

Une méthode rapide de constitution de documentation numérique

Jean-Baptiste Clais *



Statuette funéraire, Chine, dynastie Han : un bon exemple d'objet archéologique dont la datation repose sur des comparaisons stylistiques requérant un outil documentaire efficace. © RMN

En présentant le bilan de l'expérimentation – menée à partir d'exemples pris dans les arts moghol et chinois et les armes islamiques – d'une méthode rapide de constitution de dossiers documentaires de comparaison d'œuvres, l'auteur analyse cette pratique, ses principes, son fonctionnement, ses limites en regard des outils numériques existants et les perspectives qu'elle ouvre en vue de la constitution d'une documentation numérique mutualisée nationale accessible à tous les chercheurs.

Le projet du Louvre au musée Guimet

Le musée Guimet procède actuellement à une réorganisation de sa documentation. Dans le cadre de ce chantier, nous avons mis à profit une méthode productive de constitution de dossiers de comparaison d'œuvres utilisant les technologies numériques de manière à la fois simple et très peu coûteuse pour créer en quelques mois une documentation de comparaison sur l'art chinois (35 000 fiches à ce jour).

La démarche est reproductible à n'importe quel sujet d'histoire de l'art et nous semble pouvoir être particulièrement utile à nos collègues européens.

Notre méthode a donné naissance à ce jour à trois documentations. Les deux premières expérimentations réalisées en un an dans notre précédent poste concernent l'art indien d'époque moghole (9 300 fiches à l'époque, 17 000 aujourd'hui) et les armes islamiques (2 000 fiches à l'époque 2 600 aujourd'hui). La troisième, en cours de classement comprend 35 000 fiches d'objets d'art chinois classées dans plus de 2 000 dossiers et

* Jean-Baptiste Clais est conservateur au musée Guimet
jean-baptiste.clais@guimet.fr

sous dossiers. Elle a été réalisée dans le cadre de nos actuelles fonctions au sein de la section Chine du musée Guimet, en 7 mois avec une équipe constituée de quatre stagiaires à temps plein et 60 % de notre temps de travail personnel sur l'ensemble de la période. Le classement, encore en cours est mené au jour le jour par nos soins. Nous avons nommé cet outil « pARTage ».

État des méthodes existantes

Il existe actuellement deux méthodes pour la constitution de documentations de comparaison en histoire de l'art. La première est basée sur la photocopie d'ouvrage et leur classement dans des dossiers papier. L'exemple le plus abouti est la documentation du département des peintures du Louvre, alimentée depuis une cinquantaine d'années. Les fiches sont produites par des stagiaires qui photocopient ou découpent systématiquement les illustrations avec leur notice dans les publications qui parviennent au service de documentation, catalogues raisonnés, catalogues de vente, articles, créant des fiches qui sont ensuite réparties dans des dossiers par peintre ou par mouvement artistique sous la supervision du personnel de documentation, le tout piloté par un conservateur spécialement chargé de cette mission.

Cela a donné naissance à un outil formidable. Du fait de son ancienneté cette documentation a atteint un degré de pertinence considérable. Il est fort peu d'œuvres intéressantes publiées qui échappent encore à ses dossiers. De fait la seule limite d'une documentation de ce type est la bibliothèque qui l'alimente, selon qu'elle a réussi à capter une part plus ou moins importante des publications touchant à son domaine. L'intérêt spécifique de cette méthode est son caractère systématique qui fait que s'y intègrent des illustrations trouvées dans des publications que personne n'aurait le temps de dépouiller dans le cadre d'une recherche de corpus, des œuvres publiées une fois dans le guide des collections d'un château peu connu, passées une seule fois en vente il y a 60 ans... Les corpus s'enrichissent ainsi en permanence et s'accumulent dans des dossiers qui attendent qu'un chercheur ou un conservateur se penche dessus. Cette documentation a néanmoins un inconvénient matériel, son encombrement qui la rend intransportable. Il faut se rendre à Paris pour aller la consulter au Louvre. L'indexation se fait en outre sur un critère unique (pays, peintre...).

