

Les langages documentaires au service de la valorisation des collections muséales : quels enjeux ? Le cas du musée Rodin

Béatrice de Bégon de Larouzière

▶ **To cite this version:**

Béatrice de Bégon de Larouzière. Les langages documentaires au service de la valorisation des collections muséales : quels enjeux ? Le cas du musée Rodin. domain_shs.info.docu. 2013. mem_00944577

HAL Id: mem_00944577

https://memsic.ccsd.cnrs.fr/mem_00944577

Submitted on 10 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

École 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

Béatrice de Larouzière Bussereau

Le 13 janvier 2013

Les langages documentaires au service de la valorisation des collections muséales : quels enjeux ?

Le cas du musée Rodin

Jury :
AITKEN Geneviève, responsable de la documentation des collections, musée Rodin
RAIS Nadia, professeur associée, CNAM-INTD
ZYSMAN Hélène, consultante fonctionnelle GED/ECM, Smile – Open Source Solutions

Promotion 43

Remerciements

Je tiens à remercier très chaleureusement :

Le Centre des monuments nationaux, où je travaille depuis 2010, qui m'a permis de réaliser ce projet, et particulièrement, Mireille Klein, directrice scientifique, Laurent Bergeot, chef du département des ressources scientifiques, Dominique Fernandes, chef du bureau de la documentation, Karine Madec, secrétaire de documentation, et tous mes collègues de la direction scientifique, qui m'ont soutenue dans la rédaction de ce mémoire.

Nadia Raïs, professeur associé au CNAM-INTD, qui a été source d'encouragements et de conseils avisés tout au long de ce travail.

Geneviève Aitken, chargée d'études documentaires au service des collections du musée Rodin, qui m'a accueillie à bras ouverts en stage, ainsi que toute l'équipe du service des collections et du service de la recherche, de la documentation, de la bibliothèque et des archives.

Les professionnels des musées et des institutions culturelles qui m'ont reçue afin de me présenter leur travail : au château de Versailles, Christine Desgrez et Delphine Dubois, chargées d'études documentaires au service de la documentation ; au centre de recherche du château de Versailles, Isabelle Pluvieux, responsable base de données, internet et réseaux ; au musée Bourdelle, Annie Barbera, responsable de la documentation, Stéphane Ferrand, responsable du cabinet d'arts graphiques et Jean-Philippe Manzano, régisseur des œuvres ; au Centre des monuments nationaux, Pauline Bonnet, administratrice de la base de gestion des collections ;

Les professionnels du Ministère de la culture et de la communication qui m'ont reçue afin de me présenter leurs projets en cours : Laurent Manœuvre, chef du bureau de la diffusion numérique des collections (JocondeLab), Katell Briatte, chef de projet maîtrise d'ouvrage au département des systèmes d'information patrimoniaux (HADOC) et Bertrand Sajus, du département des programmes numériques (HdALab).

Ainsi que mes collèges de formation, notamment Loraine Keller pour nos discussions motivantes dans la mise en route du mémoire, tout particulièrement Amélie Vernusset pour son soutien sans faille et nos échanges constructifs lors de la rédaction du mémoire, et enfin Sylvie Bonnet, pour ses conseils avisés lors de la finalisation du mémoire.

Avec une mention spéciale à mon mari, ma famille et mes amis sans lesquels je n'aurais pu achever ce travail.

Notice

DE LAROUZIERE BUSSEREAU Béatrice. Les langages documentaires au service de la valorisation des collections muséales : quels enjeux ? Le cas du musée Rodin. 2013. p. 168. Mémoire Titre professionnel de niveau 1, CNAM-INTD. 2013

L'objectif de ce mémoire est de montrer, en s'appuyant sur l'exemple du musée Rodin, comment les musées peuvent, à l'heure actuelle, utiliser les langages documentaires et mettre à profit leur évolution dans le cadre du web de données afin de décrire les objets, enrichir et partager les données et les connaissances, et ainsi valoriser leurs collections. Ce travail porte essentiellement sur la valorisation des collections muséales, mais dans le contexte plus large de la valorisation des données culturelles et du rôle et de l'utilisation des langages documentaires. Nous donnons des éléments de compréhension à la fois sur les collections, sur les langages documentaires et sur le web de données. Notre démarche s'appuie sur les besoins et les difficultés propres à la collection du musée Rodin. Enfin nous montrons comment la gestion documentaire de cette collection bien spécifique, tout comme pour l'ensemble des collections muséales, peut et doit répondre à ces besoins et à ces difficultés en répondant aux enjeux auxquels les musées doivent actuellement faire face à l'heure du web de données.

COLLECTIONS MUSEALES ; DONNEES CULTURELLES ; INDEXATION ; LANGAGES DOCUMENTAIRES ; THESAURUS ; REFERENTIELS ; MUSEES ; VALORISATION CULTURELLE ; WEB DE DONNEES ; WEB SEMANTIQUE

The purpose of this thesis is to demonstrate, based on the example of the "musée Rodin", how museums can, at the present time, use documentary languages and benefit from their evolution in the framework of the linked data in order to describe objects, to enrich and share data and knowledge, and to gain in value their collections. This work is essentially about the promotion of the museum's collections but to a larger extent, the promotion of the cultural data and the role and the use of documentary languages. We give at the same time elements of comprehension on collections, on documentary languages and on linked data. Our approach is based on the needs and the proper difficulties of the "musée Rodin". Finally we show how the documentary management of this very specific collection, as for the whole museum's collections, can and must respond to these needs and to these difficulties, by answering to the challenges museums must currently face during linked data.

Liste des sigles

BnF : Bibliothèque nationale de France

CMN : Centre des monuments nationaux

CIDOC : Comité international pour la documentation

ICOM : International Council of Museums

IFLA: International Federation of Library Associations and Institutions

IRI : Institut de recherche et d'innovation

ISO : International Organization for Standardization

MCC : Ministère de la culture et de la communication

MuCEM : Musée des civilisations de l'Europe et de la Méditerranée

W3C: World Wild Web Consortium

Table des matières

Remerciements 2
Notice 3
Liste des sigles 4
Table des matières 5
Liste des tableaux et des figures 8
Introduction 9
Première partie : Le contexte des collections muséales et du traitement de l'information 13
1 Les collections muséales au centre de l'information 15
1.1 Repères sur les collections muséales 15
1.2 Données et ressources 16
1.3 Usages et usagers 19
1.4 Outils de gestion 20
2 Les langages documentaires 22
2.1 Rôle des langages documentaires 22
2.1.1 Définition et fonctions 22
2.1.2 Indexation documentaire 23
2.1.3 Typologies des langages documentaires 24
2.2 Les classifications 26
2.2.1 Définition 26
2.2.2 La classification décimale Dewey 27
2.2.3 La classification décimale universelle 28
2.2.4 La classification de la bibliothèque du Congrès 28
2.3 Les listes d'autorités et les listes de vedettes matières 29
2.3.1 Définition 29
2.3.2 Structure 29
2.3.3 Rameau 30
2.4 Les thésaurus 31
2.4.1 Définition et structure 32
2.4.2 Types de relations 32
2.4.3 Présentation d'un thésaurus 34
2.4.4 Normes 35

3 L'évolution des langages d'indexation et de recherche dans le cadre du web de données 37
3.1 Les taxonomies (ou taxinomies) 37
3.2 Les ontologies 38
3.3 Le web de données et le web sémantique 41
3.3.1 Repères 41
3.3.2 Nouveaux systèmes de description des données 41
Deuxième partie : Le cas du musée Rodin, structuration des données et élaboration d'un thésaurus dans le cadre d'une réinformatisation 44
1 Contexte du projet 46
1.1 Présentation des collections du musée Rodin 46
1.1.1 Renommée des collections 46
1.1.2 Gestion des collections et de l'information 47
1.1.3 Les collections sur le web 48
1.2 Analyse de l'existant et des besoins 49
1.2.1 Ressources documentaires 49
1.2.2 Diffusion de l'information 51
1.2.3 Logiciel de gestion des collections 52
1.2.4 Harmonisation des données 53
1.3 Organisation de la mission 54
1.3.1 Cahier des charges 54
1.3.2 Planning 56
2 Description et indexation d'un objet de collection muséale 58
2.1 Particularité d'une œuvre par rapport à un document 58
2.2 Réglementations et normes 59
2.2.1 Cadre juridique 59
2.2.2 Méthode de description des objets 60
2.2.3 Vocabulaires 63
2.3 Spécificité des collections du musée Rodin 64
2.4 Ce qui se fait dans d'autres musées 65
2.5 Préconisations pour une structuration des données 67
3 Élaboration d'un thésaurus 70
3.1 Méthode d'élaboration d'un thésaurus monolingue 70
3.1.1 Collecte du vocabulaire ou lexique 70
3.1.2 Listage des catégories et plan de classement 71
3.1.3 Établissement des règles d'écriture et relation entre les termes 72
3.1.4 Test et corrections 72
3.2 Complément et mise à jour 73

3.3 Difficultés et préconisations 73
Troisième partie : Quels enjeux pour la valorisation des collections muséales....75
1 Répondre aux politiques culturelles 77
1.1 Premier récolement décennal des collections....77
1.1.1 Cadre juridique 77
1.1.2 Traitement des données 78
1.2 Numérisation du patrimoine culturel 79
1.3 *Open data* 80
1.3.1 Définition de l'*Open data* 80
1.3.2 *Open data* culturel....82
2 Diffuser les données sur le web et comprendre les enjeux du web de données 85
2.1 Mise en ligne des données 85
2.1.1 Sur les sites web des musées....85
2.1.2 Sur le Portail national des collections des musées de France : Joconde....88
2.1.3 Sur des « bibliothèques » numériques internationales 89
2.2 Réutilisation et contribution des utilisateurs 91
2.3 Enjeux du web de données....93
2.3.1 Quel apport pour les musées ?....94
2.3.2 Comment cela fonctionne ?....94
3 Garantir l'interopérabilité des données....97
3.1 Structuration et harmonisation des données culturelles 97
3.1.1 Dbpedia en français dans le cadre de Semanticpdia....97
3.1.2 data.bnf.fr 98
3.1.3 HADOC....99
3.2 Gestion des vocabulaires et utilisation des référentiels....103
3.2.1 *Tagging* sémantique....103
3.2.2 Outil de gestion des vocabulaires : GINCO....104
3.2.3 Projet de référentiel des acteurs historiques....105
3.3 Bilan sur ces projets....107
Conclusion 110
Bibliographie....114
Annexes 125

Liste des tableaux et des figures

Tableau 1 : Tableau récapitulatif des typologies des langages documentaires 25
Tableau 2 : Système de classification Dewey 27

Figure 1 : Vue d'ensemble de l'information gravitant autour des collections muséales 18
Figure 2: Exemple de la classification décimale Dewey 27
Figure 3 : Notion de granularité entre les différents langages documentaires 40
Figure 4 : Processus créatif de Rodin.....65
Figure 5 : Vue d'ensemble du modèle de données du Centre Pompidou virtuel 87
Figure 6 : Extraits du site *Les Herbonautes, l'herbier numérique collaboratif citoyen*.....93

Introduction

La raison d'être d'un musée est la collection qu'il détient. Dans le Code du patrimoine, c'est parce qu'elle est conservée et organisée pour être accessible au public le plus large qu'une collection peut bénéficier de l'appellation « musée de France ». Les plus de 1 200 institutions ainsi protégées regroupent des dizaines de millions d'objets et œuvres dans tous les domaines de la création humaine et naturelle, aussi bien dans ceux de l'art, de l'histoire, de la technique ou de l'ethnologie, que dans ceux de la biologie ou de la géologie.

