent et suivent chacune des trois étapes déjà évoquées (pour un exemple d'analyse GUIMIER-SORBETS, SEIF EL-DIN sous presse).

Si, dans les premiers temps, les applications des technologies informatiques à l'archéologie avaient surtout concerné la recherche documentaire (bases de données factuelles comme référentielles), puis la recherche de structuration (utilisation des traitements statistiques pour établir des classifications et des sériations automatiques) (cf., par exemple, DJINDIAN 1991) – qui répondent en partie aux besoins de collecte et de traitement (recherche) – on constate qu'aujourd'hui ces technologies répondent aussi aux besoins de la troisième étape, celle de la publication et de la diffusion des résultats de la recherche.

Il ne peut être question ici d'envisager toutes les modalités d'intégration des nouvelles technologies aux différentes étapes du travail de l'archéologue. En particulier, nous ne reviendrons pas sur les GIS/SIG, dont le déve-

loppement durant la dernière décennie constitue pourtant un apport majeur à l'archéologie, puisqu'il permet d'intégrer l'espace parmi les critères pris en compte (pour un bilan méthodologique de l'utilisation des SIG/GIS en archéologie dans différents pays, cf. MOSCATI 1998). Nous limiterons à évoquer la représentation des informations par divers types d'images puis la publication électronique.

3.1 La représentation par des images

De tout temps, les archéologues ont complété leurs descriptions par des représentations iconiques, utilisant les techniques disponibles aux différentes époques pour rendre aussi fidèlement que possible par des dessins, des aquarelles, des photographies, etc, les aspects de la réalité qu'ils étudiaient. La numérisation permet désormais une intégration plus facile de ces documents comme forme de représentation des données archéologiques, et, à ces images traditionnelles, l'informatique permet d'en ajouter de nouvelles. En effet, qu'elles soient fixes ou animées – et alors souvent couplées à du son – ces images peuvent être d'origine numérique (lorsqu'elles sont directement saisies par un appareil photo ou une caméra numérique, par exemple), ou être produites par la numérisation secondaire d'un document analogique (un tirage, un négatif photographique, ou un dessin numérisé dans un scanner), ou encore être directement produites par des logiciels et on parle alors d'image de synthèse: le logiciel calcule les images en les déterminant point par point suivant des formules mathématiques, ce qui permet de réaliser des images en deux dimensions (2D) ou en trois (3D); dans ces dernières, toutes les faces d'un objet, même invisibles à un moment donné, sont codées et mémorisées et il est donc possible de faire tourner l'objet et de le voir sous toutes ses faces; ces images de synthèse peuvent à leur tour être intégrées à des images traditionnelles, des photographies par exemple, ce qui ouvre des possibilités de représentation et de simulation inconnues jusqu'ici, et que les archéologues commencent à utiliser quand ils en ont les moyens.

Ainsi, par exemple, l'ensemble des bâtiments de Marmaria à Delphes ont-ils pu être restitués puis présentés dans le paysage naturel, à la place des ruines qui y sont conservées actuellement². Plus ou moins développées, les utilisations de l'image de synthèse commencent à être relativement nombreuses en archéologie, comme l'a montré Archéo Virtua, le festival international du multimédia pour l'archéologie, organisé sur l'Archéodrome de Bourgogne en Mars 1999 (<http://www.archeodrome-bourgogne.com>).

Aujourd'hui, les images de synthèse sont surtout utilisées lors de la phase de diffusion des résultats de la recherche à destination du grand pu-

² Pour plus d'informations sur le travail mené par l'Ecole française d'Athènes en collaboration avec le Mécénat technologique d'EDF, voir ci-après.

