



HAL
open science

Comment définir une politique de valorisation de collections numériques issues de fonds d'archives à l'heure du web de données ? Le cas de la médiathèque du Centre National de la Danse

Amélie Vernusset

► To cite this version:

Amélie Vernusset. Comment définir une politique de valorisation de collections numériques issues de fonds d'archives à l'heure du web de données ? Le cas de la médiathèque du Centre National de la Danse. domain_shs.info.docu. 2013. mem_00945608

HAL Id: mem_00945608

https://memsic.ccsd.cnrs.fr/mem_00945608

Submitted on 12 Feb 2014

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.



CONSERVATOIRE NATIONAL DES ARTS ET METIERS

Ecole Management et Société-Département CITS

INTD

MEMOIRE pour obtenir le
Titre professionnel "Chef de projet en ingénierie documentaire" INTD
RNCP niveau I

Présenté et soutenu par

Amélie VERNUSSET

le 16 décembre 2013

Comment définir une politique de valorisation de
collections numériques issues de fonds d'archives à
l'heure du web de données ?

Le cas de la médiathèque du Centre National de la Danse

Jury :

Nadia RAÏS, Professeur associée au CNAM-INTD

Abdel MONTET, Responsable de l'informatique documentaire, Centre National de la Danse

Bertrand SAJUS, Responsable de programmes numériques, Ministère de la Culture et de la
Communication

Promotion 43

Remerciements

J'adresse mes remerciements à toutes les personnes qui m'ont aidé dans la réalisation de ce mémoire.

En premier lieu, je remercie toute l'équipe de la médiathèque du Centre National de la Danse pour leur accueil chaleureux et les échanges intéressants. Mes remerciements vont en particulier à Laurent Sébillotte pour m'avoir accueillie dans sa structure, à Abdel Montet, responsable de l'informatique documentaire en qualité de tuteur de stage et à Juliette Riandey, responsable des archives. Abdel et Juliette m'ont conseillé pendant la réalisation de mon stage et de mon mémoire.

Je remercie aussi les professionnels qui m'ont reçue pour discuter de leurs projets en cours. Bertrand Sajus, du Département des programmes numériques au sein du Ministère de la Culture et de la Communication, pour les discussions sur le web des données et les projets en cours actuellement. De la Bibliothèque Nationale de France, Jérôme Villeminez (data.bnf.fr et OpenCat) et Louis Jaubertie (la médiation numérique de Gallica). Du Ministère de la Culture et de la Communication, Katell Briatte (HADOQ et GINCO), Marie-Véronique Ieroi (LinkedHeritage et MICHAEL Culture pour la coopération avec Europeana) ainsi que Laurent Manceuvre et Carine Prunet (Joconde Lab). De la Cité de la Musique, Rodolphe Bailly (MIMO et les projets de la médiathèque de façon générale).

Je remercie également toute l'équipe de l'INTD et en particulier Nadia Rais, professeur associée à l'INTD au sein du CNAM. En tant que directrice de mémoire, elle m'a prodigué conseils avisés et encouragements tout au long de mon travail et m'a aidée à trouver des solutions pour avancer.

Enfin, je remercie mes collègues de promotion pour les discussions et échanges autour de mon travail, Béatrice De Larouzière tout d'abord mais aussi Sylvie Bonnet, Patricia Chandon Moët et Nadège Fouan.

Notice

VERNUSSET Amélie. Comment définir une politique de valorisation de collections numériques issues de fonds d'archives à l'heure du web de données ? Le cas de la médiathèque du Centre National de la Danse. 2013. p. 130. Mémoire Titre professionnel de niveau 1, CNAM-INTD. 2013

La médiathèque du Centre National de la Danse a fait le choix de développer une bibliothèque numérique pour valoriser ses collections issues principalement de fonds d'archives. Ce mémoire replace ce projet dans son contexte en proposant d'étudier les types de fonds concernés et leurs traitements, les enjeux actuels de cette valorisation et les organismes concernés par la même problématique. Dans un deuxième temps, l'auteur présente la méthodologie développée lors de son stage pour amorcer le projet. Celle-ci se base sur la sélection puis l'analyse de portails de collections culturelles numériques et une proposition de cahier des charges fonctionnel. Enfin, le mémoire ouvre des perspectives en étudiant quel pourrait être l'apport de la convergence des données, grâce au web de données, pour mieux valoriser les collections du Centre National de la Danse. La réflexion part des atouts du travail déjà existant et étudie les problématiques à résoudre. L'auteur présente ce qu'est le web de données, puis à partir d'exemples de projets actuellement en cours dans le monde de la culture et utilisant des technologies du web de données, elle propose un exemple d'application pour la future bibliothèque numérique du CND.

POLITIQUE CULTURELLE ; MEDIATHEQUE; VALORISATION; ARCHIVES; COLLECTION NUMERIQUE; BIBLIOTHEQUE NUMERIQUE; CONVERGENCE DES DONNEES; WEB DE DONNEES

The multimedia library of French National Center for Dance (CND) decided to develop its own digital library. It aims to add value to its digital collections, mainly originated from its archive founds. This work replaces the project into its context by studying the types of founds concerned and their treatment, the present stakes of the collection promotion and the cultural structures concerned. Then the author presents the methodology developed during her internship to start the project: selecting and analyzing portals of digital cultural collections and writing the functional specifications of the CND digital library. Finally the author opens perspectives on how the web of data can reinforce the promotion of the digital collections of the CND. The work starts from the forces of the actual promotion and examines the potential issues to develop the library. The author presents some key elements

to understand the web of data concept, and, based on current experimentations concerning cultural data in France, she presents an application example of how web of data can serve the CND promotional strategy.

Table des matières

Remerciements	1
Notice	2
Table des matières	4
Liste des figures	9
Introduction	10
Première partie Le contexte de la valorisation des collections numériques	14
1 Quel est le patrimoine culturel concerné par la valorisation sur Internet ?	16
1.1 Les fonds ou collections concernés	16
1.2 La typologie des fonds et leurs traitements documentaires	18
1.2.1 Le traitement documentaire des musées	18
1.2.2 Les archives et le traitement archivistique	19
1.2.2.1 La notion d'archives et le principe d'invulnérabilité	19
1.2.2.2 Zoom sur les archives audiovisuelles et privées	20
1.2.2.3 Le traitement des archives et le respect du fonds	21
1.2.3 Les bibliothèques et la notion de collection	23
2 Pourquoi valoriser des collections culturelles sur Internet ?	25
2.1 Définition du terme « valorisation »	25
2.2 Enjeux pour les organismes culturels	26
2.3 Enjeux politiques	27
2.4 Du côté des utilisateurs	28
2.4.1 Les pratiques numériques et informationnelles du public	28
2.4.1.1 Google comme point d'entrée	29
2.4.2 Les limites de l'utilisation de Google	30

3 Qui est concerné par la valorisation du patrimoine culturel ? 31

3.1 Le ministère de la Culture et de la Communication 31

3.1.1 Les politiques mises en place pour soutenir la valorisation 31

3.1.1.1 Les plans de numérisation nationaux 31

3.1.1.2 Les Services numériques culturels innovants 32

3.1.1.3 Le développement de l'éducation artistique et culturelle 33

3.2 Les organismes culturels 33

3.2.1 Un panorama des offres numériques actuelles 34

3.2.1.1 Les bibliothèques numériques 34

3.2.1.2 Les expositions virtuelles 35

3.2.1.3 Les plateformes communes 35

3.2.1.4 Les projets collaboratifs 35

Deuxième partie Quelle bibliothèque numérique pour valoriser les collections du CND ? 37

1 Contexte actuel 38

1.1 La mission du CND et de sa médiathèque 38

1.2 La genèse de la médiathèque 38

1.3 Les collections 39

1.3.1 La collection principale 39

1.3.2 Les fonds d'archives 40

1.3.3 La collection numérique 41

1.3.3.1 La politique de numérisation 42

1.4 Pourquoi une bibliothèque numérique ? 43

1.4.1 Les axes de valorisation 43

1.4.2 La valorisation actuelle des collections du CND 44

1.4.2.1 La médiation culturelle 44

1.4.2.2 L'éditorialisation 44

1.4.2.3 Les outils en ligne 46

1.4.2.4 Les partenariats 48

1.4.3 Les forces et les faiblesses de cette valorisation 48

1.4.4 Les contraintes à prendre en compte 49

1.4.4.1 Contraintes juridiques et les droits d'auteurs 49

1.4.4.2 La nature des collections elles-mêmes 49

2	<u>La méthodologie utilisée pour définir la future bibliothèque numérique</u>	51
2.1	<u>L'analyse des besoins</u>	51
2.1.1	Les besoins des trois publics cibles	52
2.2	<u>Le benchmarking de collections numériques en ligne</u>	52
2.2.1	La sélection des portails de collections culturelles	53
2.2.2	L'analyse des portails	55
3	<u>Des résultats de l'analyse aux suggestions fonctionnelles</u>	58
3.1	<u>Une petite synthèse de l'analyse comparative</u>	58
3.1.1	Tendances générales	58
3.1.2	Choix relevés comme particulièrement intéressants	59
3.2	<u>Les principales fonctionnalités retenues</u>	60
3.3	<u>Le cahier des charges fonctionnel</u>	60
3.3.1	La navigation	60
3.3.2	Affichage des ressources	61
3.3.3	Visualisation d'une ressource	62
3.3.3.1	Le cas des ressources mises en avant sur la page d'accueil	62
3.3.4	Recherche	62
3.3.5	Ergonomie	63
3.3.6	Accessibilité	63
3.3.7	Services et outils	63
3.4	<u>Points de vigilance et bilan du travail</u>	64
3.4.1	Points de vigilance pour l'interface	64
3.4.2	Bilan du travail	64
	<u>Troisième partie Le web de données, un atout pour la valorisation des données culturelles</u>	66
1	<u>Le web de données : quelques éléments de compréhension</u>	68
1.1	<u>Quel est l'apport du web de données dans la culture ?</u>	68
1.2	<u>Quelques éléments de compréhension</u>	69
1.2.1	Web de données ou web sémantique ?	69

1.2.2	Les 5 degrés d'intégration selon Tim Berners-Lee	70
1.2.3	Techniquement cela se passe comment ?	70
1.3	<u>De quoi a-t-on besoin pour y arriver ?</u>	71
1.3.1	De données interopérables	71
1.3.2	De référentiels communs	72
1.3.3	De qualifier les liens entre les ressources et les métadonnées	72
1.3.3.1	Un exemple de modèle conceptuel : le modèle FRBR	74
2	<u>Exemples d'applications actuelles dans le domaine culturel</u>	77
2.1	<u>Structuration des données et des catalogues</u>	77
2.1.1	data.bnf.fr	77
2.1.2	Le projet OpenCat	78
2.1.3	Le programme HADOC	80
2.2	<u>Les vocabulaires contrôlés et les référentiels</u>	81
2.2.1	MIMO	81
2.2.2	Joconde Lab	82
2.2.3	Outil GINCO	83
2.2.4	Autres projets	84
2.3	<u>Quel bilan pour les premiers projets ?</u>	85
3	<u>Quelle application du web de données pour la bibliothèque numérique du CND ?</u>	87
3.1	<u>Faire converger les collections</u>	87
3.1.1	Les ressources autour du ballet La Bayadère de Noureev	88
3.2	<u>Enrichir les métadonnées</u>	91
3.2.1	L'organisation des métadonnées avec le modèle FRBR	94
3.2.2	La réutilisation des métadonnées	98
3.3	<u>Harmoniser les vocabulaires contrôlés</u>	99
	<u>Conclusion</u>	103
	<u>Bibliographie</u>	108

Comment définir une politique de valorisation de collections numériques issues de fonds d'archives à l'heure du web de données ? Le cas de la médiathèque du CND.

Annexes	117
Annexe 1 Abréviations	118
Annexe 2 Liste des portails et projets étudiés (Partie 2)	119

Liste des figures

Figure 1 - Schéma d'organisation des archives	22
Figure 2 - Typologie de documents de la collection numérique du CND	42
Figure 3 - Lien entre des ressources et de leurs métadonnées	73
Figure 4 - Les Rêves dansants, sur les pas de Pina Bausch	75
Figure 5 - Schéma de la plateforme OpenCat	79
Figure 6 - Illustration des entités FRBR de La Bayadère de Petipa	95
Figure 7 - Illustration des entités FRBR de La Bayadère de Noureev	95
Figure 8 - Illustration de La Bayadère de Noureev avec quelques liens qualifiés	97

Introduction

Pour veiller à la préservation du patrimoine et au partage des connaissances, le Ministère de la Culture et de la Communication soutient activement la numérisation du patrimoine culturel par des plans de numérisation d'envergure depuis 1998. A ce jour, 2886 collections, appartenant à 880 organismes, sont référencées dans le catalogue des collections numérisées. Trois millions de documents provenant d'organismes français sont diffusés sur le portail européen Europeana. Les documents numérisés sont traités de façon structurée pour intégrer les bases de données documentaires des organismes culturels et être ensuite mis à disposition sur Internet. Si elles peuvent être consultées sur Internet, elles ne sont pas indexables par les moteurs de recherche. De son côté, le public utilise majoritairement Internet et les moteurs de recherche web, avec Google en tête, pour découvrir et rechercher des ressources et des informations de toute nature, sans se poser nécessairement la question de qui détient cette information. L'ensemble des données disponibles sur le réseau mondial constitue un immense gisement de connaissances. Ces données sont actuellement non structurées dans leur grande majorité. Dans ce cadre, les producteurs de contenus s'interrogent sur comment être trouvés et visibles sur Internet. Pour les organismes culturels, l'interrogation est déclinée de la façon suivante. Comment un organisme culturel peut-il aider à faire découvrir ses collections ? Comment valoriser ses collections en les mettant à disposition au milieu d'un océan de données ? Comment remplir ses missions de diffusion des connaissances pour des publics aux pratiques en évolution ?

C'est dans ce contexte que s'inscrit la réflexion de la médiathèque du Centre National de la Danse (CND) autour de sa future bibliothèque numérique. Le CND est un établissement public à caractère industriel et commercial (EPIC) créé en 1998 par le Ministère de la Culture et de la Communication (MCC) avec une mission de formations et services aux professionnels de la Danse, de valorisation du patrimoine et de la création artistique. La médiathèque dispose de fonds uniques, d'une grande richesse, constitués d'un important fonds documentaire spécialisé, ainsi que de fonds d'archives autour de la danse. La spécificité du sujet rend les fonds particulièrement originaux : la danse, qui relève des spectacles vivants, est un art éphémère difficilement saisissable.

Afin d'amorcer le projet de bibliothèque numérique du CND, je me suis vu confier la tâche de réfléchir à la forme que pourrait prendre ce nouveau portail alors que la médiathèque cherche à faire évoluer les modalités de valorisation de ses collections. Dans le cadre de mon stage professionnel, j'ai analysé ce qui se faisait actuellement dans le monde de la culture en terme de bibliothèque numérique pour ensuite proposer des fonctionnalités et une structuration de données adaptées à la spécificité de leurs collections numériques et de leur projet. Je suis partie pour mon analyse du travail d'organismes travaillant sur le spectacle vivant, des organismes avec lesquels le CND pourrait être amené à collaborer sur

les questions de valorisation des biens culturels ainsi que des projets menés sous l'impulsion de la Bibliothèque Nationale de France (BNF) et du Ministère de la Culture et de la Communication (MCC).

Le travail conduit dans ce stage et poursuivi dans ce mémoire porte sur l'analyse de la situation au CND et englobe une analyse de l'environnement du monde culturel. Pour cela, j'ai regardé ailleurs ce qui se fait en terme de valorisation des catalogues pour comprendre comment le monde des bibliothèques numériques est en train d'évoluer. J'ai également analysé le traitement documentaire et le traitement spécifique des fonds d'archives pour comprendre la problématique de valorisation des collections numériques suivant une logique documentaire. En effet, de nombreux documents sont issus des fonds d'archives et ont donc été traités suivant une logique archivistique. Par ailleurs, j'ai analysé les projets en cours dans le monde de la culture autour du web de données, projets qui visent à sortir de la logique des silos des bases de données à l'intérieur d'une même organisation et à enrichir le contenu grâce à des données disponibles sur Internet afin de mieux valoriser les collections. Ces problématiques étant également celle du CND, il m'a paru intéressant de regarder comment le CND pourrait intégrer le mouvement de convergence des données pour valoriser ses collections.

Ce mémoire s'inscrit dans la prolongation de cette mission et se base sur le cas concret de la médiathèque du CND. L'objectif est de proposer des pistes de réflexions pour dépasser les questions de fonctionnalités de l'interface et d'organisation des données en étudiant la question de la convergence des données et répondre aux questions suivantes. Quelles sont les enjeux actuels de la valorisation des données culturelles ? Comment la médiathèque du Centre National de la Danse peut valoriser ses collections numériques principalement issues de fonds d'archives ? Quelles sont les tendances actuelles de valorisation des collections numériques du monde culturel qui peuvent être des sources d'inspiration pour cette valorisation ? Comment le web de données, en permettant la convergence des données, peut être un formidable atout de cette valorisation ?

Pour répondre à ses questions j'ai choisi d'écarter la possibilité de réaliser une étude technique autour de la mise en place d'un projet de web de données. J'ai également choisi de ne pas traiter les questions juridiques de la mise à disposition des collections sur Internet. Ce vaste sujet nécessiterait à lui seul de faire l'objet d'une étude approfondie. J'ai choisi de présenter les enjeux de la valorisation des données culturelles et d'ouvrir quelques pistes de réflexion pour ceux qui se posent la question d'aller vers le web de données

Ce mémoire s'organise en trois parties. Nous allons d'abord poser le cadre en replaçant le travail de la médiathèque dans son contexte plus global. Pour cela nous analyserons les types de fonds concernés, les enjeux de la valorisation à l'heure de Google et les politiques

actuelles du MCC et des organismes sous sa tutelle. Nous étudierons ensuite le cas du CND en présentant le travail réalisé lors du stage. Nous approfondirons ce qui se fait en terme de valorisation des collections numériques dans des organismes culturels et les propositions retenues pour le CND afin d'alimenter sa réflexion. Enfin, nous étudierons en quoi le web de données peut être un véritable atout pour résoudre dans les prochaines années certaines difficultés de valorisation des collections numériques. Pour cela nous rappellerons ce qu'est le web de données, présenterons ensuite des exemples d'expérimentation actuellement en cours dans le monde de la culture et enfin imaginerons ce que le CND pourrait faire en terme de valorisation grâce au web de données.

Première partie

Le contexte de la valorisation des collections numériques

Dans cette première partie, nous replaçons le projet de bibliothèque numérique du CND dans le contexte global de la réflexion actuellement en cours autour de la valorisation du patrimoine culturel dans le monde de la culture et ce dans des domaines proches des activités de la médiathèque du CND, notamment avec les bibliothèques, les musées et les archives. Pour cela, nous comparerons les données culturelles concernées par la numérisation et étudierons plus en détail la problématique du traitement des collections numériques issues de fonds d'archives. Puis nous analyserons les enjeux de cette valorisation à différents niveaux : premièrement pour les collections elles-mêmes, deuxièmement pour les organismes, troisièmement au niveau politique et enfin les enjeux liés aux évolutions des pratiques numériques du public. Pour finir, nous regarderons qui est concerné par cette valorisation et présenterons les offres numériques actuelles.

1 Quel est le patrimoine culturel concerné par la valorisation sur Internet ?

La médiathèque du CND propose des collections documentaires, mais aussi des fonds d'archives constitués de fonds d'archives privées et de fonds iconographique, de partitions de danse, de manuscrits.... Elle souhaite mieux valoriser ses collections numériques issues des fonds d'archives. Nous allons comparer comment est traité le patrimoine culturel dans les trois grands types d'organismes culturels concernés par la valorisation de leurs ressources sur Internet : les musées, les archives et les bibliothèques. Nous chercherons à savoir si des expériences de valorisation dans un domaine peuvent profiter aux deux autres. Nous verrons également les problématiques soulevées par la différence entre le traitement documentaire et le traitement archivistique pour valoriser des collections numériques issues de fonds d'archives. Mais commençons déjà par quelques définitions essentielles pour savoir de quoi il s'agit.

1.1 Les fonds ou collections concernés

La numérisation du patrimoine culturel s'inscrit dans les objectifs retenus par la Commission européenne de construction d'une bibliothèque numérique européenne Europeana, point d'accès multilingue à tous les contenus culturels du patrimoine et de la création contemporaine. Sont donc concernés les archives, bibliothèques, musées, services patrimoniaux et l'audiovisuel. La majeure partie des organismes culturels est spécialisée dans un type de ressources mais complétées par d'autres. Un musée d'art, par exemple, gère des œuvres mais a la plupart du temps sa propre bibliothèque et son service d'archives. Le CND ne fait pas exception, puisqu'en plus de sa collection documentaire, il gère d'importants fonds d'archives relatifs à la danse.

Quand nous parlons de fonds d'archives, d'œuvres de musées ou de collections de bibliothèques, il est commun d'y associer l'expression « biens culturels » ou « patrimoine culturel ». Est-ce juste ? Pourquoi cette distinction entre les deux ? Revenons sur quelques définitions, afin de comprendre ce qui est conservé par les organismes culturels, au-delà des objets.

Selon l'Unesco¹ le mot « bien » s'inscrit dans un contexte plus juridique (relatif à la « propriété »), tandis que le terme « patrimoine » renvoie à la conservation et à la transmission de génération en génération. Selon l'article 1 de la Convention de 1970 et l'article 2 du Chapitre 1 de la Convention de 1995 :

« Sont considérés comme biens culturels les biens qui, à titre religieux ou profane, sont désignés par chaque État comme étant d'importance pour l'archéologie, la préhistoire, l'histoire, la littérature, l'art ou la science, et qui appartiennent aux catégories spécifiquement visées dans lesdites Conventions. ».

Jacques GUYOT et Thierry ROLLAND reviennent sur l'évolution historique et rapportent que :

« la notion de patrimoine trouve un nouvel essor lorsque s'exprime la conscience aiguë que la modernité frappe inexorablement des modes de vie, des cultures, des langues mais aussi des objets, dispositifs et machines ainsi que les savoir-faire, métiers qui les avaient construits. Les politiques mises en œuvre par l'Unesco dès sa création en 1945 sont assez symptomatiques de cette prise de conscience et d'une définition de plus en plus extensive, d'un point de vue ethnographique, de ce qui entre dans l'héritage que les sociétés entendent léguer à la postérité. »²

L'Unesco définit le patrimoine culturel selon les catégories suivantes :

- le **patrimoine culturel matériel** qui comprend :
 - le patrimoine culturel mobilier (peintures, sculptures, monnaies, instruments de musiques, armes, manuscrits)
 - le patrimoine culturel immobilier (monuments, sites archéologiques)
 - le patrimoine culturel subaquatique (épaves de navire, ruines et cités enfouies sous les mers)
- le **patrimoine culturel immatériel** : traditions orales, arts du spectacle, rituels

L'Unesco déclare que l'on entend par patrimoine culturel immatériel :

« l'ensemble des pratiques, expressions ou représentations qu'une communauté humaine reconnaît comme faisant partie de son patrimoine dans la mesure où celles-ci procurent à ce groupe humain un sentiment de continuité et d'identité. Ces pratiques concernent principalement, mais de manière non exclusive, les traditions orales, musicales ou chorégraphiques, les langues en tant que supports de ces traditions, les jeux et sports traditionnels, les manifestations festives, les savoir-faire artisanaux, les savoirs et savoir-faire liés à la connaissance de la nature ou de l'univers. »

¹ UNESCO. FAQs. Qu'est-ce qu'un « bien culturel » ? Paris, Février 2008. [Consulté le 18 novembre 2013] <http://portal.unesco.org/culture/fr/ev.php-URL_ID=36292&URL_DO=DO_TOPIC&URL_SECTION=201.html>

² [4] GUYOT p. 20-30

L'Unesco précise en outre que *« l'importance du patrimoine culturel immatériel ne réside pas tant dans la manifestation culturelle elle-même que dans la richesse des connaissances et du savoir-faire qu'il transmet d'une génération à une autre. »*

Ces définitions permettent de comprendre ce qu'il est jugé utile de transmettre aux générations futures et font résonance avec la politique du MCC, comme nous le verrons à la section 2.3. Les fonds et collections numériques relèvent donc à la fois du patrimoine matériel et immatériel.

1.2 La typologie des fonds et leurs traitements documentaires

Après avoir vu les types de biens et patrimoine culturels, voyons les organismes culturels à même de conserver et traiter ces ressources et les différences de traitement de l'information. Avec d'une part les œuvres des musées, les fonds d'archives des centres d'archives et les documents des bibliothèques et centres de ressources. Le but est ici de rappeler les différences principales entre les trois typologies de fonds pour mieux comprendre dans quel contexte sont traitées les ressources à la fois au niveau individuel et au sein d'un ensemble. Nous présenterons ensuite plus en détail le traitement archivistique et aborderons la problématique d'exploitation de collections numériques issues de fonds d'archives.

1.2.1 Le traitement documentaire des musées

Les musées sont une des composantes importantes de l'offre de collections numériques culturelles et nous verrons dans les parties 2 et 3 des exemples de valorisation et de projets intégrant le web de données venant des musées.

Le traitement des œuvres dans les musées est un traitement proche du traitement documentaire, dans le sens où ce qui est décrit est un objet avec son titre, son auteur, son sujet, sa date, etc. Cependant cette description est réalisée par un conservateur à l'issu d'un travail scientifique approfondi par lequel il apporte une connaissance et un éclairage particulier sur l'œuvre elle-même et la collection en général.

Par ailleurs, les musées ont une obligation juridique de traiter l'intégralité de leurs œuvres. Le traitement des collections est encadré par la loi « Musée » qui pose le principe d'inaliénabilité afin de garantir la pérennité des collections artistiques à travers le temps et d'échapper à la variabilité et à la relativité des jugements esthétiques :

« Les biens constituant les collections des musées de France appartenant à une personne publique font partie de leur domaine public et sont, à ce titre, inaliénables. Toute décision de déclassement d'un de ces biens ne peut être prise qu'après avis conforme d'une commission scientifique dont la composition et les modalités de fonctionnement sont fixées par décret »³.

Les musées ne sont donc pas libres de se séparer d'une partie de leurs collections ni d'en laisser une partie en friche. C'est là encore une des différences avec le traitement documentaire. Nous allons maintenant aborder le traitement archivistique pour lequel on retrouve la volonté d'échapper à la variabilité et relativité des jugements.

1.2.2 Les archives et le traitement archivistique

1.2.2.1 La notion d'archives et le principe d'inviolabilité

Pour comprendre ce que sont les archives, commençons par définir le terme en voyant ce que dit la loi du 3 janvier 1979. L'art. 1^{er} stipule :

« les archives sont l'ensemble des documents, quels que soient leur date, leur forme et leur support matériel, produits ou reçus par toute personne physique ou morale, et par tout service ou organisme public ou privé dans l'exercice de leur activité ».

Les articles suivants définissent les archives selon leur nature. Ainsi L'art. 3 précisent que :

« les archives publiques sont : 1/ Les documents qui procèdent de l'activité de l'État , des collectivités locales des établissements et entreprises publiques ; 2/ Les documents qui procèdent de l'activité des organismes de droit privé chargés de la gestion des services publics ou d'une mission de service public ; 3/ Les minutes et répertoires des officiers publics ou ministériels... ».

L'art. 9 définit les archives privées comme « l'ensemble des documents définis à l'art 1^{er} qui n'entrent pas dans le champ d'application de l'article 3 ci-dessus ».

La particularité d'un fonds d'archives est son lien avec une personne physique ou morale et son critère de constitution est la provenance comme le rappelle le code du patrimoine à l'article L.451-5 :

« Un fonds d'archives, exclusivement organisé en fonction de son origine s'oppose à la collection dont le classement est fonction de ses buts, de la commodité des utilisateurs ou des nécessités matérielles. »

³ Code du Patrimoine, article L.451-5

L'élément essentiel de tout dépôt d'archives est le fonds, dont le respect est l'une des lois de l'archivistique moderne. La première étape pour aborder un fonds d'archives est de connaître l'histoire, le fonctionnement et l'activité de l'institution dont il émane. Selon Jean FAVIER,

« le fonds doit respecter l'individualité et la composition, c'est le seul moyen d'éviter l'emploi de critères de classement qui pourraient varier selon l'époque, la mode ou les nécessités et de pouvoir utiliser un fonds d'archives en l'absence d'inventaire ou de répertoire exhaustif. »⁴

1.2.2.2 Zoom sur les archives audiovisuelles et privées

Souvent, quand on parle d'archives, on pense aux archives d'Etat ou départementales, pourtant les types de fonds d'archives sont multiples. Parmi ceux-ci, nous avons choisi d'en étudier deux plus en détail du fait de leur place particulière pour retracer les mentalités et l'évolution de la société. Il s'agit des archives privées et des archives audiovisuelles. Nous nous intéressons à ces fonds également parce que les fonds d'archives du CND sont constitués en grande partie de fonds d'archives privées et de nombreuses ressources audiovisuelles. Notons que les archives privées et les archives audiovisuelles ne s'opposent pas : les archives privées pouvant être (partiellement) audiovisuelles.

Les archives audiovisuelles

Les archives audiovisuelles constituent une part du patrimoine audiovisuel avec les productions de son enregistré, productions radiophoniques, productions cinématographiques productions télévisuelles ou autres comprenant des images en mouvement et/ou des sons enregistrés.

