



HAL
open science

Stratégies éditoriales des musées. Une approche de la médiation par l'accès ouvert aux données numérisées

Brigitte Juanals, Jean-Luc Minel

► **To cite this version:**

Brigitte Juanals, Jean-Luc Minel. Stratégies éditoriales des musées. Une approche de la médiation par l'accès ouvert aux données numérisées. Culture et Musées, 2020, Musées et mondes numériques (35), pp.49-75. <10.4000/culturemusees.4427>. <halshs-02899175>

HAL Id: halshs-02899175

<https://shs.hal.science/halshs-02899175v1>

Submitted on 9 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Stratégies éditoriales des musées et culture numérique Une approche de la médiation par les données culturelles

Brigitte Juanals, Jean-Luc Minel

Résumé:

Les politiques d'accès ouvert aux données culturelles des musées concernent désormais la mission de diffusion et de partage des collections et des connaissances. Cet article étudie les régimes d'accès et de circulation médiatique du patrimoine numérisé mis à disposition sous la forme de données culturelles en privilégiant une entrée par la médiation. Le terrain d'enquête est constitué par une sélection de vingt-et-un musées états-uniens ayant mis à disposition leurs données culturelles. Il apparaît que le choix d'une stratégie de diffusion des données culturelles ouvertes, en articulation avec un dispositif sociotechnique, implique, à une échelle transnationale, un débat sur le modèle de gouvernance permettant leur gestion pérenne. Ces données appartiennent au patrimoine culturel des institutions patrimoniales et des Etats qui les détiennent, ainsi qu'aux différentes catégories de publics qui en sont les usagers.

Cinq mots clefs

Patrimoine numérique, musées, données ouvertes, stratégies éditoriales, dispositifs sociotechniques.

Abstract :

Policies for open access to cultural data in museums now concern the mission of dissemination and sharing of collections and knowledge. This article studies access regimes and media circulation for digital heritage made available in the form of cultural data by fostering a mediation approach. The field of study consists of a selection of twenty-one US museums that have made available their cultural data. It seems clear that the choice of an open cultural data dissemination strategy, in conjunction with a socio-technical device, implies, on a transnational scale, a debate on the governance model for their perennial management. These data belong to the cultural heritage of the heritage institutions and the states that hold them, as well as to the various categories of audiences who use them.

Keywords

Digital heritage, museums, open data, semantic web, editorial strategies, socio-technical device.

La diffusion et la circulation de la culture dans le champ muséal a intégré une dimension numérique. Au sein d'un écosystème informationnel qui se complexifie dans les musées, les politiques et les pratiques éditoriales des établissements culturels se caractérisent par leur adaptation à l'évolution continue de l'offre médiatique face à la prolifération des sites web et des dispositifs numériques. Le développement de bases de données et les représentations numérisées des œuvres, associées à leur description documentaire, font partie de la gestion du patrimoine culturel et scientifique que les musées ont pour mission de préserver et de développer (Honeysett, 2012 : 49-53). Désormais, elle concerne également la mission de diffusion et de partage des collections et des connaissances auprès des publics. Les politiques d'accès ouvert aux données culturelles, qui constituent le thème traité dans cet article, en sont des évolutions récentes et centrales. Plusieurs établissements culturels états-uniens ont retenu notre attention en se distinguant par leurs annonces médiatisées et leurs politiques de mise à

disposition, en grandes quantités¹, de données ouvertes (*open data*) culturelles. Ces politiques font débat car elles soulèvent des problèmes complexes liés aux formes d'accès ouvert à la culture, aux droits de propriété intellectuelle des œuvres, ainsi qu'aux modèles documentaires, informatiques et économiques afférents.

La finalité de cette recherche est d'interroger les régimes d'accès et de circulation médiatique du patrimoine numérisé mis à disposition sous la forme de données culturelles. Le processus de construction d'une culture numérique dans ce domaine est exploré en privilégiant une entrée par la médiation. Qu'est-ce qui distingue les politiques de médiation par les données culturelles mises en place par des musées ? Quelles en sont les parties prenantes ? Quels sont les choix organisationnels et matériels de l'« ouverture » annoncée ? Quelles sont les formes de médiation aux données aménagées, les discours d'accompagnement et les dispositifs de différentes générations qui la porte ? Les politiques et les dispositifs organisationnels et sociotechniques émanant des institutions patrimoniales formatent les modes d'accès et de circulation des œuvres et des connaissances associées. Nous analysons de quelle manière et selon quels processus les choix de technologies et de langages de description des données déterminent le périmètre des espaces éditoriaux ouverts et mis en relation, les modes de représentation des connaissances et, au final, les formats de présentation et l'accès aux données culturelles pour les usagers.

UN CADRE THÉORIQUE INTERDISCIPLINAIRE POUR L'ÉTUDE DES « COMPLEXES »

Le cadre théorique interdisciplinaire conçu pour répondre à ces questions est appuyé sur les travaux réflexifs menés sur la médiation culturelle (Bordeaux & Caillet, 2013). Les objets d'étude, diversifiés et hétérogènes, englobent des politiques, des pratiques, des logiques d'acteurs, des dispositifs et des technologies ; en les considérant comme des « complexes » formés de « composants et de processus de nature hétérogène », nous étudions de quelle manière ils sont « matérialisés, institutionnalisés et opérationnalisés dans la société » (Davallon, 2004 : 34-35).