Depuis une quinzaine d'années, les musées procèdent, un à un, à l'informatisation et à la numérisation des données et des ressources permettant de décrire les collections muséales. Cette informatisation répond à la fois à une obligation juridique, l'inventaire et le récolement décennal obligatoire des collections pour lesquels un support informatisé est préférable à un support papier, à un côté pratique, une gestion scientifique et administrative des collections facilitée, et enfin à une demande nouvelle du public, le partage des connaissances avec le plus grand nombre sur le web, dans le cadre de l'ouverture des données culturelles.

Mais si pour les bibliothèques, par exemple, l'informatisation s'est appuyée sur les systèmes d'indexation et de classement et sur les langages documentaires déjà existants avec la documentation papier et normalisés au niveau national et même international, cela a été beaucoup plus compliqué pour les collections muséales dont les systèmes de classement sont restés longtemps non seulement propres à chaque musée mais aussi à chaque domaine de collection et, quelquefois même, à chaque conservateur. En effet, un conservateur au 20ème siècle pouvait rester toute sa carrière dans un même musée, être responsable d'une même collection (par exemple la collection de peinture du musée Rodin) et gérait de ce fait sa propre documentation, conservée souvent dans son bureau, et peu accessible à d'autres que lui. Certains grands musées, comme le Louvre, avaient créé des services de documentation mais souvent par domaine de collection, avec des systèmes de classement et d'indexation spécifiques, sans normalisation, et peu accessibles au grand public.

Le Ministère de la culture et de la communication (MCC) a toutefois mis en place, dès les années 1970, à disposition de tous les musées, une méthode de rédaction informatisée des notices des collections muséales, s'appuyant sur des langages documentaires normalisés tels que des thésaurus, permettant d'indexer un objet quelle que soit sa nature (archéologie, arts décoratifs, beaux-arts, ethnologie, sciences et techniques, histoire).

D'autre part, le développement des technologies liées à la mise en place du web de données implique la mise en place de nouveaux langages et de nouvelles techniques s'appuyant sur les langages documentaires traditionnels comme les thésaurus et les listes d'autorités. Or dans le cadre de l'ouverture des données culturelles, les collections muséales sont non seulement concernées par ces nouvelles réflexions sur les vocabulaires et sur les concepts, mais doivent aussi mettre en relation leurs données avec les autres données culturelles qui

ont elles aussi leurs spécificités documentaires et leurs types de données. Les données culturelles très hétérogènes mais aussi redondantes (combien de fois, par exemple, le terme « Rodin » est-il indexé de manière répétitive par les musées, les bibliothèques, les archives, et toutes les institutions culturelles) demandent toutefois, dans ce web de données qui a émergé, à être mises en cohérence et enrichies pour une meilleure lisibilité, une meilleure recherche et une meilleure utilisation de l'information par l'utilisateur (grand public, chercheurs, étudiants, etc).

Lors de mon stage au musée Rodin où ma mission a porté sur la restructuration des données et l'élaboration d'un thésaurus propre aux collections du musée, j'ai été confrontée à toutes ces problématiques : hétérogénéité des collections et des ressources ; informatisation déjà ancienne et obsolète, difficultés du projet de réinformatisation ; souhait d'une mise en valeur du travail scientifique, de l'inventaire et du récolement des collections, pour une meilleure gestion, un meilleur traitement et un enrichissement des données ; mais difficulté à structurer les données et questionnements quant à l'utilisation des langages documentaires et des préconisations du MCC qui paraissent difficiles à mettre en œuvre avec la particularité des collections du musée ; volonté d'utiliser les nouvelles technologies et la puissance du web de données, de diffuser au plus grand nombre ces données, de partager et de mettre en relation les données au niveau international ; mais difficultés dans la mise en œuvre et dans la compréhension des enjeux.

L'objectif de ce mémoire est donc de montrer, en s'appuyant sur l'exemple du musée Rodin et sur le travail que j'ai pu y réaliser, comment les musées peuvent à l'heure actuelle utiliser les langages documentaires et mettre à profit leur évolution dans le cadre du web de données, afin de décrire les objets, enrichir et partager les données et les connaissances et ainsi, valoriser leurs collections.

Ce travail porte essentiellement sur la valorisation des collections muséales mais dans le contexte plus large de la valorisation des données culturelles et du rôle et de l'utilisation des langages documentaires. Nous tâcherons pour cela de donner des éléments de compréhension à la fois sur les collections, sur les langages documentaires et sur le web de données. Notre démarche s'appuiera sur les besoins et les difficultés propres à la collection du musée Rodin. Et nous montrerons comment la gestion documentaire de cette collection bien spécifique, tout comme pour l'ensemble des collections muséales, peut et doit répondre à ces besoins et à ces difficultés en répondant aux enjeux auxquels les musées doivent actuellement faire face à l'heure du web de données.

En effet, quelle est la place de ce web de données dans la valorisation des collections muséales et comment les musées peuvent-ils en tirer profit ? Comment les musées peuvent et doivent traiter les informations, les données, les ressources liées aux collections

muséales ? Pourquoi et comment doivent-ils utiliser les langages documentaires, quels sont leurs rôles dans la description des objets, l'indexation et la recherche et quelles sont leurs évolutions et leurs utilisations dans le web de données ?

Pour cela, comment permettre une approche documentaire de l'œuvre, à partir de l'analyse d'une œuvre par un conservateur dont la mission est scientifique et non documentaire ? Comment enrichir les données et ainsi la connaissance sur les œuvres, à la fois au sein d'un musée, facilitant ainsi le travail du personnel scientifique, mais aussi à l'extérieur du musée, pour les chercheurs, les étudiants et le grand public, de tous pays ? Comment les musées, en répondant aux politiques culturelles et aux missions qui leur sont confiées, peuvent-ils utiliser et comprendre le web de données en garantissant l'échange et l'interopérabilité de leurs données afin de les diffuser et de les mettre en relation avec les autres données culturelles (autres musées, bibliothèques, archives, etc) en France et à l'international ?

Afin de répondre à ces questions, ce mémoire s'organise en trois parties. Nous allons dans un premier temps faire un point sur le traitement de l'information au sein des musées. Nous allons pour cela montrer comment les collections muséales se trouvent au centre de l'information dans les musées, puis présenter le rôle de langages documentaires et leurs différents types, et enfin montrer l'évolution de ces langages d'indexation de recherche dans le cadre du web de données dont nous donnerons une définition précise. Dans une deuxième partie, nous aborderons le cas du musée Rodin. Nous présenterons le contexte du projet, puis nous ferons un point sur la méthode de description et d'indexation d'un objet d'art et nous terminerons en présentant la méthodologie utilisée lors du stage afin d'élaborer un thésaurus propre aux collections du musée Rodin. Dans une dernière partie, nous tâcherons de répondre aux questions posées, en mettant en lumière les enjeux rencontrés par les musées, afin de valoriser leurs collections muséales : répondre aux politiques culturelles ; diffuser leurs données sur le web et comprendre le web de données ; et enfin garantir l'échange et l'interopérabilité de leurs données.

Première partie :
Le contexte des collections muséales et du traitement de l'information

Dans cette première partie, nous allons montrer comment les collections muséales se trouvent au centre du traitement de l'information au sein des musées. Autour des collections muséales à la fois régies par un cadre juridique strict, et en même temps de nature très hétérogène, gravitent l'ensemble des données produites et des ressources utilisées, des utilisateurs de plus en plus nombreux pour de nouveaux usages et des outils de gestion dont l'apport pour le traitement de l'information et la diffusion de l'information est indéniable et encore à améliorer. Puis nous présenterons de manière exhaustive le rôle des langages documentaires dans le traitement de l'information et les différents types de langages documentaires qui existent. Il est, en effet, important de revenir sur des définitions, qui dans le monde de la gestion de l'information, peuvent paraître évidentes et d'ores et déjà normalisées, mais dont le potentiel n'est pas toujours mis à profit pour les collections muséales. Nous terminerons cette première partie en montrant l'évolution des langages documentaires et l'apparition avec les derniers bouleversements technologiques de nouveaux langages d'indexation et de recherche qui ne remplacent pas les langages documentaires traditionnels, mais en sont au contraire un atout dans le web de données et le web sémantique dont nous donnerons des éléments de compréhension.

1 Les collections muséales au centre de l'information

Les collections muséales sont régies par le Code du patrimoine et revêtent un intérêt public. Les musées qui en ont la charge ont donc une mission bien particulière de conservation, de présentation et de diffusion de la connaissance auprès du grand nombre. Nous allons, dans ce cadre, et après avoir donné quelques repères juridiques sur les collections muséales, faire un point sur la spécificité et l'hétérogénéité des données produites et des ressources utilisées autour des collections, sur les usages qui doivent en être faits et pour quels usagers, et sur les nouveaux outils de gestion qui permettent aux musées d'être plus visibles sur le web.

1.1 Repères sur les collections muséales

Les collections muséales font partie des biens culturels dont la définition est donnée par l'UNESCO : *« 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. Dans le droit des biens culturels, la notion prend ainsi un sens spécifique et concerne les éléments du patrimoine culturel d'un État, sa richesse culturelle »*[1].

Plus précisément, en France, les collections muséales sont régies par le Code du patrimoine qui a fait l'objet de l'ordonnance du 24 février 2004 : *« les biens appartenant aux collections publiques et aux collections des musées de France, les biens classés en application des dispositions relatives aux monuments historiques et aux archives, ainsi que les autres biens qui présentent un intérêt majeur pour le patrimoine national au point de vue de l'histoire, de l'art ou de l'archéologie sont considérés comme trésors nationaux »*[2].