Jacques GUYOT et Thierry ROLLAND⁵ nous rappellent que les archives audiovisuelles sont utiles à des fins historiques et commerciales. Pour eux :

« La notion d'archives audiovisuelles repose sur un paradoxe. Au-delà d'une consommation de masse, ces produits prolifiques, a priori conçus comme des artefacts éphémères, fragiles également, constituent néanmoins des ressources précieuses à plus d'un titre. Elles intéressent les historiens en donnant à voir des grands moments de l'histoire collective. Les chercheurs en sciences humaines et sociales peuvent y tracer l'évolution des modes de vie, des comportements sociaux, des débats publics et controverses agitant la société ou nos imaginaires. Ce sont des indicateurs essentiels, non seulement pour les chercheurs, mais aussi pour les cinéastes, vidéastes,

⁴ [3] FAVIER. p. 3-6

⁵ [4] GUYOT. p. 3-16

documentaristes ou journalistes qui cherchent à replacer certains signes fugaces souvent saisis sur le vif dans l'histoire des mentalités. »

Dans le domaine du spectacle vivant et de la danse en particulier, les archives audiovisuelles nous paraissent essentielles pour rendre compte de l'évolution des pratiques et fixer ce qui est éphémère. Les archives audiovisuelles sont dans ce sens complémentaires des archives privées qui permettent également de rendre compte de l'histoire sociale et des mentalités.

Les archives privées

L'attention des historiens va bien au-delà des figures marquantes de l'Histoire en s'intéressant à l'ensemble d'une population. L'histoire d'une famille, d'une localité débouche sur l'histoire sociale, celles des structures économiques, des mentalités... Les archives du secteur associatif témoignent d'un phénomène à l'extension très large et apportent une contribution importante à l'histoire sociale. Les archives privées complètent nécessairement les archives publiques « *dans la mesure où elles permettent souvent de rectifier une vision du passé qui sans elles serait exclusivement administrative.* »⁶ Notons également que beaucoup de fonds privés contiennent des archives publiques en lien avec les fonctions occupées par leurs producteurs qui ont tenu à conserver leurs papiers de fonction. Ainsi la frontière entre le privé et le public est souvent aussi difficile à déterminer qu'à faire respecter. Concernant la place des archives privées, on peut dire qu'elles constituent une source historique représentative et une part importante du patrimoine national.

Par ailleurs la loi s'intéresse aux archives privées dans un esprit de sauvegarde en cherchant à concilier le respect de la propriété privée et l'intérêt général, comme nous le rappelle Françoise HILDESHEIMER⁷.

Concrètement, les archives privées peuvent être par exemple, des archives personnelles, des archives d'artistes, de scientifiques ou d'associations. Elles peuvent être constituées de documents administratifs, de communication, des correspondances privées, mais aussi de supports audiovisuels constitués d'enregistrements parfois conservés ailleurs dans des archives audiovisuelles comme celles de l'Ina ou des productions individuelles.

1.2.2.3 Le traitement des archives et le respect du fonds

Nous allons rappeler maintenant les bases du traitement des archives dans le respect du fonds afin de prendre la pleine mesure de la différence avec le traitement documentaire. Pour cela, nous nous basons sur ce que nous dit Françoise HILDESHEIMER⁸.

⁶ [5] HILDESHEIMER. p. 15-18

⁷ [5] HILDESHEIMER. p. 5-13

Le but de l'organisation des archives est de maintenir un fonds cohérent qui permet de refléter la logique interne et la diversité de l'action menée par un homme, une famille, un organisme ou une association. Un fonds d'archives se définit comme un ensemble de documents dont l'accroissement s'effectue de manière organique dans l'exercice des activités d'une personne physique ou morale. En vertu du principe de respect des fonds, les archivistes conservent intacts ces ensembles, sans les démembrer ni les mélanger entre eux. Autrement dit, chaque document doit être placé dans le fonds dont il provient et dans ce fonds, à sa place originelle.

En archivistique, la pièce d'archives ne prend sa pleine valeur qu'intégrée à un ensemble qui l'éclaire et l'authentifie. Le traitement interne des dossiers, c'est-à-dire le classement des pièces qui le constitue, est la première étape. L'ordre est le plus souvent chronologique. Certains dossiers complexes sont divisés en sous-dossiers précisément identifiés (par exemple les papiers de fonctions seront classés dans l'ordre chronologique des fonctions exercées). Les dossiers sont ensuite organisés suivant un plan de classement. On respectera toujours un classement ancien et le cas échéant, on en rétablira l'ordre s'il a été détruit.

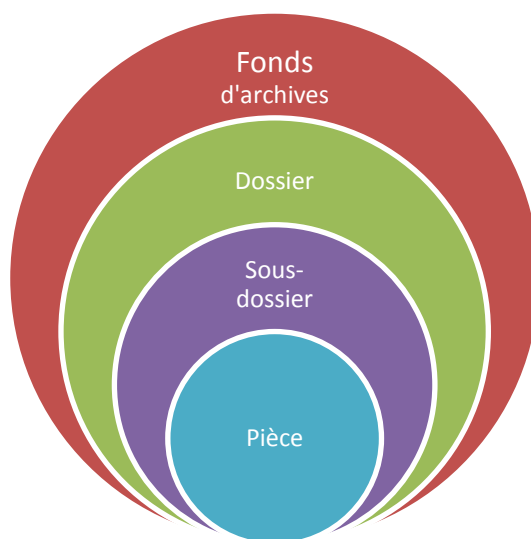


Figure 1 - Schéma d'organisation des archives

Un fonds est décrit dans un inventaire selon cette logique et en respectant une structure particulière décrite par la norme internationale EAD (Encoded Archival Description). Le choix de la profondeur de la description est variable, les pièces pouvant être décrites seulement au niveau d'un sous-dossier. La description des pièces suit une logique d'héritage, les informations communes sont décrites au niveau le plus haut suivant la granularité du fonds.

⁸ [5] HILDESHEIMER. p. 39-52

Pour résumer nous pouvons reprendre la phrase de Françoise HILDESHEIMER

« *Qui dit fonds entend donc provenance, ce qui oppose fondamentalement les archives à une notion voisine : la collection.* » Nous allons justement aborder maintenant la notion de collection.

1.2.3 Les bibliothèques et la notion de collection

Nous allons présenter le travail documentaire pour le comparer au traitement documentaire des musées et le traitement archivistique. Nous n'allons pas distinguer ici le travail des bibliothèques et des centres de ressources, mais englober les deux comme un tout. Le travail documentaire n'est pas encadré par la loi en dehors du dépôt légal des livres, périodiques, affiches, productions audiovisuelles et dernièrement du web confié en France à la BNF et l'INA. Le critère de regroupement des ressources dans le monde documentaire suit la logique de la pertinence des ressources par rapport à un public que ce soit pour la littérature blanche ou la littérature grise, les ressources pouvant être des écrits mais aussi des ressources audiovisuelles, iconographiques, etc... Les ressources rassemblées sont organisées en collections.

Qu'entend-on par collection ? Françoise HILDESHEIMER définit une collection « *comme un ensemble de documents de provenances diverses dont la réunion artificielle est le résultat du choix ou du hasard* » et nous donne l'exemple suivant : « *les dossiers de documentation contenant des documents de toutes provenances regroupés thématiquement relèvent de la notion de collection.* » Elle souligne la logique d'organisation *thématique* dans le domaine documentaire que l'on retrouve dans les plans de classement.

La description des ressources se fait document par document pour en préciser au minimum le titre, l'auteur, la date, son appartenance à un ensemble plus grand (collection pour un livre par exemple, périodique pour un article), le sujet et une présentation souvent sous la forme d'un résumé ou d'une description, le support et la technique de réalisation (pour les ressources iconographiques ou audiovisuelles notamment). Le travail de description du documentaliste n'est pas le fruit d'un travail de recherche approfondie. La valeur ajoutée du documentaliste ne réside pas tant dans la description que dans la faculté à replacer la ressource au sein des collections, la capacité à intégrer les ressources pertinentes pour compléter l'ensemble et faciliter le travail de médiation avec le public.

En résumé, nous pouvons dire que le travail documentaire dans les musées et les centres de ressources produisent des descriptions précises de chaque ressource qui sont structurées de façon proche. Les personnes sont libres d'organiser leurs collections suivant des critères qui

leur sont propres alors que les archives travaillent au niveau d'un fonds dont l'organisation doit être respectée. Dans les musées et les archives, le choix des ressources et la façon de les traiter sont encadrés par la loi. La description des collections muséales et des fonds d'archives est le fruit d'un travail de recherche approfondi.

Les points communs entre le traitement documentaire dans les musées et les centres de ressources sont importants. Cela permet de dire qu'il est pertinent de regarder ce qui se fait dans ces deux types d'organismes en terme de valorisation de collection numérique pour s'inspirer les uns les autres et échanger en terme de bonne pratique.

A l'inverse les différences entre le traitement documentaire et le traitement archivistique permettent de comprendre en quoi valoriser des collections numériques issues de fonds d'archives n'est pas simple. Les descriptions doivent être reprises depuis différents niveaux du fait du principe d'héritage du traitement archivistique. Les ressources risquent d'être insuffisamment documentées en l'état pour être exploitées efficacement. Elles devront être organisées selon une nouvelle logique, en dehors du fonds. Le fait même de présenter des pièces issues de fonds d'archives fait débat chez les archivistes. Certains pensent qu'en présentant une pièce comme un document individuel s'intégrant dans une collection organisée indépendamment de la logique du fonds, on dénature le fonds et on ne permet pas à l'utilisateur de replacer le document dans son contexte et donc de le comprendre.

Maintenant que nous avons vu ces trois grandes catégories de ressources culturelles, voyons maintenant les raisons pour lesquelles on cherche à les valoriser et quels sont les enjeux d'une telle valorisation.

2 Pourquoi valoriser des collections culturelles sur Internet ?

La première réponse qui vient à l'esprit quand on pense à cette question est « parce que Internet est la première source d'information utilisée ». Mais quel est le sens du terme « valorisation » ? Quelles sont les valeurs associées à l'exploitation du patrimoine culturel ? Quels sont les enjeux de cette valorisation pour les organismes culturels et au niveau politique ? Après avoir répondu à ces questions, nous rappellerons les pratiques numériques et informationnelles du public et les changements induits par les moteurs de recherche web et notamment Google.

2.1 Définition du terme « valorisation »

Valoriser peut être défini comme l'exploitation d'un document ou d'une ressource dans le but d'en accroître ses valeurs. Mais quelles sont justement les valeurs associées aux ressources culturelles, notamment les archives ? Dans le livre collectif « La valorisation des archives. Une mission, des motivations, des modalités, des collaborations. Enjeux et pratiques actuels »⁹ les auteurs reviennent sur le choix du terme valorisation et les notions de valeurs en archivistique. Regardons ce qu'ils nous en disent.

Les archives ont des valeurs multiples, qui se complètent et se superposent. La valeur probatoire vient du « témoignage de vérités historiques », alors que la valeur d'usage est liée aux rôles des archives dans les organisations qui les produisent et conservent. La distinction peut également se faire entre la valeur primaire et la valeur secondaire. La valeur primaire du document correspond à sa raison d'être au départ, elle évolue tout au long de son cycle de vie de sa création jusqu'à ce que les besoins pour lesquels il a été créé s'achèvent. La valeur secondaire dépend de ses « *potentialités à documenter d'autres activités pour le producteur et pour la société en général* ». La valeur d'existence trouve écho dans la dimension patrimoniale. Le patrimoine se définit par sa valeur d'exemplarité, d'identité ou de remémoration. Pour Xavier Greffe « *un objet devient patrimoine parce qu'il est lourd de sens pour une collectivité et que son existence devient le moyen de symboliser une histoire ou de faire partager des valeurs.* »¹⁰ La valeur d'existence résulte d'un processus de charge symbolique fondée sur des significations que le patrimoine incarne.

⁹ [6] HIRAUX

¹⁰ GREFFE, Xavier. La valorisation économique du patrimoine. Paris, La Documentation française, 2003. 278 p. ISBN 2110942746

La multiplicité des valeurs associées aux archives met en évidence la nature complexe des activités de mise en valeur. Au-delà de ces définitions théoriques, quels sont concrètement les enjeux de la valorisation en terme de politique publique et pour les organismes culturels ?

2.2 Enjeux pour les organismes culturels

Pour les organismes culturels, la valorisation des collections présente d'abord des enjeux de visibilité et d'accessibilité des collections elles-mêmes, afin de satisfaire les besoins du public. Cependant, les enjeux vont au-delà de l'accessibilité aux collections numériques et sont aussi liés à la visibilité de l'organisme lui-même. La numérisation des fonds puis leur mise à disposition est la première étape de cette valorisation. Elle vise un double objectif : préserver les ressources papier ou analogiques difficilement accessibles, souvent fragiles et les faire voir.

Exploiter les documents numérisés permet d'éviter les risques d'altération, de dégradation et de vols des « documents physiques ». Par ailleurs, cela permet de les rendre accessibles sans temps de préparation et de mise à disposition par le personnel et sans contrainte de visites sur place pour le public. Encore faut-il que le public trouve les collections numériques.

Aujourd'hui les ressources numériques accessibles sont « nombreuses, voir pléthoriques et les principales difficultés sont dans le repérage et la sélection » de celles-ci, commentent Jean-Philippe ACCART et Alexis RIVIER¹¹. Dans ce cadre, les organismes culturels cherchent à garantir la visibilité et l'accessibilité de leurs collections. Ils souhaitent proposer des interfaces lisibles pour que leurs collections soient vues, consultées et exploitées. Cela passe nécessairement par une adaptation aux pratiques du public et à ses outils, et ainsi aux moteurs de recherche web pour rendre leurs ressources indexables (cf. section 2.4.1). La valorisation de collections numériques permet également de valoriser le travail réalisé par les équipes en charge des collections.

D'autre part, la visibilité des collections et celle de l'organisme sont liées. Alors que la grande richesse du patrimoine culturel en France se traduit par un foisonnement des organismes culturels, de dimensions variables en terme de notoriété, taille, budget, public cible et rayonnement géographique, les organismes ont besoin d'être connus et reconnus. Ainsi, il est important pour un organisme de renforcer son rayonnement, sa légitimité et sa notoriété aux yeux de son public et de ses partenaires potentiels, voir au-delà. Renforcer la position de l'organisme peut permettre l'existence, la pérennité et le développement de ses actions

¹¹ [13]. ACCART

au-delà de la valorisation des collections et être un facteur de développement de partenariats alors que les budgets sont soumis à restriction. Cela permet en outre d'attirer le public et de renforcer sa crédibilité comme acteur clé de son domaine.

En conclusion, nous pouvons souligner que la visibilité de collections numériques et celle de l'organisme se renforcent l'une l'autre. En effet, se rendre visible par des collections numériques bien valorisées contribue à renforcer la position de l'organisme, qui en retour renforce la visibilité des collections. Les choix stratégiques des organismes culturels s'intègrent dans la politique culturelle de la France avec en premier lieu celle du Ministère de la Culture dont la plupart des organismes dépendent. Nous allons maintenant rappeler quels sont les enjeux actuels au niveau politique.

2.3 Enjeux politiques

La numérisation du patrimoine culturel est l'une des priorités depuis plus de 10 ans, notamment de la Présidence du Conseil de l'Union européenne dans le domaine culturel. En 2008 lors du lancement du projet Europeana (la plateforme européenne de diffusion de collections numériques culturelles), Christine ALBANEL, Ministre de la Culture et de la Communication sous le précédent gouvernement rappelle ainsi les principaux enjeux¹² :

« Une ère nouvelle s'ouvre pour la culture et pour ses publics, et elle s'ouvre à l'échelle du monde. L'Europe doit saisir cette chance pour mieux promouvoir sa diversité culturelle et linguistique à l'international, mais aussi pour encourager sa démocratisation au sein de ses propres frontières, notamment auprès des nouvelles générations qui ont grandi avec Internet. »

La valorisation des collections numériques s'inscrit pleinement dans la politique culturelle au niveau européen et français avec notamment :

- la démocratisation des connaissances en rendant accessible le patrimoine culturel et améliorant le partage des connaissances,
- la préservation du patrimoine culturel pour promouvoir la diversité culturelle,
- le développement de la création artistique, incluant les nouveaux usages numériques,
- l'adaptation aux nouvelles pratiques numériques, notamment pour éviter une fracture générationnelle avec les « générations 100% web ».

Cette politique participe à la stratégie de l'ouverture des données du Ministère de la Culture et de la Communication et constituant l'un des fondements d'un service public culturel

¹² [20] MCC. p. 5

numérique. Dans sa feuille de route, le ministère précise les « *10 actions pour une stratégie numérique de diffusion et de réutilisation des données publiques numériques du secteur culture* » dont le développement des technologies de réutilisation et de sémantique¹³. L'ouverture des données culturelles s'inscrit dans la politique de partage et de la diffusion des connaissances. Le mouvement des « Données ouvertes », expression française de l'*Open data*, concerne tous les domaines publics en particulier le monde scientifique et les administrations.

Au-delà du partage des connaissances et de la volonté de transparence, l'ouverture des données vise également à mettre à disposition de tout un chacun des jeux de données déjà existants afin de participer au développement économique. Les données disponibles peuvent être utilisées et valorisées par des personnes innovantes pour des usages et des publics auxquels les personnes qui géraient initialement ces données n'auraient pas su ou pu répondre. La valeur créée pourrait ensuite être source de développement économique. L'ouverture des données est liée au développement du *Big data*, parfois appelé « données massives » en français. Le *Big data* fait référence à l'analyse de données pour en extraire de la valeur, les données pouvant être privées.

Avant de regarder comment le MCC répond à ces enjeux (cf. section 3.1), voyons d'abord plus en détail ce qui a évolué du côté des pratiques informationnelles et numériques du public avec le développement d'Internet.

2.4 Du côté des utilisateurs

Pour satisfaire les besoins des utilisateurs et être sur leur chemin, il est important de connaître leurs pratiques numériques et informationnelles.

La numérisation du patrimoine culturel répond aux besoins d'un large public pour des usages variés de découverte et de connaissance du patrimoine culturel ou des usages de loisirs et de pratiques amateurs, mais aussi aux demandes de professionnels pour des usages de recherche, de connaissance scientifique ou encore des usages spécifiques concernant le tourisme, l'éducation artistique et culturelle ou l'édition.

2.4.1 Les pratiques numériques et informationnelles du public

En vingt ans, les technologies de l'information et de la communication (TIC) qui fournissent progressivement un mode d'accès élargi aux biens et services culturels sont devenues partie

¹³ [10] MCC

intégrante de la vie quotidienne des ménages. Selon l'observateur du numérique, en France 80 % des ménages sont connectés à l'internet à domicile en 2012. Même si la plus grande partie des foyers est connectée à Internet, il subsiste des inégalités selon l'âge et les niveaux de diplômes et de revenus notamment. Si 65 % des internautes utilisent Internet chaque jour, les cadres, les diplômés d'études supérieures et les étudiants sont plus de neuf sur dix à utiliser l'internet chaque jour. Dans l'Union Européenne, 70 % des particuliers utilisent régulièrement Internet, en se connectant une fois par semaine. Ce taux varie significativement selon l'âge des utilisateurs : de 93 % des jeunes de 16 à 24 ans, à 78 % des particuliers âgés de 25 à 54 ans et à 42 % des particuliers âgés de 55 à 74 ans.

2.4.1.1 Google comme point d'entrée

L'accès à Internet se fait par des moteurs de recherche web, dont le plus répandu est Google. Selon l'étude réalisée par ComScore en France en 2012, Google maintient de loin une position prédominante sur le marché des moteurs de recherche en France avec 94% des requêtes effectuées. Les autres moteurs de recherche ne dépassent pas les 2%¹⁴.

Selon une étude réalisée en avril 2013, environ la moitié des visites sur un site internet ont été effectuées depuis un moteur de recherche, 38% grâce à un accès direct (saisie de l'adresse dans le navigateur ou clic sur un favori) et le reste grâce à des liens externes (email, réseaux sociaux, forums, sites d'actualités, publicités...). On note depuis plusieurs années une baisse de l'accès direct. L'accès par liens externes devient plus important tout en restant faible. Une autre tendance est l'augmentation de l'utilisation de sites spécialisés pour effectuer des requêtes en ligne, même si elle reste marginale. On peut considérer que cet usage vient en complément des moteurs de recherche généralistes pour des demandes précises.

Les tendances chiffrées mentionnées ne comparent par la recherche d'information sur Internet avec la recherche d'information en général. Cependant, on peut affirmer que la première étape pour la recherche d'information ou de ressources passe par Internet via un moteur de recherche, le plus souvent Google. Lors d'une recherche, les utilisateurs ne se posent pas nécessairement la question de qui est susceptible de détenir l'information ou la ressource dont ils ont besoin. Ils trouvent donc en premier lieu l'information qui est sur leur chemin, c'est-à-dire dans les résultats de recherche. Jean-Philippe ACCART et Alexis RIVIER parlent du « syndrome Google » pour désigner les changements d'habitudes des utilisateurs du fait de l'influence de l'usage du moteur de recherche de l'entreprise californienne. Ils le résumant de la façon suivante : *« l'utilisateur veut tout et tout de suite : il attend des*

¹⁴ [14] comScore

réponses sans chercher dans les index et sans combiner les moteurs de recherche et sans se déplacer.» Si certaines personnes, notamment les chercheurs, consultent les bases de données spécialisées, on observe que l'utilisation de la recherche simple avec des requêtes basiques devance largement les recherches avancées. Les pratiques tendent à se rapprocher des recherches « à la Google », avec l'utilisation de facettes pour réduire les résultats. L'utilisation d'Internet et des bases de données spécialisées peut être complétée par la consultation d'un centre de ressources spécialisé.

2.4.2 Les limites de l'utilisation de Google

Le moteur de recherche Google se consacre exclusivement à l'information numérique disponible sur Internet. Or, il y a actuellement encore beaucoup d'informations non disponibles sur Internet. La non-accessibilité peut être due à la coexistence d'information papier ou analogique¹⁵ et numérique, aux limites des moteurs de recherche pour atteindre le web profond ou au fait que les données ne sont pas librement accessibles. Le classement des résultats par les moteurs de recherche met en avant les informations les plus fraîches. Par ailleurs, le résultat dépend de la compréhension de la question posée par la machine. La bonne compréhension dépend de la détection du contexte, de l'adéquation entre les termes employés par l'utilisateur et les termes référencés pour chaque ressource, termes provenant du contenu lui-même et des métadonnées et ainsi que de la non ambiguïté des termes. Nous verrons dans la troisième partie comment le web de données peut contribuer à améliorer la pertinence des résultats.

En conclusion, la valorisation des collections sur Internet a pour but de favoriser le partage des connaissances et de rendre visible des collections peu accessibles. Elle permet également de contribuer positivement à l'image d'un organisme et contribue ainsi à asseoir sa position et en retour à renforcer la visibilité des collections. Les enjeux dépassent le contexte d'un organisme pour rejoindre des enjeux politiques de démocratisation des connaissances, de préservation du patrimoine culturel, de développement de la création artistique et d'adaptation aux nouvelles pratiques numériques.

¹⁵ Voir d'objets qui peuvent être seulement représentés sur Internet comme les pièces de musée, les costumes, les décors, etc.

3 Qui est concerné par la valorisation du patrimoine culturel ?

Tous les organismes culturels sont potentiellement ou concrètement concernés par la valorisation du patrimoine culturel, avec en premier lieu le MCC. Nous allons présenter la politique qu'il mène dans ce domaine. Nous regarderons ensuite les organismes sous sa tutelle, notamment ceux en lien avec les musées, le spectacle vivant et les bibliothèques puis nous montrerons les tendances concernant les offres numériques actuelles.

3.1 Le ministère de la Culture et de la Communication

Le ministère a pour mission de « *rendre accessibles au plus grand nombre les œuvres capitales de l'humanité, et d'abord de la France* »¹⁶. La valorisation des collections numériques s'inscrit dans la politique du MCC à plusieurs titres. Parmi ses missions, nous pouvons citer sa politique de sauvegarde, de protection et de mise en valeur du patrimoine culturel « *dans toutes ses composantes* ». Il favorise la création des œuvres de l'art et de l'esprit et le développement des pratiques et des enseignements artistiques. Il contribue également avec d'autres ministères au développement de l'éducation artistique et culturelle des enfants et des jeunes adultes. Il concourt au développement des nouvelles technologies de diffusion de la création et du patrimoine culturels. Il met en œuvre, conjointement avec les autres ministères intéressés, les actions de l'État destinées à assurer le rayonnement dans le monde de la culture et de la création artistique françaises et de la francophonie.

Nous allons concrètement voir quelques exemples de politiques actuelles du ministère qui concerne le CND à travers ses missions.

3.1.1 Les politiques mises en place pour soutenir la valorisation

3.1.1.1 Les plans de numérisation nationaux

Lancé en 1996, le plan national de numérisation portait sur la numérisation des fonds iconographiques et sonores appartenant à l'État (archives nationales, bibliothèques, services de l'archéologie, monuments historiques, musées etc.). À partir de 2000, il s'est élargi aux fonds appartenant aux collectivités locales, aux fondations et associations. Depuis les années 2010, il s'est ouvert à la création contemporaine.

¹⁶ Malraux, André. Le décret fondateur du MCC. Paris. 1959

Chaque année, un appel à projet présente les priorités. En 2013 le ministère de la Culture et de la Communication, a porté une attention particulière aux éléments suivants :

- Le développement de réseaux d'opérateurs autour de contenus thématiques ou régionaux,
- l'émergence d'actions contribuant à l'éducation artistique et culturelle.

L'appel de 2013 est structuré autour de six programmes thématiques, dont l'art, l'archéologie et la création contemporaine.

Au-delà de la numérisation, le ministère soutient également l'exploitation des documents numérisés en lien avec les services numériques culturels innovants.

3.1.1.2 Les Services numériques culturels innovants

Le ministère de la Culture et de la Communication soutient depuis 2010 l'innovation numérique pour développer de nouveaux usages culturels numériques pour le grand public. Les services numériques expérimentés dans ce cadre s'appuient sur des contenus numérisés disponibles.

Le MCC définit les principaux objectifs de cette politique de la façon suivante :

- Soutenir l'innovation en matière d'usages numériques et faciliter leur expérimentation dans les institutions culturelles ou sur internet
- Promouvoir de nouvelles approches de consultation et de navigation
- Stimuler la réutilisation des ressources numériques culturelles pour tous les publics
- Intégrer des contenus et des technologies disponibles pour créer des services innovants
- Encourager de nouveaux partenariats entre opérateurs culturels, monde de la recherche et entreprises. Les expérimentations peuvent également émaner d'artistes, d'inventeurs, jeunes pousses, designers, etc...

L'importance de la valorisation du patrimoine culturel apparaît clairement comme un de ses objectifs. Actuellement une soixantaine d'expérimentations sont soutenues. Elles concernent les champs suivants¹⁷ :

- Services pour la visite, la découverte et la connaissance de musées, d'œuvres, de sites, de monuments, de territoires,

¹⁷ [12] MCC

- Services collaboratifs de visualisation et de diffusion de contenus,
- Services ludiques sur contenus culturels
- Services numériques pour l'éducation artistique et culturelle,
- Services de personnalisation (familles, enfants, groupes, multilinguisme, e-accessibilité, e-inclusion,...),
- Services de médiation numérique pour le spectacle vivant.

Parmi les types de services retenus dans l'appel à projets, nous voyons quels types de projets pourraient intéresser les organismes comme le CND avec par exemple les services de médiation numérique pour le spectacle vivant et les services numériques pour l'éducation artistique et culturelle¹⁸. Cette dernière est également une des priorités du ministère.

3.1.1.3 Le développement de l'éducation artistique et culturelle

Le développement de l'éducation artistique et culturelle est soutenu par le Ministère de la Culture et de la Communication et celui de l'Éducation nationale. Les objectifs sont l'acquisition de connaissances, la pratique et les rencontres avec les œuvres et les artistes à destination des enfants et des jeunes adultes.

A travers les politiques mises en place, le MCC soutient les actions des organismes culturels placés sous sa tutelle et que nous présentons maintenant.

3.2 Les organismes culturels

De nombreux organismes culturels sont sous tutelle du MCC, qui leur confie des missions spécifiques. Nous n'allons pas donner une cartographie précise de ces organismes, nombreux sont ceux qui sont rattachés à plusieurs ministères ou qui concernent à la fois la pédagogie, le développement artistique et la valorisation du patrimoine, nous pouvons déjà en citer les principales catégories :

- Architecture et patrimoine
- Arts plastiques
- Cinéma
- Livre et lecture
- Musées

¹⁸ [11] MCC

- Spectacle vivant

Le spectacle vivant regroupe le théâtre, la danse, l'opéra, le cirque, les arts de rue, la marionnette et les arts du geste. Parmi ces catégories, nous allons nous intéresser plus particulièrement aux musées, aux spectacles vivants et à ce qui touche les livres et la lecture avec les bibliothèques. Ces organismes ont pour points communs avec le CND le sujet (le spectacle vivant) et la proximité dans les missions et les pratiques documentaires (cf. section 1.1).