Situés dans leur contexte culturel et institutionnel, les stratégies de communication des musées (Juanals & Minel, 2016) sont ainsi analysées en interrelation avec les politiques en matière de médiation par les données culturelles ouvertes. Elles sont matérialisées dans les dispositifs organisationnels, sociotechniques et éditoriaux associés à la production ou à la collecte de données. Leur écriture et leur exploitation impliquent des choix conceptuels et techniques qui relèvent du champ disciplinaire de l'informatique et plus spécifiquement de la modélisation et de la représentation des connaissances (Shadbolt & al., 2006 ; Allemang & Hedler, 2011). Situées dans la sociologie de la traduction (Akrich, 1993 ; Latour, 1993) et la sociologie des usages (Jouët, 1993), les formes de la « médiation sociotechnique » permettent d'appréhender l'articulation entre les cadres sociaux, les technologies et les artefacts ouvrant à l'accès aux données. Nous analysons de quelle manière les modèles de description de données et les logiciels associés portent dans leur conceptualisation les politiques de médiation des musées qui les conçoivent et les mettent en œuvre. Ceci implique d'étudier les innovations techniques et sociales dans leurs interactions (Alter, 2002) et leur processus d'appropriation par un groupe social (Alter, 2000). Les systèmes informatisés et les langages de description des données sont considérés comme des « médiateurs » (Latour, 1993 : 43-44). Ils contribuent à conditionner les formes d'accès, d'usage et de circulation de données culturelles rattachées à

¹ Début 2018, les volumétries des musées étudiés s'échelonnaient en majorité entre 10.000 et 40.000 objets, 200.000 objets pour le *Yale Center for British Art* et 400.000 objets pour le MET.

différents établissements, jusqu'à déterminer les possibilités de leur mise en relation dans un espace inter-patrimonial (Juanals & Minel, 2016).

Cette recherche s'étend sur la période 2014-2018², au cours de laquelle plusieurs établissements états-uniens ont mis à disposition de grandes quantités de données culturelles « ouvertes ». Le terrain d'enquête est constitué par une sélection de vingt-et-un musées états-uniens³ sur le critère de leur engagement, dans leur communication institutionnelle et dans leurs pratiques éditoriales, dans une politique de diffusion culturelle appuyée sur des formes d'accès et d'utilisation de données ouvertes. La méthodologie prend en considération les aspects liés aux politiques, aux stratégies et aux positionnements des établissements culturels situés dans leur contexte états-unien. Les politiques de diffusion des musées consistent à rendre les données accessibles en y apportant, dans certains cas, des enrichissements. Les modes d'accès, décrits en termes « d'ouverture » des données, ont été menés et discutés (Knoblock & al., 2017) par de grands musées états-uniens dans des réseaux spécialisés – notamment le réseau AAC (American Art Collaborative), le réseau MCN (*Museum computer Network*) et le colloque annuel *Museums & the Web* – auxquels sont rattachés les établissements étudiés. Les méthodes d'enquête associent l'analyse des stratégies en matière de communication institutionnelle et de choix éditoriaux appuyés sur des dispositifs sociotechniques et des formes de médiation aux données ; elles sont complétées par le recueil et l'analyse de contenu des discours d'accompagnement afférents. Des entretiens semi-directifs ont été conduits auprès de vingt-sept professionnels rattachés aux métiers de l'informatique, de la documentation et de la conservation, dans les établissements étudiés (cf. annexe).

FORMES DE MÉDIATION ET RÉGIMES D'ACCÈS AUX DONNÉES CULTURELLES

Situées dans le cadre historique et socioculturel des Etats-Unis, les politiques de communication institutionnelle et de médiation culturelle des musées étudiés reposent sur les valeurs communes de la culture et de l'identité des Etats-nations (Bergeron, 2010 : 47). Dès l'entre-deux-guerres, la mission culturelle et démocratique de ces établissements, en complément de leur vocation patrimoniale, s'est affirmée (Ballé, 2002 : 18-19). Leurs fonctions sociales et culturelles ainsi que leurs modes de gestion actuels découlent de ces spécificités (Bergeron, 2010 : 46). Les discours institutionnels et les politiques de médiation culturelle des vingt-et-un musées étudiés s'inscrivent dans la réalisation de missions sociales, éducatives et de médiation culturelle. A la suite de la voie ouverte par l'ancien président B. Obama sur la mise à disposition des données politiques et la transparence de la vie publique, les professionnels interrogés font état d'une pression sociale et politique concernant l'accès ouvert aux données culturelles. Au sein de la société civile, ce dernier apparaît comme faisant partie de leur mission de diffusion de connaissances envers le plus grand nombre. Les arguments invoqués s'appuient sur le statut des données, qui sont publiques et d'un certain point de vue « appartiennent » à la société civile, et sur leur dimension patrimoniale, en référence à un patrimoine culturel et scientifique détenu par les musées.

De plus, dès leur création, les musées états-uniens et canadiens n'ont pas dépendu financièrement des gouvernements. Ils sont en majorité privés et financés par les citoyens, les entreprises et les fondations (Bergeron, 2010 : 47 ; Tobelem, 2013). Les établissements culturels étudiés se répartissent en trois groupes selon leurs modes de financement. Le premier groupe est formé d'institutions dont le financement est en partie assuré par le

² Recherche financée par deux projets ANR-Labex « Les passés dans le Présent », Université Paris Nanterre.

³ La liste des 21 musées et des 27 entretiens est indiquée en annexe.

gouvernement fédéral, en termes de dotation ou de ressources humaines ; à titre d'exemple, l'*American Art Museum* à Washington dépend de la *Smithsonian Institution*. Le deuxième groupe est constitué de musées dont l'essentiel du financement est assuré par des dotations régionales et complété par des donations privées. C'est le cas du *Walters Art Museum* de Baltimore ou du *Denver Art Museum*. Le dernier groupe est formé de musées dont le financement provient essentiellement de fondations privées. Le *J. Paul Getty Museum* et le *Yale Center for British Art* à New Haven en sont des exemples. Les différences de sources de financement ne sont pas un obstacle aux partenariats noués par les musées entre eux ; ainsi, le réseau AAC (*American Art Collaborative*)⁴ (voir infra) réunit des musées rattachés aux trois groupes cités.

Les modalités d'accès aux données et leur ouverture sont définies par les politiques éditoriales des établissements culturels. L'appellation « donnée culturelle » fait référence aux données sur les collections mises à disposition par les musées sous forme numérique. Il s'agit notamment d'images numérisées des œuvres, de fichiers sonores ou vidéos, de fichiers 2D ou 3D. Les métadonnées associées constituent une forme particulière de donnée qui décrit les attributs d'une donnée et lui donne une signification, un contexte et une organisation. Leur écriture est en conformité avec plusieurs standards⁵ mais s'avère variable en fonction des institutions du fait des pratiques locales de catalogages et des choix des descripteurs inclus dans les thésaurus. Ces derniers, en raison de motivations culturelles et scientifiques, sont fréquemment propres à un musée et à ses collections. Toutefois, ils peuvent aussi reprendre les thésaurus en accès ouvert et transposés dans le web de données ouvert du *Getty Research Institute* (Juanals & Minel, 2016 : 5-8).