La loi « musée », promulguée le 4 janvier 2002 et qui constitue le livre IV du Code du Patrimoine notifie également qu' *« est considéré comme musée, au sens du présent livre, toute collection permanente composée de biens dont la conservation et la présentation revêtent un intérêt public et organisée en vue de la connaissance, de l'éducation et du plaisir du public »*[3]. Elle précise que « *l'appellation "musée de France" peut être accordée aux musées appartenant à l'État, à une autre personne morale de droit public ou à une personne*

[1] Article 1 de la Convention de 1970 et article 2 du Chapitre 1 de la Convention de 1995.
[2] Code du patrimoine, article L.111-1.
[3] Code du patrimoine, article L. 410-1.

morale de droit privé à but non lucratif »[4]*.* La loi définit ainsi parfaitement la nature exacte d'un musée et les missions qui en découlent.

Cette loi pose également un principe d'inaliénabilité permettant 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 »*[5]*.*

1.2 Données et ressources

Comme il est expliqué dans le projet HADOC (*projet d'harmonisation des données culturelles*) [21, BRIATTE], réalisé dans le cadre du schéma directeur des systèmes d'information 2005-2008 du MCC, et dont nous parlerons plus longuement dans la troisième partie (cf. partie 3 section 3.1.3), les données relatives aux collections muséales les décrivent de manière pérenne et relèvent de l'information documentaire : désignation, localisation, description physique, histoire, statut, etc.

Cette information documentaire est gérée et produite par l'ensemble des métiers servant à la valorisation des collections au sein des musées : gestion des collections, bibliothèque, archives, gestion iconographique et multimédia. L'information documentaire sur les collections muséales, relevant de métiers différents, est ainsi très riche et relativement diverse.

Les ressources utilisées autour des collections muséales, quant à elles, permettent de produire cette information documentaire, comme les ressources bibliographiques, les ressources iconographiques et multimédia, les ressources terminologiques, les référentiels normatifs, etc.

[4] Code du patrimoine, article L.441-1.
[5] Code du patrimoine, article L.451-5.

Afin de donner un exemple de la variété des informations utilisées et produites sur les collections muséales, nous pouvons citer le Musée des civilisations de l'Europe et de la Méditerranée (MuCEM), ouvert en juin 2013[6]. Les collections muséales du MuCEM sont accessibles au sein du Centre de conservation et des ressources, dans un espace de 13 000 m² dont 8 000m² de réserves. Cet espace abrite les collections (objets et fonds documentaires) conservées par le MuCEM :

- 250 000 objets d'origine très diverse ;
- 130 000 tableaux, estampes et dessins ;
- 450 000 photographies ;
- 140 000 cartes postales ;
- 100 000 ouvrages et 500 revues de sciences humaines et sociales ;
- des documents patrimoniaux tels qu'une très riche collection d'impressions populaires ;
- un fonds d'archives papiers, sonores et audio-visuelles ;
- la documentation des collections du musée.

Le Centre de conservation et des ressources du MuCEM met également à disposition ses collections numériques, via son portail documentaire donnant accès à :

- des revues électroniques ;
- des bases de données documentaires.

Ces collections et ces fonds sont répartis autour de huit pôles :

- agriculture, élevage, exploitation des ressources naturelles et environnement ;
- apparences, soins du corps, sexualité ;
- industrie, artisanat, commerce ;
- arts du spectacle ;
- croyances et religion ;
- mobilités, métissages, communications ;
- vie domestique ;
- vie publique.

[6] Site internet du musée des civilisations de l'Europe et des Méditerranées (MuCEM). www.mucem.org/ [consulté le 11/10/2013]

Cet exemple, qui reflète parfaitement la réalité de l'ensemble des musées en France, montre à quel point les ressources muséales sont nombreuses, riches et hétérogènes, induisant des données tout aussi importantes et diverses, difficiles à traiter de manière homogène (cf. Figure 1). Il est toutefois nécessaire de traiter ces données de manière cohérente afin de répondre à la mission confiée aux musées de France, édictée très clairement dans la loi « musées » au livre IV du Code du patrimoine :

- conserver, restaurer, étudier et enrichir leurs collections ;
- rendre leurs collections accessibles au public le plus large ;
- concevoir et mettre en œuvre des actions d'éducation et de diffusion visant à assurer l'égal accès de tous à la culture ;
- contribuer aux progrès de la connaissance et de la recherche ainsi qu'à leur diffusion.

Le International Council of Museums (ICOM) le précise également dans ses principes déontologiques : « *Les musées assurent la protection, la documentation et la promotion du patrimoine naturel et culturel de l'humanité ; les musées contribuent à la connaissance, à la compréhension et à la gestion du patrimoine naturel et culturel* »[7].

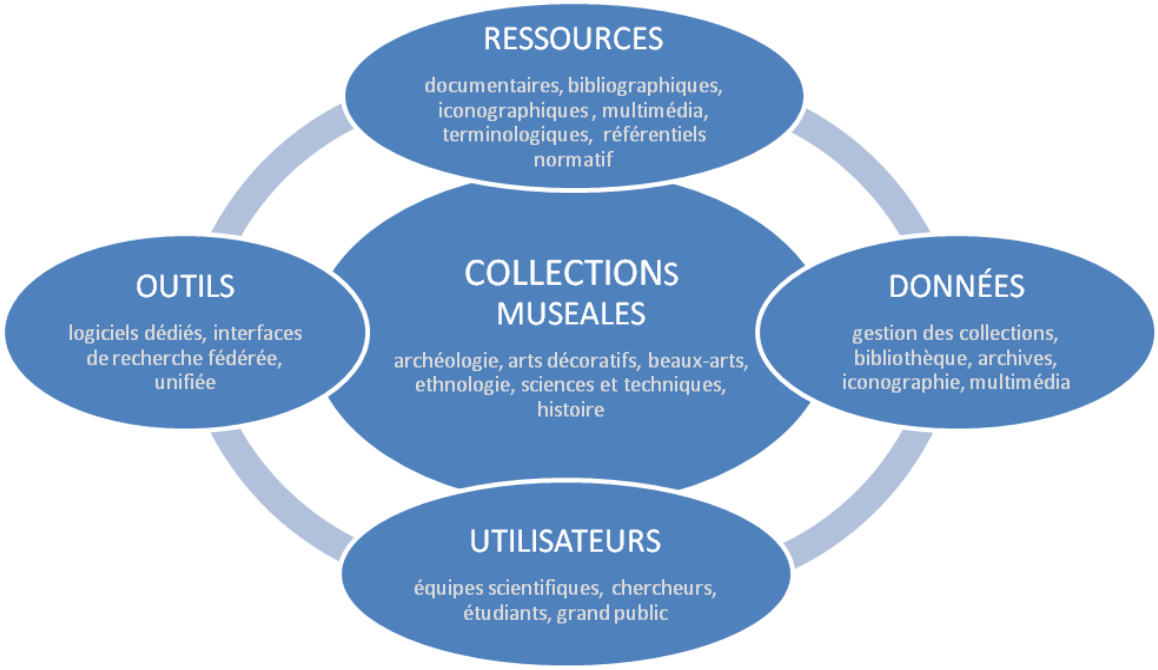


Figure 1 : Vue d'ensemble de l'information gravitant autour des collections muséales.

[7] ICOM – Code de déontologie
http://icom.museum/la-vision/code-de-deontologie//L/2/ [consulté le 11/10/2013]

1.3 Usages et usagers

Ces principes nous amènent à nous poser la question de l'usage et des usagers de ces données et de ces ressources.

En effet, l'information, qu'elle soit documentaire ou non, doit être produite et gérée pour être diffusée, utilisée, modifiée, enrichie, etc. Concernant l'information documentaire sur les collections muséales et la valorisation de celles-ci, la diffusion de l'information est d'autant plus importante.

Les données et les ressources ne sont pas destinées à être traitées et stockées uniquement au sein des musées, par et pour le personnel scientifique. Elles ont vocation à être diffusées au public, tout à la fois propriétaire et gardien du musée, selon Annie Héritier, historienne du droit, qui affirme que « *l'appartenance des musées au service public fait partie des principes essentiels non soumis à discussion* »[8].

C'est dans ce but que le MuCEM développe, à travers le Centre de conservation et des ressources, de larges activités de diffusion et d'étude de ses collections. Un espace documentaire de 330 m² est d'ores et déjà ouvert au sein de son Centre de conservation et des ressources, accessible :

- aux équipes scientifiques du MuCEM ;
- aux chercheurs ;
- aux étudiants ;
- au grand public.

L'intégralité de ces collections pourra très prochainement être consultée sur place, grâce aux salles dédiées à cet effet, qui seront accessibles au public. Une partie des collections numériques est, quant à elle, également disponible en ligne, via le portail documentaire numérique du MuCEM et donc accessible à tous.

Cette accessibilité des données et des ressources sur le web permet une diffusion plus large de l'information, rend les utilisateurs plus nombreux, et permet des usages nouveaux. Nous parlerons plus longuement de ces problématiques et de l'enjeu de l'accessibilité des données et du partage des connaissances dans la troisième partie.

---

[8] HÉRITIER Annie. Le musée face au public. A la recherche du service public culturel. Une approche historique. In Musées en mutation. Un espace public à revisiter. Dir. Martine Regourd. Paris, Éditions l'Harmattan, 2012, p.87-109. ISBN : 978-2-336-00391-7.

1.4 Outils de gestion

Les bouleversements technologiques de ces dernières années ont amené le milieu culturel à se remettre en question et à revoir ses pratiques. Les musées se sont peu à peu dotés d'outils leur permettant d'être visibles sur le Web. Il existe encore une certaine réserve sur le fait de rendre accessibles à tous, des œuvres et un savoir prestigieux, mais on observe un réel progrès dans un milieu longtemps considéré comme fermé et peu porté sur les considérations techniques et technologiques.

Depuis les années 90, les musées s'engagent donc, les uns après les autres, dans un programme d'informatisation et de diffusion des collections dont ils ont la responsabilité, se dotant d'outils professionnels performants leur permettant également d'assurer au quotidien la gestion scientifique et administrative des collections. Ces projets ont essentiellement pour but, plus que de montrer ce qui est déjà visible (par une reproduction identique de l'expérience *in situ*), de révéler ce qui est habituellement caché. Par exemple, au Centre Pompidou, sur les 99 000 œuvres du musée, 2 000 environ sont effectivement exposées. La diffusion numérique permet ainsi de rendre le patrimoine accessible à tous de manière beaucoup plus large que l'approche éditoriale traditionnelle des sites web de musée qui tend à mettre l'accent principalement sur les œuvres, ou les artistes, qui sont considérés comme d'intérêt majeur.