On peut noter que les organismes sous tutelle du MCC sont soit des lieux de visites ou consultation (comme les musées et bibliothèques), des lieux de pratique et de création artistique (par exemple le Conservatoire national supérieur d'art dramatique), des lieux d'enseignement (par exemple l'École nationale supérieure de la photographie) ou des centres nationaux (par exemple le Centre National de la Danse). Le classement selon ces domaines d'activités principaux ne signifie pas réduire ces organismes à une seule activité.

3.2.1 Un panorama des offres numériques actuelles

Nous allons maintenant présenter les principales offres numériques mises en place par les organismes culturels. Notons que ces modèles ne s'opposent pas, ils peuvent se combiner pour former une offre de service numérique spécifique à chaque projet. Nous parlerons de cette offre en employant le terme générique de « portail ».

3.2.1.1 Les bibliothèques numériques

Le terme de « bibliothèque numérique » est couramment utilisé pour désigner des outils très variés mis à disposition sur Internet pour le public. Nous l'employons ici dans le sens d'un outil permettant de mettre à disposition des collections numériques sur Internet ou un Intranet de façon organisée en vue de les mettre en valeur. Par comparaison avec les catalogues en ligne, dans une bibliothèque numérique, l'utilisateur accède au document numérique lui-même et non à sa seule notice. Par ailleurs l'interface est plus recherchée en proposant souvent différents types de navigation et pas seulement une arborescence et une recherche. Le terme « bibliothèque » s'entend au sens de rassemblement de documents sous forme de collections, les documents étant en général de différentes natures et couvrant le sujet de façon étendue. De nombreux organismes en dehors des bibliothèques en proposent sous diverses appellations. Nous pouvons par exemple citer la bibliothèque numérique de l'Institut National de l'Histoire des Arts (INHA), le portail de la marionnette, Gallica de la BNF...

3.2.1.2 Les expositions virtuelles

Les expositions virtuelles se distinguent des bibliothèques numériques par leur objectif d'aborder un thème suivant une problématique définie en proposant une sélection de documents ou d'objets. La scénographie et les textes y sont plus élaborés, afin de permettre de répondre à la problématique posée. Une exposition virtuelle peut être réalisée en écho à une exposition « réelle » comme à la BNF ou au Louvre par exemple. Elle peut aussi être indépendante de la programmation culturelle sur place, tel le Patrimoine numérisé de la ville de Besançon¹⁹.

3.2.1.3 Les plateformes communes

Les bibliothèques numériques peuvent être des plateformes communes permettant de mutualiser le développement d'un service, en rassemblant des collections qui se complètent dans une volonté de convergence. En atteignant une taille critique de documents, le portail gagne en visibilité et attractivité. Les plateformes communes peuvent correspondre à un regroupement de collections éclatées entre plusieurs acteurs sur un domaine précis, c'est par exemple le cas du Portail des spectacles et musiques du monde. A contrario, la plateforme peut aussi être un point d'accès central à des collections très diverses, provenant d'acteurs multiples, comme Gallica ou Europeana.

3.2.1.4 Les projets collaboratifs

Dans un projet collaboratif, comme son nom l'indique, l'organisme cherche à impliquer le public en lui permettant de collaborer au contenu du portail. La collaboration peut porter sur l'ajout de documents numériques ou sur les descriptions des documents numérisés (on parle de *crowdsourcing*). Le porteur du projet peut choisir de laisser le contributeur libre de publier ses contenus a priori et de les modifier si besoin par la suite ou au contraire de proposer à l'utilisateur de faire des suggestions qui seront intégrées ou non. Nous pouvons citer l'exemple de la Bibliothèque municipale de Lyon avec son portail Numelyo.

La collaboration avec le public peut également se faire hors du portail via les réseaux sociaux. Si un organisme peut utiliser les réseaux sociaux comme outils de diffusion autour de ses collections et ses projets, il peut aussi permettre aux utilisateurs de créer ses propres collections et de les partager. C'est ce que propose notamment l'équipe de Gallica via Pinterest.

En conclusion, la tendance actuelle est à des portails numériques qui permettent de proposer au plus grand nombre un catalogue des collections avec un accès aux documents

¹⁹ Lien : <http://culture.besancon.fr/?id=42> consulté le 23/11/213

numériques. Les documents sont mis en contexte et une sélection d'entre eux fait l'objet d'une mise en valeur renforcée. Pour les expositions virtuelles, ce travail d'éditorialisation est particulièrement abouti. L'autre tendance est à la convergence des données, celle-ci pouvant prendre des formes complémentaires. Une bibliothèque numérique permet de sortir des silos des bases de données, c'est-à-dire de regrouper des collections qui ont pu dans le passé être décrites dans des catalogues distincts y compris au sein d'un même organisme. Les plateformes communes vont au-delà en sortant des silos entre organismes alors que les projets collaboratifs font converger les contenus émanant d'organismes avec ceux de son public. Les bibliothèques numériques constituent une étape dans la convergence des données. Le web de données permet d'aller au-delà, pour devenir un outil de convergence dans des projets de valorisation du patrimoine culturel comme nous le verrons à la section 2 de la partie 3.

Cette première partie nous permet de situer le contexte dans lequel s'intègre la bibliothèque numérique de la médiathèque du CND. Comme on le voit, les enjeux auxquels sont confrontés les organismes culturels sont multiples et rejoignent les enjeux au niveau national. Le besoin de s'adapter aux pratiques numériques du public est un des enjeux majeurs. Nous avons vu également les types d'organismes avec lequel le CND pourrait travailler en collaboration ou échanger autour de leurs expériences respectives. Étudier le mode de traitement des fonds d'archives permet de mesurer la problématique soulevée par leur exploitation sous la forme de collections numériques. Nous allons maintenant présenter le projet de bibliothèque numérique de la médiathèque du CND, puis la méthodologie mise en place pour le travail d'analyse comparative de portails de collections et les principales fonctionnalités retenues pour répondre à la demande du CND.

Deuxième partie

Quelle bibliothèque numérique pour valoriser les collections du CND ?

1 Contexte actuel

1.1 La mission du CND et de sa médiathèque

Le Centre National de la Danse est un établissement public à caractère industriel et commercial (EPIC) créé en 1998 à partir d'un regroupement de tout ou une partie de moyens et activités de diverses structures (Théâtre contemporain de la danse, Institut de formation des enseignants de la musique et de la danse de Paris, Centre de formation des enseignants de la danse et de la musique de Lyon, secteur chorégraphique de la Cité de la musique). Les trois missions essentielles du CND sont le soutien à la création, les formations et services aux professionnels (accompagnement et information sur leur déroulement de carrière), le patrimoine et développement de la culture chorégraphique (Médiathèque spécialisée, expositions, audiovisuel, édition, recherche). Pour cela, le CND dispose de studios, certains ouverts au public pour des spectacles ou des performances, une salle de projection, une médiathèque, des collections et tout un corpus de formations, ressources et services pour les professionnels et les compagnies. L'antenne du CND à Lyon / Rhône-Alpes assure quant à lui la continuité de l'ensemble des missions dans sa région d'implantation. Environ 100 personnes travaillent au CND pour un budget proche des 10 millions d'euros.

Au sein du CND, la mission de la médiathèque est de traiter les sources spécialisées documentant la création et les pratiques de danse depuis le XXe siècle, inciter à la sauvegarde, favoriser la conservation et la diffusion d'archives essentielles pour comprendre le développement de l'art chorégraphique, mettre en perspective l'histoire de la danse pour mieux faire comprendre la création contemporaine et ainsi, favoriser le développement d'une culture du domaine auprès de larges publics. La médiathèque, dirigée par Laurent Sébillotte, est rattachée à la mission Patrimoine du CND.

1.2 La genèse de la médiathèque

A sa création en 1998, la médiathèque du CND a bénéficié du rassemblement de plusieurs établissements, dont deux centres de ressources de la Cité de la Musique²⁰ et l'ancien centre de documentation du Théâtre contemporain de la Danse²¹. L'équipe s'est constituée autour de compétences complémentaires avec des documentalistes venant du monde de la danse

²⁰ La partie Danse de la médiathèque pédagogique et la partie Danse du centre d'information Musique et Danse

²¹ Ce centre de ressources avait été supprimé pour raison budgétaire avant son intégration dans le CND qui a néanmoins bénéficié de ses collections

et d'autres de la documentation. À sa création, la médiathèque du CND a reçu le don d'une importante collection d'ouvrages et d'archives de Gilberte Cournaud, critique, libraire et galeriste spécialiste de la danse. Cette donation et les documents des anciens centres de ressources constituent le socle de la collection actuelle de la médiathèque.

L'équipe a préparé l'ouverture au public pendant deux ans. Pour valoriser ses collections, la direction de la médiathèque a choisi de développer son propre plan de classement spécialisé dans la danse en partant du fonds existant. Ce plan de classement adapté aux collections permet d'organiser les documents sur l'étendue des domaines couverts. Un plan de classement généraliste comme le Dewey n'aurait pas permis d'orienter le lecteur dans sa recherche puisque la catégorie 792.8 Danse aurait rassemblé la majeure partie des documents. Un équivalent des Principes de Classement des Documents Musicaux (PCDM) n'existant pas pour la danse, les documentalistes du CND ont innové. De la même façon, la médiathèque a développé son propre langage d'indexation sur la base de thesaurus existants pour la partie sciences humaines et complété par les descripteurs nécessaires pour la partie artistique et culturelle. Ce thesaurus, spécialisé dans la danse et riche d'environ 2000 descripteurs est organisé en 21 thèmes avec 3 niveaux hiérarchiques maximum.

1.3 Les collections

La diversité des collections du CND et la complémentarité de ses fonds témoignent de la richesse de la politique d'acquisition et de collecte de documents autour de la danse mise en place par la médiathèque du CND.

1.3.1 La collection principale

La collection principale de la Médiathèque du CND offre une documentation internationale et actualisée concernant l'histoire, l'esthétique, le contexte et la pédagogie de la danse et plusieurs disciplines comme les sciences humaines et de santé, des livres pour enfants ainsi que des périodiques nationaux et internationaux du monde culturel soit environ 36.000 ouvrages spécialisés, 39.700 fascicules de périodiques, plusieurs milliers de programmes et documents iconographiques, 1.400 disques et partitions musicales, 6.600 vidéos consultables immédiatement sur place, plus de 1.500 heures de captations des activités du CND.

La politique d'acquisition autour de la danse vise à rassembler de façon exhaustive les documents de l'édition contemporaine en langue française et le plus largement possible la littérature grise (travaux universitaires, plaquettes de présentations, programmes, etc.). La collection en langue étrangère est plus sélective. Par ailleurs, les documentalistes rassemblent également des documents physiques autour de spectacles, de compagnies et

d'artistes. Ces dossiers physiques, appelés « dossiers documentaires » sont intégrés à la collection principale. Cette politique d'acquisition complète l'apport des collections des structures d'origine et de la donation Cournand. Cela permet de couvrir l'essentiel des ouvrages parus à partir du 20^e siècle sur la danse en langue française. La médiathèque de CND est le plus grand centre de documentation sur la Danse pour le 20^e siècle et un des plus grands en Europe.

1.3.2 Les fonds d'archives

La collection principale est complétée de 80 fonds d'archives qui documentent les contextes de création, l'histoire individuelle ou collective des artistes, l'évolution des théories, de la critique ou des systèmes d'écriture dans le domaine de la danse. Ces fonds proviennent aussi bien d'artistes et de compagnies chorégraphiques, que de pédagogues, notateurs²², critiques, organismes spécialisés ou encore de plasticiens intéressés par la danse. A ces fonds, vient s'ajouter la collection de dossiers d'artistes et de lieux (plus de 10 000 entrées) rassemblés et régulièrement actualisés par les documentalistes du CND. Sur environ 80 fonds d'archives du CND, la répartition est la suivante : 44 fonds liés à des artistes ou des compagnies chorégraphiques, 8 fonds de pédagogues ou d'écoles, 10 fonds de plasticiens liés à la danse, 17 fonds de théoriciens, critiques et autres personnalités, 4 fonds de théâtres, festivals ou diffuseurs, 10 fonds d'institutions ou organismes spécialisés. Ces fonds d'archives sont majoritaires des fonds d'archives privées avec une dimension patrimoniale de par leur unicité.

Les activités archivistiques de la médiathèque se sont développées après sa création avec comme point de départ la donation en décembre 2003 du fonds Knust. Ce fonds unique porte sur une thématique importante pour le patrimoine artistique : la notation du mouvement²³. Parallèlement, dans les années 2000, les Centres Chorégraphiques Nationaux se questionnent sur le devenir de leurs archives et leur valorisation. La médiathèque du CND se propose alors d'élargir son champ d'actions en offrant d'aider les artistes et les compagnies de danse à traiter leur problématique d'archivage concernant leur travail de création, à trier, structurer le cas échéant le patrimoine documentaire et à préparer si besoin son transfert dans des institutions capables d'en assurer la conservation et la valorisation comme la médiathèque du CND. Avec ses fonds d'archives, la médiathèque dispose

²² Spécialiste de la notation chorégraphique, c'est à dire de la transcription des pas, des mouvements et des figures d'une œuvre chorégraphique à l'aide de signes conventionnels. Parmi les systèmes de notation chorégraphique prévalent la *labanotation* et le *Benesh System*. (source Larousse)

²³ Le terme d'usage pour les spécialistes est « kineptographie » ou « cinétographie »

maintenant d'une importante collection courant des années 80 aux années 2000. Cela permet de donner un panorama de la danse en France avec des artistes et compagnies importants dans le milieu mais qui ne bénéficient pas forcément d'une large couverture médiatique.

Les fonds d'archives complètent de façon étroite les documents collectés par les documentalistes du CND. Nous allons maintenant décrire la collection numérique en décrivant la nature et le mode d'acquisition qui ne diffèrent de la collection « physique ».

1.3.3 La collection numérique

Une partie des collections est sous format numérique²⁴, avec des documents physiques numérisés et des documents nativement numériques. Depuis le début des années 2010, des fonds donnés à la médiathèque l'ont été sous format électronique ou semi-électronique, avec une partie papier ou analogique et une partie numérique.

Les documents sont de nature très variés avec des documents iconographiques, des documents imprimés ou textuels et des documents audiovisuels, comme l'illustre le schéma ci-dessous.

²⁴ Nous parlerons de collections ou documents numériques pour désigner les documents numérisés ou nativement numériques sans les distinguer.

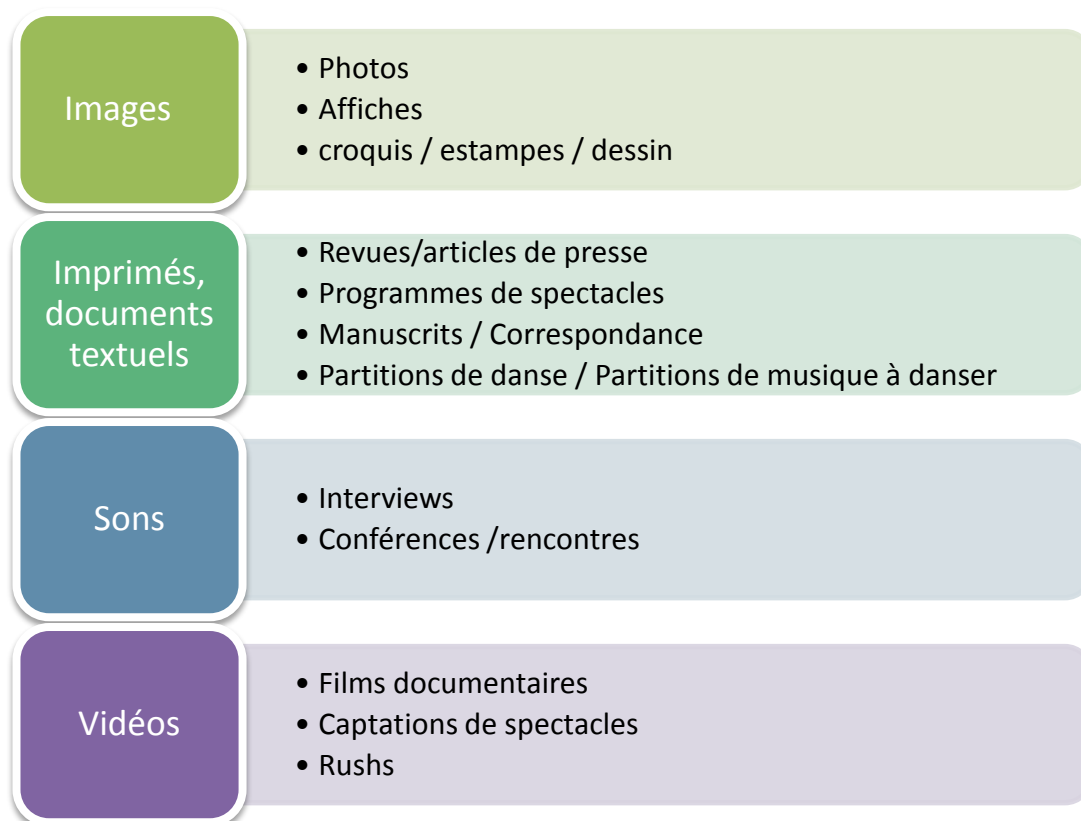


Figure 2 - Typologie de documents de la collection numérique du CND

Les documents proviennent soit des fonds d'archives, soit des acquisitions par achat ou collecte des documentalistes, soit de captations réalisées lors d'évènements au CND (pour les documents audiovisuels uniquement).

1.3.3.1 La politique de numérisation

La politique de numérisation est soutenue par les plans de numérisation du MCC (cf. partie 1 section 3.1) et rejoint les enjeux de la numérisation décrits à la partie 1, section 2. Pour le CND, la numérisation permet de garder une trace de la création en danse tout en facilitant les démarches de transmission des œuvres et des techniques. Les documents originaux sont ainsi mieux protégés et leur copie numérique potentiellement accessible sans limite et plus facilement valorisable. La médiathèque du CND, avec le soutien du Ministère de la Culture, a commencé la numérisation de ses fonds iconographiques, audiovisuels et des revues anciennes au début des années 2000. Elle participe aux programmes nationaux de numérisation concertée en arts, en partenariat avec la BNF.

Lors des premières numérisations, la médiathèque a choisi de traiter l'ensemble de ses fonds photos. La taille des fonds allant en augmentant, elle a privilégié les artistes majeurs avec des documents issus des fonds des Centres Chorégraphiques Nationaux. Au sein des fonds,

les documents sont sélectionnés pour un panorama de toutes les œuvres en privilégiant les formats non pérennes, notamment pour la vidéo. Le choix portera ainsi par exemple sur une ou deux captations ou montages par spectacles qui seront accompagnés de documents complémentaires comme des notes chorégraphiques et des programmes. Le fonds Gourreau²⁵ a vu tous ses négatifs numérisés afin de documenter 30 à 40 ans de danse en France.

1.4 Pourquoi une bibliothèque numérique ?

Le projet de bibliothèque numérique est né de la volonté de la médiathèque du CND de poursuivre sa politique de valorisation. Il s'inscrit dans la politique générale des organismes culturels et du MCC que nous avons analysée précédemment (partie 1 section 2). La médiathèque souhaite en particulier favoriser la diffusion des connaissances dans le domaine de la pédagogie et de la diffusion de la culture chorégraphique tout en s'adaptant aux pratiques numériques de son public. Pour cela, elle a choisi de diffuser plus largement ses collections avec un accès unifié et à distance de ses collections numérisées à destination de tous ses publics : les chercheurs en premier lieu mais également les amateurs de danse et les professionnels. Mettre en valeur ses collections sur Internet avec un outil répondant aux besoins des utilisateurs permettra de faciliter la consultation de documents difficilement accessibles actuellement (besoin de se déplacer, document fragile, nécessité de réserve, préparation des documents par les documentalistes...) et de les préserver. Cela permettra également d'élargir le public de la médiathèque et d'enrichir et alimenter des bases partenaires comme Gallica.

Pour bien comprendre la problématique à laquelle le CND souhaite apporter une réponse, nous allons dans un premier temps rappeler les axes de valorisation actuelles par rapport à la danse, dresser ensuite un état des lieux de la valorisation actuelle des collections, notamment numérique, puis analyser les forces et les faiblesses de la situation actuelle et présenter enfin ses contraintes.

1.4.1 Les axes de valorisation

La médiathèque souhaite faciliter la compréhension et l'exploitation des documents qu'elle traite et donner accès au plus grand nombre d'utilisateurs, sur tout le territoire national. Une de ses missions est également de faire découvrir des chorégraphes et des artistes moins connus, avec des petites compagnies.

²⁵ Le fonds Jean-Marie Gourreau est composé de documents rassemblés et produits par ce critique grand amateur de danse de 1965 à 2006.

Pour les chercheurs et les professionnels, la question des sources en danse reste cruciale pour favoriser la connaissance des artistes et des œuvres et le développement d'une culture chorégraphique. Dans le même temps, la mise en perspective de la politique publique en faveur de la danse contemporaine et l'approche historique de l'éclosion dans les années 70/80 d'une « jeune danse française » sont des problématiques que souhaitent traiter de nombreux chercheurs.

La médiathèque cherche également à permettre la diffusion et la consultation de fonds extensifs documentant de manière exceptionnelle, car très complète et détaillée, les processus de création, les contributions artistiques associées (musiques, scénographie, costumes, compagnonnages avec des artistes associés), les stratégies de communication des artistes mais aussi la réalité de la diffusion des œuvres chorégraphiques en France et à l'étranger.

La politique de valorisation de la médiathèque vise justement à répondre à ces problématiques.

1.4.2 La valorisation actuelle des collections du CND

La médiathèque du CND valorise ses collections de type médiathèque et de type archivistique pour une politique de la médiation culturelle, d'éditorialisation et d'outils informatiques, et de partenariat avec d'autres organismes.

1.4.2.1 La médiation culturelle

La médiation culturelle directe passe par l'accueil du public dans les locaux de la médiathèque avec un accès libre à la collection principale, des emprunts et un accès sur réservation aux documents plus anciens ou d'archives réservés aux professionnels et aux chercheurs. Les documentalistes aident et orientent les utilisateurs dans leurs recherches de documents sur place et à distance. Le travail de médiation est en outre complété par des ateliers présentant les services de la médiathèque, les collections ainsi que le travail de recherche bibliographique.

1.4.2.2 L'éditorialisation

Considérée comme une forme de médiation culturelle, l'éditorialisation vient compléter la médiation culturelle directe en proposant au public du contenu éditorial présentant une sélection de documents issus de ses fonds. Le travail éditorial de la médiathèque permet de mettre en valeur des documents dans les locaux du CND mais aussi sur l'Intranet de la médiathèque et son site Internet. Des expositions sont présentées en salle de lecture pour présenter des documents en lien avec l'actualité de la programmation du CND. Il n'y a pas de statistiques des visites aux expositions, il est donc difficile de savoir si les expositions

rencontrent leur public. Par ailleurs, des montages vidéo sont également diffusés dans les couloirs du CND, pendant plusieurs semaines et ce plusieurs fois par an. Ces montages sont réalisés en partie à partir de documents provenant des fonds de la médiathèque. Ils sont ensuite conservés à la médiathèque pour enrichir les fonds audiovisuels. Ces montages rencontrent leur public, du fait de leur qualité et de leur forme, adaptée au lieu de projection. Le travail éditorial des documentalistes prend également la forme de dossiers documentaires publiés sur les sites internet et intranet de la médiathèque. Ils sont regroupés par thématique²⁶. Leur richesse montre l'important travail d'éditorialisation réalisé par toute l'équipe de la médiathèque. Regardons ces dossiers de plus près.

Documents et archives et Revues électroniques

Ces sections mettent en avant des revues anciennes ou actuelles ainsi que des partitions de danse. Chaque titre est présenté de façon générale. Les documents sont accessibles depuis Internet ou l'Intranet selon les droits de diffusion de la médiathèque. Une partie de ces documents sont également mis en ligne sur Gallica, sans que le travail éditorial y soit repris.

Univers d'artistes

Deux artistes sont présentés à partir de documents issus des fonds d'archives, avec un travail de valorisation important.

Les « Carnets et notes d'Andy de Groat » présentent des cahiers et pages libres où l'artiste a réalisé un travail chorégraphique autour de quatre pièces majeures avec des maquettes, croquis et notes personnels. Ce dossier commence par une présentation détaillée de ce travail, complétée par des extraits de ses notes. Les maquettes, croquis et notes numérisées sont présentés sous forme de montage vidéo.

Le travail de l'artiste Patrick Bossatti est quant à lui présenté par une alternance de présentations textuelles et de dessins et croquis issus des fonds d'archives. Les présentations textuelles complètes, réalisées par la médiathèque sont en outre accessibles sous forme de fichiers PDF thématiques.

Dossiers et textes

Les dossiers et textes sont souvent le pendant numérique des expositions ayant lieu à la médiathèque. Ils sont réalisés autour de thématiques, par exemple la notation du

²⁶ Ces dossiers documentaires thématiques, mis à disposition sous format numérique, sont distincts des dossiers documentaires physiques intégrés à la collection principale (cf. section 1.3.1). Quand je parlerai de dossiers documentaires et de travail d'éditorialisation, je ferai référence aux dossiers documentaires thématiques.

mouvement pour le dernier. Les dossiers sont plus complets que les textes, ces derniers présentant les vitrines des expositions.

Bibliographies sélectives

Les bibliographies sélectives sont des documents PDF présentant des notices bibliographiques qui ont été sélectionnées pour des ateliers menés en médiathèque auprès de publics variés.

Un travail éditorial est également réalisé dans le cadre de la diffusion de documents dans les différents outils en ligne.

1.4.2.3 Les outils en ligne

Les outils de consultation des collections sont éclatés selon la typologie des collections. Le catalogue général recense les documents de la collection principale de la médiathèque. Les bases de données organisent les collections numériques suivant leur type, avec une base iconographique, une sonothèque et une vidéothèque (à partir de fin 2013). Les documents numériques sont intégrés au fur et à mesure de leur traitement dans les bases qui sont accessibles en intégralité depuis l'intranet de la médiathèque. Les listes d'autorités et de mots-clés, quand elles existent, ne sont pas communes aux différentes bases. Le thesaurus de la collection principale n'est pas intégré aux outils en ligne. L'utilisateur accède aux documents base par base et les moteurs de recherche sont eux-mêmes spécifiques à chaque base. Nous allons voir maintenant plus en détail la valorisation des documents numériques dans les inventaires d'archives, la base iconographique et la photothèque. Ils constituent en effet des éléments clés pour la diffusion de la connaissance en danse et l'éducation artistique.

Les inventaires d'archives

Les inventaires des fonds d'archives sont publiés fonds par fonds suivant la logique archivistique (cf. partie 1 section 1.2.2). Outre la navigation par arborescence, les utilisateurs peuvent rechercher par fonds ou sur l'ensemble des fonds. La structuration des inventaires suit le standard EAD qui a été adapté aux fonds du CND afin de proposer des champs pertinents par rapport à la nature des fonds et du sujet. Cela permet d'offrir une recherche simple complétée par une recherche avancée. Les champs de recherche proposés reposent sur la structuration des données (artiste, producteur...). Par ailleurs, les inventaires intègrent des images récupérées automatiquement de la base iconographique afin de valoriser, au plus près de leur description, les pièces d'archives numérisées. Cette mise en valeur des pièces au sein des inventaires permet aux chercheurs consultant les archives de visualiser rapidement des pièces dans le respect du fonds. Cependant cette solution ne

répond pas à la problématique de découverte des collections en plaçant le document au centre.

La base iconographique

Les documents sont décrits par une dizaine de champs sans faire l'objet d'une présentation complémentaire allant au-delà du renseignement des champs. Cela explique notamment le fait que les documentalistes ne disposent pas d'information complémentaire, la collection provenant principalement des fonds d'archives privées (cf. partie 1 section 1.2.2). Les documents ne sont pas décrits par mots-clés.

La base iconographique est structurée suivant le standard Dublin Core qualifié avec les champs suivants :

- Numéro inventaire
- Auteur du document
- Catégorie technique
- Procédé technique
- Type de document
- Format
- Type de représentation
- Période représentée
- Personnes représentées
- Œuvres représentées
- Registre de danse
- Chorégraphe

La vidéothèque

La vidéothèque sera disponible fin 2013. Elle reprendra une grande partie des fonctionnalités de Numéridanse, la plateforme commune avec les acteurs de la création en danse, mis en place en 2013 (cf. section 1.4.2.4). Elle proposera en revanche de visionner des œuvres complètes.

Les données seront également structurées suivant le format du Dublin Core qualifié. Contrairement à la base iconographique, la description de documents inclura une présentation des œuvres et des artistes ainsi qu'un champ pour les mots-clés.

1.4.2.4 Les partenariats

La médiathèque du CND participe à des initiatives de plateformes communes de valorisation du patrimoine culturel. En tant que pôle associé de la BNF pour la danse, les revues numérisées et libres de droits du CND sont référencées dans Gallica. Les inventaires d'archives sont diffusés sur les sites du Catalogue Collectif de France²⁷ et du répertoire des arts du spectacle (RASP) ainsi que celui du Patrimoine numérique²⁸. Ce dernier recense toutes les collections culturelles numérisées ou en cours de numérisation.

Par ailleurs, la médiathèque participe activement au développement du projet Numéridanse visant à offrir aux acteurs de la danse une plateforme commune de diffusion d'extraits de spectacles. Ce portail permet de présenter les extraits ainsi que les œuvres et les artistes. Alors que le CND y met à disposition des vidéo anciennes, les artistes et compagnies y publient des extraits de spectacles actuels ou récents.