Les aspects communs aux politiques et aux pratiques éditoriales mises en place sont de porter sur de grandes volumétries de données, de veiller à se limiter aux œuvres libres de droit d'auteur et de nécessiter des technologies numériques. Les motivations de tous les professionnels des établissements étudiés à s'engager dans la publication en ligne de données tiennent à la démocratisation culturelle et à l'exploration de formes renouvelées de la médiation culturelle numérique rendues possible par l'enrichissement et le croisement des données présentées. Par exemple :

« *The depth of the data we're releasing illuminates each object's unique history, from its original creator to how it arrived at the Smithsonian (...)* Now, a new generation can not only appreciate these works on their own terms, but remix this content in ways we have yet to imagine. » (20/07/2017, Courtney O'Callaghan, *Chief Digital Officer, Freer and Sackler*)

Ils font également état de motivations tenant à des stratégies de visibilité et de communication institutionnelle du fait de l'utilisation croissante du marketing culturel dans la gestion des établissements.

Les responsables de l'ouverture des données dans les musées étudiés produisent des données, les organisent, les interprètent et les diffusent. Les données publiées occupent désormais une place centrale sur leurs sites web. Elles sont destinées à une diversité d'usages, notamment culturel, éducatif et scientifique. De ce fait, les jeux de données (*data sets*) peuvent être proposés en accès libre mais comportent parfois des niveaux de restriction pour des usages commerciaux. Les différents modes d'éditorialisation auxquels ils donnent lieu déterminent leurs modes de mise à disposition.

Il ressort des expériences des institutions états-uniennes ayant fait le choix de mener leurs projets d'ouverture de données qu'elles ont été confrontées à des difficultés de différentes natures lors des étapes de modélisation et de transposition documentaire de leurs collections.

⁴ <http://americanartcollaborative.org/> [consulté le 30 janvier 2019]

⁵ Parmi les standards d'écriture des métadonnées utilisés dans les musées :: MARC, SPECTRUM, Dublin Core, DTD EAD (Encoding Archival Description).

Au début des années 2010, deux choix technologiques s’offraient aux professionnels interrogés pour affronter ces difficultés. Le premier s’appuyait sur des technologies matures, le second sur des modèles conceptuels et des technologies du Web de données ouvert.

TECHNOLOGIES MATURES ET OUVERTURE DES DONNÉES, UNE ASSOCIATION PROBLÉMATIQUE

Plusieurs institutions culturelles ont initié des programmes innovants d’ouverture de leurs données en adoptant une stratégie d’évitement des avancées conceptuelles et techniques du web de données ouvert qui auraient pourtant facilité leur partage. Leur politique est appuyée sur le choix d’une technologie informatique antérieure à la culture du web sémantique, l’interface de programmation applicative (API, *Application Programming Interface*)⁶ promue par le consortium mondial W3C (*World Wide Web Consortium*).

Le choix d’une politique d’ouverture des données fondée sur des API est une évolution récente dans les modalités d’accès aux ressources. La priorité donnée à l’accès à l’information sans prise en compte de sa conceptualisation correspond à une culture d’ingénierie informatique. Elle s’appuie sur un logiciel qui, selon les principes de la programmation orientée objet, offre des services (des « briques » de fonctionnalités) à des logiciels tiers. Conçue dans la finalité d’améliorer la diffusion des collections numériques, l’API s’inscrit dans une logique d’ouverture vers l’extérieur tout en restant paradoxalement centrée sur les ressources de l’établissement.

Afin d’appréhender les conséquences, sur les modes d’accès à l’information, du recours à cette technologie qui date des années 2000, nous avons étudié la politique d’ouverture fondée sur l’offre des API de trois institutions – *Cooper Hewitt Smithsonian Design Museum* (New York), *SFMOMA* (San Francisco), *Brooklyn Museum* (Brooklyn).

L’ouverture d’une API par le *Cooper Hewitt Museum* autorisant un accès à sa collection date de 2015⁷. L’utilisation de cet outil technique est ouverte aux développeurs internes de la *Smithsonian Institution* et à des tiers (Brenner 2015). Notre test d’utilisation de l’API a fait apparaître plusieurs difficultés : parmi plus de 80 méthodes proposées, les onze méthodes dédiées aux objets nécessitent de connaître l’identifiant de l’objet ; il n’existe pas de méthode fournissant la liste de tous les objets de la collection ; les champs décrits dans ces méthodes ne sont pas issus de vocabulaire standards et certains champs mélangent des informations ; les concordances entre les identifiants de huit entrepôts sont stockées dans un fichier de tableur dans l’entrepôt *github* du musée. Le constat est fait de la complication et de la dispersion de l’accès aux ressources tenant au choix technique de l’interface de programmation.

L’étude de l’API du SFMOMA révèle des finalités contradictoires. L’outil technique est présenté comme un instrument de travail, destiné à des artistes, permettant des visualisations esthétiques construites à partir des données (Winesmith & Carey, 2014). Toutefois, l’objectif affiché du *hackaton* « Art + Data Day », organisé par le SFMOMA en 2015 et réunissant designers, ingénieurs et conservateurs du musée, consistait à dégager des pistes pour les usages futurs de l’API :

« *For SFMOMA, the primary goal of Art + Data Day was [...] to engage participants in open dialogue about the API’s potential and the innovation it could facilitate.* » (Hogarty & al., 2015)

⁶ <https://www.w3.org/2006/webapi/admin/charter> [consulté le 30 janvier 2019]

⁷ <https://collection.cooperhewitt.org/api/> [consulté le 30 janvier 2019]

Dans cette perspective, la nécessité d’alignement des descriptions du SFMOMA avec celles d’autres institutions, l’importance des vocabulaires et de la qualité des données étaient soulignées :

« *The simple presentation mode best serves the initiate, the data hacker, artists, and designers, while a mode for delivering data in a machine-readable format—such as HAL or linked-data standard formats—enables complex connections with other museum collections and data sources.* »⁸

Toutefois, ce discours d’accompagnement ne correspond pas au dispositif mis en place : notre test d’accès à l’API a fait apparaître des vocabulaires propriétaires, sans référence à des vocabulaires standards, et le non alignement des données dans toutes les notices consultées.