blic. Certaines sont présentées lors d'expositions (par exemple, les phases de construction du Palais d'Iraq al-Amir lors de l'exposition «Jordanie, sur les pas des archéologues», à Paris, Institut du Monde Arabe, 13 Juin-5 Octobre 1997), et dans les publications électroniques en ligne ou hors ligne (sur des cédéroms essentiellement, comme la reconstitution d'une tombe à char, produit dans lequel l'auteur s'est attaché à montrer non seulement le résultat de la recherche mais aussi les étapes du travail qui y conduisait: THOUVENOT 1999). L'intérêt de ces images pour des fins pédagogiques est évident – du moins lorsque la qualité technique se double de la qualité scientifique des informations mises en oeuvre. En raison même de leur puissance évocatrice, on se doit de les utiliser avec beaucoup de rigueur et de n'y inclure que la représentation d'hypothèses très fortement étayées car le public qui les gardera en mémoire n'a pas les éléments d'appréciation nécessaires à une interprétation critique. Et, de même qu'il existe des textes de vulgarisation distincts des textes pour les spécialistes, il faudrait savoir à quel type d'images de synthèse on a affaire: des images simplifiées pour le grand public, ou des images plus complexes pour les spécialistes, images qui pourraient représenter les hypothèses alternatives de restitution.

On constate que, pour l'instant, ces images de synthèse sont plutôt destinées au grand public; le prix de leur réalisation explique en partie cet usage, mais il faut souhaiter que les spécialistes sachent aussi utiliser ce mode de représentation pour les publications destinées à la communauté scientifique: l'intérêt d'une restitution graphique de type traditionnel dans une publication d'architecture n'est plus à démontrer – alors qu'il reste limité à un angle de vue – et cela permet d'apprécier l'intérêt que présenterait ce mode de restitution plus souple et plus puissant. Ainsi le travail qui est en cours de réalisation à l'Université de Caen depuis 1995 sur la maquette virtuelle en 3D de Rome au IV^e siècle permet-il d'explorer les quartiers et même certains bâtiments de la Ville (<http://www.unicaen.fr/rome>).

L'utilisation des images de synthèse lors de l'étape de recherche est encore rare mais on peut citer quelques exemples, pour lesquels les images de synthèse ont surtout servi à tester des hypothèses. La reconstitution virtuelle en 3D du temple de Karnak (Egypte) faite par l'équipe du Mécénat technologique d'EDF en collaboration avec les chercheurs du site a permis de tester la largeur nécessaire au passage de la barque solaire dans les couloirs étroits du sanctuaire; la même équipe d'EDF, en collaboration avec l'Ecole française d'Athènes, a pu montrer, par des modélisations, que l'ordre intérieur de la Tholos de Marmaria comportait deux séries de colonnes superposées (BOMMELAER 1996). De même, au Centre d'Etudes Alexandrines (CEA-CNRS), on utilise la modélisation 3D des volumes des diverses parties de la nécropole de Gabbari à Alexandrie pour retrouver la succession des phases de creusement des différentes salles des tombes souterraines.

La valeur des images comme représentation des données archéologiques est désormais reconnue, comme la nécessité de leur conservation, de leur stockage, de leur indexation pour la mise à disposition des chercheurs pour des exploitations ultérieures. Des médiathèques existent dans des institutions de recherche archéologique et il faut insister sur la nécessité du traitement documentaire des archives scientifiques. Toutefois, on doit signaler que la recherche de photographies comme de dessins ne peut être faite qu'à travers une indexation associée sous forme textuelle (mots-clés et/ou légende en texte libre) selon des formules en partie standardisées. S'il est encore trop tôt pour que ces modélisations 3D puissent constituer une base documentaire pour la recherche, elles le deviendront quand l'ensemble disponible atteindra une masse critique; il n'existe toutefois pas encore de solution satisfaisante pour leur archivage à long terme et leur indexation documentaire (HÉNON, LÉVEILLÉ 1997).

3.2 La publication électronique en archéologie

En même temps qu'elle apportait un gain certain à la recherche archéologique, la représentation par des images de ces données a permis de les ouvrir à un public beaucoup plus large, et elles ont pris une place importante dans les publications électroniques, diffusées sur cédéroms comme sur les sites Web. Ainsi, dans la continuation des quelques vidéodisques archéologiques créés à la fin des années 80, ont été édités une série de cédéroms destinés au grand public. Tout en tirant parti de l'interactivité, ces premiers cédéroms étaient conçus comme des transpositions électroniques de produits d'information déjà existants, imprimés ou audiovisuels. Ainsi, par exemple, par ses notices et illustrations associées, le cédérom sur le Musée de l'Homme reprenait la forme des catalogues papier avec quelques possibilités d'accès hiérarchique par des menus (AA.VV. 1995a). Par ses séquences d'images animées, le cédérom consacré à la découverte du San Diego, au contraire, s'inspirait plutôt des documentaires audiovisuels (AA.VV. 1995b).