Les outils de gestion doivent permettre de traiter les données documentaires relatives aux collections et doivent également être adaptés aux différentes ressources qui sont utilisées afin de produire cette information. Les outils de gestion au sein des musées sont ainsi essentiellement dédiés aux collections, aux fonds bibliographiques, aux archives et aux fonds iconographiques et multimédia.

Certains musées, essentiellement au début des années 2000, ont fait le choix d'un seul outil permettant de gérer toutes ces ressources. C'est le cas des musées de la ville de Paris qui gèrent leurs collections, leur bibliothèque et leurs archives avec une seule application Adlib (un catalogue par ressources). Un thésaurus et des règles communes soutiennent une recherche unifiée sur les différents fonds, enrichie par des liens dynamiques entre les ressources des trois catalogues. L'application peut être dotée d'une interface API (*application programming interface*, en français interface de programmation[9]) ou d'un flux XML proposant un accès aux données depuis une application web.

Dans le courant des années 2000, les avancées technologiques ont permis aux musées de privilégier le choix d'outils dédiés à chaque type de ressources. Les données sont ainsi traitées de manière plus cohérente et selon les normes prévues pour chacune de ces

[9] API : façade par laquelle un logiciel offre des services à d'autres logiciels.

ressources. Il est, en effet, complexe de parvenir à rassembler toutes les données dans une interface commune pour les mettre à disposition du public de manière unifiée. D'autant plus que ces données, selon les ressources, ont une structure de description (format) très hétérogène : certaines respectent les normes de bibliothèques (Marc, MODS et Dublin Core), certaines des normes archivistiques (EAD), certaines correspondent à des formats définis localement et certaines ne sont même pas de nature documentaire (personnes, événements). La séparation des données en plusieurs bases permet ainsi d'avoir des données répondant aux normes spécifiques à chaque métier et donc correctement structurées et répondant mieux aux critères d'interopérabilité des données exigés par le développement du web.

Actuellement, l'enjeu pour les musées, et nous verrons cela dans la troisième partie, est de pouvoir proposer à l'utilisateur une recherche unifiée entre ces données traitées ainsi en silos[10].

Mais il est important de souligner que si ces outils de gestion et de traitement de l'information reposent, bien sûr, sur les avancées technologiques développées tout au long de ces dernières années, notamment dans le cadre du web de données, ils s'appuient toujours en grande partie, pour le traitement et la structuration des données, sur les langages documentaires traditionnels. Nous allons voir maintenant leur rôle et leurs différentes utilisations, puis leur évolution dans le web de données.

[10] En silos : données traitées dans des bases de données distinctes avec un cloisonnement de l'information rendant plus difficile l'échange d'information et impliquant des redondances d'information.

2 Les langages documentaires

Les langages documentaires jouent, en effet, un rôle important dans la gestion et le traitement de l'information. Développés, utilisés et normalisés par et pour les bibliothèques et le monde documentaire depuis longtemps et bien avant le passage au numérique, les langages documentaires sont apparus plus récemment dans le monde des collections muséales et ne sont pas toujours mis à profit et normalisés dans ce domaine. Il nous paraît donc important de revenir sur leur rôle et sur leurs différentes utilisations à travers la description des trois grandes familles de langages documentaires : les classifications, les listes d'autorité ou listes de vedettes matières, et les thésaurus.

2.1 Rôle des langages documentaires

Après avoir défini et donné les fonctions des langages documentaires, nous allons voir quel est leur rôle pour la recherche et l'indexation, puis leurs différentes typologies.

2.1.1 Définition et fonctions

Il est important de définir ici la notion de « document », au sens documentaire du terme. Cette notion renvoie, en effet, à l'objet d'étude documentaire au sens large et non pas seulement aux documents textuels. Au sein des musées, cette notion renvoie aux collections muséales et aux ressources utilisées autour de ces collections. Le document peut ainsi être un objet, un ouvrage, une image, du son, etc.

Les langages documentaires permettent justement de faciliter le repérage de ces documents et d'aider à la recherche de l'information en évitant les ambiguïtés et les pièges du langage naturel : la synonymie et la polysémie. Les langages documentaires qui sont des langages artificiels sont des ensembles de termes qui sont liés entre eux par des relations conceptuelles et lexicales et qui permettent de rendre compte des concepts contenus dans le document, mais aussi des instruments d'analyse conçus spécialement pour le traitement de l'information.

L'AFNOR définit ainsi le langage documentaire comme un *« procédé conventionnel de représentation des informations d'un document sous une forme condensée et normalisée. Langage artificiel, constitué de représentations de notions et de relations entre ces notions et destiné, dans un système documentaire, à formaliser les données contenues dans les documents et dans les demandes des utilisateurs ».*

L'association des professionnels de l'information et de la documentation en donne également une définition précise[11] : « *langage contrôlé et normalisé utilisé dans un système documentaire pour l'indexation et la recherche. Un langage documentaire permet de représenter de manière univoque les notions identifiées dans les documents et dans les demandes des utilisateurs, en prescrivant une liste de termes ou d'indices, et leurs règles d'utilisation* ».

Les langages documentaires ont donc trois fonctions principales :

- normaliser la représentation des sujets des documents ;
- permettre à l'utilisateur de "naviguer" entre des sujets voisins ;
- présenter le contenu des documents sélectionnés pour faciliter le choix définitif.

Les langages documentaires résultent d'évolutions de l'organisation des connaissances et de l'accès à l'information : accroissement du volume de l'information, modification des supports d'inscription, nouveaux modes de communication de l'information [6, MENON].

2.1.2 Indexation documentaire

Les langages documentaires, que nous allons décrire plus loin, sont ainsi des outils sémantiques, dont les termes et les règles d'utilisation permettent l'indexation documentaire. Les concepts sont représentés par des termes[12] et donne accès à un document ou à un corpus de documents.

Les principales fonctions de l'indexation documentaire sont :

- d'analyser puis de représenter le contenu d'un document par des concepts représentés par des termes ;
- de repérer et de retrouver l'information contenue dans un document par un terme comme point d'accès au document.

L'indexation est donc la représentation, par les termes d'un langage documentaire, de l'information résultant de l'analyse d'un document afin d'en faciliter la recherche. L'indexation vise ainsi à représenter le contenu des documents : thèmes, sujets, aspects, etc.

---

[11] Site internet de l'ADBS, Vocabulaire de la doc. http://www.adbs.fr/langage-documentaire-17593.htm [consulté le 15/12/2013]

[12] Concept : représentation mentale abstraite d'un objet, d'une idée conçue par l'esprit. Lié à un domaine de connaissance.

L'indexation peut être :

- manuelle : indexation par l'homme ;
- automatique : indexation par la machine ;
- en texte intégral : indexation par la machine, qui contrairement à l'indexeur humain, doit analyser mot à mot pour en extraire le ou les sujets grâce aux liens sémantiques entre les termes ;
- libre : vocabulaire large proche des utilisateurs mais choix des termes sur la base d'un seul document et problèmes de synonymie et de polysémie ;
- contrôlée : utilisation d'un langage documentaire construit sur l'indexation de textes antérieurs et cohérents sur le plan terminologique.

2.1.3 Typologies des langages documentaires

Il existe trois types de famille documentaire : les classifications, les listes d'autorité ou listes de vedettes matières, et les thésaurus [7, SERRES]. Ces langages documentaires sont toujours largement utilisés mais les deux premières restent du domaine des bibliothécaires tandis que la troisième est plutôt réservée aux documentalistes.

Regroupement selon deux critères

Les langages documentaires peuvent être regroupés :

- soit selon la structure du langage et sont répartis entre les langages à *structure hiérarchique* (classifications et thésaurus) et les langages à *structure combinatoire* (listes d'autorité ou listes de vedettes matières et thésaurus) ;
- soit selon la coordination des concepts et sont répartis entre les *langages pré-coordonnés* (classifications et listes d'autorité ou listes de vedettes matières) et les *langages post-coordonnés* (thésaurus).

La pré-coordination et la post-coordination sont deux principes opposés mais aussi complémentaires dans l'utilisation des langages documentaires. Avec les langages pré-coordonnés, les combinaisons entre les termes d'un langage documentaire s'effectuent au cours de son élaboration ou au cours de l'indexation. Avec les langages post-coordonnés, les combinaisons entre les termes s'effectuent au cours de l'indexation ou au cours de la recherche documentaire.

Distinction selon la démarche d'indexation

Il existe deux démarches d'indexation :

- L'indexation systématique, qui est utilisée avec le langage à structure hiérarchique (classifications). Les concepts sont représentés sous une forme arborescente et chaque notion est située dans une « branche » précise, à un niveau de précision déterminé (construction du langage en classes, en sous-classes, en division, etc). Cette indexation se fait en langage codé (alphabétique, numérique ou alphanumérique).
- L'indexation analytique ou alphabétique, qui est utilisée avec le langage à structure combinatoire (listes d'autorité ou listes de vedettes matières et thésaurus). Les concepts sont juxtaposés, sous forme alphabétique, et les notions sont situées dans des cercles indépendants (ex. : pêche, histoire, Bretagne). Cette indexation s'opère au niveau d'un concept, ou d'une intersection de concepts.

Distinction selon la méthode de recherche

Ces types de langages documentaires induisent différentes méthodes de recherche :

- la recherche sur un domaine : classifications ;
- la recherche sur un sujet : listes d'autorité ou listes de vedettes matières et thésaurus.

Tableau récapitulatif

Typologie langages documentaires	**Structure du langage**		**Coordination des concepts**		**Démarche d'indexation**		**Méthode de recherche**	
	Structure hiérarchique	Structure combinatoire	langages pré-coordonnés	langages post-coordonnés	Indexation systématique	Indexation analytique ou alphabétique	Recherche sur un domaine	Recherche sur un sujet
Classifications	x		x		x		x	
Listes d'autorité ou listes de vedettes matières		X	x			x		x
Thésaurus		X		x		x		x

Tableau 1 : Tableau récapitulatif des typologies des langages documentaires.

Inconvénients respectifs

Chaque type de langages documentaires peut avoir des inconvénients :

- Les classifications : ce type de langage a le désavantage de disperser les documents et les sujets, selon les disciplines et les points de vue. Mais il est possible de pallier cet inconvénient en le complétant par des index alphabétiques.
- Les listes d'autorité ou listes de vedettes matières et les thésaurus : ces types de langages qui induisent une indexation analytique rassemblent les sujets, mais dispersent les disciplines et les domaines voisins. Cet inconvénient peut être pallié en complétant les listes alphabétiques par des procédés de mise en relation (renvoi, etc).