En dernier lieu, l’API du *Brooklyn Museum*⁹, conçue selon une série de 30 méthodes ouvrant l’accès aux informations décrivant artistes, œuvres et collections, pose également des problèmes. Bien que l’étude des paramètres des méthodes révèle des principes proches d’un accès à des tables de base de données avec indication d’une clef primaire (un identifiant propre à l’institution), le modèle complet des liens entre ces tables n’est pas fourni. De plus, les vocabulaires de description sont propres au musée et les valeurs de certains champs, destinées à des lecteurs humains, s’avèrent peu exploitables automatiquement.

L’étude des modalités d’accès et d’usage des API proposées par trois grands musées illustre les types problèmes soulevés par l’usage de cette technologie. Les modes d’accès aux données des collections, bien que fondés sur des principes techniques partagés, restent limités à un seul établissement. De ce fait, ils ne permettent pas la construction d’un espace culturel inter-muséal. En effet, la sémantique de chaque méthode ou de chaque champ descriptif reste propre à chaque musée. De plus, l’alignement des données entre différentes collections est en général inexistant. Il s’avère que le choix d’une interface de programmation applicative privilégie une conception locale et « propriétaire » des données. Il reste théoriquement possible de concevoir des réalisations qui exploiteraient différentes API, mais l’effort de développement à fournir pour s’approprier les sémantiques des différentes API et de leurs modèles de données sous-jacents serait important. Cela signifie que l’institution offre la possibilité d’exploiter ses données, mais que les développements réalisés restent confinés dans l’espace informationnel du musée.

WEB SÉMANTIQUE ET PARTAGE DES DONNÉES CULTURELLES : UN DISPOSITIF INNOVANT MAIS INADAPTÉ AUX PRATIQUES LOCALES

Plusieurs musées de stature internationale se sont engagés, à partir des années 2010, dans le déploiement progressif des politiques et des technologies du web sémantique (web de données ouvert, *Linked Open Data*, LOD) (Johnson & al., 2015 : 6-7 ; Juanals & Minel, 2016). Les transpositions, par ces grandes institutions, de leurs collections ou leurs thésaurus dans le web de données ouvert ont suscité de nombreuses discussions dans les éditions du colloque annuel *Museums and the Web*. Elles constituent des décisions à long terme visant à dépasser les expériences ponctuelles lancées dans le courant des années 2000. Une tendance en cours de maturation semblait se dessiner dans les développements innovants de ces institutions patrimoniales. En mettant à disposition (en « exposant ») leurs données et en les partageant dans un espace culturel inter-muséal, elles offraient ainsi la possibilité à des acteurs tiers et aux publics de les exploiter et de les comparer à leur guise. Lancé en 2015, le projet *American*

⁸ Point de vue partagé par : *Freer and Sackler, Denver Art Museum, Art Institute of Chicago, Museum of Modern Art, San Francisco Museum of Modern Art.*

⁹ <https://www.brooklynmuseum.org/opencollection/api/docs/overview> [consulté le 30 janvier 2019]

Art Collaborative (AAC) qui réunit quatorze institutions américaines est emblématique de cette politique éditoriale (Knoblock & al., 2017).

Toutefois, nous retirons de notre travail d'enquête plusieurs analyses tenant aux difficultés et aux contraintes expérimentées par le choix de mettre en œuvre le web de données ouvert. Elles sont étayées par les entretiens menés en 2017 auprès des coordinateurs des systèmes d'information des institutions patrimoniales concernant leur expérience récente, et par les publications scientifiques des acteurs impliqués (Fink, 2018) ; elles sont complétées par l'étude des dispositifs sociotechniques construits.

L'analyse du rôle joué par les langages et les modèles conceptuels dans la détermination des formes de médiation permet de comprendre les problèmes qui se posent aux acteurs professionnels engagés dans l'ouverture de leurs données culturelles selon le modèle du web sémantique. Il apparaît que certaines des contraintes liées au choix du LOD relèvent spécifiquement de l'appropriation du langage de description RDF, d'autres du choix du modèle conceptuel CIDOC-CRM, d'autres encore de l'articulation entre ce modèle et le langage RDF. En dernier lieu, la gestion de la coexistence entre le système d'information natif et le nouvel entrepôt de données construit à partir de celui-ci génère un surcroît de travail technique.

La transposition des descriptions documentaires des collections dans les langages du LOD est indispensable pour permettre l'ancrage dans le web sémantique. Sur le plan technique, le LOD repose sur une pile de langages de description¹⁰ organisé selon le principe d'inférence. Ce dispositif permet de rechercher, dans un entrepôt de données décrites dans le format RDF, nommé *triple store*, les données répondant aux critères d'une requête exprimée dans un langage informatique d'interrogation spécifique – *Sparql*. Les langages de description du LOD sont conçus par des acteurs spécialisés en ingénierie rattachés au consortium international W3C. En termes de médiation, l'intérêt de ces langages de description et de programmation est de fournir les moyens de concevoir différents modèles conceptuels des objets d'une collection.