Enfin, la médiathèque participe à des projets d'éducation artistique en proposant des documents issus de ses collections qui sont utilisés pour alimenter des portails d'information comme le portail Histoire des Arts²⁹ développé par le MCC.

1.4.3 Les forces et les faiblesses de cette valorisation

La politique de valorisation de la médiathèque permet de mettre à disposition des services variés à destination de ses différents publics. La médiathèque capitalise sous forme de dossiers documentaires le travail réalisé pour des manifestations temporaires comme les expositions et les ateliers. Si ces dossiers permettent de mettre efficacement en valeur les trésors de la médiathèque, les dossiers eux-mêmes sont peu valorisés par leur intégration de façon linéaire sur les sites de la médiathèque. L'organisation des collections et des dossiers documentaires dans des bases distinctes selon la nature des documents limite la lisibilité et l'accessibilité des collections de la médiathèque. C'est là un des enjeux majeurs du projet de bibliothèque numérique.

Sur le terrain de la convergence des données culturelles, la médiathèque participe aux initiatives de plateformes communes mises en place par le MCC et la BNF. Elle met elle-même en place des initiatives dans le monde de la danse pour faire converger patrimoine et création, et cela de deux façons. D'une part la médiathèque du CND est le point d'appui

²⁷ Lien : <http://ccfr.bnf.fr>

²⁸ Lien : www.numerique.culture.fr

²⁹ Lien : www.histoiredesarts.culture.fr

pour la valorisation des fonds d'archives des acteurs de la création artistique en danse. D'autre part elle a mis place, en concertation avec d'autres structures, une plateforme de diffusion de vidéos de création et de patrimoine artistique. En mettant en place une bibliothèque numérique et en y intégrant toutes ses collections numériques, la médiathèque sera en mesure d'augmenter ses contributions aux plateformes communes. Sa visibilité en sera renforcée.

Une bibliothèque numérique apportera une réelle plus-value à la médiathèque du CND en terme de valorisation et visibilité.

1.4.4 Les contraintes à prendre en compte

Pour mener à bien son projet, la médiathèque devrait s'appuyer sur ses forces et dépasser les faiblesses de la valorisation actuelle de ses collections mais également faire avec deux contraintes majeures : les droits d'auteurs et la nature des collections, des archives vivantes portant sur le spectacle vivant.

1.4.4.1 Contraintes juridiques et les droits d'auteurs

Beaucoup de photographies et captations de spectacles n'ont pas été réalisées avec l'idée d'une diffusion ultérieure. Cela est particulièrement vrai pour les documents provenant de fonds d'archives rassemblés au départ par une personne physique ou morale pour son usage personnel ou interne à son organisme. Diffuser sur Internet des documents dans le respect des droits d'auteur est complexe. L'accord doit porter pour un document ou un ensemble de documents particulier(s) pour un usage défini. Par ailleurs, étant donné le type de documents et le type d'œuvres, l'accord de plusieurs voir de nombreuses personnes est souvent nécessaire (l'auteur du document, le chorégraphe, les artistes, le metteur en scène...) et ce pour des documents pouvant être anciens et pour lesquels le CND ne dispose pas de toutes ces informations.

Le problème des droits d'auteurs est complexe. Pour autant, cela n'étant pas le sujet de mon étude, j'en souligne l'importance sans y apporter de réponse.

1.4.4.2 La nature des collections elles-mêmes

Les collections de la médiathèque de la danse constituent des archives vivantes concernant des acteurs en activité. Quelle que soit la nature des documents traités, les acteurs peuvent ne pas adhérer aux choix documentaires de l'équipe. Cela est particulièrement vrai dans le traitement archivistique, les acteurs étant aussi parties prenantes de ce traitement. Il faut parfois tenir compte des enjeux de positionnement, de communication. Par exemple, un chorégraphe pourra ne pas souhaiter être considéré comme faisant partie de tel courant

artistique ou ne pas souhaiter que tel spectacle soit facilement accessible. Ce sujet rejoint la question de l'identité numérique des personnes.

Par ailleurs, il est difficile de parler d'un art contemporain pour lequel il n'y a pas de consensus concernant les critères d'appréhension, la terminologie et les façons de les mettre en perspective. La documentation nécessite de cartographier la collection en posant une organisation et une description, les deux pouvant être amenées à évoluer rapidement. Voyons quelques exemples pour mieux comprendre.

La « Danse contemporaine » commence à l'après-guerre en Europe et aux États-Unis et continue jusqu'à aujourd'hui. Mais pour combien de temps ? Le Hip-Hop continuera-t-il à être catégorisé comme une danse contemporaine indéfiniment ? Prenons un autre exemple. Un mouvement chorégraphique de danse contemporaine a été qualifié de "non-danse" par les critiques mais peut-il être repris pour décrire de façon « officielle » des œuvres ou des artistes ? Dernier exemple, les danses traditionnelles, les danses de salon et les danses de société sont considérées comme trois catégories et pourtant elles se recoupent partiellement. Comment les catégoriser ? Pour Laurent Sébillotte, directeur de la médiathèque, « le milieu de la danse a une répugnance à cataloguer et à force de ne pas s'y risquer, il y a absence de consensus ».

Par ailleurs, pour l'équipe de la médiathèque, ce travail supplémentaire viendra s'ajouter à la charge de travail déjà existante, dans une période où les restrictions budgétaires ne laissent pas entrevoir de ressources supplémentaires. Cela implique de penser un projet ambitieux par sa capacité à répondre aux objectifs identifiés et aux besoins du public tout en étant réaliste quant à la possibilité de le mettre en place et de le faire vivre, compte tenu des ressources disponibles.

Maintenant que nous avons dressé la situation propre à la médiathèque du CND, voyons la méthodologie mise en place lors de mon stage pour définir la future bibliothèque numérique.

2 La méthodologie utilisée pour définir la future bibliothèque numérique

Pour mener à bien le développement de sa bibliothèque numérique, la médiathèque a choisi de mener une étude préalable qui a fait l'objet de mon stage de 8 semaines à l'été 2013. Dans ce cadre, j'ai étudié l'état de l'art actuel des portails de collections culturels, en regardant ce qui pourrait fonctionner (et ne pas fonctionner) pour répondre aux besoins des utilisateurs. Pour cela, je suis partie des collections actuelles du CND. J'ai ensuite réalisé un premier cahier des charges fonctionnel de la bibliothèque numérique. Je présente ici premièrement l'analyse rapide des besoins du public, puis la méthodologie mise en place pour choisir les sites à étudier et le bilan de l'analyse et enfin les principaux choix fonctionnels.

2.1 L'analyse des besoins

L'analyse des besoins a pour but d'identifier, puis de définir ceux à qui s'adresse le produit ou service proposé et leurs besoins. Le choix a été fait de ne pas faire une étude approfondie des besoins explicites et implicites des différents groupes d'utilisateurs pour plusieurs raisons. Premièrement le projet s'intègre dans une politique globale de valorisation des collections culturelles. La médiathèque du CND connaît ses publics cibles, ils découlent de la mission qui lui a été confiée par le CND et par le Ministère de la Culture. Cela permet de bien cerner les besoins du public de façon générale. Ensuite le cadre de la mission était de lancer la réflexion, une étude détaillée des besoins du public était prématurée. Dès lors, il a été choisi de se concentrer sur l'analyse fonctionnelle. J'ai cependant clarifié et exprimé les besoins des utilisateurs, en me basant sur ce que nous connaissons de leurs pratiques. Cela m'a permis ensuite en analysant les sites de réfléchir à quels étaient les publics visés et leurs besoins pour chaque collection et de comparer avec le cas du CND. Lors de la rédaction du cahier des charges, cela m'a permis de bien répondre aux besoins des utilisateurs dans leur diversité.

Les publics de la médiathèque du CND peuvent être regroupés en trois catégories : les amateurs, les professionnels et les chercheurs dans le domaine de la danse. Compte-tenu des collections qui retracent le travail de chorégraphes et d'artistes moins connus, avec des petites compagnies, les amateurs intéressés par les collections du CND sont des amateurs éclairés. Les professionnels sont des danseurs, chorégraphes, notateurs mais aussi des enseignants. Ils ont souvent des casquettes multiples. Les chercheurs sont des étudiants, enseignants chercheurs et personnes effectuant des recherches sur la danse de façon large.

2.1.1 Les besoins des trois publics cibles

S'ils ont des besoins communs, chaque catégorie a également des besoins spécifiques. Les besoins communs sont les suivants :

- s'approprier facilement l'outil pour une navigation fluide lui permettant de comprendre ce qu'on lui montre et où il se trouve dans le portail,
- pouvoir suivre les actualités des collections numériques. Cela permettra de fidéliser les utilisateurs,
- pouvoir accéder rapidement au document primaire numérique.

En schématisant, on peut considérer que l'utilisateur doit pouvoir parcourir et exploiter les différents éléments proposés à ses propres fins.

Le visiteur amateur peut être demandeur d'une liberté de choix lui permettant d'explorer ces ressources selon ses affinités, ses goûts esthétiques et ses habitudes. Pour les professionnels danseurs et/ou enseignants, une visite peut être davantage axée sur les objectifs créatifs et pédagogiques. Pour les chercheurs, une visite est susceptible d'être définie par des objectifs plus explicites et ciblés.

La découverte des collections se fait à partir des suggestions de documents organisés de façon cohérente et en naviguant d'un document à un autre par sérendipité. La recherche sur des sujets précis sera surtout utile aux professionnels et très utiles aux chercheurs. En outre ceux-ci auront besoin d'accéder aux fonds rares et difficilement accessibles ainsi qu'aux travaux de recherche en lien avec les collections du CND. Concernant la pratique numérique, on peut considérer que les utilisateurs amateurs et professionnels ont plutôt un profil d'utilisation de l'outil informatique de base. Au contraire, les chercheurs sont d'avantage habitués à naviguer dans des collections et à effectuer des recherches d'information plus poussées même si l'influence de la recherche avec des moteurs basés sur la popularité des résultats se fait sentir par une moindre capacité à interroger des bases de façon avancée.

2.2 Le benchmarking de collections numériques en ligne

De façon générale, le benchmarking est une analyse comparative qui permet à une organisation d'apprendre de la performance des autres et de la sienne en les comparant tant du point de vue des approches managériales que des résultats. C'est une technique qui consiste à étudier et analyser les techniques de gestion, les modes d'organisation des autres entreprises afin de s'en inspirer et d'en tirer le meilleur. C'est un processus continu de recherche, d'analyse comparative, d'adaptation et d'implantation des meilleures pratiques pour améliorer la performance des processus dans une organisation.

Pour permettre justement un processus contenu, j'ai fait le choix de poser une méthodologie qui permette, au-delà de la mission de stage, de poser un cadre réutilisable, une fois que le projet sera plus avancé. En effet la méthode de repérage et décryptage des portails culturels permettra de suivre ultérieurement l'actualité des offres de collections numériques plus ou moins proches du CND.

2.2.1 La sélection des portails de collections culturelles

Pour mener à bien une analyse comparative, la première étape consiste à déterminer les objets de l'étude. Pour cela, j'ai travaillé en trois étapes en commençant par une phase de découverte.

J'ai choisi de recenser et tester rapidement les portails que je rencontrais et qui donnaient accès à des ressources en ligne ayant trait au monde du spectacle vivant et de la culture en général, si possible avec des types de documents variés. Je me suis ainsi fait une première idée de leurs collections et de l'organisation des portails.

Parallèlement, l'analyse du contexte général de la mission, présentée dans la première partie, a débouché sur une cartographie des organismes et collections intéressants pour mon analyse : des organismes du spectacle vivant, mais également des musées, bibliothèques et centres d'archives ainsi que des lieux de pratique et de création artistique, des lieux d'enseignement et des centres nationaux.

Suite à cette première approche, j'ai recherché des sites en partant d'une recherche approfondie. J'ai utilisé Google et Exalead pour des recherches avec les termes de « bibliothèque numérique », « expositions virtuelles » et « collections numériques ». J'ai également recherché des articles scientifiques faisant référence à des sites de collections numériques s'adressant à des documentalistes, des archivistes mais également des chercheurs, sociologues et historiens. L'objectif était de recenser les portails mentionnés, non seulement par les personnes qui les mettent en place, mais également par leurs utilisateurs.

Ayant trouvé un nombre restreint de sites dans le domaine des arts du spectacle vivant, j'ai complété cette recherche de deux façons. J'ai consulté les sites recensés dans les signets de la BNF sous le thème « Arts », avec les catégories « Danse, ballet » et « Arts du spectacle » et ses sous-catégories « Théâtre », « Théâtre de rue », « Cirque », « Marionnettes » et « Mime ». J'ai cherché dans Gallica les documents rattachés au sujet « Danse », « Cirque » et « Mime » et repéré les bibliothèques associées. Cette recherche à partir des services de la BNF m'a permis de valider la liste des sites recensés sans toutefois la compléter.

En dernier lieu, pour m'assurer d'avoir un panorama correspondant à mes besoins, j'ai défini ceux-ci de la façon suivante :

- Des collections traitant du spectacle vivant
- Des sites plaçant le document au centre, à l'inverse des sites s'appuyant sur des documents pour illustrer un propos
- Des sites offrant une approche très grand public, avec un mode de recherche simple
- Des sites plus pointus dédiés à un public plus spécialisé
- Une taille de collections variée avec une typologie de documents variés
- Des moyens humains / financiers variés : les projets florissants pouvant être plus ambitieux et originaux et les projets à moyens limités me rapprochant du projet du CND
- Une complémentarité des portails les uns avec les autres
- Son intérêt particulier, relevé lors de la phase de découverte

La réalisation d'une grille d'analyse succincte m'a permis d'objectiver le choix des sites.

Finalement, les sept portails retenus pour l'analyse sont les suivants :

- **Le portail des arts de la marionnette (PAM)**
Portail retenu pour la similitude du projet avec celui du CND dans le domaine du spectacle vivant et son envergure restreinte
- **Le site Rue et Cirque de Hors les murs**³⁰
Portail retenu pour le domaine du spectacle vivant
- **Le site de la médiathèque de la Cité de la Musique dédié aux médiathèques et conservatoires**³¹
Portail retenu pour le domaine du spectacle vivant et la nature des documents avec des documents audiovisuels, iconographiques et pédagogiques et les liens entre des documents de différentes natures
- **La bibliothèque numérique de l'INHA** (Institut National de l'Histoire de l'Art)
Portail retenu pour les suggestions de documents et l'utilisation d'un logiciel de gestion de collection numérique envisagé par la médiathèque du CND

³⁰ Le site Rue et Cirque était hors service pendant l'été et n'a donc pas été intégré à l'étude.

³¹ Ce site payant est accessible depuis les médiathèques et conservatoires clients.

- **La bibliothèque numérique de la bibliothèque municipale de Lyon**
Site retenu pour le fonds d'archives et le fonds de photographies en direction des amateurs offrant des outils collaboratifs
- **La partie Œuvres du site Musée du Louvre**
Portail retenu pour la navigation fluide où le document est mis au centre
- **La Digital Public Library of America (DPLA)**
Portail retenu pour une ouverture à l'international et pour la valorisation de fonds importants et hétérogènes avec une navigation très simple, avec de grands moyens et une mise en ligne très récente

Les liens vers les sites retenus figurent en annexe 2.

2.2.2 L'analyse des portails

Une fois les sites retenus, le travail d'analyse a débuté par la réalisation d'une grille d'analyse. Avoir une telle grille permet une approche systématique et méthodique indispensable pour objectiver le regard. En partant de grilles existantes, j'ai développé mes propres critères en les structurant par catégorie. J'ai ensuite retravaillé la grille de façon itérative au fur et à mesure des analyses. La grille d'analyse est jointe en annexe 3.

Concrètement, pour réaliser ce document, je suis partie de plusieurs grilles d'audit de sites web. La principale difficulté rencontrée a été de me placer dans le champ de l'analyse comparative à but prospectif. En effet les grilles sont souvent réalisées dans le cadre d'audit ou dans l'analyse comparative pour choisir des solutions informatiques. Or, les objectifs ne sont pas les mêmes.

L'audit constitue un outil d'amélioration continue, permettant de faire le point sur l'existant afin d'en dégager les points faibles et/ou non conformes puis de corriger les écarts et dysfonctionnements constatés³². Dans le choix de solutions logicielles, il s'agit d'évaluer si la solution testée répond aux besoins en proposant les fonctionnalités prioritaires et n'a pas de restrictions d'usage ou des dysfonctionnements la rendant inadaptée, dans le but de définir la solution répondant le mieux à ses besoins.

Quant à l'analyse comparative prospective recherche elle les solutions intéressantes ou à éviter dans chaque objet d'étude. Le travail se fait ainsi sans chercher à juger ni départager les services proposés. Ayant moi-même réalisé par le passé des audits et des évaluations de solutions logicielles, il m'a été nécessaire de veiller à garder le cap de la comparaison et de

³² Source : Wikipedia. Audit. 01/09/2013. [Consulté le 17/09/2013],
<<http://fr.wikipedia.org/wiki/Audit>>

l'étude des moyens mis en œuvre pour valoriser les collections numériques. Cela est vrai pour la réalisation de la grille mais aussi pour l'analyse elle-même.

Concrètement, pour développer les critères, je me suis appuyée sur le projet du CND et ses besoins, ses collections et choix de valorisation. Le fait d'avoir déjà fait une première découverte et pré-analyse des sites m'a permis de penser ce qu'il y avait de commun et de dissemblable entre le projet du CND et ceux étudiés. J'ai déterminé les critères retenus en les structurant autour des huit thèmes.

1. La partie **Identification** cadre le projet, l'organisme dont il dépend, le type de collections et les publics cibles.
2. La partie **Contenus** définit la collection en terme de couverture géographique, temporelle, de taille, de nature des documents de façon précise.
3. La partie **Questions juridiques** présente les licences d'utilisation pour les documents et les notices.
4. La partie **Ergonomie** détaille la façon dont est organisée la navigation avec le critère retenu pour la navigation principale, la zone la plus importante sur la page d'accueil, comment certains documents sont mis en avant, les choix de navigation (par exemple par thème, par époque...) et comment fonctionne la navigation dans les collections.
5. La partie **Moteur de recherche** présente les types de recherche possibles et les options de tri, affichage et relance.
6. La partie **Documents** étudie comment sont présentés les documents, la qualité et la structuration des descriptions, la présence d'une vignette, le choix d'affichage des notices abrégées et complètes. Cette partie a été répétée quand des documents de différentes natures étaient traités différemment.
7. La partie **Outils / Services** présente les outils offerts : les espaces personnels, les outils d'interaction et les outils d'export.
8. La partie **Points notables** développe les points faibles et les points forts retenus par rapport aux propres besoins du CND.

Les critères ont été complétés par la suite par itération au fur et à mesure de l'analyse des portails. La description de la navigation dans les portails a été réalisée sous forme de schéma. Cela a permis de regarder la fluidité de la navigation et de me rendre compte de façon objective si le document primaire était mis au centre. Cela m'a également permis distinguer les choix de sites pour un seul type de navigation ou au contraire une pluralité de navigation, que ce soit selon les choix de l'utilisateur ou de types de documents.

J'ai également relevé la structuration du contenu et le choix du vocabulaire pour décrire les documents ainsi que les termes choisis dans la navigation, notamment pour mettre en avant les documents. Des captures commentées ont également complété le travail d'analyse. Le but était double. Il s'agissait à la fois d'illustrer les points relevés mais aussi de garder trace de la version d'étude, les sites web évoluant rapidement.

J'ai également intégré ponctuellement dans l'analyse les autres sites. Cela s'est fait sur la forme d'une analyse générale des points intéressants présentés en compléments. La liste des portails étudiés de façon ponctuelle est présentée à l'annexe 2.

3 Des résultats de l'analyse aux suggestions fonctionnelles

3.1 Une petite synthèse de l'analyse comparative

Lors de la synthèse de l'analyse comparative, j'ai relevé quelques tendances générales des portails ainsi qu'une sélection de solutions particulièrement intéressantes pour le projet de la médiathèque.

3.1.1 Tendances générales

La synthèse présentée concerne principalement les interfaces et leurs fonctionnalités. En ayant sélectionné les portails pour leurs spécificités propres et pour leur complémentarité, il m'a été ensuite difficile de dégager des tendances générales en terme de structuration des données et de politique éditoriale.

La tendance générale est à la similarité entre la navigation principale (c'est-à-dire en naviguant dans le site) et la navigation dans les résultats de recherche, pour lesquels les modes d'affichage sont adaptables. Les modes d'affichage classiques par liste ou mosaïque sont les plus courants mais peuvent aussi être complétés par des frises et/ou des cartes.

Les facettes sont largement utilisées pour permettre d'élargir et de rétrécir les critères de sélection, y compris dans le cas de recherches implicites. Les facettes sont implémentées de différentes façons : les critères pour chaque champ peuvent être classés par ordre alphabétique (ou de grandeur) ou par nombre d'occurrences de documents répondant à chaque critère. La liste des critères peut être complète ou abrégée. Un moteur de recherche peut être intégré dans les facettes avec parfois une auto-complétion. Dans ce cas, les critères trouvés sont complétés de leur typologie (par exemple la qualité de la personne) et le nombre d'occurrences.

Une large place est faite sur les pages d'accueil à la mise en avant de documents, que ce soit de façon individuelle ou par collection sous forme de « Une », « Zoom », « Expositions virtuelles » et autre « Dossiers thématiques ».

Les sites sont majoritaires structurés par thème ou type de documents, parfois les deux. On note pourtant que le Portail des Arts de la Marionnette a choisi une structuration sous forme d'arborescence dans sa base de données alors que le site de la Cité de la Musique propose une navigation mélangeant un type de documents (« Vidéos »), quatre styles musicaux correspondants à différents publics cibles, et un thème (« Instruments »).

La plupart des sites proposent des liens entre les ressources ou des liens lançant des requêtes implicites depuis la page d'accueil mais aussi dans les notices elles-mêmes. L'INHA a choisi d'insérer les liens sur des mots au milieu de textes explicatifs, alors que la plupart des sites ont choisi de mettre les liens en dehors des textes rédigés, permettant de clairement identifier les zones de liens.

Les résultats proposés se présentent en général sous forme de notices abrégées avec une vignette du document. On peut également trouver de vignettes accompagnées d'un titre uniquement, la notice s'affichant seulement à la demande.

3.1.2 Choix relevés comme particulièrement intéressants

Dans le cadre de la réflexion sur la valorisation des collections numériques culturelles sur Internet, je présente ici la méthodologie mise en place dans le cadre de ma mission au CND. Je ne présente pas ici l'ensemble des choix relevés mais une sélection ainsi que leurs éventuelles limites par rapport aux collections du CND.

Le Louvre a choisi de mettre en valeur les photos de ses œuvres avec une double présentation des documents. Dans la partie haute de la page, les documents s'affichent en grande taille un par un ; dans la partie basse, les mêmes documents sont présentés dans une liste de résultats classique avec vignette et notices brèves permettant d'accéder rapidement à un document.

Pour donner plus de vivacité et de modernité aux notices, certains sites ont choisi une mise en forme sortant du classique « champ : description » en remplaçant le nom du champ par une icône ou une forme rédigée (comme « Mis en scène par... »). C'est par exemple le cas du site des Opéras en actes.

Le Portail des Arts de la Marionnette met en avant dans les notices des ressources similaires ainsi que des ressources externes au portail. La Cité de la Musique et le Regroupement des Opéras de France ont choisi de mettre en avant non seulement des documents mais également des artistes ou œuvres auxquels sont rattachés des ressources. Cela permet une approche plus proche de la démarche encyclopédique que documentaire et permet de faciliter la découverte des collections avec une utilisation plus instinctive.

L'organisation des résultats de recherche sur une carte ou une frise permet de découvrir les documents d'une façon différente mais aussi d'appréhender l'étendue géographique et temporelle des collections proposées.

3.2 Les principales fonctionnalités retenues

Le travail sur les fonctionnalités a été complété par des suggestions de choix éditoriaux et des exemples. Cela permet de mettre en contexte les fonctionnalités et de nourrir la réflexion autour de la future bibliothèque numérique. Le choix d'une politique éditoriale sera une des prochaines étapes dans le développement du projet. Sa définition sera le fruit d'une réflexion de l'ensemble de l'équipe de la médiathèque.

Le choix s'est porté sur un portail mettant en valeur des ressources issues des collections pour les trois publics sans créer d'espaces spécifiques à chaque catégorie. La navigation dans les collections sera la même que la présentation des résultats de recherche. L'utilisateur pourra choisir entre plusieurs modes de présentation. Le document sera mis en valeur avec un accès rapide à sa visualisation, la notice servant d'appui, avec des documents mis en avant dès la page d'accueil et servant d'illustration et d'accroche.

3.3 Le cahier des charges fonctionnel

Le cahier des charges fonctionnel a été élaboré à partir de l'analyse comparative, avec comme objectif de mettre en avant une partie des collections. La priorité est en effet de proposer du contenu aux trois publics cibles pour leur permettre de partir à la découverte des collections et d'élargir leurs connaissances sur la danse. Le cahier des charges décline les fonctionnalités permettant de répondre à cet objectif. Celles-ci sont complétées par des propositions de choix éditoriaux. Les choix proposés se veulent réellement applicables au projet de la médiathèque compte-tenu des collections et des ressources disponibles. Cependant le but de ce travail est une appropriation du projet par l'équipe afin de faire avancer la réflexion et de lancer le projet. Ce cahier des charges n'a donc pas vocation à être implémenté tel quel.

3.3.1 La navigation

Dans la navigation, la mise en avant de documents et la possibilité de naviguer d'un document à l'autre ont été privilégiées. La navigation dans les collections se fera donc à plusieurs niveaux, afin d'offrir plusieurs points d'entrée dans les collections et satisfaire le besoin de naviguer librement dans les collections. L'utilisateur pourra par exemple naviguer par :

- type de documents
- recherche
- artiste
- œuvre

- suggestion de documents

Les suggestions porteront sur des documents individuels et sur des ensembles de documents sélectionnés pour répondre aux besoins des différents publics et seront accessibles depuis la page d'accueil. Une ressource pourrait être mise en avant en proposant un « Document du jour », « Document du mois » ou un « Zoom ». Des documents seront présentés également dans des ensembles cohérents, par exemple sous forme de :

- Dossiers thématiques
- Découvertes d'œuvres ou d'artistes majeurs
- Découvertes par type de danses
- Tour du monde de la danse en X photos, vidéos, documents
- Chronologie documentée de la danse
- Insolites
- Expositions virtuelles
- « Best of » pour chaque fonds d'archives

Le choix des documents pourrait être celui de l'anniversaire d'un événement ou d'un artiste, du dernier don dans les archives mais aussi aléatoire pour les photos/vidéos ou encore le fruit d'un travail documentaire.

3.3.2 Affichage des ressources

L'affichage principal sera organisé par liste. Le document sera mis en avant avec la vignette qui occupera une place importante. Une notice abrégée documentera le document. Elle sera complétée par une notice détaillée qui s'affichera à la demande, avec la ressource. Cet affichage par liste viendra compléter un affichage document par document dans la partie haute de la page en offrant une meilleure visibilité.

L'utilisateur pourra restreindre la recherche par des facettes, un moyen adapté aux pratiques des publics cibles. L'utilisateur pourra ajouter ou retirer un choix facilement et sélectionner plusieurs choix d'un même champ (par exemple sélectionner deux personnes dans le champ Chorégraphe). Chaque intitulé sera complété du nombre d'occurrences correspondantes et les choix classés par nombre d'occurrences descendantes. Les choix sans occurrence seront masqués. Le nombre de choix sera limité aux 5/7 premiers choix avec la possibilité d'en afficher plus. Les facettes pourront être complétées d'un moteur de recherche permettant de trouver les critères de choix si les listes sont longues. Les résultats de critères indiqueront le nombre d'occurrences de documents répondant à ce critère et la typologie (si la personne est un metteur en scène, un chorégraphe par exemple). Le choix des critères des facettes sera adapté à la structure des collections pour éviter d'avoir des listes de critères longues et beaucoup de critères avec seulement quelques occurrences.

L'affichage par liste pourra être complété par les solutions suivantes. L'affichage par mosaïque présentera les documents avec leurs vignettes et leur titre. L'affichage par œuvre ou par artiste permettrait d'accéder à un page recensant de façon organisée tous les documents accessibles autour d'une œuvre ou d'un artiste. Cet affichage pourrait être limitée aux œuvres et artistes pour lesquels assez de documents sont disponibles. L'affichage chronologique par frise pourrait compléter les modes de navigation si cela s'avère pertinent par rapport à la collection.

3.3.3 Visualisation d'une ressource

La ressource s'affichera avec le document en prévisualisation ou avec le *player* chargé pour l'audio et la vidéo et une notice brève. L'accès à la notice complète se fera à la demande afin de ne pas surcharger d'information l'utilisateur, notamment amateur. Le mode plein écran sera accessible. La notice pourra par exemple s'ouvrir avec un affichage par onglets, le premier onglet mettant en avant le document lui-même et l'accompagnant de quelques champs explicatifs et deux autres onglets proposant les ressources complémentaires et la notice complète. Lorsqu'une ressource est ouverte, il sera possible de passer directement à la ressource précédente/suivante de la liste ou de revenir à la liste.