C'est pour pallier cette limite qu'a été engagé le processus de constitution de bases de données de comparaison informatiques⁽¹⁾. Mises en ligne, elles sont accessibles à tous. Alimentées correctement elles peuvent devenir un excellent outil comme la documentation



Figurine d'immortel, ivoire, Chine, dynastie Ming : fiche extraite de la base « pARTage », réalisée par traitement du catalogue d'exposition, *30 siècles d'ivoire du musée des arts asiatiques Guimet*, P. Ickowicz (dir.).
© RMN

papier précitée. Elles permettent d'indexer finement les œuvres et donc de faire des recherches multicritères. Ces bases ont toutefois une limite : le coût humain et financier de leur réalisation qui rend très hypothétique le développement massif de ces outils à moyen terme. Ces outils requièrent en effet des opérateurs compétents longuement formés, ils ne peuvent être alimentés massivement par des stagiaires. Le contenu étant éditorialisé, car publié en ligne, les informations doivent être vérifiées auprès des institutions détentrices. Il faut en outre gérer la question de l'autorisation de publication en ligne des photographies. Il faut concevoir les *thesaurii*. Remplir correctement une fiche dans un tel outil demande *a minima* dix bonnes minutes sans parler des tâches précitées. À cela s'ajoute la question de la conception et de la maintenance informatique. Une base de données est programmée avec un langage informatique donné, pour un environnement donné (Windows, MacOS...). L'histoire de la micro-informatique montre que tant les langages de programmation que les environnements évoluent régulièrement. Si les entreprises privées sont aujourd'hui dotées

de services informatiques importants et intègrent dans leurs budgets des programmes réguliers de mise à niveau ou de renouvellement régulier des outils informatiques, force est de constater que dans de nombreux musées particulièrement en France, les moyens ne sont pas comparables. Ainsi, nombre de musées nationaux français continuent d'utiliser pour leur gestion de collection une base de données obsolète dont la conception remonte aux années 1980. On conserve en l'état l'outil faute de moyens pour le faire migrer dans un support moderne. La durée de vie des outils informatiques se moule sur le rythme des musées et non sur celui de l'obsolescence technologique. Il y a fort à parier que cette situation n'évolue pas dans les années à venir. Pour y remédier il faudrait doter les services informatiques des musées de ressources financières beaucoup plus importantes, ce que l'actuelle politique de réduction des dépenses publiques rend hautement improbable.

Pour résumer, ces deux types d'outils représentent deux options. D'un côté une grande simplicité de conception et d'alimentation et l'ampleur des données collectées, contrebalancée par une faible indexation qui implique un temps de consultation important et les limites pratiques de la consultation. De l'autre la pertinence du résultat obtenu dans la recherche, une large diffusion via la mise en ligne qui implique un coût de réalisation important, la qualité des informations du fait de l'éditionnalisation du contenu, qui implique par définition, la prise en charge d'un sujet d'une taille limitée, et enfin un risque d'obsolescence technologique.

Notre méthode

Si l'on voulait réunir les qualités des deux outils, il faudrait un système fonctionnant sur informatique (ce qui le rend duplicable et consultable sur plusieurs sites) et qui puisse être alimenté par des opérateurs peu formés (stagiaires, *volunteers*), pour un coût matériel très limité, qui ne nécessite pas d'investissement informatique majeur et qui ne présente pas de risque d'obsolescence technologique.

Nous avons cherché à développer un système répondant à ces pré-requis dans notre précédent poste de conservateur au département des arts de l'Islam du Louvre où nous avons été affectés à l'art moghol et aux armes islamiques. La documentation de comparaison du département dans ces deux domaines étant peu fournie, il nous était nécessaire de créer un outil opérationnel le plus rapidement possible. La question de la productivité a donc été au centre de notre réflexion. Nous avons choisi de reproduire le principe de la documentation des peintures du Louvre en version numérique en réalisant des scans d'ouvrages ou de captures d'écrans de sites Web (base de données en ligne des musées étrangers, sites d'antiquaires...). Le procédé



Épingle à cheveux, ivoire, Chine, dynastie Shang : fiche extraite de la base « pARTage », réalisée par traitement d'une capture d'écran de la banque d'images en ligne de la MRNGP.