Le choix d'un modèle conceptuel est central car les concepts constituent les points d'entrée dans les collections ; les conservateurs qui en ont la charge y sont particulièrement attentifs. Toutefois, l'exploitation d'un modèle conceptuel se heurte à plusieurs difficultés qui sont autant de nature conceptuelle que liées aux usages. Le principal écueil dans l'élaboration et le nommage de concepts est de parvenir à identifier un ensemble de concepts situé à un niveau élevé de généralité. De plus, ils doivent être articulés avec ceux correspondant aux besoins scientifiques des modes de description de collections spécifiques. En étant le fruit d'histoires singulières, ces concepts « locaux » sont ancrés dans une culture et un contexte particuliers nécessitant une sémantique adaptée (Juanals & Minel, 2016 : 5-8). Dans le domaine patrimonial, la norme (ISO 21127:2006) CIDOC-CRM, « CIDOC object-oriented Conceptual Reference Model (CRM) », constitue une référence (Le Bœuf, 2014) : elle joue un rôle central de « colle sémantique » dans la médiation de l'information sur le patrimoine culturel en permettant d'unifier de manière cohérente, dans un langage formel commun, des sources d'information locales qui pourraient de ce fait être partagées. Co-développée par l'influente organisation ICOM (*International council of museums*) et le CIDOC *Documentation Standards Group*, cette norme s'adresse de manière explicite aux conservateurs, aux chercheurs et aux universitaires du secteur culturel et patrimonial. .

Toutefois, tous les acteurs interrogés dans les musées ont souligné la grande difficulté, pour leurs équipes d'ingénierie en place, de l'étape de conversion des données natives de leurs institutions patrimoniales dans le format RDF du fait de sa complexité.

¹⁰ Langages RDF, RDFS et OWL. <http://www.w3.org/RDF/> [consulté le 30 janvier 2019]

« *We were interested by the AAC project, but we did not have the competent technical staff to contribute to this project.* » (25/07/2017, Jennie Choi, *General Manager of Collections Information Digital, Metropolitan Museum of Art*).¹¹

L'une des principales critiques adressée à la norme CIDOC-CRM concerne sa posture théorique éloignée de la culture de métier et des pratiques courantes de catalogage des objets du domaine patrimonial dans les musées :

« *The CRM was created with the notion that it is best applied by curators who know the most about the objects in the museum. Is this notion realistic for most museums?* » (Fink, 2018 : 34, 21/07/2017, E. Fink, fondatrice et responsable scientifique du *American Art Collaborative, AAC*).

Ce constat a conduit l'AAC à proposer un modèle conceptuel commun simplifié avec un nombre limité de concepts issus du CIDOC-CRM afin de s'adapter aux réalités des pratiques et des savoirs de métier des musées partenaires. Dans le même esprit, afin de contourner la complexité de mise en œuvre du CIDOC-CRM, le *Getty Research Centre* explore un modèle plus simple, le CDWA, sur l'argument qu'il est déjà d'un usage courant dans les musées américains (31/7/2014, J. Cobb et P. Harpring, *Managing Editor, Getty Vocabulary Program*).

En dernier lieu, l'usage de ce dispositif sophistiqué a fait apparaître des problèmes d'articulation entre le format RDF et la norme CIDOC-CRM, lesquels sont issus de domaines scientifiques et de métiers différents. E. Fink tire les conséquences de cette situation :

« *In some instances, the CRM did not seem ideally suited for RDF and LOD overall. Since the CRM was not created for LOD, some aspects of the CRM do not follow common protocols*» (Fink, 2018 : 34, 21/07/2017, E. Fink, AAC)¹².

Une analyse en représentation des connaissances ouvre à une interprétation du problème posé en termes de modélisation : le CIDOC-CRM est issu de la conception de la programmation orientée objet dont les notions d'héritage et de taxinomie forment le socle conceptuel, alors que le LOD a pour socle les notions d'inférence et de type, chères à l'intelligence artificielle. De ce fait, les deux approches construisent des modèles conceptuels très différents qui peuvent donner lieu à des inconsistances dans la représentation des connaissances lorsque l'on associe ces deux approches¹³ (Allemang & Hendler, 2011).

En dernier lieu, des problèmes de coexistence de systèmes d'information sont apparus : ils ont amenés le *Getty Research Institute* et le *Yale Center for British Art* à construire des entrepôts de données en format RDF à partir de leurs systèmes d'information natifs mais à conserver ce dernier, notamment dans le but de gérer des données de type administratif. Cette situation nécessite une chaîne de traitement régulièrement mise en œuvre pour prendre en compte les ajouts et corrections effectuées sur les données qui décrivent les collections.

De plus, ces systèmes d'informations natifs s'appuient, pour la plupart d'entre eux, sur des technologies anciennes qui ne donnent pas la possibilité de créer des champs pouvant encoder des valeurs complexes. E. Fink synthétise les retours de terrain des institutions partenaires de l'AAC :

« *Collection information systems (CISs) (...) do not allow data to be exported in ways that easily relate to the schema of the CRM.(...). Partners also lamented the fact that in preparing their data for this project, they had difficulty inserting identifier codes into their CIS to*

¹¹ Point de vue partagé par : *Freer and Sackler, Princeton University Art Museum, Denver Art Museum, Art Institute of Chicago, Museum of Modern Art, San Francisco Museum of Modern Art.*

¹² Point de vue partagé par : *The Walters Art Museum, Yale Center for British Art, Smithsonian American Art Museum.*

¹³ A titre d'illustration, il est possible de déclarer dans le langage RDFS que B est une sous-classe (subClassOf) de A et réciproquement que A est une sous-classe de B, ce qui est inconsistant dans un langage de programmation orienté objet.

structure and link fields to various thesauri, classifications, periods, and other references. » (E. Fink, 2018 : 33 ; 21/07/2017, E. Fink, AAC)¹⁴

Ces expériences font apparaître les potentialités et les difficultés rencontrées pour traiter informatiquement, en s'appuyant sur un modèle qui associe graphe et taxinomie, des descriptions approximatives de données cataloguées selon une logique arborescente et destinées à une interprétation humaine.

L'HYBRIDATION SOCIOTECHNIQUE COMME MODÈLE UNIFICATEUR

La mise en œuvre de la médiation aux données culturelles nécessite de dépasser le clivage conceptuel et technique entre les deux grands dispositifs sociotechniques analysés. L'interface de programmation applicative est familière aux équipes informatiques mais sa logique d'accès, locale et fermée sur un établissement, autorise mais complique l'accès à l'information. Le web de données ouvert présente l'apport d'une logique d'accès standardisée et ouverte à l'interrelation des ressources entre établissements, mais les langages conceptuels et de programmation sont difficiles en termes de compréhension et d'usage par les équipes informatiques.