2.2 Les classifications

Les classifications, langages documentaires encyclopédiques des bibliothèques, sont les plus anciens des langages documentaires. Les classifications sont apparues dès la fin du 19ème siècle et résultent du développement éditorial de cette époque qui provoque un accroissement important des collections. Après en avoir donné une définition, nous présenterons trois types de classification : la classification décimale Dewey, la classification décimale universelle et la classification de la bibliothèque du Congrès.

2.2.1 Définition

C'est un langage documentaire encyclopédique essentiellement utilisé pour le classement dans les bibliothèques et les centres de documentation, afin de permettre un rangement cohérent des ouvrages sur les rayons, faciliter l'accès thématique aux ouvrages, et simplifier la gestion de l'accroissement.

Les classifications sont fondées sur la représentation structurée des connaissances. Chaque domaine à couvrir est ainsi découpé en sous-domaines, chaque sous-domaine étant lui-même décomposé, jusqu'à atteindre le niveau de finesse suffisant. Les notions et leurs relations sont représentées par un système de notation. Ce système lui donne des capacités d'extension qui permettent de suivre les évolutions amenées par les progrès scientifiques.

Comme nous venons de le voir plus haut, c'est un langage à structure hiérarchique et pré-coordonné qui induit une indexation systématique (concepts représentés sous une forme arborescente), en langage codé, et qui permet de regrouper les sujets par thématiques. L'indexation se fait par choix successifs, du général au particulier, par rétrécissement croissant du champ couvert. Ce langage documentaire est toujours utilisé aujourd'hui dans de nombreuses bibliothèques pour l'accès par thématique aux ouvrages.

2.2.2 La classification décimale Dewey

La classification décimale de Dewey (CDD), développé en 1876 par Melvil Dewey, bibliographe américain, est très utilisée dans le monde anglo-saxon. Ce système permet de classer l'ensemble d'un fonds documentaire dans une bibliothèque, de connaître le sujet d'un document et de le localiser dans la bibliothèque. Il est utilisé dans plus de 200 000 bibliothèques dans le monde et largement aussi sur Internet.

Le système de classification Dewey répartit les connaissances humaines en 10 grandes classes :

000	Généralités	***500***	Sciences naturelles et mathématiques
100	Philosophie et psychologie	***600***	Technologie
200	Religion	***700***	Arts beaux-arts et arts décoratifs
300	Sciences sociales	***800***	Littérature et techniques d'écriture
400	Langues	***900***	Géographie et histoire

Tableau 2 : Système de classification Dewey.

Ces classes se subdivisent en 100 divisions, puis 1 000 sections et ensuite une multitude de sous-sections.

Des normes très précises fixent la composition de l'indice d'un document. L'indice doit notamment être composé de trois chiffres au minimum (classe, division et section). Plus la notion à exprimer est fine, plus l'indice est long. Le chiffre zéro désigne les généralités.

Exemple[13] :

611.347

600 **TECHNOLOGIE (Sciences appliquées)** [classe]
610 Sciences médicales Médecine [division]
611 Anatomie humaine [section]
611.3 Organes de l'appareil digestif [sous-section]
611.34 Intestins
611.347 Gros intestin

611.347 est l'**indice** de l'ouvrage et en indique le sujet.

Figure 2: Exemple de la classification décimale Dewey.

[13] Site internet Chercher pour trouver : L'espace des élèves, École de bibliothéconomie et des sciences de l'information (EBSI), Université de Montréal. http://www.ebsi.umontreal.ca/jetrouve/ [consulté le 15/12/2013]

2.2.3 La classification décimale universelle

La classification décimale universelle (DCU) a été développée en 1905 par des juristes belges, fondateurs de l'Institut International de Bibliographie, à partir de la classification décimale Dewey. Elle a connu depuis plusieurs éditions, existe dans de nombreuses langues et représente un système général de classement largement répandu, même s'il est aujourd'hui en recul. Le développement en est assuré par le consortium UDC international[14].

Ce système reprend et améliore le système de classification Dewey avec également une répartition des connaissances humaines en 10 classes, divisées en sous-classes, etc. Des ajustements ont été faits : la classe 4a été intégralement transférée en tête de la classe 8 (linguistique), les indices n'ont pas d'obligation de longueur minimum et un enrichissement de notation par plusieurs signes est possible, permettant de combiner les indices entre eux, avec un système élaboré de ponctuation permettant d'exprimer de manière très fine des catégories thématiques.

Les capacités de ce système sont très riches mais par conséquent très complexes, ce qui peut expliquer qu'il soit finalement moins utilisé que d'autres systèmes de classifications, moins riches mais plus simple d'utilisation.

Exemple [7, SERRES] :

373.62 (441.1/.5) : Enseignement professionnel technique en Bretagne

2.2.4 La classification de la bibliothèque du Congrès

La classification de la bibliothèque du Congrès, en anglais *Library of Congress Classification*[15], a été mise au point par la bibliothèque du Congrès à Washington pour les bibliothèques universitaires. C'est le système utilisé par toutes les bibliothèques nord-américaines.

Ce système répartit également la connaissance en classes et sous-classes : 21 grandes catégories, elles-mêmes subdivisées en sujets plus précis. Ces catégories sont définies par une lettre, suivies par 1 ou 2 lettres, puis par une série de chiffres.

Exemple :

NA1-9428 : Architecture

---

[14] Site internet du consortium UDC international.
http://udcc.org/ [consulté le 15/12/2013]

[15] Site internet de la bibliothèque du Congrès, Plan de classification de la bibliothèque du Congrès.
http://www.loc.gov/catdir/cpso/lcco/ [consulté le 15/12/2013]

2.3 Les listes d'autorités et les listes de vedettes matières

Les listes d'autorités sont apparues au début du siècle et ont profité de l'accélération de la publication et de la circulation de l'information sous forme imprimée. Les catalogues imprimés, avec une indexation matière permettent la diffusion des informations bibliographiques et favorisent l'émergence des premiers « réseaux de bibliothèques » [6, MENON]. C'est un système qui s'est beaucoup développé ces dernières décennies. La liste de vedettes matières est un type de liste d'autorités. Nous allons ici en donner une définition, puis en montrer la structure et en décrire un exemple avec le répertoire RAMEAU.

2.3.1 Définition

C'est un langage à structure combinatoire et pré-coordonné, de type encyclopédique, qui induit une indexation analytique (concepts juxtaposés et notions situées dans des cercles indépendants ; indexation située au niveau du sujet).

Une vedette, appelée aussi entrée, est un terme qui représente le sujet de la notice bibliographique et sert au classement et à la recherche dans un catalogue, une bibliographie ou un index.

Une vedette matière est une vedette représentant le ou les sujets contenus dans un document. Les notices bibliographiques sont ainsi classées par ordre alphabétique des vedettes identifiant les sujets. Mais si l'indexation se fait en langage naturel, le classement et la recherche manquent de cohérence. Il est donc nécessaire de faire appel à des listes d'autorité de vedettes matières.

Une liste d'autorité est ainsi une liste des vedettes ou des termes qui doivent être obligatoirement et nécessairement utilisés dans le catalogage ou l'indexation et destinés à définir la ou les notions exprimées dans un document.

2.3.2 Structure

Une liste d'autorité est donc une liste normalisée de termes, appelés vedettes matières, ayant un caractère obligatoire. Il est impossible de créer soi-même des vedettes matières. Ces termes sont rangés par ordre alphabétique comme dans un dictionnaire.

Une vedette matière peut ne comporter qu'un seul mot ou une seule expression et est généralement complétée par d'autres éléments.

Les vedettes sont organisées en :

- vedettes principales ou têtes de vedette, qui expriment l'essentiel du sujet ;
- vedettes secondaires ou sous-vedettes :
    - de sujet,
    - géographique,
    - chronologique,
    - de forme.

Les listes d'autorité de vedettes matières sont le langage le plus proche du langage naturel mais la syntaxe est compliquée et l'usage est réservé aux professionnels.

C'est un système encore aujourd'hui très utilisé par les catalogueurs de grandes bibliothèques dans le monde : la liste de vedette matières de la Bibliothèque du Congrès américain (Library of Congress Subject Headings) dans les pays anglophones, le Répertoire de vedettes matière de l'Université Laval à Québec (Laval RVM), ou RAMEAU dans le monde francophone.

2.3.3 Rameau

Le Répertoire d'Autorités Matières Encyclopédique Alphabétique et Unifié (RAMEAU) a été constitué dès les années 1980 par la Bibliothèque nationale de France (BnF), à partir du Répertoire de vedettes matière de l'Université Laval, lui-même traduit de la liste de vedette matières de la Bibliothèque du Congrès[16]. Ce langage est largement utilisé, outre la BnF, par les bibliothèques universitaires, de nombreuses bibliothèques de lecture publique ou de recherche, ainsi que plusieurs organismes privés.

Il permet la construction de vedettes matières, selon des règles syntaxiques précises : il se compose ainsi d'un vocabulaire de termes reliés entre eux et d'une syntaxe indiquant les règles de construction des vedettes matière à l'indexation. C'est un langage très riche avec près de 200 000 entrées[17].

La spécificité de RAMEAU est la possibilité d'évolution et d'enrichissement par différentes instances (mises à jour du langage, etc). En effet, le langage RAMEAU est enrichi au fur et à mesure, à partir des propositions formulées par le réseau de ses utilisateurs. RAMEAU est

[16] Site internet de la BnF, Rameau-Langage d'indexation Rameau. http://rameau.bnf.fr/informations/intro.htm [consulté le 15/12/2013]
[17] Site internet de la BnF, RAMEAU-Statistiques. http://rameau.bnf.fr/informations/chiffres.htm [consulté le 15/12/2013]

donc alimenté à la fois par la BnF qui fournit les mises à jour trimestrielles, à partir de son fichier d'autorité matières mais aussi par le réseau des utilisateurs Rameau habilités à proposer des créations et des modifications. L'aspect organisationnel, collectif, de Rameau est très important dans le paysage des bibliothèques françaises.

Le langage RAMEAU rend possible plusieurs modes d'interrogation par sujet dans les catalogues :

- en tant que langage contrôlé associant formes retenues (ou vedettes) et formes rejetées (ou termes exclus), il permet une recherche par mot-clé ;
- en tant que langage pré-coordonné régi par une syntaxe, il permet une interrogation par feuilletage d'index ;
- en tant que langage hiérarchisé, il permet une navigation sémantique dans le vocabulaire en suivant les relations génériques, spécifiques et associatives entre les termes.

