



**JOURNEE PROFESSIONNELLE « INFORMATISATION, NUMERISATION ET  
MISE EN LIGNE DES COLLECTIONS DES MUSEES DE FRANCE »  
Paris, 8 juin 2012**



Mise en ligne : septembre 2012

**Bases de données des musées : usages et contributions**

**Bernadette DUFRENE, université Paris 8, Laboratoire Index-Paragraphe**

Parmi les missions du musée, l'étude et la recherche sur les collections constituent une priorité du vaste domaine de la conservation. Depuis les années 1980-90, l'informatisation progressive de l'inventaire et la création simultanée de bases de données – ensemble de données relatif à un domaine de connaissance structuré en vue d'en permettre une exploration et une exploitation par des utilisateurs- ont transformé tant les pratiques professionnelles (celles des musées et celles de la recherche externe au musée) que les pratiques amateur en mettant à disposition des corpus de textes et d'images. Ce mouvement s'inscrit dans ce qu'on a pu appeler la « redocumentarisation du monde »<sup>1</sup> pour signifier ce que l'informatique liée à la diffusion sur des réseaux changeait aux conditions et aux finalités de la documentation : les métadonnées y acquièrent un rôle central tant du point de vue du repérage de l'information que du point de vue de l'influence qu'exerce désormais l'indexation sur la description des patrimoines. C'est un point de vue excentré, non pas celui des producteurs de corpus numériques mais celui de l'utilisateur à la fois comme chercheur et comme amateur, que l'on m'a demandé de représenter à cette table ronde. J'aurais recours aux termes de Pascal (qui lui-même se réclamait de saint Augustin) pour poser la question des usages des bases de données : comme la grâce de Dieu, les bases de données sont nécessaires, mais la question est de savoir si elles sont suffisantes et dans quelles conditions elles sont efficaces. Je vais tenir sur cette question les deux positions pour les confronter, la position du chercheur « jésuite » - les bases de données sont nécessaires et suffisantes - et la position du chercheur « janséniste » - les bases de données sont certes suffisantes mais encore faut-il qu'elles soient efficaces !

---

<sup>1</sup> Jean-Michel Salaün, « La redocumentarisation, un défi pour les sciences de l'information », « études de Communication n° 30, Entre information et communication, Les nouveaux espaces du document, Université de Lille 3, décembre 2007

## I - Des bases de données nécessaires et suffisantes

Le chercheur jésuite va dire que les bases de données sont nécessaires et suffisantes : dès qu'un « centre de calculs », pour reprendre l'expression de Bruno Latour, dresse avec ses outils, ses statistiques, ses normes descriptives, ses procédures et ses dispositifs la « carte » d'un territoire de l'information, alors il donne une configuration à ce « territoire » qui permet aux utilisateurs de se le représenter avec beaucoup de précision, de lui donner une identité, d'en rendre les données (au sens d'informations numérisées) commensurables. C'est d'un double point de vue, politique et scientifique, que les bases de données peuvent apparaître nécessaires et suffisantes.

### La dimension politique

Les cartes des patrimoines telles que nous les connaissons en France sont le résultat de deux opérations : la création de l'Inventaire général par Malraux en 1964 et son informatisation. Dans la perspective de Malraux, la création de l'Inventaire s'inscrit certes dans une continuité, la volonté de recenser les richesses artistiques de la France mais comporte aussi un objectif inédit : « un inventaire agissant comme un explorateur chargé de repérer les objets qui matérialisent les nouvelles valeurs dont chaque collectivité se dote pour exister. Par collectivité, nous n'entendons pas seulement les nouveaux territoires constitués : régions, parcs naturels, agglomérations, voisinages, mais les communautés d'intérêt que sont les disciplines, les métiers, les associations, les entreprises et les castes des différents niveaux du savoir »<sup>2</sup>. C'est alors un champ immense qui s'ouvre : la volonté d'exhaustivité explique la mise en place de taxinomies indissociable du recours à l'informatique.