Faisant ce constat, le *Yale Center for British Art* a conçu un dispositif hybride combinant des modes d'accès différents, mais cette situation complique la gestion des données :

« Le *Yale Center for British Art* a fait le choix du LOD, d'un *SparqlEndpoint* et d'un entrepôt *github*¹⁵ et d'une API d'interface. La grande difficulté est de synchroniser ces différents entrepôts de données avec le système de gestion TMS qui gère les données natives». (7/07/2018, E. Delmas, *Yale Center for British Art*)

Un dispositif sociotechnique en déploiement confirme l'orientation vers la construction de technologies hybrides visant à dépasser les difficultés rencontrées. Le standard *International Image Interoperability Framework (IIIF)*, initié en 2011 par un groupement de bibliothèques et d'universités¹⁶, rend possible la visualisation, dans une même interface, de documents numérisés disséminés dans différentes collections de bibliothèques. Il s'est révélé adapté à la visualisation d'œuvres numérisées détenues par les musées, incitant plusieurs institutions états-uniennes¹⁷ à rejoindre le consortium international IIIF créé en 2015¹⁸.

Sur le plan technique, la mise en œuvre d'IIIF repose sur deux interfaces de programmation applicative (API) : l'API *Image* définit un service Web qui fournit une image en réponse à une requête précisant ses caractéristiques ; l'API *Presentation*, nécessaire pour associer des métadonnées aux images, spécifie, sous la forme d'un document au format JSON-LD¹⁹, la structure et la présentation d'un objet numérisé et de ses contenus associées. L'exploitation de ces API implique l'utilisation de visionneuses d'images qui sont pour la plupart d'entre elles s-*Diva.js*, *Mirador*, *OpenSeadragon*, *Universal Viewer* – des logiciels libres. Plusieurs

¹⁴ Point de vue partagé par : *The Walters Art Museum, Smithsonian American Art Museum*.

¹⁵ Service web d'hébergement gratuit pour les logiciels libres lancé en 2008, racheté en 2018 par la société Microsoft.

¹⁶ *British Library*, Université de Stanford, Bibliothèque Bodléienne (Université d'Oxford), Bibliothèque nationale de France, *Nasjonalbiblioteket* (Bibliothèque Nationale de Norvège), Laboratoire National de Los Alamos, Université de Cornell.

¹⁷ *Art Institute of Chicago, Carnegie Museum of Art, Cooper Hewitt Smithsonian Design Museum, Frick Collection, J. Paul Getty Trust, Harvard Art Museums, Hill Museum and Manuscript Library (HMML), National Gallery of Art, Paul Mellon Centre, Yale Center for British Art*.

¹⁸ <https://iiif.io/community/groups/museums/#about> [consulté le 30 janvier 2019]

¹⁹ Standard du W3C (14/01/2014). Un graphe RDF peut être « sérialisé » en JSON-LD.

démonstrateurs (dont le projet Bibliissima avec la BNF en France), ont été mis en place par les membres fondateurs du Consortium²⁰.

Lors de notre enquête, tous les professionnels interrogés ont manifesté leur intérêt pour ce standard. L'accord de partenariat, conclu en novembre 2017, entre le consortium *American Art Collaborative* (cf. supra) et le consortium IIF a pour projet d'exploiter les collections numérisées de huit des musées²¹ partenaires de AAC :

« *The American Art Collaborative (AAC), fourteen art museums working together to implement linked open data, got terrific support from the IIF Consortium. Art museums come from a strong visual culture so IIF excited the museums from the first demonstration.* » (R. Allen, *Deputy Director, Smithsonian American Art Museum*, Communiqué de presse du consortium IIF, 1/11/2017).

Dans la même perspective, l'accord conclu en juin 2017 entre le *Getty Museum* et le *Yale Center for British Art* les a amenés à diffuser, sur leurs sites web, respectivement 30 000 et 70 000 images compatibles dans le standard IIF :

« *By adopting the IIF, our images can now travel beyond the confines of our own website and become fully interoperable with images from other collections, greatly enhancing the ability to pursue research in the digital environment.* » (Matthew Hargraves, Conservateur en chef des collections, *Yale Center, Yale News*, 2/06/2017 ; 19/07/2017, E. Glass).

Le standard IIF a été initialement conçu pour visualiser sur une même interface web des images conservées par des institutions patrimoniales disposant de différents fragments ou de différentes versions d'une même œuvre ou d'un même artiste. Plus largement, la visualisation d'œuvres facilite des comparaisons scientifiques. Ses fonctionnalités ouvrent également la voie à des expérimentations d'usages à destination du grand public (Roddis, 2018) visant à favoriser l'apprentissage par l'observation, (*slow looking*) (Tishman, 2017).

Ainsi, l'hybridation sociotechnique, en associant API et représentation de données liées, pourrait, au-delà de son objectif initial de fournir des outils de partage interopérables aux historiens de l'art, chercheurs et érudits, proposer des outils de médiation en direction du grand public.

LA CONSTRUCTION D'UN ESPACE CULTUREL INTER-MUSEAL : UN CHOIX STRATEGIQUE

Cette enquête sur les stratégies des établissements culturels états-uniens en matière de gestion des données ouvertes rejoint les résultats de nos terrains de recherche sur les politiques numériques en France et en Europe (Juanals & Minel, 2016, 2018). Ces travaux nous conduisent à distinguer deux orientations possibles concernant la vision stratégique d'une politique d'établissement dans ce domaine : la première voie privilégie une stratégie fédérative et la seconde, une stratégie agrégative.