Les cédéroms à vocation culturelle destinés au grand public ont connu un relativement grand développement entre 1994 et 1997, en Europe et tout particulièrement en France et en Italie, tandis qu'aux Etats Unis les produits se développaient directement sur Internet. La majorité des cédéroms culturels a été consacrée à des musées et/ou à des artistes connus d'un très vaste public (GUIMIER-SORBETS 1998a): dans cette perspective dictée par les lois du marché, l'art antique et les civilisations anciennes ont été davantage traités que l'archéologie proprement dite, et c'était souvent sans le concours des spécialistes du domaine. La qualité de ces produits a progressivement augmenté, d'abord fondée sur un enrichissement de la forme des données, puis de la qualité scientifique de leur contenu (par exemple COPPENS 1996), enfin grâce à l'adjonction de fonctionnalités permettant de nouveaux

usages: dans les meilleurs produits, l'utilisateur dispose de divers moyens d'accès aux informations, il peut les réélaborer pour son usage personnel et enregistrer le résultat sur son disque dur (GUIMIER-SORBETS 1998b; pour un exemple de cédérom aux fonctionnalités étendues cf. AA.VV. 1995c). Ainsi données multimédia et outils de recherche sont directement associés, ce qui permet à l'utilisateur de tirer parti des nouvelles fonctionnalités offertes par ces produits d'information.

Parallèlement, les jeux sur cédéroms se sont développés et les meilleurs produits français du genre ludo-culturel ont été édités par la Réunion des Musées nationaux (RMN, Direction des Musées de France): après une énigme au Château de Versailles, les titres suivants ont concerné une tombe de l'Égypte pharaonique puis la Cité interdite de Pékin³. Ces produits ont connu un réel succès commercial et, par leur qualité tant fonctionnelle que scientifique, ils ont constitué un excellent support de diffusion, vers le grand public, des connaissances issues des milieux de la recherche.

Depuis deux années environ, on constate un ralentissement de la production de ces cédéroms grand public mais un intérêt toujours croissant pour la diffusion d'informations sur Internet. Qu'en est-il de la publication électronique à destination des spécialistes? Elle a connu un développement plus lent et progressif, tant sur cédérom que sur Internet (ROSS, HIGGS 1993). Des recherches et des expérimentations ont été rapidement suivies des premières réalisations. Les expérimentations ont porté, là encore, sur la nature des informations, leur type de représentation, et les diverses méthodes d'accès. Parmi ces travaux français, on peut citer la publication archéologique sur CD-ROM, éditée par P. Arcelin et l'équipe de Lattes (ARCELIN 1997; ARCELIN, RICHET 1990), cédérom expérimental démontrant l'intérêt de diverses formes de publications; on peut citer aussi l'expérimentation sur des textes de publications traditionnelles relatives au site de Delphes qui montrait l'accès à des textes variés, des images et des notices de base de données à travers une interrogation combinée en langage naturel et par des mots-clés⁴. D'autres travaux en cours dans le même centre de recherche portent maintenant sur les types de liens (fonctionnels pour la navigation comme sémantiques) qui sont nécessaires pour permettre à des chercheurs une consultation plus riche à travers un ensemble de textes et d'images relatifs à un même thème⁵.

³ Aux éditions RMN, Canal+ Multimédia, Cryo Interactive Entertainment, Versailles, *complot à la cour du Roi-Soleil*, 1996; *Égypte 1156 av. J.-C., l'énigme de la tombe royale*, 1997 et *La Cité interdite*, 1998.