La convention RAMEAU de 2011 associe les principaux acteurs du réseau RAMEAU à travers les instances signataires : la Bibliothèque nationale de France (BnF), d'une part, le ministère de l'Enseignement supérieur et de la recherche, le ministère de la Culture et de la Communication et l'Agence bibliographique de l'enseignement supérieur (ABES), d'autre part.

Au niveau international, RAMEAU se développe et rayonne également à travers différents programmes. Il a notamment été traduit en 2008 en SKOS (*Simple Knowledge Organisation Systems*, en français Système simple d'organisation des connaissances) (cf. section 2.4.4, section 3.3.2 et partie 3 section 2.3.2) afin de pouvoir être utilisé dans l'univers du Web et des nouvelles technologies. Ce programme s'inscrit dans le projet data.bnf.fr dont nous parlerons plus longuement dans notre dernière partie.

2.4 Les thésaurus

Le thésaurus est un langage documentaire apparu dans les années 1960 avec l'informatisation documentaire. La définition officielle en anglais décrit le thésaurus comme *« controlled and structured vocabulary in which concepts are represented by terms, organized so that relationships between concepts are made explicit, and preferred terms are accompanied by lead-in entries for synonyms or quasi-synonyms »*[18]. Nous allons ici en donner la définition et la structure, présenter les différentes relations possibles entre les

[18] Norme internationale ISO 25964 « Thésaurus et interopérabilité avec d'autres vocabulaires », 2013.

termes, puis donner des indications de présentation et enfin faire un point sur les nouvelles normes et standards en cours.

2.4.1 Définition et structure

Pour le définir plus précisément, le thésaurus est un vocabulaire contrôlé et dynamique de termes représentant des concepts : termes préférentiels (descripteurs) et termes non préférentiels (non descripteurs). Ces termes obéissent à des règles terminologiques propres et sont reliés entre eux par des relations sémantiques. C'est un langage à structure combinatoire, à la fois pré-coordonné et post-coordonné, qui induit une indexation analytique et qui sert à contrôler la recherche.

Avec les thésaurus, l'indexation s'effectue au niveau de chaque concept élémentaire, chacun de ces concepts étant représenté par un ou plusieurs termes, appelés termes préférentiels. Un terme préférentiel est un terme servant à indexer un document au niveau du concept (et non du mot). Un thésaurus est constitué de :

- termes retenus : terme préférentiel ;
- termes rejetés : terme non préférentiel ;
- relations entre les termes (concepts) ;
- représentation de ces concepts sous différentes formes.

Un thésaurus permet de représenter les relations entre les concepts et facilite ainsi la recherche ce qui est sa fonction primordiale.

2.4.2 Types de relations

Il existe trois types de relations : les relations d'équivalence (entre les termes), les relations hiérarchiques (entre les concepts), les relations associatives (entre les concepts).

Les **relations d'équivalence** éliminent la synonymie et fonctionnent dans les deux sens :

- du terme non préférentiel vers le terme préférentiel : le signe utilisé est « EM » (employer) ;
- du terme préférentiel vers le terme non préférentiel : le signe utilisé est « EP » (employé pour).

Exemple : basse-taille
EM bas-relief
bas-relief
EP basse taille

Les **relations hiérarchiques** expriment les rapports de supériorité et de subordination entre les concepts (se rapprochant ainsi des classifications). Les relations hiérarchiques sont constituées de deux types de relation, identifiant le lien entre une catégorie de notions et chacun de ses cas particuliers :

- du général au spécifique : le signe utilisé est « TS » (terme spécifique) ;
- du spécifique au générique : le signe utilisé est « TG » (terme générique).

Les relations génériques et spécifiques constituent l'ossature d'un thésaurus, dirigent et contrôlent son utilisation.

Exemple : bas-relief

TG relief
TS relief écrasé
TS relief méplat

Un terme préférentiel peut renvoyer à un grand nombre de termes spécifiques, mais en principe, à un seul terme générique, celui qui est immédiatement supérieur. La nouvelle norme sur les thésaurus, que nous allons voir un peu plus loin, autorise la relation poly-hiérarchique, qui permet de relier plusieurs concepts appartenant à des catégories différentes à un concept, et de caractériser ainsi plus finement.

Les **relations associatives** expriment des analogies de signification entre les concepts, et permettent ainsi d'élargir la recherche en formant un réseau sémantique plus large et en enrichissant les possibilités de navigation entre les concepts.

- symbolisées par le signe « TA » (terme associé).

Exemple : livre

TA reliure

Les thésaurus comportent de brèves explications, *les notes d'application*, précisant les modalités d'emploi d'un terme préférentiel et permettant éviter les ambiguïtés de la polysémie :

- symbolisées par le signe « NA » (note d'application).

2.4.3 Présentation d'un thésaurus

Il existe différents types de présentation des thésaurus, généralement combinés : une présentation alphabétique, une présentation hiérarchique et une présentation graphique.

La **présentation alphabétique** peut être présentée de deux manières :

- Soit sous forme de liste alphabétique principale, correspondant à la partie principale du thésaurus. Les termes préférentiels sont ainsi présentés sous forme de liste alphabétique, chacun étant accompagné de l'indication de toutes ses relations, dans un ordre défini. C'est une présentation claire mais limitée, car elle ne permet qu'une vue partielle de la hiérarchie entre les termes.
- Soit sous forme de liste alphabétique permutée : les termes préférentiels composés apparaissent autant de fois qu'il y a de mots constituant l'expression, et à tour de rôle à leur place alphabétique. L'index permuté permet de retrouver un terme préférentiel composé à chacun de ses mots, et rassemble tous les mots ayant une racine commune. C'est un outil très pratique pour la recherche des termes préférentiels.

La **présentation hiérarchique** est un regroupement des termes préférentiels par champs sémantiques, autour d'un thème, d'un domaine. Elle met en valeur les termes préférentiels avec leurs relations génériques, spécifiques, etc. Elle permet de visualiser les différents niveaux de la hiérarchie, matérialisés par un décalage vers la droite des termes spécifiques, ou par une disposition sous forme d'arbres.

La **présentation graphique est également possible pour** exposer le thésaurus. L'ensemble des termes préférentiels est découpé en champs sémantiques, chaque groupe sémantique est doté d'une grille, et chaque terme préférentiel est assigné à une place déterminée par rapport aux deux axes (horizontal et vertical). Les relations entre les termes préférentiels sont mises en évidence par la disposition graphique (système de coordonnées, de cercles concentriques symboliques indiquant les différents niveaux) et par les flèches entre les termes préférentiels, permettant de représenter les relations. Les champs sémantiques sont appelés micro-thésaurus ou terminogrammes.

2.4.4 Normes

La conception des thésaurus s'appuie sur des standards et des normes.

La **norme internationale ISO 25964** ***« Thésaurus et interopérabilité avec d'autres vocabulaires »*** est parue en totalité en 2013. Elle est composée de deux parties distinctes :

- La partie 1, ISO 25964-1:2011, *« Thésaurus pour la recherche documentaire »,* est parue, en anglais, en août 2011. Un livre blanc en propose une présentation francophone [8, DALBIN, YAKOVLEFF, ZYSMAN]. Cette partie couvre les thésaurus mono et multilingues.
- La partie 2, *« Interopérabilité avec des vocabulaires contrôlés »*, est parue en 2013. Cette partie met en correspondance avec d'autres types de vocabulaires : les thésaurus mais aussi les classifications, taxonomies, vedettes matières, ontologies et terminologies [5, DALBIN].

Cette norme fournit des directives et des recommandations applicables à la conception et la gestion des thésaurus monolingues et multilingues, ainsi qu'à leur mise en forme afin d'optimiser leur exploitation dans des applications et sur les réseaux.

Elle remplace les normes internationales (ISO 2788:1986 : monolingue ; ISO 5964:1985 : multilingue) et françaises (NF Z47-100:1981 : monolingue ; NF Z47-101:1990 : multilingue).

Cette norme répond non seulement à une évolution des usages des thésaurus, mais surtout à un besoin de normes et de standards adaptés à la représentation, au transfert, à l'accès et à la réutilisation des connaissances, dans un contexte de multiplication des données et des flux d'informations sur les réseaux et entre les systèmes d'information. Dans ce cadre, la nouvelle norme recentre les thésaurus sur les concepts et accentue les directives sur la formalisation, la sémantisation, le multilinguisme et l'interopérabilité.

Les évolutions principales de la norme portent ainsi sur :

- les concepts : distinction formalisée entre les concepts et les termes les représentant ;
- les relations : distinction plus fine des types de relations et du formalisme imposé par le modèle de données ;
- le regroupement de concepts : regroupement à partir d'une classification, par domaines ou thématiques ; regroupements par facettes ;
- le multiculturalisme et le multilinguisme : ajustements et précisions ;
- l'intégration des thésaurus à des applications ;

- les formats d'échange et protocoles ;
- la conception, la gestion et la maintenance.

C'est une norme qui positionne clairement le thesaurus par rapport à d'autres types de vocabulaires contrôlés, qui est adaptée aux évolutions des systèmes d'information et qui est en phase avec le web sémantique et ses standards [12, RABAUT, ZYSMAN].

Le **langage formel *SKOS*[19]** est en grande partie compatible avec la norme ISO 25964. Ce système, développé et recommandé par le Word Wild Web Consortium (W3C)[20], l'organisme de standardisation des langages informatiques utilisés sur web, est construit sur la base du langage RDF[21] et spécifie le format et les spécificités fonctionnelles permettant la publication et l'utilisation des systèmes d'organisation de concepts, tels que les thésaurus, dans le cadre du web sémantique (cf. section 3.3.2 et partie 3 section 2.3.2).

Ces nouvelles normes et standards montrent à quel point les langages documentaires sont encore d'une actualité « brûlante » dans le traitement numérique de l'information et du web sémantique et ne sont ni statiques, ni obsolètes, mais offrent de nouvelles possibilités d'utilisations et d'évolutions d''indexation et de recherche permettant l'amélioration de l'organisation des connaissances sur le web de données.

[19] SKOS (Simple KnowledgeOrganization System) : langage de représentation des vocabulaires structurés et d'intégration sur le web sémantique.
[20] Site internet du W3C.
http://www.w3c.fr/ [consulté le 15/12/2013]
[21] RDF (*Resource Description Framework*) : langage à la base de toute indexation sur le Web ; norme fondamentale du web sémantique.