L'informatisation sert le projet de l'Inventaire non seulement parce qu'elle mobilise le calcul numérique et qu'elle favorise le stockage de données mais aussi parce qu'elle va permettre le développement de la description dans le cadre de l'Inventaire comme l'a bien montré Jean-Marie Pérouse de Montclos : «La création de l'inventaire général fut le moment où l'on a fait un effort sans précédent pour faire rentrer les descriptions du patrimoine dans le champ des sciences Le recours à l'informatique a eu pour effet, paradoxal, de produire à l'origine, c'est-à-dire dès 1964, une amplification du genre descriptif. L'ordinateur n'acceptant que des mots, il fallait pour l'utiliser paraphraser toutes les images, exercice qui fut gracieusement appelé verbalisation.

Cependant, effet second, et celui-là de longue durée, la création de lexiques hiérarchisés qui ont fondé les bases de données actuellement en activité au ministère de la culture »<sup>3</sup>. C'est donc d'une réflexion à la fois scientifique - sur la nature des patrimoines et les terminologies propres à les représenter - et politique - sur la nécessité de dresser des cartographies du patrimoine - que sont issues les premières bases de données. Les bases de données de l'Inventaire permettent de dresser une carte

---

<sup>2</sup> Michel Melot, *Mirabilia Essai sur l'Inventaire général du patrimoine culturel*, Paris, Gallimard, 2012, p.26

<sup>3</sup> Jean Marie Pérouse de Montclos, « La description », in Pierre Nora, *Science et conscience du patrimoine*, Paris, 1994, p.196

des patrimoines, qui est aussi une construction symbolique des territoires locaux, régionaux ou nationaux : la rubrique « accès géographique » en témoigne.

La période qui va de la fin des années 1960 aux années 1990 constitue de ce point de vue une étape importante dans la normalisation des descriptions, l'établissement de classements et de typologies. L'intervention du législateur permet d'uniformiser les pratiques : en ce qui concerne l'inventaire des biens de musées, l'arrêté du 25 mai 2004 donne une définition très précise des rubriques : la première vise à définir le statut juridique des œuvres ou des biens appartenant aux collections des musées, la deuxième prévoit la description du bien, la troisième des rubriques complémentaires « quand l'information est pertinente pour le bien (auteur; collecteur, fabricant, commanditaire, propriétaire lorsqu'il s'agit de collections scientifiques et techniques; date ou époque; date de récolte, lorsqu'il s'agit de collections scientifiques et techniques; fonction d'usage; provenance géographique) ». A celles-ci peut s'ajouter une rubrique « Observations » réservée à des informations dignes d'intérêt mais non absolument nécessaires (première date de présence attestée dans le musée en cas d'origine inconnue, utilisateur illustre, premier et dernier propriétaire, anciens numéros d'inventaire, autres numéros d'inventaire, mentions en cas de vol ou de disparition).

#### **La redocumentarisation du patrimoine**

On voit comment les interactions entre l'informatique et le projet politique de l'Inventaire à la fois se confortent et orientent notre rapport au patrimoine : la série l'emporte sur l'unicum, la description normalisée sur la description approfondie de l'œuvre. L'adoption de la base de données comme forme symbolique privilégiée entraîne donc le risque de la « réduction typologique »<sup>4</sup>.

Cependant cette question est occultée durant la décennie où se développe le web de l'information (1993-2005) si l'on considère que les deux sommets mondiaux de Genève (2003) et de Tunis (2005) qui popularisent les notions de « société de l'information » voire de « société de la connaissance » (Unesco) en sont les points culminants. C'est l'époque où les grands musées des pays développés constituent leurs bases de données ; ils le font à partir de la documentation telle qu'elle est pratiquée dans le champ du patrimoine : établir des notices i.e. décrire les œuvres pour les cataloguer, les indexer, mais aussi établir des dossiers d'œuvres, les documenter pour en expliquer le contexte d'apparition voire la fortune critique ; ces opérations relèvent de ce que les sciences de l'information nomment « la documentarisation » dont « l'objectif est d'optimiser l'usage du document en permettant un meilleur accès à son contenu et une meilleure mise en contexte ».