Dans la stratégie fédérative, revendiquée par le projet AAC (Fink, 2018 : 21), chaque institution patrimoniale est responsable de l'administration de ses données conservées dans des entrepôts indépendants. Les faiblesses associées à ce choix résident dans l'obligation, pour chaque institution, de développer ses propres outils de conversion de données. De plus, du fait de la dissémination des données culturelles dans différents entrepôts, les performances de recherche d'information sont susceptibles de se dégrader. En revanche, ce choix garantit la pérennité de l'écosystème informationnel, même en cas de disparition de l'un de ses acteurs,

²⁰ *Stanford University Libraries, Princeton University Libraries, British Library, Europeana.*

²¹ *Walters Art Museum, National Portrait Gallery, Archives of American Art, Smithsonian American Art Museum, Dallas Museum of Art, Indianapolis Museum of Art, Crystal Bridges Museum of Art, Colby College Museum of Art.*

et la mise à jour des données dans la mesure où elle est réalisée par les institutions qui en sont détentrices.

Le choix d'une stratégie agrégative apporte une réponse aux difficultés rencontrées par les équipes des musées dans l'appropriation informatique et documentaire de dispositifs sociotechniques innovants. Elle est mise en place notamment par le projet *Wikidata*²². Selon cette orientation, les données des institutions patrimoniales sont collectées et centralisées dans un entrepôt unique, leur gestion étant assurée par une organisation indépendante des institutions patrimoniales. Conçu à l'origine comme un entrepôt de référence, *Wikidata* prend désormais en charge la collecte et le stockage des données, ainsi que la mise à disposition d'outils de recherche d'information (Deshayes, 2018 ; Molinié, 2018). Ce choix, qui est celui du MET, garantit de bonnes performances dans la recherche d'information par les publics ; elles sont associées à une économie de moyens dans la conversion et l'administration des données collectées. Toutefois, il pose une question de gouvernance concernant la pérennité, en termes financier et technologique, de l'organisation en charge de la gestion et de la diffusion des données. De plus, cette stratégie, pour les institutions patrimoniales, revient à déléguer la conservation et l'accès à leurs données culturelles à une organisation externe rattachée à une fondation privée. A titre de comparaison, l'Union européenne, avec la fondation *Europeana*, a fait le choix d'une organisation à la fois agrégative et fédérative.

CONCLUSION

L'analyse des régimes d'accès aux données culturelles révèle des tensions de différentes natures. En premier lieu, elle fait apparaître l'ancrage national, culturel et historique, des politiques et des pratiques éditoriales des musées états-uniens, concomitamment à l'histoire singulière des collections d'un établissement et de ses personnels. En second lieu, les innovations technologiques et documentaires confrontent les musées à la nécessité de faire des choix en tenant compte de leurs cultures organisationnelles et de leurs savoirs de métier. Les problèmes, de diverses natures, rencontrés par les équipes des musées états-uniens dans l'ouverture de leurs données, font apparaître que la culture numérique recouvre des métiers, des connaissances et des savoir-faire différents en fonction des professionnels impliqués. Le choix d'une solution sociotechnique hybride ou celui de la délégation de leur gestion informatique et documentaire, constituent des orientations possibles vers des systèmes documentaires ouverts et interopérables. Elles rendent réalisables leurs politiques d'innovation concernant l'ouverture de leurs données culturelles mais ont des conséquences sur les conditions de leur conservation.

Les politiques d'accès ouvert aux données culturelles des musées concernent désormais la mission de diffusion et de partage des collections et des connaissances. Il apparaît que le choix d'une stratégie de diffusion et de gestion des données culturelles ouvertes, en articulation avec un dispositif sociotechnique, est devenu une question politique et économique. Elle implique, à une échelle transnationale, un débat sur le modèle de gouvernance permettant une gestion pérenne des données culturelles ouvertes, qui appartiennent au patrimoine culturel des institutions patrimoniales et des Etats qui les détiennent, ainsi qu'aux différentes catégories de publics qui en sont les usagers.

²² Le projet *Wikidata* est hébergé par *Wikimedia Foundation*, financièrement soutenue par des contributeurs privés, dont *Google Inc.*

BIBLIOGRAPHIE

- Akrich (Madeleine). 1993. « Les formes de la médiation technique ». *Réseaux*, n°60.
- Allemang (Dean), Hedler (Jim). 2011. *Semantic Web for the Working Ontologist : Effective Modeling in RDFS and OWL* : Morgan-Kaufman.
- Alter (Norbert). (2000). *L'innovation ordinaire*, Paris, PUF, 2000.
- Alter (Norbert) (dir.). 2002. *Les logiques de l'innovation*, Paris, La Découverte, Recherches.
- Ballé (Catherine). 2002. « Le musée, une organisation en changement », in Fourteau, Claude (dir.), *Les institutions culturelles au plus près du public*, Musée du Louvre – *La Documentation Française*, Paris.
- Bergeron (Yves). 2010. « Métissages entre musées d'art et musées de société dans les musées nord-américains », *Cultures & Musées*, 16, p. 45-63.
- Bordeaux (Marie-Christine), Caillet (Elizabeth). 2013. « La médiation culturelle : pratiques et enjeux théoriques », dans Gottesdiener (Hana) et Davallon (Jean) (sous la direction de), « La muséologie : 20 ans de recherche », *Culture & Musées*, 20, p. 139-161.
- Brenner (Sam). 2015. «Reconsidering searching and browsing on the Cooper Hewitt's Collections website. », *MW2015: Museums and the Web 2015*. Consulté le 15 décembre 2018, <http://mw2015.museumsandtheweb.com/paper/reconsidering-searching-and-browsing-on-the-cooper-hewitts-collections-website/>
- Davallon (Jean). 2004. « Objet concret, objet scientifique, objet de recherche », *Hermès*, 1 (38), p. 30-37.
- Deshayes (Benoit). 2018. «Wikimedia, Wikidata et institutions patrimoniales (GLAM)», Séminaire du centre de recherche Dominique-Vivant Denon du musée du Louvre et du labex Les passés dans le présent.
- Fink (Eleanor). 2018. « American Art Collaborative (AAC), Linked Open Data (LOD) Initiative, Overview and Recommendations for Good Practices ». http://americanartcollaborative.org/wpcontent/uploads/2018/03/AAC_LOD_Overview_Recommendations.pdf.
- Hogarty (Sarah Bailey), Winesmith (Keir), Hrudka (Matthew), Schechter (Beth). «Art + Data: Building the SFMOMA Collection API», *Museums and the Web*, <http://mw2015.museumsandtheweb.com/paper/art-data-building-the-sfmoma-collection-api>
- Honeysett Nik (Getty). 2012. « The Future of Museum Scholarly Catalogues », *The Bulletin of the ASIST*, Volume 38-3, février-mars 2012, [En ligne]. <http://www.asis.org/Bulletin/February2012/index.html>. Consulté le 27 janvier 2019.
- Johnson (Larry), Becker (Samantha Adams), Estrada (Victoria), Freeman (Alex). 2015. *The NMC Horizon Report: 2015, Museum Edition*. Austin, Texas: The New Media Consortium [En ligne]. <http://cdn.nmc.org/media/2015-nmc-horizon-report-museum-EN.pdf>. Consulté le 29 janvier 2019.
- Jouët (Joël). 1993. « Pratiques de communication et figures de la médiation », *Réseaux*, 11(60), p. 99-120.
- Juanals (Brigitte), Minel (Jean-Luc). 2016. «La construction d'un espace patrimonial partagé dans le web de données ouvert. Entre interopérabilité et stratégies institutionnelles », *Communication*, 34(1), p. 2-18.
- Juanals (Brigitte), Minel (Jean-Luc). 2018. «Politique de publication en ligne des données (open access) et circulation des connaissances sur les RSN : le cas du Louvre», Séminaire du centre de recherche Dominique-Vivant Denon du musée du Louvre et du labex « Les passés dans le présent ».