© Musée Guimet/J.-B. Clais

consiste à retoucher les fichiers produits de manière à créer *in fine*, un document numérique analogue à celui produit dans une documentation papier, une fiche avec la photographie d'un objet, son texte et un élément d'identification de la source. Le logiciel avec lequel nous travaillons pour la retouche d'image, Photo-Filtre, est gratuit.

Cette approche n'a rien d'innovante, au contraire, elle est rustique, à la portée de tous. Tous les historiens d'art classent dans leur disque dur des scans d'objets, des photographies... La spécificité de notre travail est d'avoir pu optimiser le mode de production de manière à obtenir un rendement considérable dans la production. Une recherche dans une documentation de ce type implique de visionner un très grand nombre d'images à la chaîne. Un soin particulier est donc apporté à la disposition des éléments dans la fiche pour limiter la fatigue visuelle lors d'une consultation. Autant que faire se peut, image et texte sont donc répartis en deux blocs clairement séparés, les éléments qui pourraient parasiter visuellement la fiche sont éliminés.

Nous procédons de même avec les catalogues d'exposition, de vente et les différents ouvrages que nous scanons. Le but est de réunir une image et son texte sur une seule fiche dans laquelle l'image est la plus grande possible et le texte lisible.

Maximisation du rendement de la production de fiches

Nous avons appliqué une démarche de division du travail en séquence. Le point essentiel est de travailler en grande masse de manière à développer des gestes réflexes optimisés pour chacune des phases du travail. Une méthode a été créée pour les captures d'écran, une autre pour les scans. Ces deux méthodes sont des bases que

l'on adapte ou modifie en fonction des spécificités de la source traitée (mise en page du livre, organisation de la page Web), et des outils à disposition du documentaliste (photocopieur-scanner, scanner à plat...)⁽²⁾. Deux principes doivent guider le documentaliste lorsqu'il adapte la méthode à une source nouvelle.

Le premier principe est de réduire le nombre de gestes. Le moindre geste en plus grève la productivité de quelques secondes par fiches, ce qui, au demeurant peut sembler dérisoire, mais en grande masse s'avère très important. Par exemple, les images produites par le copieur du département des arts de l'Islam ne pouvaient être retouchées directement dans PhotoFiltre, du fait du format de fichier envoyé par le copieur, les barres d'outils de PhotoFiltre n'étaient pas activées lorsque l'image s'ouvrait. Il fallait cliquer sur le bouton « couleur indexées » dans le logiciel pour les activer. Cela prenait deux à trois secondes par fiche, mais à l'échelle de 20 000 scans réalisés⁽³⁾, cela correspond à 11 h de travail... La productivité de notre méthode tient à l'élimination progressive d'un très grand nombre de micro-actions parasites comme celle-ci. Le plus gros gain de temps a été obtenu en supprimant toute saisie textuelle. Le processus se limite à des copier-coller d'images, donc des clics de souris et des raccourcis clavier.

Le second principe consiste à optimiser physiquement ses gestes, trouver la bonne posture de travail, celle qui nous demande le moins de mouvement importuns ou pénibles lors de l'utilisation du matériel. Il faut aussi tester les interactions avec le logiciel, pour voir ce qui est le plus efficace : clic droit + choix dans un menu ou raccourci clavier (CTRL + une touche généralement plus rapide). Optimiser les gestes lors du scan des pages permet aussi de produire plus vite des scans

bien droits, ce qui évite dans les phases ultérieures le travail de redressement des images particulièrement chronophage. Nous n'avons pas la place ici de détailler toutes les petites optimisations qui permettent à une échelle massive d'économiser du temps. Chacun devra tâtonner. Évidemment plus on a été confronté à des sources de formats divers, plus notre maîtrise du logiciel augmente, plus il devient facile de trouver des micro-optimisations.

Éléments d'évaluation de la productivité

La variation du temps de traitement nécessaire d'une source à l'autre rend difficile l'établissement d'un temps idéal de traitement. Nous pouvons néanmoins donner les chiffres moyens de notre productivité lors de nos différentes expériences.