Des liens permettront de valoriser le contenu et d'aider à la navigation : chaque notice pointerait vers les artistes principaux (interprètes, chorégraphes...), les compagnies, les œuvres et des ressources complémentaires. Le lien affichera, si possible, une page artiste ou œuvre et sinon relancera une recherche implicite.

3.3.3.1 Le cas des ressources mises en avant sur la page d'accueil

Depuis la page d'accueil, le lien vers une ressource individuelle pourra afficher celle-ci seule ou dans la collection à laquelle elle appartient avec, par exemple, une liste proposant des liens vers des requêtes implicites (comme voir d'autres vidéos sur la même époque, du même chorégraphe, autour de la même œuvre).

Un ensemble de ressources s'affichera en donnant la possibilité de naviguer entre celles-ci.

3.3.4 Recherche

La recherche secondera la navigation dans les collections, elle ne sera pas le mode d'exploitation privilégié. L'utilisateur pourra élargir et rétrécir ses critères de recherche et devra toujours être en capacité de savoir où il est dans le site et ce qu'il voit.

La mise en forme de la recherche sera la même que dans la navigation principale avec les facettes. La recherche simple pourra être complétée par une recherche avancée se limitant à

quelques champs. Les résultats s'afficheront par critère de pertinence, date d'ajout du document et date du document/du spectacle.

3.3.5 Ergonomie

Côté rédactionnel, afin de donner une image moderne et une plus grande fluidité de lecture, les textes seront courts, avec des phrases dynamiques. Les notices seront mises en forme de façon à ne pas donner l'impression de lire une notice de bibliothèque. Les liens seront mis en avant en dehors des textes rédigés.

Le graphisme du site permettra à l'utilisateur de savoir ce qu'il visualise. Un jeu d'icônes sera utilisé de façon homogène pour symboliser les types de ressources et les champs de la notice.

Dans une liste de ressources, les critères de sélections des ressources seront visibles au-dessus de la liste des résultats ou des facettes que ce soit pour les requêtes implicites ou les résultats des recherches de l'utilisateur.

3.3.6 Accessibilité

Le CND étant un établissement public à caractère industriel et commercial (EPIC), la bibliothèque numérique devra répondre au référentiel général d'accessibilité pour les administrations (RGAA). Le portail suivra les recommandations du RGAA et celles à venir du Ministère de la Culture.

3.3.7 Services et outils

La bibliothèque numérique pourra mettre en avant l'important travail d'éditorialisation réalisé par l'équipe avec un ensemble de produits incluant des dossiers documentaires, des bibliographies, etc. Ce travail documentaire sera complété par des fonctionnalités permettant de faciliter l'utilisation et permettre aux utilisateurs d'interagir. Plusieurs outils seront proposés (classés par ordre de priorité) :

1. Exportation en PDF des notices et des ressources
2. Exportation vers Zotero des notices
3. Flux RSS avec un flux général et/ou par rubrique/thème/type de documents
4. Espace utilisateur authentifié avec sauvegarde et envoi de panier de recherches et sauvegarde et gestion de notices, envoi par mail.

Le site proposera aux utilisateurs de contribuer à la description des documents de ses collections en leur proposant de contribuer à la description d'œuvres orphelines et de suggérer des modifications ou améliorer des notices. Ces fonctionnalités pourront être développées dans un deuxième temps.

3.4 Points de vigilance et bilan du travail

3.4.1 Points de vigilance pour l'interface

Les utilisateurs s'attendent en général à trouver un ensemble de fonctionnalités classiques dans un portail. Or leur mise en place peut s'avérer difficile compte-tenu des collections du CND. Je souligne ci-dessous quelques points à surveiller.

La représentation des résultats de recherche sur une carte ou une frise permettrait de montrer la couverture géographique et temporelle de collections. Cependant, de par la nature, les objectifs et l'origine des fonds, la couverture est incomplète et la répartition peu homogène avec beaucoup de documents sur les années 80-90 et moins sur d'autres périodes. Une solution serait de proposer une frise pour découvrir un fond particulier ou une collection particulière.

Le critère géographique peut concerner les lieux de production ou l'origine des danses, parfois multiples. Il serait intéressant d'étudier si ce critère constituerait une aide ou une source d'incompréhension pour l'utilisateur.

Par ailleurs, les collections du CND couvrent beaucoup d'artistes pour lesquels peu de documents sont disponibles. Une navigation par artistes (ou un champ Artistes dans les facettes) peut s'avérer difficile à mettre en place efficacement. Cependant une absence de navigation par artistes pourrait être incompréhensible du point de vue des utilisateurs. Ce point de vigilance devra être testé lors de l'implémentation. Par ailleurs, le fait de ne pas avoir forcément de documents concernant les artistes les plus connus pourraient également susciter des interrogations pour le public ou être vu comme un manquement.

3.4.2 Bilan du travail

En terme de bilan, je peux dire que la méthodologie mise en place m'a permis de mener à bien ma mission : proposer à l'équipe de la médiathèque des pistes pour amorcer le lancement du projet de bibliothèque numérique.

La grille d'analyse a été un bon outil de travail. Avoir une grille détaillée aide à regarder un portail dans son ensemble, tant du côté fonctionnel que du côté de la structuration des données. Cela permet également de sortir du rôle d'audit et de comparatif entre les sites.

L'analyse des portails a été un moyen de travail pour arriver à des propositions de fonctionnalités intéressantes. Je n'ai donc pas réalisé de synthèse complète de cette analyse. Je me rends compte a posteriori que cela aurait permis de capitaliser le travail réalisé et de mieux l'exploiter lors des prochaines phases du projet. Cela a été partiellement compensé par un tableau général reprenant côte à côte les résultats de l'analyse de chaque portail. Ce

document pourra servir de référence à l'équipe du CND avant de se reporter, si besoin, au document détaillant l'analyse d'un portail particulier.

Dans la réalisation du cahier des charges fonctionnel, je me suis principalement concentrée sur l'interface de la future bibliothèque numérique. Je n'ai pas étudié en détail la partie structuration des données en dehors de la préconisation d'utiliser un modèle conceptuel, le modèle FRBR (cf. partie 3 sections 1.3.3.1 et 3.2.1). J'ai réalisé quelques maquettes permettant de visualiser et matérialiser les solutions proposées.

La principale difficulté rencontrée pour proposer des idées concrètes est venue du fait qu'il fallait imaginer la future bibliothèque numérique. En effet, ma mission a constitué une phase préliminaire pour permettre à l'équipe de s'approprier le projet, pour qu'elle puisse ensuite construire collectivement sa politique éditoriale.

En conclusion, pour transformer ses collections numériques en véritable bibliothèque numérique, le CND a de sérieux atouts. Un important travail d'éditorialisation permet déjà de faire découvrir des œuvres, des artistes et des documents particulièrement intéressants. Il pourra servir de point d'entrée dans les collections. Et c'est justement un des besoins clé du public, notamment amateur et professionnel. Le public a besoin de trouver les ressources de la médiathèque depuis des moteurs de recherche web. Une fois sur le portail, il peut rebondir vers d'autres documents, d'autres découvertes ou recherches. Pour arriver à ce résultat, il est nécessaire d'avoir des ressources bien structurées, richement décrites. Cela permet un bon référencement par les moteurs de recherche. Cela permet également d'offrir une bibliothèque numérique riche de liens entre les ressources, proposant des sélections de documents régulièrement actualisées grâce aux recherches implicites (cf. section 3.2).

La médiathèque traite les fonds suivant les standards actuels des bibliothèques et des archives. Néanmoins tous ses documents numériques ne sont pas richement décrits au sens documentaire, notamment ceux qui proviennent des fonds d'archives. Les descriptions ne sont pas reliées par des référentiels communs et le thesaurus n'est pas utilisé pour les collections numériques. Par exemple les dossiers documentaires ne sont pas reliés aux documents numériques qui les composent. Tous les documents numériques ne sont pas prêts à être efficacement exploités dans la future bibliothèque numérique. Un travail important devra être réalisé au fur et à mesure de la diffusion des documents. Comment y arriver avec les petits moyens de la médiathèque du CND ? Beaucoup de bibliothèques spécialisées et organismes culturels en général se pose la même question. Une des pistes actuellement étudiée et expérimentée est la convergence des données, avec le web de données.

Troisième partie

Le web de données, un atout pour la valorisation des données culturelles

De plus en plus d'expérimentations liées au web de données ou au web sémantique sont mises en place dans des projets de valorisation de collections culturelles numériques. De ces expérimentations, des projets opérationnels commencent à voir le jour. Nous pouvons citer par exemple *data.bnf.fr* de la BNF ou le Centre Pompidou Virtuel. Quel est l'intérêt pour des organismes culturels de rejoindre le mouvement ? On pourrait le résumer ainsi : partager des données enrichies afin d'améliorer la diffusion de collections culturelles en ligne. Comment la médiathèque du CND pourrait elle aussi profiter de cette évolution technologique pour mieux accroître sa visibilité et rendre ses collections numériques plus accessibles ? L'objet de cette dernière partie est d'apporter des pistes de réflexion pour y répondre.

Nous commencerons par présenter quelques éléments permettant de comprendre sur quoi reposent les projets se basant sur le web de données (section 1). Mais si comprendre comment les données devraient être structurées, reliées, explicitées est une chose, imaginer comment il est possible de le mettre en place en est une autre. C'est pourquoi j'ai choisi de présenter des exemples concrets de projets actuellement en cours dans le monde culturel (section 2). Nous pourrons alors voir comment le web de données pourrait permettre de répondre à certaines problématiques de la future bibliothèque numérique du CND (section 3).

1 Le web de données : quelques éléments de compréhension

1.1 Quel est l'apport du web de données dans la culture ?

Le web de données offre des nouvelles possibilités de valorisation des contenus et le développement de services innovants. Ces deux sujets sont au centre des enjeux actuels que nous avons développés précédemment (cf. partie 1 sections 2.2 et 2.3). Le web de données s'appuie sur ce qu'est au cœur du travail documentaire : les vocabulaires, les schémas de données et les métadonnées. Muriel AMAR et Bruno MENON résume :

« cet environnement permet de mettre en avant les contenus pertinents, en améliorant la précision des moteurs de recherche et la qualité des résultats présents. Il facilite l'exploration de grands corpus, en alimentant les facettes et en fournissant des outils de navigation, et enrichit leur compréhension. Face à ces multiples possibilités, les enjeux de la création d'un web de concepts organisés concernent alors tous les acteurs du marché de l'information et de la connaissance. »³³

L'enjeu est également stratégique. Dans la période actuelle de restriction budgétaire, les organismes se doivent de montrer qu'ils servent de plus en plus efficacement les missions qu'on leur a confié alors que le travail s'accroît, à personnel constant ou en baisse. Les bibliothèques, archives ou musées voient un accroissement de leur collection et une diversification de leurs services. Les technologies du web de données sont un moyen et une grande opportunité dans ce contexte pour tendre à la diffusion et à la réutilisation des données culturelles.

De plus les technologies et outils du web de données offrent la promesse d'accroître la visibilité en ligne et les opportunités de lier les notices entre elles dans un contexte où prévalent les moteurs de recherche web comme Google. Or actuellement, peu de ressources issues des bases de données des organismes culturels sont indexées par les moteurs de recherche et quand elles le sont, l'ordre des résultats n'est pas forcément adapté. De fait les utilisateurs risquent de ne pas trouver la ressource qu'ils cherchent s'ils ne savent pas au préalable quel organisme est dépositaire de la ressource, voir dans quelle base de données elle se trouve (cf. partie 1 section 2.4).

Concrètement le web de données offre notamment les moyens de faciliter l'accès à l'information, de relier des ressources entre elles et de les rendre visibles sur le web par les moteurs de recherche et en poussant les contenus vers les utilisateurs. Le web des données

³³ [15] AMAR. p. 22-23

permet de sortir des silos documentaires, d'utiliser les données des autres organismes et de faire profiter de ses données. Nous verrons à la section 2 de façon concrète, à partir de projets expérimentaux actuels, comment cela est mis en place.

1.2 Quelques éléments de compréhension

1.2.1 Web de données ou web sémantique ?

Tim Berners-Lee a commencé à parler du concept de web sémantique entre 1999 et 2001. Le concept a repris et développé le World Wide Web Consortium (W3C). Ce dernier en parle en 2001 en ces termes : *« l'information reçoit une signification bien définie, permettant ainsi une meilleure collaboration entre hommes et machines [et] ajoute de la structure aux contenus significatifs des pages Web »*.

Aujourd'hui, le W3C le voit comme un *« cadre permettant aux données d'être partagées et réutilisées par-delà les limites applicatives, organisationnelles ou communautaires »*. Depuis d'autres définitions et expressions sont apparues. Tim Berners-Lee lui-même préfère le terme de « web de données ».

Le web des données, appelé parfois web 3.0, s'inscrit dans la suite des évolutions du web, qui se sont développées de façon parallèle. Pour comprendre, revenons brièvement sur celles-ci. Avec le web 1.0, appelé également le web documentaire, les contenus sont organisés par page, avec des liens hypertextes entre elles. Il n'y a pas de traitement automatique sur le sens des documents ou leur contexte. Seuls l'encodage et les chaînes de caractères sont pris en compte par les machines.

L'idée d'enrichir ses traitements pour apporter une connaissance supplémentaire et donner du sens au contenu naît. Mais comment faire ? Deux idées apparaissent. La première est d'impliquer les utilisateurs pour qu'ils donnent du sens au contenu : c'est le web 2.0, appelé parfois web contributif. La deuxième approche est de chercher à présenter et expliciter le sens de façon à permettre aux outils de s'en servir. L'objectif est double : représenter les données et concevoir les outils capables de les interpréter. C'est l'idée du web sémantique ou web 3.0.

Mais comment faire ? L'idée émerge de pouvoir enrichir le web d'une connaissance supplémentaire en partant du sens de ces contenus, afin de permettre des traitements plus précis et performants. La question est alors de savoir comment apporter la connaissance et en imprégner les processus.

Faut-il alors parler de « web sémantique » ou de « web de données » ? Au départ, l'expression employée par Tim Berners-Lee est web sémantique, avec la perspective

d'expliciter le sens des contenus. En réfléchissant à comment y arriver, la perspective a évolué vers la possibilité qui nous est offerte d'ouvrir les silos de données de toutes tailles, depuis l'application individuelle de notre carnet d'adresses, jusqu'aux immenses bases de génomique, et de les échanger, de les relier, de les mélanger selon nos besoins. Cela correspond au « web de données » mais l'expression web sémantique est restée dans la pratique même si elle ne correspond pas aux possibilités réelles de mises en place. Voyons maintenant plus en détail ce qu'entend Tim Berners-Lee par « web de données ».

1.2.2 Les 5 degrés d'intégration selon Tim Berners-Lee

L'intégration optimale au web de données est pour Tim Berners-Lee des données accessibles, structurées, dans un format non-propritaire, identifiées et liées entre elles sémantiquement. Il définit des niveaux permettant d'évaluer l'intégration au web de données. Ces niveaux, appelés aussi « 5 étoiles », sont :

- * données accessibles,
c'est-à-dire mises à disposition sur internet,
- * * données structurées,
c'est-à-dire organisées selon un schéma de structuration défini,
- * * * format non propriétaire,
avec un schéma de structuration ouvert et connu,
- * * * * données identifiées,
avec des vocabulaires ouverts permettant d'identifier les contenus,
- * * * * * liées entre elles sémantiquement,
grâce à leur identification.

Notons cependant que garder ses données fermées n'empêche pas un organisme d'appliquer les principes du web de données. Il est possible de les rendre uniquement visibles sur Internet sans les ouvrir. L'ouverture limitée aux bases d'un organisme est également réalisable. Dans ce cas, l'application est limitée aux données internes à la structure et à des données ouvertes disponibles à l'extérieur. Si cette pratique fait débat comme allant contre la philosophie même du web de données, elle est néanmoins répandue dans le monde de la culture.

1.2.3 Techniquement cela se passe comment ?

La représentation formelle de la connaissance de façon compréhensible pour la machine se fait avant tout avec des langages et des outils³⁴. Cette représentation formelle nécessite d'identifier les éléments de connaissances (documents, personnes, sujets...) et de déclarer

³⁴ [16] Athena

les liens entre les ressources. L'identification se fait au moyen d'un identifiant appelé URI pour *Uniform Resource Identifier*. Les liens sont faits sous forme de déclaration qui s'exprime par un triplet sujet/prédicat/objet en suivant le langage RDF (*Resource Description Framework*). Le langage RDF sert de base à des extensions en fonction de besoins particuliers (cf. section 1.3.3.1 sur les modèles conceptuels FRBR).

Pour décrire les données, il faut des référentiels qui expriment les concepts. Leur représentation dans le web de données peut se faire par des ontologies au format OWL ou par le système d'organisation des connaissances SKOS (*Simple Knowledge Organization System*).

Pour exploiter les données, c'est-à-dire les rechercher, en ajouter, modifier ou supprimer, des requêtes sont nécessaires. Cela se fait avec le standard SPARQL qui est à la fois un protocole, un langage de requêtes et un formalisme pour l'expression des résultats.

Par ces explications un peu techniques, nous voyons que le web de données n'est pas une « boîte noire » mais une succession de couches technologiques permettant de répondre chacune à un besoin précis.

1.3 De quoi a-t-on besoin pour y arriver ?

1.3.1 De données interopérables

Le but étant que de pouvoir lier et réutiliser les données, l'interopérabilité est indispensable. Elle se définit comme la capacité à s'assurer que le système, les procédures et la culture d'une organisation sont gérés de façon à maximiser les opportunités d'échanges et de réutilisations de l'information, aussi bien en interne qu'avec l'extérieur. Paul Miller définit six types d'interopérabilité, nous allons retenir les deux principaux dans notre contexte, à savoir l'interopérabilité technique et l'interopérabilité sémantique.

L'interopérabilité technique est facilitée par l'utilisation de standards communs par exemple pour les types de fichier et les métadonnées. L'interopérabilité sémantique est facilitée tant par la qualification des liens entre les métadonnées et que par l'utilisation de vocabulaires communs pour la description des données, des thesaurus et des référentiels communs par exemple. Nous verrons dans les points suivants comment les standards de description, les modèles conceptuels et les référentiels communs permettent de rendre interopérables les données.

1.3.2 De référentiels communs

Les vocabulaires scientifiques et techniques garantissent la qualité des données et permettent dans le contexte du web de données une navigation intelligente. Actuellement, les référentiels ne sont pas les mêmes selon les pays, les langues et les organismes qui gèrent les données. Si on veut les exposer sur Internet, pouvoir croiser les données entre elles, il faut pouvoir s'adapter à une variété de terminologies sans perdre pour autant la sienne. C'est pourquoi des référentiels communs sont indispensables.

Prenons un exemple basique pour illustrer cela. Selon les domaines, on désigne les objets par leur usage ou par leur forme. On va alors parler d'« assiette à soupe » ou d' « assiette creuse » pour parler du même objet. Dans ce cas précis, l'utilisateur sait vraisemblablement qu'il s'agit du même objet. Exprimerait-il alors sa recherche avec une requête incluant les deux expressions ? Peut-être pas... (cf. partie 1 section 2.4.1). Les machines, elles, ont besoin qu'on explicite le lien d'équivalence entre les deux expressions pour rendre les données sémantiquement interopérables. Cela est valable également pour les personnes, comme nous le verrons à la section 3.3 dans un cas d'application du CND.

1.3.3 De qualifier les liens entre les ressources et les métadonnées

Si nous avons besoin de données interopérables et de référentiels communs, nous avons également besoin de qualifier les liens entre les ressources entre elles ainsi qu'entre les ressources et les métadonnées. Traditionnellement les métadonnées sont organisées selon des standards, par exemple par exemple l'Unimarc ou le Dublin Core dans les bibliothèques. Comme nous l'avons vu, le standard de description conforme au web de données est le modèle RDF. Pourquoi les standards habituels ne suffisent-ils plus ? Pour y répondre, revenons sur le concept de métadonnées et de ressources.

Les métadonnées permettent de décrire et donner accès à une ressource qui peut être un texte, un objet physique, un document multimédia, un logiciel, une personne, un concept ou encore des collections de ceux-ci. Les ressources peuvent aussi être des substituts à d'autres ressources : par exemple une image numérique d'un livre, une photographie d'une œuvre. Les métadonnées d'un substitut ne doit pas décrire la ressource originale elle-même. Par ailleurs, les ressources peuvent être liées à d'autres ressources dans un rapport autre que le rapport de ressource-substitut. La relation peut relever de la création, de l'usage ou de la mise à disposition (édition).

Les métadonnées elles-mêmes peuvent être de différentes natures :

- **Administration** pour gérer et administrer les ressources (acquisition, droits, localisation)
- **Description** pour décrire et identifier les ressources
- **Préservation** pour gérer la préservation des ressources (données de migration et condition)
- **Technique** : comment un système fonctionne ou comment les métadonnées se comportent (formats, cryptage, mots de passe)
- **Utilisation** : le niveau et le type d'utilisation des ressources (information de suivi des utilisations et des utilisateurs)

Pour mieux, comprendre, voyons le schéma ci-dessous (en anglais). Il distingue les types de liens entre une ressource et ses ressources dérivées et les types de métadonnées associées.



Figure 3 - Lien entre des ressources et de leurs métadonnées³⁵

Une métadonnée est donc une information structurée au sujet de n'importe quel type de ressource. Elle sert à identifier, décrire, gérer et rendre accessible cette ressource.

³⁵ Source : Athena [16]

L'utilisation d'un modèle conceptuel de structurer les métadonnées et les relations permet de rendre compte de cette complexité, ce que ne peut pas faire un standard de description traditionnel.

Les ressources et leurs relations entre elles étant variées, il paraît aujourd'hui illusoire de trouver un modèle unique qui reflèterait cette richesse. La solution est donc de développer, pour des domaines particuliers, des modèles conceptuels standardisés qui sont des extensions du modèle RDF et de les associer à l'utilisation de vocabulaires communs.

Notons que si les modèles conceptuels sont antérieurs au web de données et au modèle RDF, ils sont devenus un élément clé de la mise en œuvre de projets autour du web de données.

Pour les musées un nouvel modèle est développé au MCC : le modèle HADOC. Nous verrons ce projet à la section 2.1.3). Pour les notices bibliographiques, il s'agit du modèle FRBR que je vous propose de regarder un peu plus en détail maintenant. Il sera ensuite repris pour imaginer son implémentation au CND. Notons toutefois que les modèles conceptuels sont en cours d'évolution actuellement.

1.3.3.1 Un exemple de modèle conceptuel : le modèle FRBR

Regarder le fonctionnement du modèle conceptuel FRBR nous permettra de donner un exemple de liens entre des ressources et de la qualification de ces liens. Nous ne cherchons pas ici à expliquer le modèle FRBR dans toute son étendue.

Sous le modèle FRBR, les métadonnées sont organisées en « entités » organisées en 3 groupes, le premier décrit l'œuvre, le deuxième les personnes et le troisième les concepts et sujets.

Le groupe 1 :

- Œuvre : la création intellectuelle ou artistique
- Expression : la réalisation d'une œuvre
- Manifestation : la matérialisation d'expression d'œuvres
- Item : exemplaire d'une manifestation

Le groupe 2 :

Il rassemble les entités responsables du contenu ou de la production matérielle ou de la gestion juridique des produits.

- Personne
- Collectivité
- Famille

Le groupe 3 :

Il rassemble les entités envisagées comme sujet d'une œuvre.

- Concept
- Objet
- Evènement
- Lieu

L'on peut comparer chaque entité à une petite brique. Pour décrire un document, on peut décrire les briques pertinentes pour notre ressource puis les assembler ensemble en explicitant le rapport entre les briques. Prenons un petit exemple :

Une œuvre : Les Rêves dansants, sur les pas de Pina Bausch

Trois personnes : Pina Bausch, Anne Linsel et Rainer Hoffmann

Si je dis une œuvre « Les Rêves dansants, sur les pas de Pina Bausch » a été réalisé par les personnes Anne Linsel et Rainer Hoffmann et a pour sujet de la personne Pina Bausch, je qualifie les liens entre mes quatre entités (1 œuvre + 3 personnes).

Ceci peut être schématisé de la façon suivante :

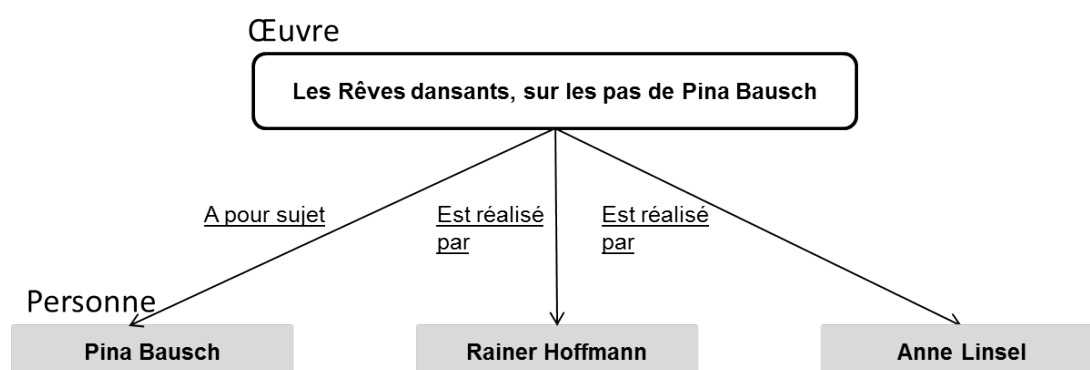


Figure 4 - Les Rêves dansants, sur les pas de Pina Bausch

Les entités peuvent être liées à des référentiels communs. Par exemple la personne Pina Bausch peut être répertoriée dans un référentiel commun lui-même accessible, structuré et dans un format non-propritaire, par exemple le format SKOS.

Comment définir une politique de valorisation de collections numériques issues de fonds d'archives à l'heure du web de données ? Le cas de la médiathèque du CND.

Une même entité peut être utilisée pour décrire plusieurs ressources. L'entité personne Pina Bausch est dans notre exemple Sujet, mais elle peut également être Auteur ou Interprète d'une autre Entité œuvre.

2 Exemples d'applications actuelles dans le domaine culturel

Les trois structures à l'avant-garde en France dans le domaine sont le MCC, la BNF et le Centre Pompidou. Des collaborations ont également lieu entre Europeana, le MCC et la Cité de la Musique, notamment. Les projets décrits dans cette section sont des exemples d'applications actuelles du web de données dans le monde de la culture. Certains sont déjà opérationnels, d'autres le seront prochainement. J'ai choisi de vous présenter les projets pour lesquels j'ai pu rencontrer les personnes impliquées. Je mentionnerai simplement quelques autres projets. Les projets présentés sont complémentaires les uns des autres. Cela est le reflet de la volonté d'expérimentation et de travail en collaboration dans les organismes culturels.

Les projets s'articulent autour d'une part de la structuration des données et des catalogues et d'autre part les vocabulaires contrôlés et les référentiels.

2.1 Structuration des données et des catalogues

Pour la structuration des données et des catalogues, trois projets visent des objectifs différents : expérimenter en partant de l'existant (*data.bnf.fr*), partager des données entre organismes (OpenCat) et définir un nouveau modèle conceptuel (HADOCC).

2.1.1 data.bnf.fr

Le projet *data.bnf.fr* est lancé en 2011 pour expérimenter les possibilités du web de données pour s'adapter aux nouveaux usages en exploitant les données existantes de la BNF. Quatre objectifs sont visés.

- Sortir des silos des bases de données en reliant les données des différentes bases et proposer un point d'entrée aux catalogues des collections physiques et numériques, provenant des bases bibliographiques mais aussi archivistiques de la BNF,
- Proposer une recherche organisant les résultats par œuvre, artiste et sujet et non plus par document.
- Rendre les données visibles sur Internet, les bases de la BNF n'étant pas indexées par les moteurs de recherche.
- Rendre les données compréhensibles par une machine.

La BNF dispose de listes d'autorité Auteur, Œuvre et Sujet de qualité. La structure de la base data.bnf.fr repose sur celles-ci. Pour être sûr de garantir la qualité et l'homogénéité des données par rapport aux listes d'autorité, le périmètre initial des notices à intégrer est restreint.

Actuellement, environ 60 millions de notices bibliographiques sont recensées dans data.bnf.fr. Elles sont originellement décrites suivant le format MARC ou le modèle EAD. Pourtant en se basant sur les listes d'autorité et la structuration des notices d'origine, les liens entre les ressources et la nature de ces liens ont pu être exploités pour créer des notices conformes au modèle FRBR avec aujourd'hui environ 200.000 auteurs et 20.000 œuvres et sujets.

Cela permet de créer des pages d'œuvres, d'auteurs et de sujet en organisant les données par événements, éditions ou type de documents à l'intérieur d'une page d'œuvre. La structuration des données permet également de les présenter de différentes façons, par exemple en les plaçant sur une carte ou une frise.

Par ailleurs, les données sont lisibles par les machines et réutilisables. Pour cela *data.bnf.fr* regroupe et expose en ligne ses données provenant de sources hétérogènes et les converties au format RDF. Elles sont indexées par les moteurs de recherche et reliées à d'autres ressources qu'elles soient internes à la BNF ou externes (Sudoc, VIAF, Wikipedia...). Pour faciliter la diffusion et la réutilisation des données, les jeux de données sont disponibles au téléchargement en RDF.