⁴ Cette expérimentation utilisait le logiciel Spirit (société T.GID), logiciel documentaire très puissant permettant une recherche en langage naturel avec traitement linguistique. Cf. GUIMIER-SORBETS 1993.

⁵ Cette expérimentation, complémentaire de la précédente, utilise le logiciel Acrobat (société Adobe), standard du marché de l'édition électronique. Son objectif est de montrer

La revue *Gallia-Informations*, chronique des fouilles en France publiée par le CNRS, vient de sortir un premier numéro sur cédérom. Sous sa forme électronique, la revue offre à la fois les données de type traditionnel (notices textuelles, cartes, photographies...) mais aussi des moyens d'accès beaucoup plus variés et puissants que la version papier des numéros précédents, puisque la recherche peut se faire par des grilles chrono-culturelles, par une interrogation croisée à partir du thésaurus, ou par une recherche dans le texte intégral. Elle offre aussi les index cumulés des numéros précédents et l'ensemble constituera progressivement un ensemble cumulatif. Ce produit est un excellent exemple des possibilités nouvelles d'intégration des données et des outils de recherche offrant un large choix de modes d'accès aux informations.

Une constatation similaire peut être faite à propos des informations disponibles actuellement sur le Web. Les ressources informationnelles utiles aux spécialistes en archéologie sont beaucoup plus réduites que celles qui sont destinées au grand public, mais on trouve aussi bien des informations diffusées de façon ponctuelle sur des sites que des publications déjà organisées en série, comme les périodiques électroniques: ainsi en 1998, on a recensé sur Internet une quarantaine de revues destinées a priori aux spécialistes, dont 35 sont issues de l'édition traditionnelle et cinq ont été entièrement conçues pour le réseau (Zaïd 1999). Le dispositif d'information archéologique sur Internet se met en place, et on y retrouve les diverses modalités tant techniques que professionnelles utilisées par la plupart des autres communautés scientifiques. Et d'ores et déjà les ressources informationnelles disponibles sur Internet sont tellement abondantes qu'il est nécessaire, pour les trouver, de passer par des outils de recherche: des guides/annuaires cartographiant ces ressources ou encore des «moteurs de recherche» spécialisés ou généralistes, travaillant à partir de l'indexation soit manuelle soit automatique de chaînes de caractères, sur des sites particuliers ou sur l'ensemble des sites préalablement repérés. De plus telle qu'elle est organisée aujourd'hui, l'information utile pour les spécialistes n'est pas seulement constituée de données publiées, quelle que soit leur forme de représentation, mais aussi des outils de recherche qui leur sont associés et des liens hypermédia qui proposent un sous-réseau de consultation spécialisée. Ainsi, il est possible de se brancher sur un site Web aussi bien pour consulter les informations qu'il publie que pour utiliser les liens qu'il propose vers d'autres ressources, elles-mêmes constituées de données et d'outils de recherche. Là encore, on retrouve l'intégration des données et des outils de recherche qui nous paraît être une caractéristique majeure de l'évolution de cette dernière décennie.

à quelles conditions l'édition électronique peut constituer une réelle valeur ajoutée à un ensemble de textes rédigés et publiés dans une perspective d'édition traditionnelle.

Toutefois, si les technologies offrent des solutions en amélioration constante pour la majorité de nos besoins d'information, il reste un besoin encore mal couvert alors qu'il est fondamental pour une science d'érudition comme l'archéologie: c'est celui de l'archivage des données publiées sous forme électronique. On sait que nous travaillons tous nécessairement sur les données publiées, données récentes comme données anciennes. Le papier, quoique fragile, reste consultable sans besoin d'appareillage, mais on connaît les difficultés d'accès aux ouvrages rares et anciens, et aux archives de fouille. La numérisation offre évidemment l'avantage majeur de la dématérialisation du document, qui autorise des accès distants et multiples dans de bonnes conditions de conservation, mais la pérennité des conditions de consultation des documents électroniques ne présente pas aujourd'hui les mêmes garanties que celles du papier. Si les outils présents sur Internet (par exemple, les robots de recherche) permettent une bonne maîtrise des flux d'informations à l'échelle internationale, ils n'assurent pas la gestion des stocks de données. Il faudra donc que notre communauté scientifique complète le dispositif technique et documentaire pour la nécessaire conservation des archives numériques.