En ce qui concerne la capture de site Web, un opérateur efficace produit selon la configuration du site Web entre 120 et 200 fiches par jour⁽⁴⁾. Pour les scans, la productivité dépend très fortement du matériel de scan. Lors de notre premier essai, au département des arts de l'Islam du Louvre, en un an en travaillant seul 80 % de notre temps sur ce projet, 11 300 fiches d'objets ont été terminées et classées, 6 800 scans restant en attente de traitement⁽⁵⁾. Il faut préciser que nous travaillions alors avec un matériel particulièrement peu adapté. Lorsque nous avons entrepris le même travail au musée Guimet pour la documentation de la section Chine que nous venions d'intégrer, nous avons pu bénéficier d'un matériel beaucoup plus adapté commercialisé récemment⁽⁶⁾ et de notre expérience d'utilisation de PhotoFiltre. Ainsi lors d'un test, travaillant seul sur des scans de catalogues de ventes récents, à raison de six heures par jour pendant trois semaines nous avons pu produire 2 500 fiches.



Capture d'écran d'une fiche d'objet présentée sur le site Web du musée Guimet, dans son état natif.
© Musée Guimet/J.-B. Clais



Capture d'écran d'une fiche d'objet présentée sur le site Web du musée Guimet après traitement.
© Musée Guimet/J.-B. Clais

Ainsi, qu'il s'agisse de produire des scans ou des captures d'écran de sites Web, le rendement moyen d'un opérateur bien équipé et formé qui travaille 80 % de son temps sur ce projet, tourne autour de 130 fiches par jour, soit plus ou moins 30 000 fiches par an.

Le facteur humain

Qualifier le travail de scans de tâche ingrate est un euphémisme. Plusieurs conditions nous semblent toutefois pouvoir rendre un tel travail supportable :

- l'opérateur doit avoir un besoin ou une envie d'apprendre sur le sujet qu'il fiche, ce qui donne du sens à son travail. Il nous semble donc impératif de réserver ce travail à des étudiants ou des professionnels de documentation ayant une formation ou une passion dans le domaine traité ;
- ce travail ne doit pas constituer l'unique activité d'une personne. Il nous semble que personne ne devrait consacrer plus de 80 % de son temps à cette activité ;
- il ne faut surtout pas segmenter l'activité. Chaque participant doit pouvoir traiter une source de A à Z, autant que possible. Il ne faut surtout pas cantonner une personne au scan et une autre au traitement d'image par exemple ;
- il faut, autant que possible, éviter que ce travail soit solitaire. Le travail en équipe en *open space* améliore considérablement les conditions de travail. En effet le scan ou la retouche d'image ne requérant que peu de concentration intellectuelle, il devient possible aux opérateurs de discuter.

Nous avons appliqué ces principes au musée Guimet. Les stagiaires qui proposent de collaborer au projet se voient libérés pendant 20 % de leur temps pour les projets professionnels de leur choix avec le soutien de leur maître de stage. Nous les invitons à découvrir les différents métiers du musée et à préparer le terrain

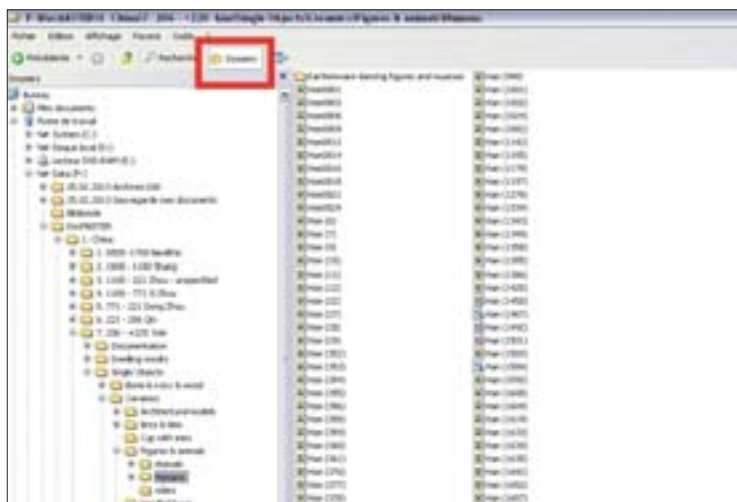
pour d'autres stages où ils apprendront autre chose. Le reste du temps, ils travaillent en équipe de deux ce qui leur permet aussi de s'entraider. Il en résulte une amélioration significative tant du bien-être de chacun d'entre eux que de la productivité.