Les buts poursuivis ont été atteints et cela vient confirmer les choix stratégiques effectués. Initialement un projet expérimental puis data.bnf.fr est devenu un service à part entier dans l'offre de la BNF. En 2012, plus de 80% des visites du site proviennent d'un moteur de recherche, avec un taux de rebond vers une autre application de la BNF d'environ 70%. Le projet a reçu en 2013 le Data Intelligence Awards et, conjointement avec Gallica, le Stanford Prize for Innovation in Research Libraries (SPIRL).

2.1.2 Le projet OpenCat

Le projet OpenCat vient dans le prolongement de data.bnf.fr. C'est un prototype de catalogue fonctionnant dans le web de données. Ce projet est un prolongement de data.bnf.fr. La BNF expose depuis 2011 ses données pour permettre à d'autres de les réutiliser, les bibliothèques de lectures publiques par exemple. Partant du constat que celles-ci n'avaient pas forcément les moyens de se lancer dans l'aventure, la BNF en partenariat avec la société Logica et la bibliothèque municipale de Fresnes a développé une plateforme

commune permettant de consolider les données de différentes provenances. Celles de data.bnf.fr, du catalogue de Fresnes et d'autres sources sont présentées ensemble et organisées par œuvre, auteur ou sujet Rameau.

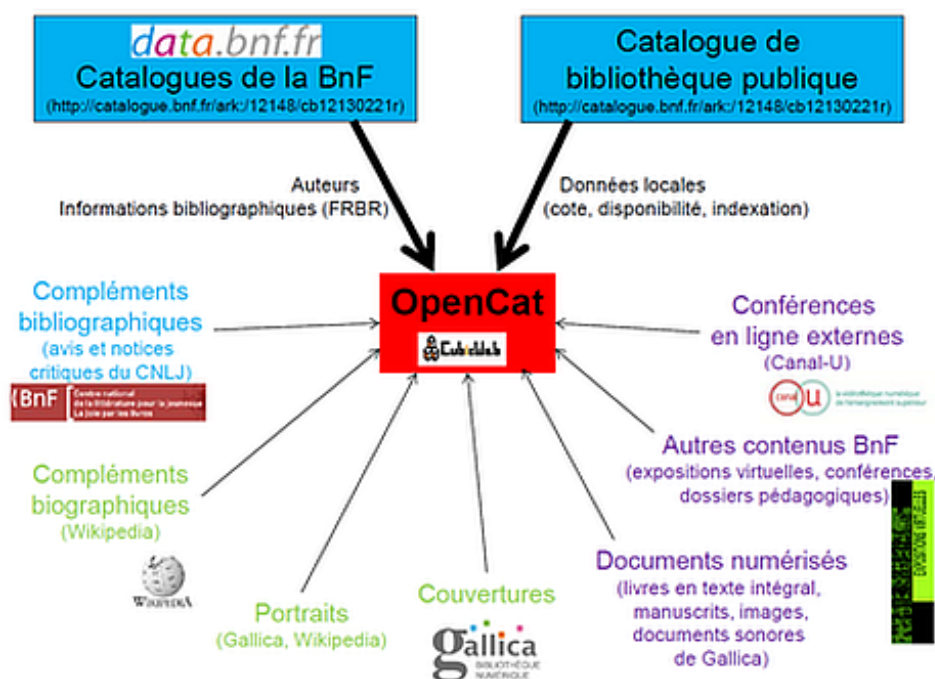


Figure 5 - Schéma de la plateforme OpenCat³⁶

L'objectif du projet est de montrer qu'il est possible de mutualiser le catalogue et de proposer une interface utilisateur différente du catalogue de bibliothèque. La nouvelle interface vise à la recherche et proposer des ressources directement exploitables en ligne en plus des notices bibliographiques.

Prenons un exemple. Si je recherche dans le catalogue OpenCat de Fresnes « Maupassant », on me propose l'auteur Guy de Maupassant (1850-1893). La page correspondante affiche une courte présentation de l'auteur issue de data.bnf.fr, la liste des documents disponibles à la bibliothèque de Fresnes, des liens vers d'autres ressources comme les livres numérisés de Gallica et les expositions virtuelles de la BNF. Les documents disponibles s'affichent également sous forme de frise.

OpenCat a été développé sur la plateforme Open Source CubicWeb et peut donc être librement réutilisé. Chaque bibliothèque peut choisir les données qu'elle souhaite lier et

³⁶ Source : MCC [23]

adapter l'interface et la façon de naviguer dans les données de son choix. La prochaine étape du projet est la mise à disposition de la plateforme afin de permettre aux bibliothèques qui le souhaitent de tester OpenCat avec ses propres données.

La réutilisation des notices bibliographiques est particulièrement intéressante pour les collections de bibliothèques généralistes. En effet, celles-ci n'adaptent pas les descriptions à un public ou des besoins particulier(s).

Le projet étant récent, nous ne pouvons pas encore tirer de bilan.

2.1.3 Le programme HADOC

Le MCC initie en 2008 le programme HADOC (HARmonisation de la production des DONnées Culturelles). Le point de départ est le constat de la grande hétérogénéité des données documentaires produites par le ministère et ses partenaires, hétérogénéité qui pose des difficultés en terme de consolidation, de diffusion et de réutilisation de ces données.

Ce projet, transverse à l'ensemble des directions du ministère, a pour objectif de fournir un cadre normatif à la production des données culturelles et d'en outiller la mise en œuvre. Il vise à accroître la qualité des données et à optimiser les processus de production. Pour cela, trois leviers d'actions sont identifiés :

- l'harmonisation des données et la mise en œuvre de référentiels partagés
- l'harmonisation des processus de production de ces données et la réalisation d'une nouvelle chaîne de production permettant de remédier à l'obsolescence fonctionnelle et technique de la chaîne de production actuelle
- la généralisation de pratiques normalisées.

Pour l'harmonisation des données, le choix est fait de proposer un nouveau modèle conceptuel qui devrait permettre de répondre aux besoins des bibliothèques, des musées et des archives. Pourtant ce modèle harmonisé ne se limite pas à un modèle d'échange.

Pour favoriser l'harmonisation des processus de production de données, une modélisation en UML complète la structuration des données et propose un modèle d'implémentation pour les applications informatiques. Le modèle est adaptable aux contextes métiers et aux contraintes locales.

Le modèle permet en outre dès la phase de production une structuration de l'information interopérable avec les modèles d'échange et de diffusion actuellement utilisés dans le domaine culturel, par exemple les modèles CIDOC (ou CIDOC CRM) et ses extensions,

utilisés dans les musées, le modèle EDM ou le schéma XML LIDO (tous deux développés par Europeana).

La première version du modèle est sortie en juillet 2013. Nous ne pouvons donc pas encore donner d'exemples concrets d'utilisations ou d'applications ou dresser un premier bilan. Nous pouvons souligner quelques points. Parmi les points positifs, premièrement nous pouvons citer son champ d'application à toutes les données culturelles. Cela permet une harmonisation des données entre les bibliothèques, les musées et les archives. Deuxièmement la mise à disposition d'un modèle d'implémentation pour les outils informatiques. Alors que les éditeurs de logiciels ne proposent pas encore de solutions compatibles avec les nouveaux modèles conceptuels, cette offre peut les inciter à implémenter le modèle HADOC. Par contre, il n'existe pas actuellement de version RDF du modèle HADOC, dans sa version actuelle. Cela limite donc fortement la compréhension des machines pour les données décrites sous le modèle HADOC.

2.2 Les vocabulaires contrôlés et les référentiels

Pour les vocabulaires contrôlés et les référentiels, trois projets visent des objectifs complémentaires : faire converger des collections avec un thesaurus multilingue (MIMO), donner du sens aux données avec le tagging sémantique (Joconde Lab) et offrir un outil de gestion des thesaurus (GINCO).

2.2.1 MIMO

Le projet du Musical Instruments Museum Online (MIMO) est un projet Europeana qui court de 2009 à 2011. Onze partenaires y participent, dont la médiathèque de la Cité de la Musique. Chaque partenaire apporte ses ressources (soit au total 45.000 images d'instruments de musique, 1.800 fichiers audio pour découvrir le son des instruments et 300 clips vidéo.) ainsi que leurs métadonnées.

Le problème à résoudre était simple : pouvoir rechercher dans une langue des photographies d'instruments de musique de tous les musées partenaires. La solution adoptée a été de définir une classification et un thesaurus communs et multilingues³⁷, d'ajouter le terme voulu dans une langue et d'utiliser les termes traduits dans la recherche.

³⁷ Les langues en question sont l'allemand, l'anglais, le français, l'italien, le néerlandais et le suédois.

La classification MIMO des instruments de musique est basée sur le plan de classement numérique de référence dans la musique développé par Erich Moritz von Hornbostel et Curt³⁸. Y ont été ajoutés les instruments de musique électronique.

Le thesaurus MIMO mis au point organise en trois niveaux hiérarchique les termes génériques et spécifiques. Les termes scientifiques et ceux employés par le grand public ont été également ajouté afin de faciliter les recherches des différents publics.

Des listes d'autorités de facteurs d'instruments complètent également les vocabulaires. Les données sont intégrées dans une base de données dédiée (« MIMO-DB ») au format standard LIDO³⁹ puis générées au format HTML et RDF pour leur intégration dans Europeana.

Les données exposées sont ouvertes et réutilisables. La Cité de la Musique les utilise dans son portail *education.citedelamusique.fr*, une offre payante de contenu en ligne. Le projet MIMO a renforcé l'interopérabilité technique et sémantique dans le domaine des instruments de musique et ce dans plusieurs langues et en proposant un outil pour les alignements.

2.2.2 Joconde Lab

Joconde Lab est un projet expérimental de tagging sémantique appliqué au corpus de la base Joconde du ministère de la Culture et de la Communication. Il est dirigé par la délégation générale à la langue française et aux langues de France en collaboration avec le service des musées de France et le département des programmes numériques. Actuellement en cours de développement, le projet devrait sortir début 2014.

Ce projet reprend l'expérimentation menée dans le cadre du portail Histoire des Arts HDA-Lab en l'appliquant à une autre échelle avec les 400.000 notices d'œuvres des musées de France, notice regroupée dans la base Joconde. Cette base de données, créée en 1975 et gérée par le MCC, répertorie dans un catalogue collectif les collections des musées français. Son catalogue a reçu plus de 1 million de visites du monde entier en 2012.

Concrètement le projet consiste à appliquer des liens conformes aux technologies du web de données entre les mots-clés documentaires de la base Joconde et les entrées de DBpedia en français. L'un des principaux objectifs de ce liage consiste à internationaliser l'accès à

³⁸ La version révisée par Jeremy Montagu en 2009 du plan de classement numérique de référence dans la musique développé par Erich Moritz von Hornbostel et Curt

³⁹ Un schéma XML pour le moissonnage des données des musées développé par l'International Council of Museums ICOM.

Joconde par ces mots-clés. Les liens avec Dbpédia⁴⁰ permettent en effet, d'obtenir automatiquement des équivalents de ces mots-clés dans diverses langues.

La particularité de la base Joconde est d'agrèger les notices d'œuvre de tous les musées de France. Chaque établissement utilise ses propres vocabulaires contrôlés comme des thesaurus ou des listes simples. Les termes sont ensuite alignés sur les thesaurus de la base Joconde, thesaurus en cours d'alignement avec les termes de DBpédia. Le travail se fait de façon semi-automatique : une première passe avec l'outil informatique permet de suggérer rapidement des alignements de termes qui sont ensuite validés manuellement par une personne qui en assure la cohérence intellectuelle. Par effet collatéral, cet alignement permet de se rendre compte des termes non présents dans DBpédia et donc des lacunes de Wikipédia sur des sujets pour lesquels le personnel des musées seraient compétent pour apporter leurs contributions.

2.2.3 Outil GINCO

L'outil de Gestion Informatisée de Nomenclatures Collaboratives et Ouvertes GINCO est un projet du MCC. Cet outil est développé en code ouvert qui permet de concevoir, produire et diffuser ces vocabulaires. GINCO s'appuie sur les normes et standards les plus récents que constituent la norme ISO 25964-1:2011 et les langages SKOS. Il permettra notamment de mettre en œuvre la différenciation entre concept et terme. GINCO offrira les fonctionnalités classiques d'un outil de gestion de thésaurus :

- création de termes et de concepts,
- gestion des relations d'équivalence entre termes, de hiérarchie (mono- ou poly-hiérarchie) et d'association entre concepts
- gestion de tableaux de concepts (facettes) et de groupes de concepts (par thématique, domaine etc.),
- typage des relations et des notes.

L'outil permettra les éditions alphabétiques et hiérarchiques ainsi que l'import et l'export de vocabulaires au format SKOS/RDF.

⁴⁰ DBpedia est un projet d'exploration et extraction automatiques de données dérivées de Wikipédia. Son principe est de proposer une version structurée et sous forme de données normalisées au format du web sémantique des contenus encyclopédiques de chaque fiche encyclopédique. DBpedia vise aussi à relier à Wikipédia (et inversement) des ensembles d'autres données ouvertes provenant du Web des données [source : Wikipedia]

Une première version de l'application est en voie d'achèvement et offre d'ores et déjà l'ensemble des fonctionnalités nécessaires à la production d'un vocabulaire. Dans les prochaines versions, seront ajoutées les fonctions collaboratives, les fonctions d'alignement entre vocabulaires ou vers des ressources externes ainsi que la prise en compte de la dimension multilingue dans les éditions.

2.2.4 Autres projets

D'autres projets sont actuellement en cours, dans des organismes français ou au niveau européen.

Au niveau des organismes français, l'on peut citer le Centre Pompidou Virtuel sorti en 2012. L'objectif fixé est de devenir le premier centre de ressources au monde sur l'art moderne et contemporain. Le Centre Pompidou virtuel est de donner accès à l'ensemble des contenus culturels produits par le Centre Pompidou à destination de ses publics : images et dossiers des œuvres de la collection, mais aussi dossiers pédagogiques, interviews vidéo d'artistes et de commissaires, captations de colloques et de conférences, archives, etc. L'ensemble de contenus est organisé selon une architecture sémantique avec une navigation dans des concepts avec une interface simple et intuitive⁴¹.

Au niveau européen, nous pouvons citer deux projets développés par Europeana et Michael Culture Association. Le projet Europeana Photography (fin du projet en 2015) vise à mettre à disposition des photographies visant à combler le manque de ressources iconographiques datant du début de la photographie dans Europeana. Par ailleurs, les métadonnées d'Europeana de ce projet seront rendues disponibles dans les 11 langues parlées par les partenaires.

Le projet Linked Heritage actuellement en cours vise à contribuer au développement dans Europeana de la quantité des contenus, provenant aussi bien du secteur privé que du secteur public, à améliorer la qualité du contenu en terme de richesses de métadonnées, réutilisations potentielles et unicité ainsi qu'à améliorer la recherche, l'extraction et l'utilisation des contenus d'Europeana.

⁴¹ N'ayant pu rencontrer les personnes impliquées dans ce projet, je ne développe pas ce point.

2.3 Quel bilan pour les premiers projets ?

Ces projets constituent les premiers projets de mise en place du web de données dans le monde de la culture. Nous pouvons remarquer qu'ils visent à mieux répondre aux besoins des utilisateurs et proposent des solutions pour les producteurs ou gestionnaires de collections culturelles.

Du côté des utilisateurs, les projets cherchent à améliorer l'accès aux collections et s'adapter aux besoins des utilisateurs avec les éléments suivants :

- rassembler des ressources issues de différentes bases et directement exploitables (data.bnf.fr, OpenCat et MIMO via Europeana),
- adapter les vocabulaires à différents utilisateurs et niveaux de langue (MIMO et Joconde Lab),
- proposer des notices multilingues (Joconde Lab et MIMO).

Parallèlement les projets visent à pallier aux problématiques techniques et aux moyens limités des organismes culturels, notamment les petites structures avec les solutions suivantes :

- Obtenir des données compatibles avec le web de données en partant de la structuration actuelle des données (au format RDF pour data.bnf.fr et OpenCat et au format SKOS pour les vocabulaires contrôlés pour Joconde Lab).
- Proposer un modèle conceptuel harmonisé pour les données des musées, bibliothèques et archives (HADOC)
- Proposer des solutions informatiques « web de données » pour aider à la valorisation des données de petites structures (OpenCat, Ginco et Joconde Lab) ou une support à une implémentation logicielle (HADOC).

Nous pouvons souligner que les projets menés en collaboration avec différents organismes sont nombreux. Certains sont des projets menés à l'échelle européenne et internationale. Il serait intéressant de regarder ce qui se fait dans les autres pays.

Chaque projet apporte des solutions à une partie du problème. Cela illustre la dynamique et la complémentarité des projets pour explorer le web de données dans toutes les directions. Cela montre aussi la complexité du sujet. Les changements ne peuvent se faire que par étape. Cela montre également la pluralité des approches.

Il est pour le moment trop tôt pour dresser un bilan de l'implémentation de la plupart des projets et de leurs adaptations par le public et les autres organismes. Nous pouvons déjà

émettre le souhait que de plus en plus d'organismes s'emparent de ces projets pour les implémenter, les adapter. Plus les expériences seront riches par leur nombre et leur diversité, plus les retours d'expériences seront nombreux, plus les projets pourront progresser afin de répondre aux besoins des utilisateurs et aux problématiques des organismes culturels. Le Centre Georges Pompidou, avec son Centre Pompidou Virtuel propose d'ores et déjà un projet opérationnel.

3 Quelle application du web de données pour la bibliothèque numérique du CND ?

Après avoir vu le fonctionnement du le web de données et passé en revue les projets actuellement mis en place dans le monde de la culture, voyons maintenant comment la future bibliothèque numérique de la médiathèque du CND pourrait tirer parti de la mise en réseau des données.

Nous allons en particulier étudier comment le web de données pourraient permettre d'améliorer ou faciliter les points suivants :

- Créer des liens entre les ressources des collections numériques,
- Donner un accès unifié à des ressources situées dans des bases distinctes de la médiathèque, du CND et à l'extérieur du CND,
- Mieux décrire les documents,
- Améliorer la mise en contexte,
- Accroître la pertinence et la facilité d'utilisation de la recherche grâce à la structuration des données

La valorisation des collections implique de reprendre l'ensemble de la chaîne de traitement documentaire et de réfléchir à la politique de valorisation et de collaboration du CND dans son ensemble. L'objectif de mon propos est d'apporter des pistes pour alimenter la réflexion sur la future bibliothèque numérique. Pour cela, nous présenterons des idées de convergence en se plaçant au niveau des collections, ensuite à celui des descriptions des collections, enfin au niveau des vocabulaires contrôlés. Pour illustrer notre propos et simplifier la présentation, nous présenterons la convergence à travers l'exemple d'une œuvre précise : le ballet La Bayadère de Noureiev.

3.1 Faire converger les collections

Offrir un point d'accès unifié aux collections numériques de la médiathèque est l'objectif premier du projet de bibliothèque numérique du CND. Le web de données contribue à mettre en place une exploitation transversale des fonds du CND et à sortir de la logique des silos, d'abord en interne puis avec les autres. Nous savons techniquement comment cela pouvait se faire à travers les exemples des projets data.bnf.fr et OpenCat. Regardons quelles collections pourraient converger et être reliées pour en faciliter la découverte et l'exploitation.

Les collections numériques du CND n'étant pas exhaustives sur le sujet de la Danse, la bibliothèque numérique du CND pourrait mettre en avant ses ressources numériques tout en proposant des ressources complémentaires à celles-ci et ainsi offrir plus d'information quant à une époque, un mouvement artistique, une œuvre, un artiste, une compagnie. Les ressources peuvent être de différentes natures, par exemple des livres ou articles de presse, des photos ou des vidéos de personnes, spectacles, des costumes, des partitions, de la musique, des peintures ou dessins représentant des artistes ou des spectacles, des programmes de spectacles passés ou futurs mais aussi des pages ou sites Internet... La liste est infinie.

Quelles pourraient être ses ressources complémentaires ? Tout d'abord, celles que la médiathèque du CND conserve en propre, soit : les dossiers documentaires présentant des artistes, la notation en danse ou encore des mouvements artistiques ainsi que les ressources de la collection principale de la médiathèque pour les ouvrages, articles de presse ou vidéos. Elles pourraient aussi venir du CND lui-même hors médiathèque, notamment les présentations de spectacles, artistes ou compagnies, les ressources professionnelles. D'autres organismes possédant des collections complémentaires peuvent apporter leur concours comme la Cité de la Musique, le Centre National du Costume de Scène et les Opéras de France ainsi que les fondations autour d'un artiste ou une compagnie, les Centre Chorégraphiques Nationaux.... Pour des ressources avec des documents élargissant le sujet, il est possible de se tourner vers la BNF avec Gallica, mais aussi Wikipédia pour la mise en contexte et l'INA pour les ressources audiovisuelles diffusées à la radio et télévision françaises.

3.1.1 Les ressources autour du ballet La Bayadère de Noureev

Prenons notre exemple concret : le ballet La Bayadère. La Bayadère est un ballet en trois actes et sept tableaux chorégraphié par Marius Petipa sur une musique de Léon Minkus. Sa création a eu lieu le 23 janvier 1877 au Théâtre Bolchoï Kamenny de Saint-Petersbourg. Une des productions les plus connues est celle du danseur et chorégraphe Noureev présentée au public au Palais Garnier le 8 octobre 1992. Quelles sont les ressources et métadonnées disponibles au CND et ailleurs autour de cette œuvre ?

Les fonds d'archives

La médiathèque a dans ses fonds d'archives une quarantaine de documents concernant cette œuvre, dont :

- des extraits de la Bayadère antérieurs à la version de 1992 et l'enregistrement d'une émission de télévision, issus des dépôts de documents audiovisuels de la Rudolf Nureyev Foundation,
- des captations du spectacle de 1992 et un documentaire en hommage à Rudolf Noureev diffusé sur la télévision française avec des extraits des répétitions de la production de 1992, issus du Fonds audiovisuel Douce François-Freitas,
- des partitions chorégraphiques basées sur la musique de Minkus et des partitions d'extraits musicaux de la Bayadère de Minkus, issus de la collection de partitions musicales
- Des programmes de spectacles ayant produit la Bayadère ou des extraits de la Bayadère, issus de la collection de programmes.

La collection principale

Dans sa collection principale, le CND a également d'ouvrages, articles, films et programmes concernant cette œuvre et l'artiste Noureev, nous pouvons en citer quelques-uns :

- Le programme détaillé édité par l'Opéra de Paris
- Des DVD de spectacles incluant la Bayadère
- Des DVD de reportage sur le spectacle ou les artistes de l'opéra de Paris ayant dansé La Bayadère
- Un DVD réalisé par Charles Picq "Le tour du monde en 80 danses", l'article parle de l'œuvre la Bayadère
- Un article pédagogique accompagnant le DVD de Charles Picq
- Des livres en langues française et anglaise sur les décors de scène, sur les ballets

Dossiers documentaires :

Dans les dossiers documentaires du CND :

- Un montage vidéo à partir des croquis et notes d'Andy de Groat sur le ballet La Bayadère⁴².

En dehors de la médiathèque :

⁴² Lien :

http://mediatheque.cnd.fr/spip.php?page=andy_de_groat_video&id_article=327&id_video=333

En dehors de la médiathèque du CND, d'autres ressources sont disponibles sur Internet. Nous pouvons notamment citer :

- La page Wikipedia : http://fr.wikipedia.org/wiki/La_Bayad%C3%A8re permettant de replacer l'œuvre dans son contexte de création
- La page de la fondation Noureev : <http://www.noureev.org/rudolf-noureev-ballets-grands-roles/bayadere-rudolf-noureev-solor-petipa>
- Les photos des représentations à l'Opéra de Paris conservées avec les archives de l'organisme à la BNF. Sur Gallica, nous pouvons retrouver des recueils de photographies : <http://gallica.bnf.fr/ark:/12148/btv1b9080506q.r=noureev.langFR>
- Les émissions audiovisuelles françaises archivées par l'Ina traitant de l'œuvre, de Noureev, par exemple les extraits du spectacle disponible sur le site En Scène : <http://fresques.ina.fr/en-scenes/fiche-media/Scenes10100/la-bayadere.html> ou encore le reportage de France 2 sur l'exposition "Mille et une nuits" au Centre national du costume de scène de Moulins, incluant des extraits du spectacle "La Bayadère" par NOUREEV à Opéra Garnier : <http://www.ina.fr/video/3706385001039/musee-du-costume-mille-et-une-nuits-video.html>.
- La collection Noureev du CNCS (Centre National du Costume de Scène) où la fondation Noureev a déposé tous les costumes⁴³. Notons que sur la collection présentée en ligne, il n'est pas indiqué quels costumes sont liés à quels spectacles.

Par ailleurs, nous pourrions aussi rechercher des informations dans les collections de la médiathèque du CND et à l'extérieur du CND sur Noureev lui-même, les autres artistes, l'œuvre d'origine de Marius Petipa, les ballets russes, les ballets en général, mais aussi les dans les collections muséales les tableaux et peintures qui ont inspiré les décors ou représentant des artistes ou le spectacle lui-même.

Si grâce au web de données, la bibliothèque numérique relie toutes ses ressources, d'abord en interne puis avec les ressources extérieures au CND, l'expérience utilisateur sera grandement enrichie. Il pourra découvrir les collections dans toute leur richesse en passant d'un document à l'autre avec des informations d'un document pour mettre en contexte d'autres documents.

⁴³ Lien : <http://www.cnsc.fr/expositions/collection-noureev/espace-permanent-rudolf-noureev/41>

3.2 Enrichir les métadonnées

Si la bibliothèque numérique peut profiter de la convergence des ressources, elle peut également profiter de la description de celles-ci pour enrichir la présentation des œuvres, des artistes, des contextes de création ou de production, des mouvements artistiques... Voyons comment.

Les descriptions des ressources se trouvent avec les ressources que nous venons de présenter de façon non-exhaustive. Nous n'allons pas les représenter ici. Mais imaginons que toutes ses ressources de la médiathèque du CND, comme du CND et de l'extérieur soient décrits selon le même modèle conceptuel ou des modèles interopérables. Le CND pourrait profiter de ces métadonnées pour enrichir les descriptions de ses ressources numériques. Comment les métadonnées pourraient-elles être réutilisées ?

Comme nous l'avons vu dans la première partie de ce mémoire, les descriptions des pièces de fonds d'archives répondent à une logique d'inventaire et ne sont pas identiques à celles dont le CND a besoin pour sa bibliothèque numérique.

Revenons à notre ballet la Bayadère. La vidéo de captation du spectacle d'octobre 1992 est décrite de la façon suivante :

Archives liées à Rudolf Noureev, Fonds audiovisuel Douce François-Freitas, 1991-1992, Captations de spectacles chorégraphiques

Intitulé : Bayadère (La) (02/10/1992).

Dates extrêmes : 2 octobre 1992

Description matérielle :

- métrage ou nombre de documents : 2
- support : Casette vidéo
- données techniques : VHS (NTSC)
- durée 130

Présentation du contenu : contient la première et la soirée qui s'en suit

Chorégraphes : Noureev, Rudolf (1938-1993)

Noms d'organismes : Ballet de l'Opéra de Paris

Chorégraphies : Bayadère (La)

La vidéo du même spectacle proposée par l'Ina est décrite de la façon suivante (avec le contexte tronqué ici) :

Résumé : Ballet de Marius Petipa, sur une musique de Léon Minkus, remonté par Rudolf Noureev. La Bayadère raconte l'histoire de la bayadère (mot qui désigne une « danseuse » ou « servante » du temple) Nikiya et du valeureux guerrier Solor.

Type de média : Intégrale - Vidéo - Captation

Date de diffusion : 08 octobre 1992

Œuvre(s) : La Bayadère (Léon Minkus - Compositeur(trice) / Rudolf Noureev - Chorégraphe / Marius Petipa - Chorégraphe / Alexandre Tarta - Réalisation)

Artistes et personnalités :

- Léon Minkus - Compositeur(trice)
- Rudolf Noureev - Chorégraphe
- Marius Petipa - Chorégraphe
- Ezio Frigerio - Décorateur(trice)
- Michel Quéval - Direction musicale
- Franca Squarciapino - Costumier(ère)
- Isabelle Guérin - Danseur
- Laurent Hilaire - Danseur
- Elisabeth Platel - Danseur
- Vinicio Cheli - Lumières

Thèmes : Danse > Danse classique

Lieux : Europe > France > Ile-de-France > Paris > Opéra Garnier

Contexte :

Le ballet mythique dans la version de Rudolf Noureev, enregistré à l'opéra de Paris, Palais Garnier. Un classique.

La Bayadère raconte l'histoire de la bayadère (mot qui désigne une « danseuse » ou « servante » du temple) Nikiya et du valeureux guerrier Solor qui se sont mutuellement juré une fidélité éternelle. Le grand brahmane, également amoureux de Nikiya, apprend ses relations avec Solor. De plus, Dugmanta, rajah de Golconda, souhaite fiancer Solor à sa fille Gamzatti - Hamsatti dans la production originale.

[...]

Les ombres de Nikiya et de Solor sont réunies et emportées vers l'Himalaya.

Des informations contenues dans cette notice pourraient être réutilisées pour décrire les ressources du CND dans la bibliothèque numérique. Et ce n'est qu'un exemple. La Cité de la Musique pourrait enrichir d'informations sur la musique de Minkus, l'Opéra de Paris sur l'établissement où a eu lieu le spectacle, etc.