La redocumentarisation liée au numérique a deux conséquences majeures : du point de vue de l'accès, elle donne un rôle de premier plan aux métadonnées qui vont permettre de repérer le document en en assurant l'indexation ; du point de vue du contenu, reposant sur l'hypertextualité, elle rend commensurables des données, fait converger des documents : la forme de document numérique que

---

<sup>4</sup> Jean-Michel Leniaud, « Typologie de l'inventaire et du classement », in Pierre Nora, op.cit., p. 321-330.

constitue la base de données non seulement structure un champ de connaissance particulier mais, reposant sur l'hypertextualité, la présence de liens navigationnels, mais encore permet d'établir de nouvelles configurations symboliques.

L'analyse de la base de référence Joconde 7 peut servir d'illustration : Joconde se présente comme le « catalogue collectif des collections des musées de France, accessible sur Internet au public le plus large ». La base qui recouvre environ 500 000 notices et s'enrichit constamment révèle les processus de fabrication de la mémoire par les bases : celle-ci ne tient pas seulement au traitement quantitatif mais au fait qu'il soit appliqué à des formes sémiotiques. De ce point de vue, la base de données n'est pas un instrument neutre : d'abord parce qu'elle opère une sélection mais aussi parce, tout en rendant intelligible un patrimoine, elle le qualifie : ainsi les champs « domaine » et « dénomination » servent le classement de l'objet et en même temps en spécifient le statut dans le champ du patrimoine, souvent en reproduisant les hiérarchies de l'histoire de l'art. Il est d'ailleurs intéressant de noter l'effort de réflexivité des administrateurs de la base en ce qui concerne à la fois la structuration et le vocabulaire : le blog de Joconde explicite la méthodologie et anticipe les questions ou les difficultés, par exemple en ce qui concerne la détermination des domaines et le classement des objets dans ceux-ci <sup>5</sup>. Il est aussi intéressant de constater la marge d'interprétation du documentaliste : selon les musées, un même objet pourra être classé dans le domaine « ethnologie » ou « arts décoratifs ». Le traitement de l'information repose donc non sur une « écriture de choses » mais sur une « écriture de mots », la désignation des choses et des idées par les mots : toute réflexion sur l'écriture numérique du patrimoine doit tenir compte de cet aspect qui permet à la fois l'alimentation et la consultation des bases de données et que matérialise, pour Joconde, la rubrique « vocabulaires scientifiques » et, pour l'Inventaire, le thésaurus de la désignation qui « développe sous une forme méthodique les 1135 termes utiles à la désignation des œuvres architecturales et les 2529 utiles à la désignation des œuvres mobilières ». C'est à partir de ces nœuds documentaires et des hyperliens qui leur sont associés que peut se déployer la connaissance : les parcours thématiques ou les expositions virtuelles qu'offre Joconde le montrent.

Ainsi conçue la base de données répond aux modalités du document qu'a analysées le collectif Pédaouque : comme forme de classement rendant visible et lisible une catégorie patrimoniale, « aux modalités anthropologiques », comme forme de description normalisée rendant intelligible une œuvre, « aux modalités intellectuelles » et comme forme communicationnelle relevant de logiques patrimoniales, « aux modalités sociales »<sup>6</sup>.

La forme de document numérique que constitue la base de données apparaît en ce sens comme un des principaux outils de la mémoire ; elle est particulièrement révélatrice du rôle que jouent les TIC dans le champ du patrimoine comme techniques d'inscription et comme techniques symboliques :

---

<sup>5</sup> <http://portail-joconde.over-blog.com/article-comment-utiliser-les-champs-domaine-et-destination-utilisation-108639104.html>

<sup>6</sup> Cf Jean-Michel Salaün, article cité supra

« En précisant des normes et des standards pour assembler des formes sémiotiques, les technologies de l'information mobilisent le calcul numérique non au profit d'une formalisation du raisonnement, mais au projet d'une transmission de la mémoire »<sup>7</sup>.