Knoblock (Craig A.), Szekely (Pedro), Fink (Eleanor), Degler (Duane), Newbury (David), Sanserson (Robert), Blanch (Kate), Snyder (Sara), Chheda (Nilay), Jain (Nimesh), Krishna Ravi (Raju), Sreekanth (Nikhila Begur), Yao (Yixiang). 2017. « Lessons Learned in Building Linked Data for the American Art Collaborative », *ISWC*, Vienna, Austria.

Latour (Bruno). 1996 [1993]. *La clé de Berlin. Petites leçons en sociologie des sciences*. Paris : Le Seuil, coll. Points sciences.

Le Bœuf (Patrick). 2014. « Le Modèle conceptuel de référence du CIDOC : de la sémantique des inventaires aux musées en dialogue », *Culture & Musées*, 22, p. 89-111.

Molinié (Christelle). 2018. « Ouverture des données et des contenus, valorisation des collections au musée Saint-Raymond, musée des Antiques de Toulouse », Séminaire du centre de recherche Dominique-Vivant Denon du musée du Louvre et du labex « Les passés dans le présent ».

Roddis (Tristan). 2018. « Making metadata into meaning : Digital storytelling with IIIF », *MWI8*, Vancouver. Consulté le 27 Décembre 2018.

Shadbolt (Nigel), Hall (Wendy), Berners-Lee (Tim). 2006. « The Semantic Web Revisited » *IEEE Intelligent Systems*, p. 96- 101, consulté le 13 décembre 2018.

Tishman (Shari). 2017. *Slow Looking: The Art and Practice of Learning Through Observation*. London: Routledge.

Tobelem, (Jean-Michel). 2013. « Démocratiser la culture. Le cas des États-Unis d'Amérique. Un contexte en évolution » in *Démocratiser la culture. Une histoire comparée des politiques culturelles*, sous la direction de Laurent Martin et Philippe Poirrier, Territoires contemporains, nouvelle série, 5. Consulté 25/12/2018, http://tristan.u-bourgogne.fr/CGC/publications/Democratiser_culture/JM_Tobelem.html

Winesmith (Keir), Carey (Anna). 2014. « Why build an API for a Museum Collection ? » *SFMOMA*, <https://www.sfmoma.org/read/why-build-api-museum-collection/>. Consulté le 29 janvier 2019.

ANNEXE

Liste des entretiens

1. Rachel M. Allen, Deputy Director, Smithsonian American Art Museum, Washington
2. Jennifer Berry, Rights & Reproductions Coordinator, the Freer and Sackler, Washington
3. Kate Blanch, Systems Manager, Data & Digital Resources, The Walters Art Museum, Baltimore
4. Allegra Burnett, Web and Social Media Content Manager, Museum of Modern Art, New York
5. Jennie Choi, General Manager of Collections Information Digital, Metropolitan Museum of Art, New York
6. Joan Cobb, Software Architect, Getty Vocabulary Program, Getty Research Institute, Los Angeles
7. Grace Cory, Digital Asset Manager, Freer and Sackler, Washington
8. Emmanuelle Delmas, Collections Data Manager, Yale Center for British Art, New Haven
9. Eleanor Fink, Art and Technology, Information Sciences Institute, fondatrice et responsable scientifique du projet American Art Collaborative, Washington
10. Georgina Goodlander, Web and Social Media Content Manager, Smithsonian American Art Museum, Washington
11. Cathryn Goodwin, Manager of Collections Information, Princeton University Art Museum,
12. Princeton.
13. Fairlight Gutierrez, Digital Marketing/Content Strategy, Editor, Denver Art Museum, Denver.

14. Patricia Harpring, Managing Editor, Getty Vocabulary Program, Getty Research Institute, Los Angeles.
15. Lindsey Housel, Manager of Digital Initiatives, Denver Art Museum, Denver.
16. Craig Knoblock, Director of Information Integration, University of Southern California
17. Liz Nelly, Museum Computer Network, Art Institute of Chicago, Chicago.
18. Courtney O'Callaghan, Chief Digital Officer, Freer and Sackler, Washington
19. Peter Samis, Associate Curator, Interpretive Media, San Francisco Museum of Modern Art, San Francisco.
20. Stanley Smith, Head of Collection Information and Access, J. Paul Getty Museum, Los Angeles.
21. Sara Snyder, Webmaster at the Archives of American Art, Smithsonian American Art Museum, Washington.
22. Kiser Spencer, Developer, Metropolitan Museum of Art, New York
23. Janet Strohl-Morgan, Associate Director for Information and Technology, Princeton University Art Museum, Princeton.
24. Pedro Szekely, Research Associate Professor, University of Southern California, Los Angeles.