Classement et consultation de la documentation

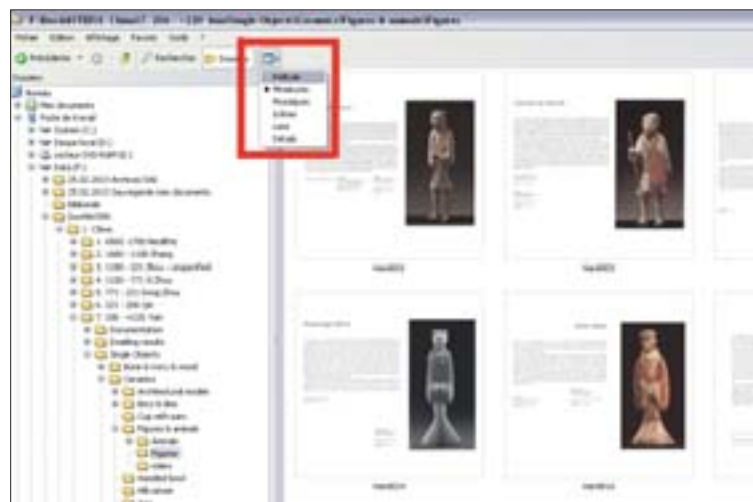
La méthode de consultation des fiches est on ne peut plus rustique, elle utilise l'explorateur Windows. Les fiches sont rangées dans des dossiers et sous-dossiers, organisés par dynasties/périodes, matériaux, sous-groupes. On navigue dans l'arborescence des dossiers via le volet latéral gauche « dossier ». Une fois le dossier choisi, on utilise les outils d'aperçu Windows pour observer le contenu. L'option affichage « vignette » dans le menu du bouton « affichage » d'abord. Si l'on souhaite voir une fiche avec le texte lisible, il faut utiliser l'option affichage « pellicule » dans le menu du bouton « affichage » ou faire un clic droit avec la souris sur le fichier qui nous intéresse et cliquer sur « aperçu » dans le menu déroulant.

Nous utilisons ici les outils de base de Windows XP, mais il est possible d'utiliser n'importe quel programme de type « visionneuse » et les outils de Windows 8 s'avèrent encore plus efficaces, notamment grâce à l'affichage en « très grande vignette ».

Comme une documentation papier, ce système présente, pour l'instant, la caractéristique de ne proposer qu'une indexation unique. De la sorte, la recherche doit se faire systématiquement, dossier après dossier, ce qui prend du temps. Cela reste néanmoins tout à fait fonctionnel. Nous avons récemment effectué pour un collègue indien une recherche iconographique sur les objets figurant des éléphants dans notre base d'objets d'Inde moghole. Il ne nous a fallu qu'une petite



Capture d'écran pointant le positionnement du bouton « dossier » dans l'explorateur de fichier Windows.
© Musée Guimet/J.-B. Clais



Capture d'écran pointant le positionnement de l'option affichage « vignette » dans le menu du bouton « affichage » de l'explorateur de fichier Windows.
© Musée Guimet/J.-B. Clais

matinée pour parcourir l'intégralité des 9 300 fiches disponibles à l'époque.

Les fiches peuvent être très simplement copiées ou déplacées d'un dossier à un autre en quelques clics. Il est donc possible, nous l'avons expérimenté avec les moghols, de créer des duplicata de base organisés selon des critères différents (matériau pour l'une, typologie fonctionnelle d'objet pour l'autre). Néanmoins, à terme il serait souhaitable de trouver des modes de classement consensuels au niveau international en mettant à contribution des spécialistes pour la création des arborescences. C'est en effet à eux en dernier ressort que s'adresse cet outil.

Ajoutons un dernier point. Nous nous abstenons actuellement d'ajouter des métadonnées/tags à ces fichiers pour des raisons de budget. Nous avons voulu réaliser cette documentation à coût zéro en matière de logiciel. Si nous recevions des fonds pour acheter un logiciel avancé de reconnaissance de caractères, celui-ci pourrait scanner un à un les fichiers et ajouter en métadonnée tous les mots qui y apparaissent. Cela permettrait alors de faire des recherches par mot clé dans cette documentation. Nous aurions ainsi les avantages d'une base de données classique, mais aussi un outil doté d'un contenu très important du fait de la productivité et de la rusticité de notre méthode.