Cette mémoire n'est ni uniforme ni univoque : le jeu de la sélection versus de l'oubli rend les bases de données nécessaires et suffisantes. Il ne reste plus qu'au « centre de calcul » d'agir avec d'autres « centres de calcul », d'enrichir la granularité pour que les mailles du filet permettent une pêche de plus en plus miraculeuse, pour rester dans le registre religieux !

Pour autant sont-elles efficaces ?

## **II. Des bases de données efficaces ?**

Le chercheur janséniste va objecter que si la base de données est nécessaire, elle n'est pas à elle-seule suffisante. Pour qu'elle soit « efficace », c'était déjà le mot d'Augustin, il faut que l'usage n'en soit pas local ou limité à une communauté ; il faut que la structure des métadonnées soit une matrice suffisamment englobante et ouverte pour pouvoir accueillir non seulement le plus grand nombre de données mais encore une modification de l'ensemble de la configuration, par l'intégration des prélèvements venus directement du terrain de la recherche (qui peuvent modifier profondément les protocoles).

### **Web 2.0 et nouveaux usages ?**

Avec l'avènement du web 2.0 dont le fameux article de Tim O'Reilly a constitué le point de départ, les notions de participation et de contribution ont été largement vulgarisées ; en 2007, l'expression « économie de la contribution » est popularisée par le philosophe Bernard Stiegler qui voit dans le web 2.0 l'émergence d'un nouveau statut pour le public, non plus celui de consommateur mais celui de producteur ou de co-producteur de contenu. La sociologie des usages guide la réflexion que l'on peut avoir sur les bases de données et leurs utilisations dans ce nouveau contexte. Elle révèle un paysage fortement contrasté : la moitié des internautes est constituée selon les enquêtes de Forrester Research Inc. (2007, 2009, 2010) <sup>8</sup> d'utilisateurs n'ayant un usage que de consultation, tandis que l'autre moitié a des usages qui relèvent de la participation (20%) ou de la contribution (20%).

Le web contributif regroupe pour O.Ertzcheid les créateurs de contenus et ceux qui visent à initier des débats sur différents types de réseaux sociaux. Au sein de ces usages contributifs, les créateurs de

---

<sup>7</sup> Bruno Bachimont, « Du texte à l'hypotexte : les parcours de la mémoire documentaire », in Charles Lenay et Véronique Havelange (dir), *Mémoire de la technique et techniques de la mémoire*, Paris, 1999, p.195-225.

<sup>8</sup> Enquêtes citées par Olivier Ertzcheid, [http://affordance.typepad.com/mon\\_weblog/2010/10/accroche-toi-au-pinceau-enleve-echelle.html](http://affordance.typepad.com/mon_weblog/2010/10/accroche-toi-au-pinceau-enleve-echelle.html)

contenus sont 40%, et les conversationnalistes 60%. O.Ertzcheid fait remarquer à propos des adjectifs "participatif", "collaboratif", "contributif", qu'il s'agit de trois termes « souvent indistinctement et abusivement employés ». S'appuyant sur le trésor de la langue française, il propose les définitions suivantes : « la contribution peut être définie comme la "part apportée à une œuvre commune" : en l'occurrence, cette œuvre commune sera constituée par le web ; la collaboration est "la participation à l'élaboration d'une œuvre commune". La proximité sémantique des deux termes est évidente, même si dans le contexte du web 2.0 il est possible d'envisager des formes de collaboration non-nécessairement contributives. La collaboration relèverait alors davantage de l'engagement, et la contribution, de l'action... La participation est l'action de participer à quelque chose" en – deuxième sens – "manifestant une adhésion, une complicité, une conscience d'ordre intellectuel »<sup>9</sup>.