3 L'évolution des langages d'indexation et de recherche dans le cadre du web de données

Il serait, en effet, logique de penser que l'explosion des technologies liées à internet et les mutations informationnelles auraient pu avoir comme conséquence la disparition des langages documentaires. Mais, en réalité, comme nous venons de le voir, les langages documentaires sont, encore aujourd'hui, tout à fait adaptés pour être utilisés dans le cadre du web de données qui en exploite les possibilités sémantiques ;

Malgré tout, la complexité technique et conceptuelle de l'interopérabilité a conduit les spécialistes à envisager également de nouveaux moyens pour l'organisation du savoir sur le web et à penser de nouveaux systèmes de représentation et de structuration, tels que le développement de taxonomies et la constitution d'ontologies. Ces nouveaux systèmes de gestion des connaissances constituent un tremplin pour assurer l'avenir des langages documentaires dans le contexte du web sémantique qui constitue le socle du web de données.

3.1 Les taxonomies (ou taxinomies)

Les termes taxonomie ou taxinomie sont équivalents et peuvent être utilisés de la même manière.

Ces termes sont peu cités dans la littérature des sciences documentaires et de l'information mais sont en revanche largement utilisés dans le monde industriel. Les taxonomies sont, en effet, apparues avec le développement du web marchand et des sites intranet d'entreprises. Ont alors été créés des cadres de présentation et de découverte des ressources qui sont les taxonomies. Comme les classifications, le système des taxonomies est un langage documentaire de type hiérarchique.

Il est malgré tout difficile de définir exactement ce que sont les taxonomies. Le dictionnaire de français Larousse définit ce terme comme *« science des lois de la classification ; classification, suite d'éléments formant des listes qui concernent un domaine, une science ».* Wikipédia précise que les taxonomies ont pour objectif de décrire les organismes vivants et de les classer, et cette définition s'étend maintenant à d'autres sciences telles les sciences humaines et sociales, les sciences de l'information ou l'informatique. Mais ce n'est pas suffisant pour comprendre ce que sont les taxonomies et il apparaît nécessaire de se pencher davantage sur leurs fonctions plutôt que sur leur structure pour déterminer leur nature.

Bruno Menon, dans un article sur les langages documentaires paru en 2007 dans la revue *Documentaliste-Sciences de l'Information*, tente ainsi d'en donner une définition ou plutôt une fonction : *« une taxonomie est un cadre d'organisation pour des ressources numériques de toutes natures, destiné à en permettre une présentation ordonnée et y donnant accès par navigation hypertextuelle*. Bruno Menon précise également que ces ressources numériques n'étant pas toujours documentaires, les taxonomies ne sont effectivement pas toujours des langages documentaires.

Les taxonomies décrivent essentiellement ce qui est générique-spécifique. Cette relation crée une structure arborescente, hiérarchique, qui organise des concepts[22]. Les taxonomies permettent, de cette manière, de structurer une grande variété d'objets : des objets concrets, mais aussi des événements.

Frédéric Saint-Dizier éclaircit aussi la nature des taxonomies en précisant dans son article « Taxonomie » dans le dictionnaire de Sémantique *Sémanticlopédie* que les taxonomies *« se caractérisent par deux propriétés fondamentales : la transitivité et l'héritage de propriétés. La transitivité permet de dire que : si C est un sous-type de B et B un sous-type de A alors C est un sous-type de A (…) ».* Quant à *« l'héritage descendant des propriétés, elle induit que si B est un sous-type de A, alors, sauf mention contraire, il hérite de toutes ses propriétés ».*

Les systèmes d'ontologies les plus récents développent une architecture de représentation de connaissances qui nécessite un approfondissement de la notion de taxonomie.

3.2 Les ontologies

L'ontologie, au singulier, est une discipline philosophique, définie comme la théorie de l'être (dictionnaire de français Larousse) et qui étudie les propriétés générales de l'être. Au pluriel, les ontologies visent à fonder une démarche de représentation des connaissances à partir de ressources diverses utilisées initialement dans des contextes comme l'intelligence artificielle, l'ingénierie linguistique ou la gestion des connaissances. Les ontologies sont devenues une composante centrale dans bon nombre d'applications et sont appelées à jouer un rôle central dans la construction du Web de données[23].

Une littérature de plus en plus importante leur est consacrée. Les ontologies y sont décrites comme constituants des systèmes à base de connaissances. Bruno Menon précise que les

---

[22] SAINT-DIZIER Patrick. Taxonomie. In D. Godard, L. Roussarie et F.Corblin (éd.), Sémanticlopédie/ dictionnaire de sémantique, GDR Sémantique & Modélisation, CNRS, 2006. http://www.semantique-gdr.net/dico/ [consulté le 09/12/2013]

[23] Site de la BnF, Élaboration d'ontologies. http://www.bnf.fr/fr/professionnels/modelisation_ontologies/a.elaboration_ontologies.html [consulté le 15/12/2013]

ontologies étant un *« trait d'union entre théorie et ingénierie, l'approche méthodologique ou ingénierie ontologique, y occupe également une place importante ».*

En réalité, parmi tous les types d'organisations des connaissances qui ont été développés jusqu'ici pour permettre à l'humain de communiquer aux machines le sens des mots, les ontologies informatiques sont les plus puissantes.

Basées sur des langages spécifiques, RDF et OWL[24] (cf. section 3.3.2), elles permettent de décrire l'environnement sémantique d'un concept de la façon la plus détaillée et formalisée. Il s'agit concrètement d'un référentiel de vocabulaire partagé, qui présente une vision consensuelle des experts d'un domaine autour des concepts manipulés dans ce domaine. De plus, les ontologies présentent un degré bien plus sophistiqué que les autres modèles pour la description des relations entre les concepts, puisque celles-ci peuvent être décrites en détail dans la modélisation.

Grâce aux programmes de recherche déjà menés pour développer cette technologie, il existe aujourd'hui des ontologies stables et largement étendues dans des domaines aussi variés que la médecine, la géopolitique, le droit, la finance, etc. L'étendue des travaux réalisés aujourd'hui permet d'extraire des sous-parties d'ontologies généralistes existantes (un ou plusieurs champs sémantiques) pour constituer ou compléter sa propre ontologie de domaine.

On peut se demander quelle est réellement la structure des ontologies et quelles sont les différences entre les ontologies et les autres langages documentaires tels que les thésaurus et les taxonomies. Les ontologies, en effet, incluent les langages documentaires mais leur développement s'appuie également sur d'autres pratiques.

Quoi qu'il en soit les ontologies, comme les langages documentaires traditionnels, sont développées autour de l'indexation et des métadonnées mais cherchent en outre à articuler les éléments terminologiques et/ou conceptuels d'un ou de plusieurs domaines de connaissance ou d'activité.

[24] OWL : *Web Ontology Language.*

Afin de replacer tous ces éléments dans le contexte actuel, il est possible, comme le dit Bruno Menon citant Gail Hodge, de parler de « *système d'organisation des connaissances* » incluant :

- les listes de termes : listes d'autorité, glossaires, dictionnaires ;
- les schémas de classement : classifications, plans de classement et taxonomies ;
- les réseaux de concepts : thésaurus, réseaux sémantiques et ontologies.

Et finalement, comme le fait remarquer Béatrice Pierre dans son mémoire sur l'avenir des langages documentaires dans le cadre du Web sémantique, il n'y a qu'une notion de granularité entre chaque système. Nous reprenons ici, afin de le démontrer, le schéma récapitulatif de Fred Leise[25] dont Béatrice Pierre s'est inspirée et qu'elle a complété, afin d'en synthétiser les nuances :

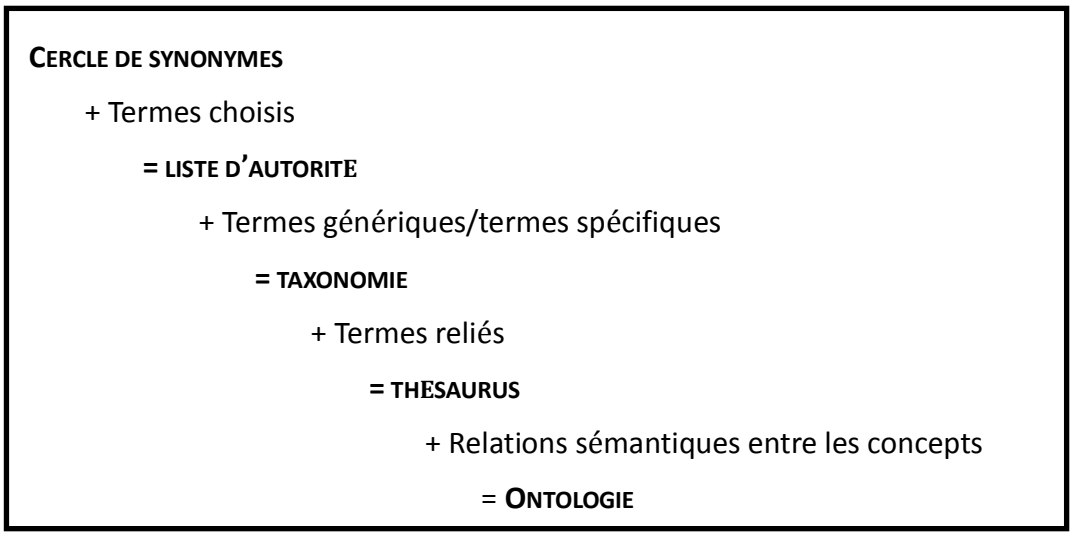


Figure 3 : Notion de granularité entre les différents langages documentaires.

C'est donc sur ces langages documentaires, qui sont des outils sémantiques incontournables pour le traitement de l'information, l'indexation et la recherche, que repose le web de données et le web sémantique dont nous allons voir maintenant les fonctionnalités pour lesquelles il a fallu faire évoluer les fondamentaux du web.

[25] LEISE Fred. Controlled vocabularies: an introduction.2005, p 124.

3.3 Le web de données et le web sémantique

3.3.1 Repères

Le web de données, en anglais *linked data* (données liées) allie les principes fondamentaux du web (identifiants URI[26], protocole HTTP[27]) aux principes du web sémantique. Son but est de traiter l'infobésité de l'information sur internet (*big data*) et de la rendre exploitable en garantissant :

- la publication de données structurées sur le web ;
- l'interopérabilité entre les données, quels que soit leur source, système ou application d'origine ;
- la mise en relation de ces données.

C'est ainsi un moyen de mettre en relation la multitude des données produites en silos isolés et de constituer un réseau global d'information, cohérent, disponible en ligne et facilement réutilisable.