4. LOGIQUES DES PRATIQUES SOCIALES DES ORGANISATIONS ET DES INDIVIDUS

Pourquoi l'usage de la publication électronique sur Internet n'est-il pas plus développé dans notre communauté scientifique? Même si des informations potentielles existent, dispersées sur de nombreux serveurs institutionnels comme sur des pages personnelles, une étude réalisée récemment sur les données relatives au Moyen Age montre que «90% des informations trouvées n'aident nullement le chercheur dans l'exercice de sa discipline» (DUCOURTIEUX 1998). On ferait vraisemblablement la même constatation pour l'ensemble de l'archéologie: si on trouve si peu de données intéressantes pour les chercheurs, c'est que la communauté des archéologues n'utilise pas – ou très peu – ce *medium* pour publier ses résultats, alors qu'il présente, malgré ses inconvénients, des avantages indéniables face aux limites et à la lourdeur des publications traditionnelles.

Puisqu'on a constaté la permanence des besoins de traitement d'information des archéologues et la progression constante des technologies qui peuvent y apporter des solutions de plus en plus adaptées, on pourrait tirer des conclusions très optimistes quant à leur utilisation par les archéologues, comme on l'avait fait pour l'essor des bases de données dans les années 70 et encore au début des années 80. Mais il convient, cette fois, de ne pas négliger les logiques sociales – qui n'ont guère changé depuis cette époque – et d'analyser leur fonctionnement. Or on sait qu'aujourd'hui les seules banques de données archéologiques un peu importantes sont des bases référen-

tielles, qu'il s'agisse de catalogues de bibliothèques, de photothèques ou d'archives, et que ces entreprises sont assurées par des institutions qui y voient un intérêt pour la gestion automatisée de leur fonds ainsi qu'une «vitrine» valorisante pour leur notoriété.

Les bases de données factuelles larges et ouvertes à tous n'existent pratiquement pas, alors que se multiplient les fichiers documentaires réalisés par un chercheur pour la durée de son étude. Les bases dont certains d'entre nous avaient rêvé autrefois – et dont nous avons commencé la réalisation – n'existent plus, ou presque plus, même si un manuel d'archéologie récent rappelle leur intérêt et les commente en ces termes «il faut saluer l'abnégation, en quelque sorte, des chercheurs qui oeuvrent ainsi pour le bien commun» (JOCKEY 1999)⁶. Les bases factuelles actuellement vivantes sont généralement l'émanation d'un groupe de personnes collaborant à une même recherche et se partageant la constitution et les bénéfices scientifiques de l'exploitation de l'outil documentaire commun: le partage de ces ressources informationnelles est une part du ciment qui lie de petites communautés d'intérêt scientifique, à l'intérieur d'un même laboratoire ou, plus souvent, dans ces réseaux informels largement internationaux qu'on nomme maintenant les «collaboratoires». Ils tirent parti des technologies de l'Intranet, réseau d'accès réservé, pour partager des informations à distance *via* Internet: voilà un type d'usage de l'Internet qui devrait se développer en Archéologie. Mais quel sera l'avenir de la publication électronique destinée aux spécialistes? Nous allons essayer de montrer que, si les logiques individuelles n'y conduisent pas aujourd'hui, les logiques des institutions, au contraire, y conduiront peut-être, et qu'alors la position de l'ensemble de la communauté pourrait s'en trouver modifiée.