Enfin, comme nous dépouillons toutes les publications dans un domaine, nous voyons réapparaître les objets, passés en vente, publiés dans un catalogue d'exposition... Lorsque cela se produit, nous créons un sous-dossier pour les fiches correspondant à cet objet que nous nommons du numéro d'inventaire de l'objet. Ainsi se créent naturellement, sans travail spécifique, la partie communicable de ce que pourrait être un dossier d'œuvre. De la sorte, il devient possible avec cette documentation de régler le problème de l'accès aux dossiers

d'œuvres. Tout ce qui relève du « scientifique » dans un dossier d'œuvre, au terme de la création d'une documentation de ce type, se trouve accessible dans les dossiers numériques et tout ce qui doit rester confidentiel (documents relatifs à l'achat, au prix, au vendeur) reste dans le dossier papier. Ce suivi biographique des objets permet parfois de repérer des restaurations discutables opérées entre deux publications ainsi que des objets volés ayant été fichés préalablement ⁽⁷⁾.

Politique d'alimentation

L'intérêt d'une documentation comme celle-ci dépend des sources choisies pour son alimentation. Notre objectif est de parvenir pour l'art chinois à un degré de pertinence analogue à celui de la documentation des peintures du Louvre pour la peinture européenne.

Néanmoins atteindre ce niveau de pertinence requiert des années. C'est pourquoi nous avons choisi de commencer par alimenter notre outil avec les collections des grands musées, en dépouillant leurs bases de données. Nous avons déjà extrait les données des bases de données en ligne du LACMA, du British Museum, du Walter Art Museum, du Minneapolis Institute of Arts, du Metropolitan Museum of Arts, de l'Asian art Museum de San Francisco. Nous allons entreprendre de scanner de manière systématique les publications de collection en plusieurs volumes de la Cité Interdite et du musée du Palais de Taipei. De la sorte, nous aurons un état relativement complet des objets chinois importants publiés par les plus grandes institutions d'art asiatique dans le monde.

L'étape suivante sera de scanner l'intégralité des ouvrages contenant des objets d'art chinois présents dans la bibliothèque du musée Guimet. Celle-ci présente l'avantage d'être organisée dans ses magasins par date d'entrée dans les collections. Nous pourrions ainsi commencer à scanner les ouvrages les plus récents donc



Capture d'écran montrant le mode d'affichage « vignette » de l'explorateur de fichier Windows.

© Musée Guimet/J.-B. Clais



Capture d'écran montrant le mode d'affichage « pellicule » de l'explorateur de fichier Windows.

© Musée Guimet/J.-B. Clais

produire des fiches à jour en termes de bibliographie. Ensuite viendront progressivement s'intégrer dans les dossiers des objets issus de publications anciennes moins connues, des rapports de fouilles archéologiques, le dépouillement du marché de l'art... Chaque ouvrage scanné est listé dans un tableau de type Excel. Le temps nécessaire au traitement de la bibliothèque complète de Guimet est difficile à évaluer. Selon les moyens alloués et les progrès techniques, on peut supposer que ce travail prendra entre dix et vingt ans. Mais ce temps de traitement tient à ce que peu de musées travaillent sur l'art chinois. Dans le domaine européen, même si la bibliographie est considérable, le nombre de musées existant permettrait de répartir la tâche.