Seulement les informations proviennent de différents sites et toutes les informations ne sont pas intéressantes pour la bibliothèque numérique du CND. Comment faire pour utiliser seulement l'information pertinente alors qu'elle peut provenir de différentes sources ? En utilisant un modèle conceptuel de structuration de données, comme nous allons le voir maintenant avec le modèle FRBR.

3.2.1 L'organisation des métadonnées avec le modèle FRBR

Nous cherchons à comprendre l'apport d'un modèle conceptuel pour notre bibliothèque numérique et non de présenter le choix d'un modèle particulier et de son implémentation au CND. Nous prenons ici le modèle FRBR à titre d'exemple. Le modèle a été présenté à la section 1.3.3.1.

Revenons à notre exemple, le Ballet La Bayadère de Noureev et voyons plus précisément comme pourrait être décrit suivant le modèle FRBR la notice bibliographique de la cassette vidéo avec les captations du spectacle conservée dans le fonds d'archives Douce François-Freitas⁴⁴.

L'œuvre de Noureev est inspirée de l'œuvre de Petipa. Nous avons donc deux œuvres avec un lien « d'inspiration »

Œuvre de Petipa (entité du groupe 1) :

- L'œuvre originale la Bayadère, de Marius Petipa
- Son expression : la représentation au Théâtre Bolchoï Kamenny de Saint-Pétersbourg le 23 janvier 1877
- Une de ses manifestations : programme édité pour l'occasion
- Un de ses items : le programme conservé dans un fonds d'archives du théâtre⁴⁵, par exemple.

Nous pouvons schématiser les entités du groupe 1 de la façon suivante :

⁴⁴ Pour cela, nous n'allons pas présenter comment cela devrait être implémenté mais présenter comment cela fonctionne en simplifiant le nombre d'entités.

⁴⁵ Ceci est une supposition, je n'ai pas vérifié cette information.

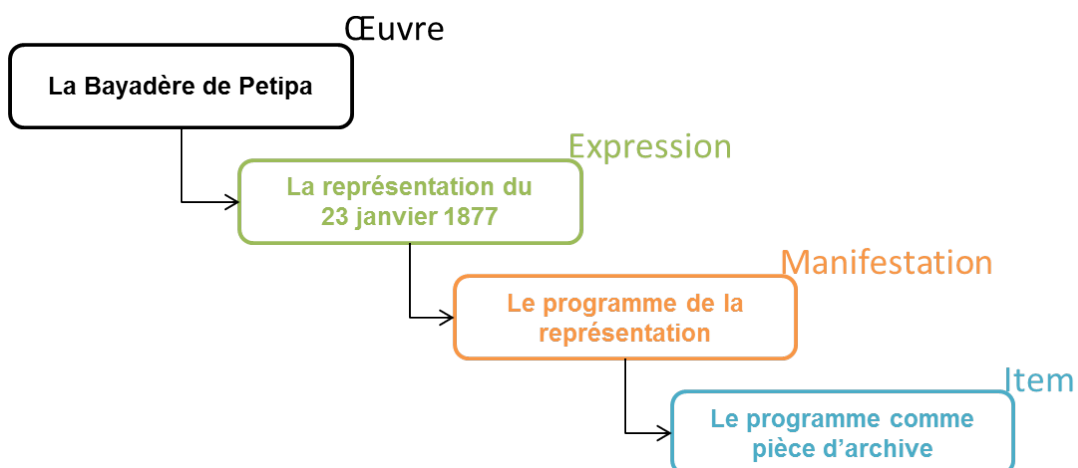


Figure 6 - Illustration des entités FRBR de La Bayadère de Petipa

Œuvre de Noreev (entité du groupe 1) :

- L'œuvre originale la Bayadère, de Noreev, œuvre inspirée de celle de Petipa.
- Son expression : la représentation au Palais Garnier le 8 Octobre 1992
- Une de ses manifestations : les captations du spectacle réalisé par une personne ou une équipe particulière
- Un de ses items : la cassette vidéo VHS conservé dans un fonds d'archives de la médiathèque du CND.
- Nous pouvons schématiser ces entités du groupe 1 de la façon suivante :

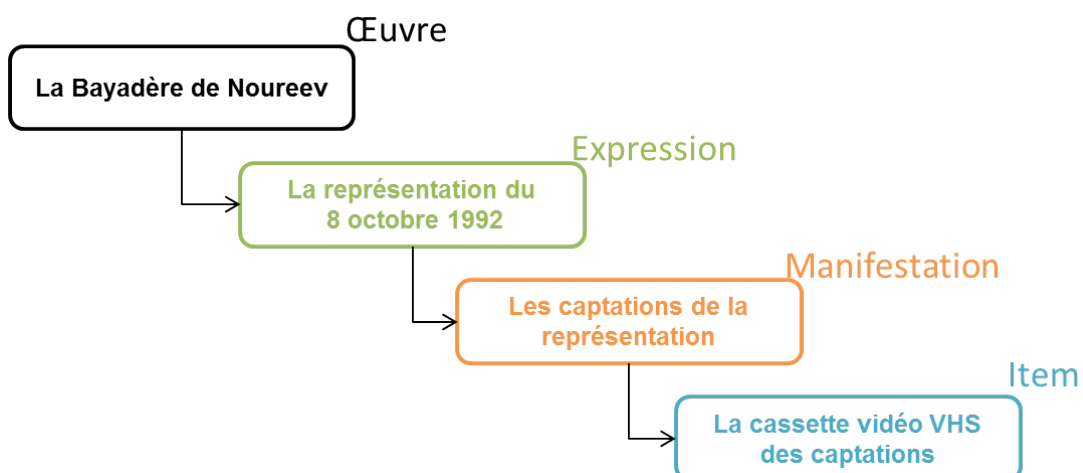


Figure 7 - Illustration des entités FRBR de La Bayadère de Noreev

Entités du groupe 2 :

Dans notre exemple nous aurions :

- Personne : Pepita, Noureev, Douce François-Freitas ...
- Collectivité : le Théâtre Bolchoï Kamenny, le CND...

Le groupe 3 :

Les entités envisagées comme sujet d'une œuvre. Nous n'allons pas détailler ici les entités car cela soulève la question des référentiels communs et des vocabulaires contrôlés dont nous parlerons à la section 3.3. L'idée est de comprendre comment les liens peuvent être faits et qualifiés.

- Concept
- Objet
- Evènement
- Lieu

La notice bibliographique de la cassette vidéo avec les captations du spectacle conservée dans le fonds d'archives Douce François-Freitas est schématisée à la figure Figure 1 - Schéma d'organisation des archives. Là encore tous les liens ne sont pas indiqués pour des raisons de clarté, il manque les liens vers les personnes artistes, liens qui peuvent être qualifiés suivant le rôle des personnes.

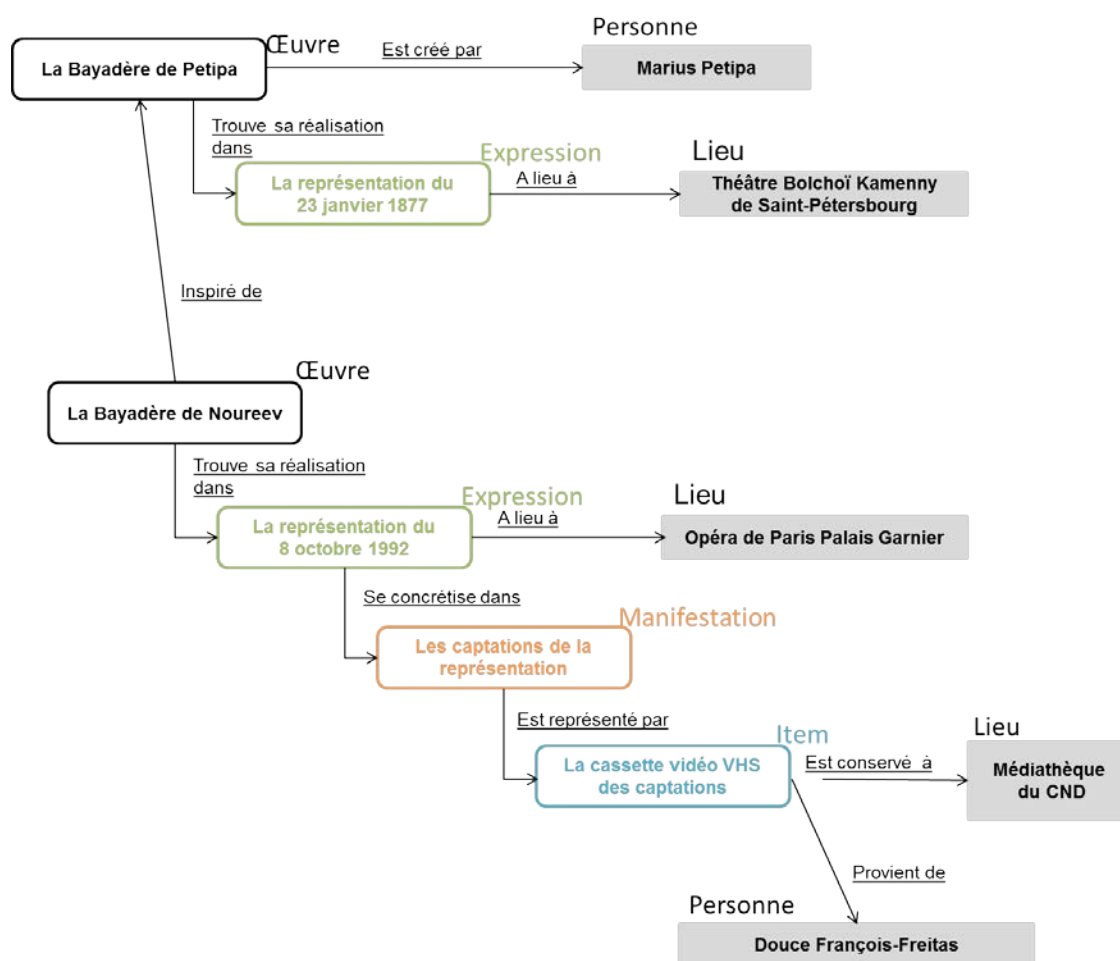


Figure 8 - Illustration de La Bayadère de Nouriev avec quelques liens qualifiés

Ce schéma qui paraît bien compliqué à première vue, peut se lire comme suit :

L'œuvre La Bayadère de Nouriev est inspiré de l'œuvre La Bayadère de Petipa, a été donnée en représentation le 8 octobre 1992 à l'Opéra de Paris au Paris Garnier, il existe des captations de cette représentation sous forme d'une cassette vidéo qui est conservée à la médiathèque du CND comme pièce d'archives provenant du fonds Douce François-Freitas. L'œuvre La Bayadère de Petipa a été donnée en représentation le 23 janvier 1877 au théâtre Boschoï Kamenny de Saint Petersburg.

Revenons au modèle FRBR pour aller plus loin. Les métadonnées sont organisées par entité, chaque entité étant décrite par une liste de champs⁴⁶. Nous présentons ici quelques champs particulièrement pour décrire une œuvre :

- titre de l'œuvre

⁴⁶ Les champs correspondent aux « attributs » dans le modèle FRBR

- forme de l'œuvre
- date de l'œuvre
- complétude visée
- public visé
- contexte de l'œuvre

3.2.2 La réutilisation des métadonnées

Cette organisation en briques d'information facilite la réutilisation des données. En effet, elle permet de réutiliser les métadonnées d'une ressource brique par brique selon les besoins et les liens qualifiant les relations entre les briques d'information.

Revenons à notre exemple. Si nous voulons décrire les captations du spectacle de 1992 issues du Fonds audiovisuel Douce François-Freitas, nous pouvons réutiliser des entités décrivant l'œuvre la Bayadère de Petipa et son expression au Bolsoï. Nous pouvons également utiliser la description de l'Ina concernant l'œuvre Noureev et sa manifestation au Palais Garnier en 1992. Nous pouvons aussi utiliser les liens qui qualifient les relations entre les entités.

Par ailleurs, les descriptions de l'œuvre de Noureev et sa manifestation peuvent être utilisées comme sujet des documentaires sur le spectacle proposés par le CND sous forme de DVD dans la collection principale.

En conclusion, la réutilisation de métadonnées avec un modèle conceptuel permet d'enrichir les descriptions des ressources en choisissant finement les métadonnées que l'on souhaite réutiliser et en capitalisant les descriptions par entité pour ne pas dupliquer l'information. Est-ce que l'utilisateur souhaite et a besoin d'une description fine des ressources ?

En effet, un des objectifs de la bibliothèque numérique du CND est de présenter ses collections en plaçant le document au centre. L'utilisateur n'est pas intéressé par une notice bibliographique présentant une liste infinie de champs ni même une vingtaine de champs. L'intérêt d'une telle structuration réside dans l'exploitation des collections. Elle permettra de naviguer par liens entre les documents, de proposer des recherches implicites et des facettes permettant d'élargir, d'affiner sur un sujet, un lieu, un artiste, une œuvre, etc. Cela permettra en outre d'améliorer la recherche, notamment pour créer des recherches implicites permettant des sélections de documents régulièrement renouveler. Cela permettra d'exploiter le travail des dossiers documentaires pour les ressources en lien avec un dossier.

3.3 Harmoniser les vocabulaires contrôlés

Avant de reprendre les descriptions d'un document pour décrire un autre document, notamment quand il provient d'un autre organisme, il faut s'assurer que les ressources parlent bien de la même chose. Techniquement, nous avons vu à la section 1.2.3 que cela fonctionnait avec les URI. Mais comment faire pour être sûr que l'on prenne le bon URI ? En alignant ses vocabulaires contrôlés sur des référentiels communs et en utilisant des identifiants uniques. Nous prendrons comme exemple le référentiel commun VIAF. Nous aurions également pu prendre comme référence l'ISNI (International Standard Name Identifier) qui a pour objet l'identification de tout contributeur de contenus culturels, scientifiques et techniques à travers le monde.

Nous avons détaillé à la section 1.3.2 l'intérêt des vocabulaires contrôlés et des référentiels communs. Voyons maintenant concrètement l'apport pour le CND.

La médiathèque utilise déjà des listes d'autorité pour les personnes, les organismes et les œuvres aussi qu'un thesaurus pour les sujets (dans la collection principale uniquement). Revenons à notre exemple de la Bayadère et regardons ce que l'on trouve dans les référentiels VIAF au sujet de l'œuvre, de l'expression de l'œuvre et des auteurs Nureev et Minkus. Pour simplifier, prenons l'exemple de la liste d'autorité personnes, le raisonnement restant le même pour les autres listes d'autorité.

L'entité recherchée	L'entité trouvée dans le VIAF
Noureev (le danseur et chorégraphe)	<p>Variation des formes du nom :</p> <ul style="list-style-type: none">• Nureyev, Rudolf, 1938-1993• Noureev, Rudolf Gametovitch, 1938-1993• Nure'ev, Rudolf (Rudolf Hametovič), 1938-1993• Nureev, Rudol'f Ch. 1938-1993• Noureïev, Rudolf• Nureyev, Rudolf• Нуреев, Рудольф Гаметович, 1938-1993• Нуреев, Рудольф, 1938-1993• Nureyev, Rudolph• Nurejev, Rudolf, 1938-1993• Nureyef, Rudolf, 1938-1993• Nureyev, Rudolph Hametovitch, 1938-1993• Нуреев, Рудольф Хамитович 1938-1993 <p>Des exemples d'œuvres associées :</p> <ul style="list-style-type: none">• La bayadère• Romeo and Juliet• Nureyev: An Autobiography with Pictures• van. Noerejev, 1994:• Nureyev: dancing through darkness

L'entité recherchée	L'entité trouvée dans le VIAF
Minkus (le compositeur)	<p>Variation des formes du nom :</p> <ul style="list-style-type: none"> • Minkus, Ludwig, 1826-1917 • Minkus, Léon • Minkus, L, (1826-1917) <p>Des exemples d'œuvres associées :</p> <ul style="list-style-type: none"> • Don Quixote • Bayadere • Paquita • Don Quixote. Grand pas de deux • La source • Don Quijote • Don Quixote. Sel. • Don Quichotte • Ballet Russe de Monte Carlo • Ra bayadēru La Bayadere

L'on remarque dans ces deux exemples que le référentiel propre en plus de la version retenue les différentes orthographes des noms, années de naissance et de mort ainsi et les liens entre les auteurs et les œuvres existent déjà et sont qualifiés. En liant les données à des référentiels communs, cela permet de désambigüiser les noms et les sujets et d'enrichir ses données avec leurs variations. La recherche dans la bibliothèque numérique et sur Internet est alors plus efficace.

En liant ses données à des référentiels communs, l'interopérabilité sera renforcée et les résultats de recherches gagneront en pertinence. Par ailleurs, en exposant ses propres données, le CND pourra faire profiter d'autres organismes de ses connaissances sur le monde de la danse.

Les référentiels communs sont complémentaires des thesaurus et autres vocabulaires contrôlés qui permettent de décrire finement les ressources en fonction des publics. Ils concernent les « objets métiers » ayant un haut niveau de partageabilité. Souvent les thesaurus particuliers peuvent être ensuite alignés sur des référentiels afin de créer des ponts entre les données internes à une structure et l'extérieur, notamment les moteurs de recherche. C'est ce que nous avons vu avec le projet Joconde Lab. En exposant leurs thesaurus, les organismes peuvent profiter des compétences des uns et des autres. Par exemple, en mettant à disposition son thesaurus de la danse, d'autres structures seraient susceptibles d'affiner leur indexation sur une base commune et enrichir le thesaurus. Par ailleurs, le CND pourrait bénéficier des vocabulaires contrôlés de structures voisines par exemple pour le vocabulaire spécialisé dans la description de ressources audiovisuelles mais

également dans la description de spectacles vivants comme les déplacements sur scène, la musique, les costumes, etc.

En conclusion nous voyons à travers l'exemple du ballet la Bayadère les possibilités que représente le web de données pour la valorisation des collections numériques du CND. La médiathèque du CND pourrait lier ses ressources entre ses propres bases ainsi qu'avec des bases du CND lui-même et de l'extérieur. En sortant des silos des bases de données (en intra), il est possible de présenter ses collections en augmentant la visibilité et l'accessibilité de la valorisation déjà existante, notamment éditoriale.

Cela permet également d'augmenter son attractivité en proposant une plus grande richesse de collections grâce à un accès à des documents complémentaires. D'autres organismes pourraient également utiliser des documents du CND dans leurs bibliothèques numériques et ainsi susciter des visites sur la bibliothèque numérique du CND.

D'autre part, l'adoption d'un modèle conceptuel compatible avec le modèle RDF permettrait de rendre les données compréhensibles par une machine et d'en accroître l'exploitabilité pour les utilisateurs. Par ailleurs, partager des métadonnées et se lier aux référentiels communs, améliore la description des données et leur mise contexte. Cela permet en outre d'augmenter leur exploitabilité et de les rendre accessibles dans une interface facilement utilisables grâce aux sélections, liens et recherches implicites. Tous ces éléments permettraient d'accroître la visibilité des collections sur Internet et la visibilité du CND et de sa médiathèque tout en s'adaptant aux usages du public.

Si la description des possibilités offerte est séduisante, les explications techniques montrent qu'il ne s'agit pas de réaliser une simple conversion des données. La convergence des données fait partie d'une politique globale de valorisation et nécessite de repenser le processus de production des données. Comment le CND pourrait-il rejoindre le web de données ? Nous pouvons essayer d'apporter quelques éléments de réponses tout en se gardant de prévoir un plan d'actions. En effet, la mise en place, si elle se fait, dépendra de l'avancement des projets portés par le MCC et la BNF, du développement d'outils informatiques compatibles et de la politique des autres organismes culturels. Par ailleurs, le choix de rejoindre le web de données et la stratégie de convergences ne peut être que le fruit d'une réflexion globale de la politique de la médiathèque. Nous pouvons toutefois présenter quelques pistes.

Classiquement, la première étape consiste à définir un périmètre restreint, tout en se basant sur les atouts identifiés. Le choix pourrait porter sur des données internes à la médiathèque. La description approfondie de données pour une réutilisation en interne porterait sur une

sélection restreinte de documents, par exemple les documents liés aux dossiers documentaires. Cela permettrait de mettre en avant le travail déjà existant de description et de mise en contexte des documents. Par ailleurs, des partenariats sur des projets ciblés pourraient être mis en place, notamment autour de référentiels, de vocabulaires contrôlés et du tagging sémantique.

La deuxième étape pourra être de choisir un nombre limité de partenaires avec lesquels mettre en place une ouverture des ressources, des métadonnées en commençant par la mise en place de référentiels communs. Pour cela, une partie des organismes dont nous avons étudié les solutions actuelles pour être ses partenaires, par exemple la Cité de la Musique, le Rassemblement des Opéras de France...

Quelle que soit la stratégie choisie, les projets actuels autour du web de données présentés à la section 2 pourraient permettre au CND d'avancer dans la mise en place du web de données et de tester, avec ses collections, les possibilités réelles offertes pour la valorisation de ses propres collections. Les projets présentés montrent en effet qu'ils permettent de répondre aux besoins de la bibliothèque numérique du CND. Ils permettent un accès simplifié aux ressources et une meilleure ergonomie en offrant des liens entre les ressources, des résultats de recherche pertinents et une visibilité sur Internet. C'est notamment le cas pour les projets data.bnf.fr et MIMO. Par ailleurs des projets visent à mettre à disposition des outils nécessaires à l'intégration des technologies du web de données. Le projet GINCO pourrait permettre au CND d'adapter son thesaurus au web de données et de le partager. Le projet OpenCat pourrait offrir une solution pour lier ses données avec celles d'autres organismes, notamment de la BNF et cela, sans utiliser la partie catalogage commun qui intéresse moins le CND de fait des spécificités de ses collections numériques. La méthodologie et les outils mis en place dans le cadre du projet MIMO et Joconde Lab pourraient également permettre au CND d'enrichir et d'aligner le thesaurus du CND avec d'autres. Les projets présentés seront testés, complétés et suivis par d'autres. Il sera donc intéressant de suivre le bilan et l'évolution de ces projets ainsi que les projets qui ne manqueront pas d'être mis en place dans le futur.

Conclusion

Pour répondre à la problématique posée en introduction, concentrons-nous d'abord sur le projet de bibliothèque numérique du CND pour ensuite voir quels enseignements en tirer pour d'autres organismes culturels.

Les enjeux de la valorisation des collections numériques sont d'abord politiques. Au niveau national, il s'agit de répondre aux besoins de préservation et de diffusion de la connaissance et de favoriser la création artistique. La valorisation est également un sujet potentiellement stratégique pour chaque organisme culturel. En effet, il leur est nécessaire de rendre accessibles leurs collections, de se faire connaître ou de renforcer leur position en se plaçant sur le chemin de leur public, c'est-à-dire sur Internet. Pour y arriver, cela implique de s'adapter à ses usages.

Alors, concrètement, comment la médiathèque du Centre National de la Danse peut valoriser ses collections numériques principalement issues de fonds d'archives ? Cela passe par la diffusion des documents numériques sur des portails documentaires sous forme de collections. Comme les fonds d'archives sont traités suivant une logique archivistique, le traitement des documents numériques devra être adapté à un usage autre que l'inventaire d'archives garantissant le respect de son organisation par le producteur du fonds. Actuellement, les documents ne sont pas tous traités pour être consultés en dehors de leur contexte de fonds. Le passage à une exposition sous forme de collections numériques fait d'ailleurs toujours débat chez les archivistes. Il sera intéressant de suivre l'évolution de la discussion et ses aboutissements.

Des réponses peuvent venir des tendances actuelles de valorisation des collections numériques du monde culturel. Quelles sont-elles ? Pour s'adapter à l'évolution des pratiques numériques et informationnelles du public, les organismes ont besoin de proposer de naviguer de façon intuitive dans les collections et de mettre en avant des documents servant de points d'entrée dans les collections. Les portails peuvent prendre différentes formes : des bibliothèques numériques, des plateformes communes mais également des expositions virtuelles. On peut citer également l'exemple d'organismes qui intègrent leurs collections numériques au site Internet général de leur structure. La Cité de la Musique propose, par exemple, d'écouter des extraits de musique lors de la consultation du programme des concerts, tandis que Le Louvre propose de découvrir les œuvres « incontournables » pour préparer sa visite du musée. Une des tendances est également d'intégrer les technologies du web de données à sa stratégie de valorisation.

Des projets actuels dans ce domaine sont porteurs d'avenir pour une meilleure valorisation des collections. Le web de données permet en effet de lier ses collections mais également

ses métadonnées entre ses propres bases de données mais aussi avec celles de l'extérieur. Cela permet un enrichissement mutuel mais également de mieux répondre au public en terme de richesse de descriptions, mise en contexte et organisation des collections. Cela permet en outre de voir ses contenus indexés par les moteurs de recherche et compréhensibles par une machine.

Le fait de proposer un portail de collections en se basant sur les technologies du web de données permet donc de mieux valoriser ses collections mais aussi de les rendre visibles au milieu d'un océan de données.

La problématique de la médiathèque du CND rejoint la problématique de nombre d'organismes culturels. L'étude menée au CND peut offrir à d'autres organismes culturels des pistes de réflexion pour l'élaboration de leur stratégie de valorisation des collections numériques. Pour se lancer dans un projet de bibliothèque numérique, la première étape est d'analyser les points forts et les points faibles de la valorisation actuelle de son établissement compte tenu des besoins de structuration et d'organisation des contenus. Nous présentons ici quelques points à étudier. Les collections doivent couvrir largement et de façon unique le sujet dédié pour atteindre une taille critique intéressante en vue d'une bibliothèque numérique. Si ce n'est pas le cas, des expositions virtuelles ou une contribution à une plateforme commune sont plus appropriées.

Il est par ailleurs important de souligner que la diffusion des collections seule ne permet pas de répondre aux besoins du public en termes d'aide à la découverte et de mise en contexte des collections. La structuration des documents et leurs descriptions sont nécessaires pour créer des chemins de navigation et une recherche pertinente. Il est également utile d'avoir des documents accompagnés de contenus éditoriaux pour guider la découverte.

Enfin, le développement de bibliothèques numériques intégrant des technologies du web de données nécessite de mettre en place des projets difficilement à la portée d'un organisme de petite taille. Il sera indispensable de nouer des partenariats et de s'appuyer sur les projets de structures plus importantes comme le MMC et la BNF.

La clé de la réussite d'un projet réside aussi dans sa capacité à rencontrer son public. En effet, si le public ne s'approprie pas le portail de collections numériques, aussi intéressant soit-il, le projet ne vivra pas. Dans les tendances actuelles de la valorisation des collections, deux approches permettent de créer du lien avec le public en lui donnant la possibilité de s'approprier le projet. Il s'agit de la médiation numérique et du *crowdsourcing*. La médiation numérique permet aux organismes et aux usagers de faire la promotion des collections sur les média sociaux, soit en signalant ou rediffusant du contenu existant soit en créant des

sélections et autres parcours de découverte. Le *crowdsourcing* permet d'améliorer la description de documents et peut s'avérer particulièrement utile pour les documents dits « orphelins ». Un groupe de travail composé d'acteurs du MCC et d'organismes culturels vient d'être créé pour adapter les pratiques collaboratives du web 2.0 au niveau de granularité du web de données et permettre une collaboration au niveau de la donnée. Ce projet fait partie de la feuille de route stratégique annoncée le 7 novembre 2013 par la Ministre Aurélie Filippetti lors de la journée « Automne Numérique ».

La médiation numérique et le *crowdsourcing* permettent de créer une dynamique vertueuse de pérennisation d'un portail. Il est notable que les expériences sont particulièrement positives quand elles s'appuient sur une petite communauté dynamique. Pour les organismes culturels spécialisés, l'émergence d'une telle communauté paraît tout à fait réaliste.

Toutefois, dans les projets de valorisation de collections numériques, une contrainte importante existe : le respect des droits d'auteurs, qui courent jusqu'à 70 ans après la mort de l'auteur. Dans ce domaine le MCC défend une exception culturelle limitant l'ouverture des données au respect des droits d'auteurs. Cette contrainte est particulièrement complexe dans le domaine du spectacle vivant, notamment dans les documents d'archives où il y a un empilement d'auteurs. Les fonds récents sont souvent ceux qui intéressent le plus le public. Ils sont également les plus riches : les moyens technologiques accessibles permettant de capter le son et le mouvement sont récents.

Pourtant, cette contrainte n'empêche pas des avancées importantes dans la diffusion des contenus culturels sur Internet et dans l'ouverture des données. D'une part, les projets peuvent être adaptés pour répondre au besoin du public en offrant une visibilité sur Internet à une partie des collections, qui pourra s'accroître au fur et à mesure de l'acquisition de documents avec droits de diffusion et du traitement juridique des documents anciens. En mettant en place des projets de valorisation sur Internet, la diffusion des collections numériques se répandra et l'impact de cette diffusion pourra être évalué. Cela contribuera sûrement à faire évoluer les pratiques et les regards, notamment ceux des auteurs ou de leurs ayant-droits.

D'autre part, le MCC a réaffirmé son engagement d'ouvrir les données culturelles lors de la journée du 7 novembre 2013. L'ouverture progressive des données culturelles renforcera les possibilités de réutilisation des documents autorisés et des métadonnées dont l'exposition dépend des organismes culturels. La convergence des données, facilitée par l'utilisation des technologies du web de données, renforcera la visibilité des collections numériques sur Internet et enrichira l'expérience utilisateur sur les portails de collections numériques.