Aujourd'hui, les informations publiées sur des pages individuelles du Web ne sont presque jamais qualifiées et leur valeur n'est pas reconnue par la communauté scientifique. Elles sont rarement signées sur les serveurs, et encore moins soumises à l'expertise d'un comité scientifique reconnu. Lorsqu'un chercheur les utilise, il ne pense pas à en signaler la source – la même pratique avait déjà été constatée pour les informations issues des bases de données – et l'auteur risque ainsi de se trouver dépossédé de ses découvertes à moins qu'elles n'aient d'abord été publiées sur un support traditionnel; et l'intégrité d'une publication seulement électronique des résultats n'est pas facilement garantie. De plus, les institutions ne tiennent pas compte de ces formes de publication dans les procédures d'évaluation pour la carrière des chercheurs. Pour toutes ces raisons, il est facile de comprendre que, dans la

⁶ L'emploi du terme «abnégation» montre bien que la communauté scientifique ne prend pas en compte ce type de travail dans l'évaluation de la notoriété et de la carrière de ceux qui réalisent ces bases documentaires, dont la qualité même, pourtant, est fonction de la valeur scientifique des données qu'elles renferment.

situation actuelle, les logiques individuelles ne conduisent pas à publier des résultats d'abord sur Internet.

La logique des pratiques éditoriales des institutions pourrait être différente. En effet, si elles considèrent que leur renom international ne peut plus se passer d'une présence sur le Web, ces institutions vont être amenées à réfléchir au contenu de leurs sites. La rentabilité d'un site Web, c'est sa notoriété, c'est-à-dire le nombre de consultations, évaluées – au moins de façon superficielle – par le compteur des visiteurs. Il faut donc assurer le plus grand nombre possible de connexions, et on commence à voir quels moyens sont mis en oeuvre, stratégies «superficielles» comme l'indexation par les métadonnées qui assure le plus large repérage possible par les moteurs de recherche les plus en vogue⁷, ou encore l'attrait des images. Mais si ces moyens peuvent accroître le nombre de connexions, ils ne sont pas suffisants pour attirer les «internauts» professionnels dont on attend la reconnaissance. Seule, en définitive, la qualité du contenu informationnel et sa valeur d'usage font la réelle valeur d'un site Web, et donc sa rentabilité. Pour cela, il ne suffit pas d'y placer des informations intéressantes et validées, il faut qu'elles soient à jour et qu'elles changent suffisamment souvent pour que l'utilisateur potentiel revienne les consulter régulièrement. Et il est évident qu'après la phase actuelle où l'on trouve surtout des informations sur l'organisme éditeur du site, il va falloir y mettre des informations qui ont une valeur d'usage particulière et on peut donc estimer que c'est par là que les informations professionnelles sur Internet vont se développer en se diversifiant.

Ainsi, par exemple, le site de l'Ecole française d'Athènes (EFA), offre, en plus des informations relatives à ses chantiers de fouilles et à ses publications, la consultation en ligne de son catalogue de bibliothèque; la prochaine étape, à l'étude, concerne les chroniques de fouilles actuellement publiées sous forme traditionnelle dans la revue de l'EFA, le *Bulletin de Correspondance hellénique*. Cet exemple, choisi parmi d'autres, montre comment l'institution offre des services de plus en plus étendus, en ouvrant la consultation des catalogues de ses fonds documentaires: on retrouve les bases de données référentielles déjà évoquées; l'institution pourrait offrir aussi un accès distant à des documents numérisés, ouvrages rares ou archives scientifiques, et c'est toute la question des bibliothèques et fonds d'archives numériques. Les institutions vont-elles se contenter de signaler les documents qu'elles possèdent et en réserver la consultation à ceux qui se rendent sur place, ou bien vont-elles offrir le service d'une consultation distante de ces ressources documentaires? Par la consultation des chroniques de fouilles, l'EFA ferait une

⁷ Une indexation trop large va entraîner du «bruit», c'est-à-dire va conduire jusqu'au site des usagers qui ne seront pas intéressés, mais si le nombre de visiteurs est l'objectif principal du site, on comprend le risque de dérive entraîné par de telles pratiques.

transposition électronique – et cumulative – de ressources déjà publiées de façon traditionnelle. Mais il est possible d’imaginer que ces institutions souhaiteront, toujours dans la même logique, publier en outre des informations originales, dont la qualité sera validée par leur renom scientifique. La réputation d’une institution tient à la qualité et au volume de ses publications scientifiques, on peut penser que les publications électroniques joueront un rôle au même titre que les publications papier.