Durabilité

La solution que nous utilisons – des fichiers images classés dans des dossiers – est la plus durable qu'il soit possible d'imaginer à ce jour d'un point de vue technique. Elle ne peut connaître l'obsolescence technologique parce qu'elle ne repose pas sur un programme mais sur l'un des seuls principes à n'avoir jamais évolué depuis l'apparition de la micro-informatique en 1979⁽⁸⁾. La migration des données d'un ordinateur ancien à un ordinateur récent se fait par un simple « copier/coller ». Le format de fichier choisi (.tif) est un format pérenne utilisé dans l'archivage photographique. Reste la question du stockage d'une quantité croissante de données. Cela ne pose pas non plus de problème. Toutes nos documentations réunies tiennent sur une clé USB. Elles ne représentent que 163 Go de données pour 59 000 fichiers (scans non traités compris) soit 2,9 Mo par fichier en moyenne. En estimant que les scans seront à termes plus nombreux que les captures de sites Web – plus légères, le poids moyen des fichiers devrait donc être plus important, de l'ordre de 3 à 3,2 Mo. Les disques durs portables aujourd'hui accueillent déjà 2 Tera de données. Il faudrait donc environ 740 000 de nos fiches pour en remplir un. Au vu du rythme d'accroissement de nos données et de celui de l'augmentation des capacités de stockage des disques durs portables, on peut estimer que notre documentation tiendra toujours à l'intérieur d'un disque dur portable ou d'une clé USB.

Par ailleurs, nous réalisons régulièrement des copies de notre documentation sur des supports de données mobiles pour multiplier les copies hors-site, dans l'hypothèse d'un problème majeur qui détruit l'intégralité de notre système informatique, sauvegardes comprises.

Perspectives

Si nous pouvons copier cet outil sur une simple clé USB, il nous semble évident qu'il a vocation à être mis à disposition des professionnels. Comme tous les outils de documentation des musées, il ne peut être toutefois mis en ligne sur Internet sous peine de violer les

lois sur le copyright. Au regard de la nouvelle formulation de l'exception pour les bibliothèques, archives et musées, il semble tout à fait légal d'en prêter des copies à des documentations de musées ou centres de recherche partenaires, si la consultation se fait sur place et qu'aucune exploitation commerciale de l'outil n'est faite⁽⁹⁾.

Mais au-delà de ce processus, il nous semble que des objectifs plus ambitieux pourraient voir le jour. Il serait intéressant de lancer une collaboration nationale et internationale visant à répartir le travail de scan au sein d'un réseau d'institutions partenaires. Le partage en ligne du tableau listant les ouvrages scannés permettrait à chaque musée d'apporter une contribution non-redondante. On peut imaginer par exemple que pour le suivi du marché de l'art, les musées américains scannent les catalogues de vente américains, les chinois les catalogues chinois... Chaque participant transmettrait sa production au fur et à mesure à un musée centralisateur dans un domaine précis. En France, chaque grand département scientifique pourrait devenir le pilote dans son domaine de spécialité. Chaque année à date fixe, une réunion des musées contributeurs dans chaque thématique aurait lieu. Un pré-classement des fiches préalablement transmises par les contributeurs aurait été réalisé par le musée pilote. Les spécialistes présents se chargeraient de contrôler les dossiers correspondant à leur domaine de spécialité pour proposer des précisions, de nouvelles hiérarchies de dossiers ou des réattributions de fiches. PhotoFiltre permet en effet de rajouter du texte sur les fiches. Il est ainsi possible d'ajouter par exemple dans une police colorée spécifique des annotations telles que : « 12 décembre 2012, monsieur X, institution X, signale que l'objet est Jin et non pas Yuan comme précédemment indiqué ». Au terme de la réunion, chaque



Épingle à cheveux, ivoire, Chine, dynastie Shang : fiche extraite de la base « pARTage », réalisée par traitement d'une capture d'écran de la banque d'images en ligne de la MRNGP.

© Musée Guimet/J.-B. Clais

participant repartirait avec une copie à jour de la documentation pour son institution. Le musée pilote se chargerait d'envoyer des copies aux institutions qui n'auraient pas pu se faire représenter.

En ce qui concerne notre domaine particulier, l'art chinois, une des perspectives serait à terme de marquer ou d'isoler une copie des fiches des objets datés avec certitude, pour fournir un matériau le plus exhaustif pour la vérification de la datation des corpus. À mesure que progresserait le dépouillement des sources peu explorées (particulièrement les micro-collections et cabinets de curiosités des innombrables demeures historiques en Europe), des objets peu connus ayant un historique intéressant viendraient s'ajouter aux corpus. Le volume des pièces intéressantes connues et empruntables augmenterait ce qui soulagerait les collections des grands musées trop souvent sollicitées pour des prêts. Pour les étudiants de master, cette base proposerait des corpus prêts à l'emploi. Ils perdraient donc moins de temps à constituer leur corpus, ce qui leur permettrait de consacrer plus de temps aux lectures d'articles et à la réflexion *stricto sensu*.