Les enjeux de la valorisation des collections numériques en France rejoignent ceux des autres pays. Des projets sont menés à l'échelle européenne, notamment à travers Europeana. Développer des ponts entre les collections européennes et proposer des contenus multilingues renforcera également la diffusion des connaissances et contribuera aux développements de la création artistique.

Bibliographie

Les références bibliographiques ci-dessous sont retranscrites selon la norme Z44-005 et organisées par thème puis par ordre d'auteur ou par ordre antéchronologique pour la catégorie Les politiques culturelles. Les thèmes sont organisés de la façon suivante :

Compréhension du sujet

Les archives et les bibliothèques

Les politiques culturelles

Les pratiques numériques et informationnelles

Le web de données

Les projets actuels

Portails numériques

Les projets autour du web de données

Data.bnf.fr

OpenCat

HADOC

MIMO

Joconde Lab

Ginco

Compréhension du sujet

Les archives et les bibliothèques

[1] Archives de France. Les archives privées. Manuel pratique et juridique. Christine Nougaret ; Pascal Event (sous la direction de). Paris : La Documentation française, 2008. p. 206. 9782110068521

Ce livre est une référence en archivistique. Il présente l'évolution historique des archives et la place récente des archives privées ainsi que le cadre juridique (chapitre 1 et 2). Le chapitre 5 présente le traitement archivistique et les typologies d'archives privées.

[2] CHABIN, Marie-Anne. L'opposition millénaire archives/bibliothèques a-t-elle toujours un sens à l'ère du numérique ?. In *BBF*. 2012. t 57, n° 5. p. 26-30. 12928399 [consulté le 19 novembre 2013] < <http://bbf.enssib.fr/consulter/bbf-2012-05-0026-006> >

Cet article permet d'approfondir la réflexion sur la distinction entre les archives et les bibliothèques. Il compare l'évolution des fonctions et des missions de ces deux types d'organismes. Il est particulièrement intéressant en complément des ouvrages présentant l'archivistique mentionnés plus haut, en y apportant un regard critique.

[3] FAVIER Jean. Les archives. 5e édition corrigée. Presses Universitaires de France, Paris : 1991. p. 127. Que sais-je ?. 9782130441441

Cet ouvrage dresse un panorama des archives et de leurs traitements. Il est utile comme première approche ou comme rappel avant de consulter des ouvrages plus pointus.

[4] GUYOT Jacques, Les archives audiovisuelles. Histoire, Culture, Politique. Thierry Rolland. Paris : Armand Colin, 2011. p. 192. 2200255954.

Cet ouvrage traite des archives audiovisuelles dans son ensemble. Je me suis principalement appuyée sur la première moitié de cet ouvrage. Il replace les archives audiovisuelles dans l'évolution historique des archives en France (chapitre 1), puis traite les spécificités de ce type d'archives (chapitre 2). Ce livre permet de comprendre les spécificités de l'archivistique audiovisuelle. Toutefois, il traite des archives audiovisuelles sous l'angle de fonds constitués de films et programmes audiovisuelles et non de tout type de document sur support vidéo ou audio.

[5] HILDESHEIMER, Françoise. Les archives privées. Le traitement des archives personnelles, familiales, associatives. Paris : Editions Christian, 2000. p. 93. 2864960435.

Ce livre s'adresse au grand public pour expliquer ce que sont les archives privées et comprendre les spécificités de leur traitement. Il est également utile pour ces explications générales concernant l'archivistique et le respect du fonds.

[6] HIRAUX Françoise (sous la direction de). La valorisation des archives. Une mission, des motivations, des modalités, des collaborations. Enjeux et pratiques actuels. Françoise Mirquet. Louvain-la-Neuve : Academia, 2012. , 192 p. Publications des archives de l'université catholique de Louvain. 9782806100597.

Ce livre est issu d'une dizaine de contributions sur la valorisation des archives à la dixième journée des Archives de l'université catholique de Louvain (2010) rassemblant des professeurs d'histoire, de sciences de l'information et de la communication, d'archivistique, mais aussi de directeur d'archives et d'archivistes. Je me suis particulièrement intéressée à l'introduction et au chapitre 1 qui traitent des enjeux et de la multiplicité des valeurs des archives.

[7] MOUREN Raphaële (sous la direction de). Manuel du patrimoine en bibliothèque. Paris : Éditions du Cercle de la Librairie, 2007. p. 407. Collection Bibliothèques. 9782765409496.

Le premier chapitre est particulièrement intéressant par rapport à notre problématique. Il permet de situer ce qu'est un bien patrimonial et présente la place des bibliothèques dans la conservation et la valorisation de ce patrimoine. Le cinquième chapitre présente en outre les politiques actuelles de valorisation (numérique et physique) et les politiques nationales. En outre il montre la complémentarité des fonds classiques de bibliothèque.

Les politiques culturelles

[8] FILIPPETTI, Aurélie. Ministre de la Culture et de la Communication. Conférence de presse. In Journée Automne Numérique. 7 novembre 2013, Levallois-Perret. 7 novembre 2013. [consulté le 01 décembre 2013] <<http://www.culturecommunication.gouv.fr/Espace-Presses/Communiqués-de-presses/Aurelie-Filippetti-a-lance-le-5-octobre-dernier-l-Automne-numerique-un-evenement-destine-a-favoriser-la-creation-et-l-education-artistique-a-l-heure-du-numerique>>

Le texte de la conférence de presse fait le point sur les priorités du MCC pour la création et l'éducation numériques. Il rappelle les priorités du MCC concernant l'ouverture des données culturelle et l'intégration des technologies du web de données.

[9] Ministère de la Culture et de la Communication. Groupe de travail métadonnées culturelles. Paris : 01 septembre 2013. , [consulté le 17 septembre

2013]. <<http://cblog.culture.fr/projet/2013/09/01/groupe-de-travail-metadonnees-culturelles>>

Ce document annonce le groupe de travail nouvellement créé pour soutenir une politique culturelle numérique opérationnelle autour des métadonnées dans le cadre du Web sémantique. Ses travaux donneront lieu à un guide de recommandations et de positionnement stratégique du MCC qui sera rendu public au cours du premier trimestre 2014.

[10] Ministère de la Culture et de la Communication. Feuille de route Open Data du Ministère de la Culture et de la Communication [PDF]. Paris : avril 2013, [consulté le 17 septembre 2013]. <<http://cblog.culture.fr/wp-content/uploads/2013/07/Feuille-de-route-open-data-MCC.pdf>>

Ce document présente la feuille de route du MCC concernant l'ouverture des données culturelles.

[11] Ministère de la Culture et de la Communication. Appel à projets 2012 « Services numériques culturels innovants » : priorité au tourisme et au patrimoine. Paris, 09 avril 2012. [consulté le 17 septembre 2013]. <<http://www.club-innovation-culture.fr/appele-a-projets-2012-services-numeriques-culturels-innovants-priorite-au-tourisme-et-au-patrimoine>>

Ce document présente les priorités pour l'appel à projets de services numériques innovants.

[12] Ministère de la Culture et de la Communication. Culture Labs. Paris, 2012. [consulté le 23/11/2013] <<http://culturelabs.culture.fr>>

Ce site du MCC traite du soutien à l'innovation numérique du MCC afin de développer des usages culturels numériques innovants. Une page rappelle les objectifs, une autre présente les projets en cours. Ce site permet de suivre les projets actuels et futurs.

Les pratiques numériques et informationnelles

[13] ACCART, Jean-Philippe. Mémento de l'information numérique. Alexis Rivier. Paris : Editions du Cercle de la Librairie, 2012. 184 p. Collection Bibliothèques. 978-2765413325

Ce manuel dresse un panorama du monde numérique en quatre parties : le champ de l'information numérique, les ressources numériques accessibles, les possibilités et méthodologies de recherche ainsi que de l'utilisation de l'information récoltée. Je me suis appuyée sur cet ouvrage pour présenter les évolutions des pratiques numériques et

informationnelles ainsi que l'influence de Google dans la production et la recherche d'information.

[14] comScore. Communiqué de presse. comScore publie une étude 2013 France Digital Future in Focus – Le Digital en France. Paris : 14 mars 2013. [consulté le 18 novembre 2013]

<https://www.comscore.com/fre/actualites_et_evenements/Press_Releases/2013/3/comScore_publie_une_etude_2013_France_Digital_Future_in_Focus_Le_Digital_en_France2>

Ce communiqué présente une étude chiffrée datant de 2013 concernant les pratiques numériques en France.

Le web de données

[15] AMAR, Muriel. Dossier et Bruno Menon (Université Paris 8). Dossier : Web sémantique, web de données : quelle nouvelle donne ?. Bruno Menon. **In** Documentaliste – Sciences de l'Information, Décembre 2011, Volume 48: n°4. p.20 à 61. 00124508. [Articles téléchargeables en PDF, consulté le 15 novembre 2013] <<http://www.adbs.fr/revue-docsi-volume-48-n-4-decembre-2011-dossier-web-semantique-web-de-donnees-quelle-nouvelle-donne--111766.htm>>

Ce dossier présente une vue d'ensemble sur les enjeux, les technologies, les standards et les projets en cours mais aussi un élargissement aux implications dans d'autres domaines et une bibliographie. Cependant, il reste assez complexe, concernant la partie technologique.

[16] Athena WP3 Working Group "Identifying standards and developing recommendations". Digitisation: standards landscape for european museums, archives, libraries. Rome : Ministero per I Beni e le Attività Culturali, 2011. Chapitre 1, p.7-20. eContentPlus. [disponible en PDF, consulté le 15 novembre 2013]

<<http://www.athenaeurope.org/getFile.php?id=435>>

Ce livret présente l'ensemble des formats standards utilisés dans les musées, centres d'archives et bibliothèques européens. Le premier chapitre présente l'intérêt des modèles conceptuels et les relations entre les ressources et les métadonnées.

[17] Bibliothèque Nationale de France. Modèles FRBR, FRAD et FRSAD. 30 mars 2012. [consulté le 19 juillet 2013]

<http://www.bnf.fr/fr/professionnels/modelisation_ontologies/a.modele_FRBR.html>

Présente le modèle FRBR et ses dérivés ainsi que le rapprochement avec le modèle en cours dans les musées, le CIDOC-CRM.

[18] ISAAC, Antoine. Sémantique et interopérabilité [PDF]. In Journée d'études ANFOR / BNF. 28 Mars 2008, Paris. p. 28 [consulté le 16 novembre 2013]

<www.bnf.fr/documents/isaac.pdf>

La présentation a eu lieu dans le cadre d'une journée d'études visant à réfléchir au rôle des référentiels dans l'interopérabilité des systèmes d'information et dans le web sémantique. Elle présente de façon succincte l'intérêt du web sémantique et surtout le fonctionnement des référentiels et de leur alignement. Ses schémas facilitent la compréhension.

[19] WENZ, Romain. « Hypertextualisation » La quête du lien sémantique en bibliothèque. Revue de la BNF, 2012/3 n° 42, p. 36-41. 9782717725346

L'auteur développe l'opportunité que représentent les liens pour les bibliothécaires en explicitant la différence entre les liens hypertextes et les liens dans le web de données. Il présente le fonctionnement des liens pour le projet data.bnf.fr.

Les projets actuels

Portails numériques

[20] Ministère de la Culture et de la Communication. Culture et Recherche n°118-119 automne-hiver 2008-2009. Numérisation du patrimoine culturel [PDF]. Ministère de la Culture. p. 48 [consulté en ligne le 15 novembre 2013]

<<http://www.culture.gouv.fr/culture/editions/r-cr/cr118-119.pdf>>

Ce numéro, consacré à la politique de numérisation du patrimoine culturel, présente les enjeux de la numérisation ainsi que de la valorisation des collections numériques pour les organismes culturels. Il s'appuie sur de nombreux exemples de projets en cours en 2008 en France et au niveau européen. Les projets exposés concernent des portails nationaux, européens et d'organismes culturels ou des plateformes communes. Un article (p. 25) est consacré à la création contemporaine, sujet rarement traité dans la littérature et donc particulièrement intéressant.

Les projets autour du web de données

Data.bnf.fr

[21] BNF. data.bnf.fr. Juillet 2011. [consulté le 9 septembre 2013]

<<http://data.bnf.fr/about>>

Cette page reprend les éléments clés d'information du projet data.bnf.fr, avec une présentation du projet, les chiffres clés actualisés et des informations sur les

technologies utilisées dans le but d'informer le grand public. Cette page présente des liens vers des documents complémentaires et permet de suivre les communications de la BNF sur le projet.

[22] BOULET Vincent. S'appuyer sur la structure des données et les liens : le format MARC comme tremplin pour le Web de données : l'exemple de data.bnf.fr [PDF]. In IFLA WLIC 2013. 8 août 2013, Singapour. p. 17 [consulté le 17 septembre 2013]
<<http://library.ifla.org/250/1/222-boulet-fr.pdf>>

Cet article explique concrètement comment la BNF a transformé ses données du format MARC au format FRBR en se basant sur les listes d'autorité, les classifications et thesaurus et comment les liens ont été qualifiés. Il donne des exemples concrets d'alignements de référentiels entre ceux utilisés à la BNF (p.7-8). Ce document est technique et s'adresse avant tout à des personnes capables de lire données écrites au format MARC et RDF.

OpenCat

[23] BNF. L'expérimentation OpenCat. Paris : 21 octobre 2013. [consulté le 22 septembre 2013] <http://www.bnf.fr/fr/professionnels/modelisation_ontologies/a.opencat.html>

Cette page présente les enjeux et le fonctionnement de l'outil OpenCat. Elle propose également un lien pour tester l'outil avec ses propres collections.

HADOC

[24] BRIATTE Katell. Harmoniser la production des données culturelles ou Comment un projet documentaire devient un projet structurant [PDF]. In Séminaire VocabNomen. 04 mai 2010, Poitiers. p. 62 [consulté le 2 octobre 2013] <http://blogpeda.ac-poitiers.fr/vocabnomen/files/2010/05/1-Harmoniser-les-donnees-culturelles_K.-BRIATTE_20100504.pdf>

L'auteur présente le point de départ du projet HADOC et les problématiques ainsi que les solutions proposées.

[25] Ministère de la Culture et de la Communication. Présentation du programme [HADOC]. Paris : 2 juillet 2013. [consulté le 1 octobre 2013]
<<http://www.culturecommunication.gouv.fr/index.php/Etudes-et-documentation/HADOC/Presentation-du-Programme/Presentation-du-programme>>

Le MCC présente le projet par le MCC de façon officielle.

MIMO

[26] Europeana. MIMO - Musical Instrument Museums Online. La Haye. [consulté le 27 juillet 2013] <<http://pro.europeana.eu/web/mimo>>

Europeana présente brièvement le projet avec les chiffres clés.

[27] BAILLY, Rodolphe. MIMO multilingual thesaurus [PDF]. In Linked Heritage Seminar. 18 avril 2013. [consulté le 29 juillet 2013] <www.linkedheritage.eu/getFile.php?id=467>

Cette présentation a été donnée lors d'un séminaire sur la terminologie et le multilinguisme pour des professionnels européens et français du milieu de la culture et de la recherche autour de la thématique du multilinguisme. L'auteur y explique la problématique de départ du projet MIMO, la mise en place de la solution et l'outil proposé.

Joconde Lab

[28] Ministère de la Culture et de la Communication. Lettre d'information Joconde. Paris : 11 décembre 2012. [consulté le 15/11/2013]

<http://www.culture.gouv.fr/documentation/joconde/fr/apropos/joconde-info-11.pdf>

Cette lettre d'information du portail Joconde présente le projet en cours Joconde Lab.

GINCO

[29] Ministère de la Culture et de la Communication. Le projet GINCO. Paris. [consulté le 1 octobre 2013] <[www.culturecommunication.gouv.fr/Etudes-et-](http://www.culturecommunication.gouv.fr/Etudes-et-documentation/HADOC/Referentiels2/Les-vocabulaires-scientifiques-et-techniques/Le-projet-GINCO)

[documentation/HADOC/Referentiels2/Les-vocabulaires-scientifiques-et-techniques/Le-projet-GINCO](http://www.culturecommunication.gouv.fr/Etudes-et-documentation/HADOC/Referentiels2/Les-vocabulaires-scientifiques-et-techniques/Le-projet-GINCO)>

Le MCC présente le projet par le MCC de façon officielle.

Annexes

Annexe 1 Abréviations

BNF : Bibliothèque Nationale de France

CND : Centre National de la Danse

INA : Institut National d'Audiovisuel

INHA : Institut National de l'Histoire des Arts

MCC : Ministère de la Culture et de la Communication

DPLA : Digital Public Library of America

Annexe 2 Liste des portails et projets étudiés (Partie 2)

Analyse complète

Portail de la Cité de la Musique pour les médiathèques et les conservatoires

[consulté le 01 décembre 2013]

<<http://media.citedelamusique.fr>>

Ce portail donne accès à plus de 45 000 documents pour illustrer l'histoire de la musique et la diversité des styles musicaux. Il propose des enregistrements de concert, des guides d'écoute multimédias, des dossiers pédagogiques et des ressources du musée de la Cité de la Musique. Une partie du portail est en accès gratuit en mode découverte.

Portail des Arts de la Marionnette (PAM)

[consulté le 01 décembre 2013]

<www.artsdelamarionnette.eu/>

Le PAM est une plateforme commune de l'Institut International de la Marionnette, réalisée en collaboration avec plus de 20 partenaires dont la BNF et des compagnies. Le portail présente une première approche des pratiques traditionnelles et de la création contemporaine dans tous les pays du monde. Les documents rassemblés vont des collections des institutions du patrimoine à celles de la création contemporaine à travers six siècles d'histoire de la marionnette en France.

Bibliothèque numérique de l'INHA (Institut National de l'Histoire de l'Art)

[consulté le 01 décembre 2013]

<<http://bibliotheque-numerique.inha.fr/collection>>

La bibliothèque numérique de l'INHA propose la consultation de documents dont les originaux sont conservés à la Bibliothèque de l'Institut National d'Histoire de l'Art (collections Jacques Doucet), la Bibliothèque centrale des musées nationaux (BCMNM) et le Service des collections de l'École nationale supérieure des beaux-arts (ENSBA). Les documents, au nombre d'environ 13 000, sont de nature variée avec notamment des photographies, des estampes, des imprimés, des dessins, des manuscrits.

Bibliothèque numérique de Lyon Numelyo

[consulté le 01 décembre 2013]

<<http://numelyo.bm-lyon.fr/>>

La bibliothèque numérique de la bibliothèque municipale de Lyon donne accès aux différentes ressources numérisées (plus de 60 000 livres, sans compter des centaines de documents de la presse lyonnaise, d'images, manuscrits, etc. progressivement à la disposition du public.).

Musée du Louvre – Collections et départements

[consulté le 01 décembre 2013]

<<http://www.louvre.fr/departements>>

Les « collections et départements », intégrés au site Internet du musée du Louvre, présentent les collections organisées en département avec une brève introduction et une sélection d'œuvres majeures photographiées.

Digital Public Library of America

[consulté le 01 décembre 2013]

<<http://dp.la/>>

La « bibliothèque publique numérique d'Amérique » est décrite comme la plus vaste entreprise de numérisation et de diffusion du patrimoine culturel et scientifique américain. Elle offre l'accès à plus de deux millions de documents issus des plus riches bibliothèques, musées et archives des États-Unis.

Étude complémentaire

Musée Albert Kahn - archives de la planète

[consulté le 01 décembre 2013]

<<http://albert-kahn.hauts-de-seine.net/archives-de-la-planete/presentation/>>

Les Archives de la planète représentent un témoignage photographique et cinématographique unique du quotidien des habitants du monde au début du XXe siècle.

L'histoire par l'image

[consulté le 01 décembre 2013]

<<http://www.histoire-image.org>>

Portail « Fresques de l'Ina »

[consulté le 01 décembre 2013]

<<http://fresques.ina.fr>>

Site est à destination des enseignants, il apporte du contenu pédagogique (la partie Pédagogie est soumise à abonnement)

« Mallettes pédagogiques » du Centre de Développement Chorégraphique Les Hivernales

[consulté le 01 décembre 2013]

<<http://www.hivernales-avignon.com/education-artistique/mallettes-pedagogiques>>

Pour « la Danse en 10 titres, 10 dates et Le Tour du Monde en 80 danses »

Portail des Opéras de France

Portail non accessible sur Internet actuellement (version bêta)

Histoire par l'image

[consulté le 01 décembre 2013]

<<http://www.histoire-image.org>>

Réalisé par l'organisme de Réunion des musées nationaux à destination des enseignants et des élèves proposent des analyses d'œuvres pour aller au-delà de la simple légende.

Opéras en actes

[consulté le 01 décembre 2013]

<www.cndp.fr/opera-en-actes/accueil.html>

Réalisé par le Centre national de documentation pédagogique (CNDP)

Histoire des arts

[consulté le 01 décembre 2013]

<<http://www.histoiredesarts.culture.fr>>

Réalisé par le ministère de la Culture, il propose près de 5000 ressources commentées pour l'enseignement de l'histoire des arts : dossiers pédagogiques, expositions virtuelles, vidéos... ainsi que des repères chronologiques par domaines artistiques (architecture, bande dessinée, littérature, jardins, musique, peinture, théâtre...)

Archives nationales d'Outre-Mer

[consulté le 01 décembre 2013]

<www.archivesnationales.culture.gouv.fr/anom/fr/index.html>

Bibliothèque de Documentation Internationale Contemporaine (BDIC)

[consulté le 01 décembre 2013]

<www.bdic.fr>

La bibliothèque de documentation internationale contemporaine constitue, au sein du Grand Paris, une mémoire historique des 20e et 21e siècles.

Pinterest pour les collections de la BNF sur Gallica

[consulté le 01 décembre 2013]

<www.pinterest.com/gallicabnf>

Outil collaboratif pour la réalisation et le partage de collections de documents issus de Gallica.

Annexe 3 Grille d'évaluation de collections numériques culturelles

Identification	
Titre	
URL	
Organisme	
Type	
Taille	
Date de création	
Pays	
Dates du site	
Création	
Mise à jour	
Consultation	
Site des collections numériques	
Types de collection	Bibliothèque numérique <input type="checkbox"/> Fonds d'archives privées <input type="checkbox"/> Fonds d'archives publiques <input type="checkbox"/> Collection de musée <input type="checkbox"/> <input type="checkbox"/>
Public cible	Grand public/amateurs <input type="checkbox"/> professionnels <input type="checkbox"/> Enseignants <input type="checkbox"/> chercheurs <input type="checkbox"/>
Résumé de présentation	

Contenus	
Couverture géographique	Régionale <input type="checkbox"/> Nationale <input type="checkbox"/> Européenne <input type="checkbox"/> Mondiale <input type="checkbox"/>
Période couverte	
Taille de la collection	Nombre de documents : Nombre de fonds / collections :
Types de documents	
Images fixes (hors revues, livres et livrets)	Photos <input type="checkbox"/> Dessins/croquis/gravures <input type="checkbox"/> Affiches <input type="checkbox"/> Reproduction d'objets <input type="checkbox"/> <input type="checkbox"/>
Documents sonores	Reportage <input type="checkbox"/> Entretien <input type="checkbox"/> Captation <input type="checkbox"/> <input type="checkbox"/>
Images animés	Films de fiction <input type="checkbox"/> Films documentaires <input type="checkbox"/> Montage pédagogique <input type="checkbox"/> Spectacles filmés <input type="checkbox"/> Rush / Captation <input type="checkbox"/> <input type="checkbox"/>
Documents textuels numérisés	Livre <input type="checkbox"/> Manuscrits <input type="checkbox"/> Notes <input type="checkbox"/> Périodiques <input type="checkbox"/> Article de presse <input type="checkbox"/> <input type="checkbox"/>
Autres documents <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Questions juridiques	
Accès	Libre <input type="checkbox"/> Libre après inscription <input type="checkbox"/> Restreint <input type="checkbox"/> détails.....
Licence du site	usage exclusivement privé <input type="checkbox"/> <input type="checkbox"/>
Gestion des droits liés aux	

Questions juridiques	
Documents et notices	
Ergonomie	
Structuration du site par	Public cible <input type="checkbox"/> Classement des documents <input type="checkbox"/> Utilisation <input type="checkbox"/> Thématique <input type="checkbox"/> <input type="checkbox"/>
Zone la plus importante (accueil)	contenus présélectionnés <input type="checkbox"/> Recherches implicites <input type="checkbox"/> moteur de recherche <input type="checkbox"/> détails.....
Navigation dans...	Le temps <input type="checkbox"/> Les lieux <input type="checkbox"/> Les évènements <input type="checkbox"/> Les personnages <input type="checkbox"/> Les oeuvres <input type="checkbox"/> Les espaces dédiés à chaque public <input type="checkbox"/> Les thèmes <input type="checkbox"/> Les types de documents <input type="checkbox"/> <input type="checkbox"/>
Mise en avant d'un document	Oui <input type="checkbox"/> Non <input type="checkbox"/> Critère Navigation
Sélections proposés	Oui <input type="checkbox"/> Non <input type="checkbox"/>
Appellations	Dossiers <input type="checkbox"/> Thèmes <input type="checkbox"/> Expositions virtuelles <input type="checkbox"/> Zoom <input type="checkbox"/> Actualités <input type="checkbox"/> <input type="checkbox"/>
Description générale	La sélection est valorisé en.....
Critères de sélection	
Types de documents	

Navigation	
Depuis un document	
Critère du Précédent/Suivant	Date de création <input type="checkbox"/> Date de mise en ligne <input type="checkbox"/> Titre <input type="checkbox"/> <input type="checkbox"/>
Visualisation par	Document par document <input type="checkbox"/> Mosaïque <input type="checkbox"/> Liste <input type="checkbox"/> <input type="checkbox"/>
Information sur le document	Par survol/Clic <input type="checkbox"/> Toujours affiché <input type="checkbox"/> Affichage minimal <input type="checkbox"/> Affichage détaillé <input type="checkbox"/> <input type="checkbox"/> Détails.....
Suggestion d'autres documents	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Depuis une sélection	
Accès aux documents	Premier document <input type="checkbox"/> Liste <input type="checkbox"/> Mosaïque <input type="checkbox"/> <input type="checkbox"/>
Visualisation par	Document par document <input type="checkbox"/> Mosaïque <input type="checkbox"/> Liste <input type="checkbox"/> <input type="checkbox"/>
Information sur le document	Par survol/Clic <input type="checkbox"/> Toujours affiché <input type="checkbox"/> Affichage minimal <input type="checkbox"/> Affichage détaillé <input type="checkbox"/> <input type="checkbox"/> Détails.....
Suggestion d'autres sélections	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Suggestion d'autres documents	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....

Moteur de recherche	
Type de recherche	Recherche simple <input type="checkbox"/> Recherche avancée <input type="checkbox"/>
Simple : présentation	
Critères de recherche avancée	
Affichage des résultats de moteur de recherche	Liste <input type="checkbox"/> Mosaïque <input type="checkbox"/> Frise <input type="checkbox"/> <input type="checkbox"/>
Tri des résultats (précisez le tri par défaut)	Date de création <input type="checkbox"/> Date de mise à jour <input type="checkbox"/> Type de documents <input type="checkbox"/> Titre <input type="checkbox"/> Auteurs <input type="checkbox"/> <input type="checkbox"/>
Filtrage des résultats	Non <input type="checkbox"/> Par facette <input type="checkbox"/> Autre <input type="checkbox"/> détails...(relance auto, clic).....
Elargissement de la recherche	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....

Documents	
A compléter type par type	DUPLIQUER CE TABLEAU POUR DECRIRE CHAQUE TYPE DE DOCUMENTS
Description au niveau d'un ensemble	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Description individuelle	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Affichage de la description	
Vignette	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Qualité de la description	Bon <input type="checkbox"/> Moyen <input type="checkbox"/> Mauvais <input type="checkbox"/> détails.....
Ergonomie	Bon <input type="checkbox"/> Moyen <input type="checkbox"/> Mauvais <input type="checkbox"/> détails.....
Nom du player requis	Aucun <input type="checkbox"/> <input type="checkbox"/>

Outils / Services	
Espace personnel	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Interaction	
Vote sur le contenu	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Documents les plus consultés	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Commentaire sur le contenu	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Suggestion pour le site	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Description de documents orphelins	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Partage	Réseaux sociaux <input type="checkbox"/> Mail <input type="checkbox"/>
Suivre l'actualité	
Actualités du site visible sur l'accueil	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Hors du site	Twitter <input type="checkbox"/> Facebook <input type="checkbox"/> Flux RSS <input type="checkbox"/> <input type="checkbox"/> Détails
Produit documentaire	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails...(ex. Newsletter).....
Export	
Notices	Oui <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Résultats de recherche	Oui avec vignettes <input type="checkbox"/> Oui sans vignettes <input type="checkbox"/> Non <input type="checkbox"/> détails.....
Format	PDF <input type="checkbox"/> Mail <input type="checkbox"/> Zotero <input type="checkbox"/> détails.....

Points notables	
Point faibles	
Point 1	
Point 2	
Point 3	
Point 4	
Point forts	
Point 1	
Point 2	
Point 3	
Point 4	

Schémas de navigation dans le site

Captures d'écran