A ce moment là – et à ce moment là seulement – les stratégies individuelles risquent de changer: les institutions auront besoin d’«auteurs électroniques», la communauté mise en confiance par le support et la validation institutionnels utilisera les données – en citant leurs auteurs – puis, souhaitons-le, elle prendra ces publications en compte dans les évaluations de notoriété et de carrière. Le «nouveau» (?) *medium* sera utilisé, pour ses qualités propres, en complémentarité des publications traditionnelles. Restera alors à en assurer l’archivage et à en garantir ensuite les conditions de consultation... mais nous n’en sommes pas encore là!

Les technologies numériques et les progrès réalisés par les industries de la langue devraient pouvoir nous rendre deux séries de services pratiques. En premier lieu, nous connaissons tous le problème des ouvrages épuisés, toujours utiles et difficiles à trouver en bibliothèque, et qu’on ne peut pas rééditer, par manque d’un nombre suffisant d’acheteurs potentiels dans un laps de temps compatible avec le coût de stockage. Dans le respect des droits d’auteur et de copie, on pourrait imaginer de numériser de tels ouvrages et de rendre ces fichiers accessibles sur le réseau, gratuitement ou non, pour une lecture à l’écran mais aussi une impression à la demande.

En second lieu, les progrès des industries de la langue pourraient nous aider à mieux maîtriser les flux sans cesse croissants d’informations exprimées dans un nombre toujours croissant de langues différentes. En effet, un des points forts de notre discipline, qui constitue aussi une de ses difficultés, réside dans le fait que chacun publie dans sa langue et que les langues à usage scientifique sont de plus en plus nombreuses, au fur et à mesure que des pays riches en matériel archéologique étudient eux-mêmes leur patrimoine. Ainsi, par exemple, pour la compréhension de la civilisation grecque, il n’est plus possible d’ignorer les nombreuses publications éditées récemment sur la Macédoine; or, malgré les résumés en anglais, ces publications sont difficiles à comprendre, pour ceux qui ne connaissent pas le grec moderne, car ce ne sont pas de simples catalogues illustrés mais bien des études assorties de raisonnements interprétatifs, riches mais aussi plus difficiles à comprendre complètement. Et, à l’inverse, nos publications non-anglophones sont de moins en moins utilisées par la communauté anglo-saxonne. Les technologies mises au point par les industries de la langue, en plein développement dans le cadre de l’Union européenne ou dans celui de la veille technologi-

que, permettent non pas d'obtenir des traductions automatiques (dont les résultats restent encore très éloignés de nos exigences) mais des aides efficaces pour le filtrage des notions recherchées, quelle que soit la langue utilisée. Il faudrait coupler ces outils déjà existants et en cours d'amélioration à des dictionnaires conceptuels multilingues comprenant le vocabulaire et les expressions propres à l'archéologie, et il faudrait ensuite les tenir à jour.

Mais, quelles que soient nos précautions d'analyse, les technologies vont continuer à progresser et leurs usages à moyen terme sont impossibles à imaginer avec vraisemblance. Toute prévision est donc vouée à l'échec, nous le savons désormais. Cela ne nous dispense pas de rester vigilants et d'assurer la «veille» nécessaire pour un usage optimisé. Et il faut aussi garder à l'esprit que l'utilisation d'un outil n'est pas neutre et que les méthodes employées sont à la fois la cause et l'effet de nouvelles problématiques de recherche: pour une part, l'évolution de notre discipline est aussi fonction de l'évolution générale des «nouvelles technologies de l'information».

ANNE-MARIE GUIMIER-SORBETS
UMR «Archéologies et Sciences de l'Antiquité»
Maison René Ginouvès
CNRS-Université de Paris I-Université de Paris X