Comme les autres champs sociaux, le champ du patrimoine est traversé par ces nouvelles formes de coopération. D'après O.Ertzcheid, on peut considérer que le web participatif, ce sont les comportements en ligne se résumant à des opérations documentaires apparentées à l'indexation, au commentaire ou à l'amélioration (wikis par exemple) de ressources existantes ; quant au web contributif, il ne concerne que les comportements en ligne les plus qualifiés en termes de compétences ou de niveaux d'interaction et s'inscrivant dans une logique de production de contenus originaux<sup>10</sup>. C'est donc tout un faisceau d'usages allant de coopérations fortes (blogs, commentaires, collections) à des coopérations faibles (celle des joiners) qui se déploie sur le web et qui affecte les pratiques patrimoniales des professionnels et des amateurs.

#### **Les nouvelles matrices documentaires**

Le web participatif et collaboratif par les coopérations qu'il instaure bouscule la structuration des connaissances et les vocabulaires employés et donne lieu à de nouvelles configurations. Il faut donc - dirait notre chercheur janséniste - prévoir une « matrix » étonnamment évolutive et souple, *comprehensive* (au sens anglais du terme). Des coopérations naît en effet la nécessité de construire la veille et l'évolution du système. L'intégration des nouvelles données de la recherche sur l'art, l'élargissement des perspectives, l'introduction de la controverse caractérisent la documentation collaborative qui fait intervenir aussi bien des historiens d'art que des conservateurs, des restaurateurs ou des spécialistes de l'image. Le projet Cranach financé par la Mellon Fondation<sup>11</sup> en est représentatif : la documentation collaborative repose certes sur une infrastructure technique importante - elle a comme objectif « de développer de nouvelles sortes d'outils de recherche pour faciliter la transmission des informations relatives à l'histoire de l'art et à la conservation des œuvres »<sup>12</sup> - mais c'est la conception de l'interface qui oriente la perception du patrimoine. La structuration de la base de données est préservée : elle repose sur une notice classique, mais elle comporte trois autres rubriques : « art historical information

---

<sup>9</sup> Ibidem

<sup>10</sup> Ibidem

<sup>11</sup> [http://www.lucascranach.org/aboutus\\_en.html](http://www.lucascranach.org/aboutus_en.html) voir aussi l'ensemble des projets : <http://mac.mellon.org/issues-in-conservation-documentation/pilot-projects>

<sup>12</sup> Traduction du texte : « to facilitate transmission of art historical and conservation information across institutions and international borders in order to advance scholarship and learning », [http://www.lucascranach.org/aboutus\\_en.html](http://www.lucascranach.org/aboutus_en.html)

» « material/technique » « conservation/restoration », « images ». L'interface ouvre sur une pluralité d'approches et de lectures possibles de l'œuvre et condense différents usages professionnels : conservation et régie des œuvres, étude scientifique, critique. La masse des données et leur circulation, la pluralité des approches créent de nouvelles conditions pour le savoir, le diffractent en une multiplicité d'angles.

Un autre élément d'évolution lié au web 2.0, ce sont, à l'opposé des taxinomies et des vocabulaires contrôlés, les folksonomies, à savoir la possibilité donnée aux internautes de classer de manière ouverte, libre et spontanée des documents ou des informations ; dans le champ du patrimoine, il s'agit d'introduire d'autres descriptions que celles des institutions patrimoniales en amenant les internautes à tagger les œuvres. C'est l'expérience menée sous le nom de Steve Museum<sup>13</sup> : le projet se présente comme le projet « de professionnels des musées et de tous ceux qui pensent que le tagging social peut offrir d'authentiques nouveaux modes de description du patrimoine culturel, encourageant l'implication du visiteur à l'égard des collections ». Les résultats attendus de ce tagging social se situent sur des plans différents : culturel, social et patrimonial. Culturel parce qu'en taguant les œuvres, l'internaute entre dans un processus d'appropriation, social car cette indexation populaire élargit le public des amateurs, patrimonial car ce sont les axes de description d'une œuvre qui sont engagés et donc la reconnaissance de la légitimité d'une pluralité de lectures et d'intérêts. C'est donc une modification en profondeur du rapport au patrimoine que l'expérience engage dans le sens de ce qu'en France on considérerait comme une démocratisation de l'accès à l'art. L'expérience s'appuie sur une méthodologie rigoureuse qui vise aussi bien les professionnels que les amateurs prêts à se livrer au jeu : « Our activities include researching social tagging and museum collections ; developing open source software tools for tagging collections and managing tags ; and engaging in discussion and outreach with members of the community who are interested in implementing social tagging for their own collections... The goal of this ongoing project is to enhance the functionality of the steve software tools, to encourage the use of social tagging methods in museums of all sizes and types, to develop innovative, "next-generation" tagging interfaces, and to explore cross-collection searching and browsing via the aggregation of tags »<sup>14</sup>.