Nous avons en outre commencé à intégrer dans la base les fichiers pdf d'articles scientifiques correspondant aux dossiers et sous-dossiers créés ; les essais thématiques présents dans les ouvrages scannés sont aussi intégrés dans les dossiers. Nous aurons ainsi à long terme à disposition chaque corpus et les études qui lui ont été consacrées – bien que sur ce point, il faille admettre que le travail sera beaucoup plus long et sans doute moins prioritaire.

Conclusion

Le projet initié au Louvre et mûri au musée Guimet a montré qu'il est aujourd'hui possible de constituer très vite pour un coût très limité une documentation numérique de comparaison sur n'importe quel sujet de l'histoire de l'art. L'ensemble du procédé décrit ici peut être par exemple appliqué aux arts occidentaux.

Nous invitons nos collègues et tutelles à se saisir de la méthode proposée pour travailler de concert à la constitution d'une documentation numérique mutualisée nationale, outil collaboratif, partagé entre institutions et accessible dans nos centres de documentation à tous les chercheurs. Vu la productivité et le très faible coût de production, elle pourrait être mise en place par des institutions pilotes dans chaque domaine et croître rapidement. Une première étape pourrait être par exemple le scan systématique de tous les catalogues de ventes. Dans ce domaine, un ou deux documentalistes aidés de quatre stagiaires pourraient sans difficulté traiter en flux continu tous les catalogues de vente publiés dans une année et même entamer une

campagne de scans rétrospectifs⁽¹⁰⁾. Vu la productivité de la méthode, l'outil serait rapidement utile et surtout pourrait être copié et distribué à tous les musées de France qui en feraient la demande.

Notes

(1) Nous parlons ici de bases de données de comparaison, de bases de corpus et non de bases de gestion de collection qui sont un autre sujet. Le meilleur exemple à ce jour de base de données de corpus existant est la base La Fayette qui répertorie les peintures des écoles américaines dans les collections françaises.

(2) Pour plus de précisions sur ces 2 méthodes, contacter l'auteur.

(3) Il faut distinguer en matière de scans, le nombre de scans produits du nombre de fiches produites, dans un ouvrage, un objet et son texte peuvent occuper trois pages alors que dans un autre une page peut accueillir six objets et leur texte.

(4) Ajoutons au sujet des captures en ligne que nous déconseillons de faire des captures des sites de ventes. Les images sont en basse résolution, mais surtout l'indexation laisse tellement à désirer que l'on peut ne pas trouver la moitié des pièces passées dans une vente. En termes de qualité d'image et de captation systématique des informations, il est plus efficace de scanner les catalogues papier.

(5) Après traitement de ces scans nous estimons que la base « moghols » devrait croître jusqu'à 15 000 fiches et la base « armes islamiques » jusqu'à 2 500 fiches.

(6) Nous avons pu faire procéder à l'acquisition d'un scanner à angle qui permet de scanner chaque page à plat sans déformation sur l'image obtenue. Ce scanner envoie directement les images en format tif dans un dossier où elles sont prêtes à être traitées et déjà nommées.

(7) Voir par exemple, *Sotheby's Londres Arts of the islamic world*, 2006. sale L06220, lot 142 « two mughal calligraphic tiles » puis de Guise, L. (éditeur) *Islamic Arts Museum Malaysia*, Vol. II, (contributeurs : Adline Abdul Ghani, Allan T. Najim, Assim Zuhair Mahmood Qisho... et al.). 2^e édition revue et augmentée, Kuala Lumpur, 2009. « Pair of calligraphic tile, inv. 2008.1.142 ».

(8) Nous avons pu constater ce point dans le cadre d'une activité bénévole que nous menons dans un musée de l'informatique où la question de la pérennité des données et des programmes se pose.

(9) Sur ce point voir l'article L122-5, alinéa 8 du *Code de la propriété intellectuelle*.

(10) Un tel outil donnant un état de tous les objets publiés dans les ventes pourrait en outre faciliter la traque par l'OCBC des objets volés.