BIBLIOGRAPHIE

- AA.VV. 1995a, *Musée de l'Homme, Anthropologie, Préhistoire, Ethnologie*, Muséum d'Histoire naturelle, ODA Laser Edition. CD-ROM.
- AA.VV. 1995b, *Le Trésor du San Diego, à la recherche du galion perdu*, Carré Multimédia, Club d'Investissement Média, Millemédias. CD-ROM.
- AA.VV. 1995c, *L'Art du Moyen Age, Occident, Byzance, Islam*, RMN, Carré Multimédia, Gallimard. CD-ROM.
- ARCELIN P. 1997, *La Publication archéologique sur CD-ROM, exemples pratiques d'écriture électronique*, Paris, Ministère de la Culture.
- ARCELIN P., RICHET C. 1990, *L'édition scientifique en archéologie métropolitaine: un état de la question*, «Les Nouvelles de l'Archéologie», 41, 10-13.
- BÉGUIN D. 1996, *Les antiquisants face à l'informatique et aux réseaux*, Novembre (<http://www.sciences-sociales.ens.fr/atelier/articles/ArticleInternetnov96.html>).
- BOMMELAER J.-F. (sous la dir.) 1996, *Marmaria, le sanctuaire d'Athéna à Delphes*, Athènes, 108-109.
- COPPENS Y. 1996, *Aux origines de l'Homme*, Microfolies. CD-ROM.
- DJINDJIAN F. 1991, *Méthodes en Archéologie*, Paris, Armand Colin.
- DUCOURTIEUX C. 1998, *MENESTREL ou Médiévistes sur l'interNET, Sources, Travaux, Références En Ligne*, «Le Médiéviste et l'ordinateur», 37, Hiver, 4-7.
- GUIMIER-SORBETS A.-M. 1993, *Des textes aux images. Accès aux informations multimédias par le langage naturel*, «Documentaliste, Sciences de l'Information», 30, 127-134.
- GUIMIER-SORBETS A.-M. 1996, *Le traitement de l'information en Archéologie: archivage, publication et diffusion*, «Archeologia e Calcolatori», 7, 985-995.

- GUIMIER-SORBETS A.-M. 1998a, *L'édition électronique dans le domaine de l'art: production, publics, usages*, «Revue du Musée des Arts et Métiers», 24, 15-22.
- GUIMIER-SORBETS A.-M. 1998b, *Fonctionnalités, modes d'accès, interfaces: pour une meilleure appropriation par l'utilisateur du contenu des cédéroms culturels*, in *Bibliothèque publique d'Information, Centre Georges Pompidou: Rencontres médias 2 (1997-98)*, Paris, J. Clément éd., 143-157.
- GUIMIER-SORBETS A.-M., SEIF EL-DIN M. (sous presse), *Les peintures de la nécropole de Kom el Chougafa à Alexandrie: éléments de méthode pour la lecture iconographique et l'interprétation du style «bilingue»*, in *Actes du VIIe Colloque international pour la Peinture murale antique (Saint-Romain-en-Gal, 6-10 Octobre 1998)*.
- HÉNON P.-M., LÉVEILLÉ J. 1997, *s.v. Images de synthèse*, in *Dictionnaire encyclopédique de l'information et de la documentation*, Paris, 282-283.
- JOCKEY P. 1999, *L'Archéologie*, Paris, 276.
- MOSCATI P. (ed.) 1998, *Methodological Trends and Future perspectives in the Applications of GIS in Archaeology*, «Archeologia e Calcolatori», 9.
- ROSS S., HIGGS E. 1993, *Electronic Information Resources and Historians: European Perspectives*, London.
- THOUVENOT S. 1999, *Sur le chemin d'une dame* (diaporama animé interactif présenté sur cédérom au festival Archéo Virtua).
- ZAÏD N. 1999, *La Publication archéologique sur Internet*, «Archéologia (Dijon)», 352, janvier, 20-29.

ABSTRACT

After giving a summary of the development of computer applications in archaeology from the 1960s up to the present, the Author dwells upon a diagram outlining the main processes which characterise the work of the archaeologist and specifically those in which computer methods have assumed an important role. The Author identifies the following essential stages for the use of computers in archaeology: data gathering, data structuring, also for purposes of interpretation, and data diffusion, in order to make results known. This last stage seems to be particularly influenced by the developments brought about in recent years by the introduction of multimedia systems which have made the electronic publication of archaeological data possible by employing digital supports like CD-ROM as well as by creating Internet sites.