La question de la description et de l'intégration en son sein des vocabulaires issus des données du web 2.0 (tagging social ou wiki) est aussi au cœur de la réflexion sur le web sémantique et les moteurs de recherche sémantiques : comme le montre l'exemple du Centre Pompidou Virtuel, ce qui est en jeu, c'est un changement de paradigme, du « paradigme de la conservation » qui est celui des bases de données à « un paradigme de la diffusion » : en ce sens, outre le fait de s'interroger sur les mots clefs, le projet implique, selon un de ses responsables, « de réorganiser l'ensemble des processus de production, d'organisation, et de diffusion des contenus numériques afin de construire un centre de ressources de référence dans le domaine de la création moderne et contemporaine, et plus

---

<sup>13</sup> <http://www.steve.museum/>

<sup>14</sup> Traduction : « Nos activités incluent la recherche sur le tagging social et les collections muséales; le développement d'outils de logiciel libre pour le tagging des collections et la gestion de ces tags; et la discussion et la prise de contact avec les membres de la communauté qui sont intéressés par la mise en œuvre du tagging social pour leurs propres collections....Le but de ce projet en cours est d'améliorer la fonctionnalité des outils du logiciel de Steve, d'encourager l'utilisation de méthodes de tagging social dans les musées de toutes tailles et de tous types, de développer des interfaces de marquage novateurs, "de nouvelle génération" et d'explorer la recherche croisée sur les collections et la navigation via l'agrégation des tags. »

largement, du mouvement des idées contemporaines. Chaque contenu doit être identifié, indexé, traduit, rendu libre de droits, interopérable, et archivé de manière pérenne. Pour cela, le Centre Pompidou virtuel innove en adoptant un modèle de données basé sur le Web sémantique. Les ressources ne sont pas organisées suivant une hiérarchie rigide, mais permettent à l'internaute de naviguer par le sens. Les contenus sont ainsi naturellement décloisonnés et traités de manière homogène, afin de permettre leur organisation en fonction des besoins de chaque utilisateur, et non en fonction d'une logique dictée par des usages ou des structures définis<sup>15</sup> a priori ».

## Conclusion

La réflexion sur les bases de données ne peut être abstraite d'un questionnement plus global sur les usages du web, des usages de consultation aux usages participatifs voire contributifs. Comme le soulignait Milad Doueihi à propos des blogs, « on s'éloigne du type hiérarchique d'organisation et de présentation de l'information pour passer à un modèle plus sémantique et ontologique, fondé sur la prolifération de catégories et de mots-clés, qu'ils soient produits par les utilisateurs ou fixés centralement »<sup>16</sup>. On a vu que l'enjeu pour le patrimoine, c'était l'hybridation des axes de description. Les usages du web les plus récents dessinent progressivement une autre conception de la communauté patrimoniale qui inclut désormais, outre les représentants des diverses sciences du patrimoine (conservateurs, chercheurs, restaurateurs, spécialistes de l'imagerie etc.), les amateurs dont la simple présence et a fortiori l'activité produisent une reconfiguration des rapports. Le dialogue qui s'instaure par ces formes de coopération entre la communauté scientifique et les amateurs, a une conséquence d'ores et déjà perceptible : il institue le patrimoine moins comme objet de consommation que comme bien commun revendiqué.

---

<sup>15</sup> Emmanuelle Bernès, <http://www.semweb.pro/talk/2546>

<sup>16</sup> Milad Doueihi, *La grande conversion numérique*, Paris, Le Seuil, 2008, p.109.