Le principe du web sémantique, appelé aussi « *Web 3.0* », a été lancé en 2011 par l'anglais Tim Berners Lee, inventeur du Word Wilde Web (WWW)[28] et directeur du W3C (cf. section 2.4.4) qui explique que le web a besoin de sémantique et qui cherche à relier le web au monde réel par la sémantique. Le web sémantique répond aux objectifs du web de données en rendant les données plus facilement réutilisable et interprétables automatiquement par des machines. On permet ainsi de simplifier les échanges de données entre les acteurs et de traiter le *big data*.

3.3.2 Nouveaux systèmes de description des données

Le web originel repose sur trois notions fondamentales :

- l'identifiant URI qui permet de nommer et d'indiquer n'importe quelle ressource sur le web et de lui donner un chemin d'accès (ex : http://www.musee-rodin.fr) ;
- le protocole HTTP qui permet de transférer les données ;

[26] URI (*Uniform Resource Identifier*) : norme du Word Wide Web permettant d'identifier une ressource sur le web.
[27] HTTP (*HyperText Transfer Protocol*) : protocole de communication client-serveur développé pour le web.
[28] WWW : littéralement « toile » (d'araignée) mondiale et communément appelé web. Le web est un système hypertexte fonctionnant sur Internet dont il est une des applications.

- le langage HTML *(Hypertext Markup Language*) qui est un langage de balisage permettant de définir l'emplacement des données sur une page web et de les publier, chaque page étant reliée aux autres par des hyperliens.

Le web de données repose également sur les deux premières notions mais pour le mettre en œuvre, il a été nécessaire de faire évoluer le langage HTML. En effet, le web sémantique relie non pas des pages mais des données et s'adresse en premier lieu à des machines. Il a donc fallu travailler sur la structuration documentaire des données (séparer le contenu et la forme dans les pages) et créer un nouveau langage pour la description des données :

- Le langage XML (*eXtensible Markup Language*) qui permet de représenter une information structurée.

Et afin de représenter des domaines de connaissances, sous forme de modèles, et de les lier à des ressources, ceci permettant de créer des interactions entre les données, il a fallu créer encore deux autres « briques » :

- Le langage RDF (*Resource Description Framework*, en français Cadre de description des ressources) qui permet de décrire les données ; le triplet RDF est la norme fondamentale pour le web sémantique depuis 2004 [13, SAJUS] :
  - R comme Ressources : tout type de ressources possédant un URI ;
  - D comme Description : caractéristiques et concepts des ressources permettant de les catégoriser et de les relier ;
  - F comme Framework : modèle et syntaxe de ces descriptions, programme qui permet de lire ces ressources ;
- SparQL (*Protocol Knowledge Organization System*) qui permet de retrouver les données, et qui est un langage d'interrogation s'appliquant aux données RDF.

Enfin, pour bien interpréter les données, il est nécessaire de créer des schémas qui expliquent le sens, les catégories, les relations, les contraintes des données et de leur structures tels que :

- OWL (*Web Ontology Language*) qui est un langage de représentation des connaissances construit sur le langage RDF et qui permet de construire des ontologies web structurées ;
- SKOS (*Simple Knowledge Organization System*) qui permet de représenter des vocabulaires contrôlés et structurés comme les thesaurus et de les intégrer sur le web sémantique.

Ces langages ou « vocabulaires » permettent aux machines d'appréhender le sens des concepts en les renseignant sur leurs définitions, leurs usages et les relations sémantiques

qui les lient entre eux. Le tri des données disponibles sur le web, rendu plus efficace, garantit ainsi un accès facilité et rapide à une information pertinente pour l'utilisateur[29].

Profitant de ces dernières normes et technologies, de nombreux projets sont menés, initiés principalement par les États ou des organismes de recherche, visant aussi bien à encourager la découverte scientifique qu'à permettre le partage de données. Nous verrons plus en détail dans notre troisième partie quels sont les enjeux du web de données pour les musées et quels sont les projets initiés par le MCC.

Cette première partie nous a permis de dresser un tableau théorique du contexte des collections muséales, des langages documentaires et des dernières avancées technologiques fondamentales pour la gestion de l'information et de la documentation. Dans notre deuxième partie, la description d'un cas pratique telle que la collection du musée Rodin va nous permettre de mesurer toute la problématique de la structuration des données et de l'utilisation des langages documentaires pour les collections muséales dans le cadre d'un projet de réinformatisation des données.

[29] KELLER Loraine, INTD, promotion 43, projet Web « Antalaji ».
http://user12.plateforme06-yws.fr/ [consulté le 15/10/2013]

Deuxième partie :
Le cas du musée Rodin, structuration des données et élaboration d'un thésaurus dans le cadre d'une réinformatisation

Un vaste chantier de réinformatisation est en cours actuellement au musée Rodin impliquant les collections du musée, mais aussi les ressources gravitant autour de ces collections (ouvrages, archives et images). Ma mission de stage a porté, dans le cadre de ce projet de réinformatisation et du changement d'outil de gestion des collections, sur la restructuration des données et sur l'élaboration d'un thésaurus propre aux collections du musée. Nous allons présenter dans cette partie le contexte du projet, puis nous ferons un point sur la méthode de description et d'indexation d'un objet d'art et nous terminerons en présentant la méthodologie utilisée afin d'élaborer un thésaurus propre aux collections du musée Rodin, les difficultés rencontrées et les préconisations proposées.

1 Contexte du projet

Afin de bien comprendre le contexte de la mission qui m'a été confiée, et d'en établir le périmètre, nous procéderons, après une présentation des collections du musée Rodin, à une analyse de l'existant et des besoins, puis nous expliquerons ensuite comment la mission a été organisée.

1.1 Présentation des collections du musée Rodin

Au cœur du musée Rodin se trouvent des collections mondialement connues, dont le traitement, la gestion documentaire et la mise en ligne sur le web, sont des priorités dans le cadre du projet de réinformatisation.

1.1.1 Renommée des collections

Le musée Rodin a été créé en 1916 à l'initiative de Rodin, suite aux trois donations consenties par l'artiste à l'État. C'est un musée national, sous tutelle du MCC, devenu établissement public en 1993. Les missions du musée sont la sauvegarde, l'étude, la mise en valeur et la diffusion de l'œuvre et des collections de Rodin.

Rodin a non seulement fait don à l'État de son œuvre mais aussi de ses droits de propriété artistique : le musée Rodin est investi de la qualité d'ayant droit de Rodin au sens des dispositions légales sur la propriété littéraire et artistique. Aux termes du décret n° 93-163 du 2 février 1993, le musée a notamment pour mission de *« faire connaître l'œuvre de Rodin et de faire respecter le droit moral qui y est attaché » (article 2)*[30].

Le musée a ainsi l'opportunité de pouvoir vendre des épreuves originales, ce qui fait une de ses particularités par rapport aux autres musées : *« L'article R 122-3 du Code de la Propriété Intellectuelle prévoit que sont considérées œuvres d'art originales, les éditions de sculptures dans la limite de 12 exemplaires numérotées et épreuves d'artistes confondus. Le musée Rodin met en pratique cette limitation prévue par le décret n° 93-163 du 2 février 1993, décret relatif au musée Rodin version consolidée au 7 décembre 2005 au terme duquel les éditions originales du musée Rodin sont limitées à 12, numérotées de 1/8 à 8/8 et de I/IV à IV/IV, y compris les éditions originales existantes »*[31].

[30] Décret n°93-163 du 2 février 1993 relatif au musée Rodin.
[31] Site internet du musée Rodin.
http://www.musee-rodin.fr/fr/professionnels/respect-du-droit-moral [consulté le 15/10/2013]

Le musée Rodin est constitué de deux sites : celui de la rue de Varenne, structuré autour de l'hôtel Biron et de son jardin, un hôtel particulier construit au début du XVIIIe siècle, que Rodin occupait, comme locataire, depuis 1908 et où sont présentées les œuvres majeures de Rodin ; le site de Meudon, constitué autour de la villa des Brillants, résidence de Rodin qui en a également fait don à l'État, de la tombe de l'artiste et du musée où sont présentés des plâtres soulignant ainsi l'aspect atelier de l'artiste.

Les collections présentées par le musée Rodin comprennent :

- les œuvres de Rodin : 6 745 sculptures, 6 200 dessins et 200 peintures ;
- ses collections d'antiques : 6 418 objets ;
- la totalité de ses manuscrits et archives (fonds historiques) ;
- sa collection de photographies : 12 000 ;
- ses meubles et le mobilier muséographique : 256.

La renommée internationale de l'artiste, auteur de sculptures universellement connues, telles *Le Penseur*, devenu une icône de l'art occidental, ou encore *Le Baiser,* contribue à la notoriété des collections.

Des projets successifs ont été menés ces dernières années : en 1997, construction des réserves à Meudon ; en 2005, aménagement de l'ancienne chapelle de Paris afin de créer une salle d'exposition temporaire, des bureaux pour l'administration et une bibliothèque accessible au public ; en 2007, projet de réinformatisation ; en 2012, ouverture du chantier de l'hôtel de Biron et nouvelle impulsion au projet de réinformatisation ; envisagé pour fin 2014, présentation d'un parcours et d'une muséographie renouvelés et mise en ligne d'un portail documentaire et d'une recherche fédérée.

1.1.2 Gestion des collections et de l'information

Les lieux de gestion de l'information autour des collections muséales du musée Rodin sont concentrés au sein du service des collections et du service de la recherche, de la documentation, de la bibliothèque et des archives.

Au sein du service des collections, la gestion de l'information est sous la responsabilité de conservateurs, d'attachés de conservation, de régisseurs des œuvres et, depuis peu, d'un chargé d'études documentaires des collections. Le poste de chargé d'études documentaires des collections a été créé en 2009 dans le cadre du projet de réinformatisation. Il n'existait pas auparavant malgré des ressources documentaires importantes et une informatisation des données dès les années 1990 ; celui-ci est en charge de la gestion de la documentation

des collections (les dossiers d'œuvres) et de l'administration de la base de gestion des collections.

Au sein du service de la recherche, de la documentation, de la bibliothèque et des archives la gestion de l'information est sous la responsabilité de conservateurs, d'attachés de conservation et, depuis peu, d'un bibliothécaire et d'un archiviste. Ce service s'occupe de la gestion du fonds historique (antérieur à la mort de Rodin : manuscrits et archives), des dossiers documentaires, de la bibliothèque et des archives institutionnelles (postérieures à la mort de Rodin) et du fonds photographique. Le bibliothécaire a été recruté en 2007 et l'archiviste en 2009 afin de participer notamment à la réinformatisation des systèmes d'information du musée Rodin. Ce sont également des créations de poste, comme celui de chargé d'études documentaires des collections.

Le webmaster, qui avait dirigé la refonte complète du site internet, a quitté le musée Rodin en 2012. En remplacement, un responsable des systèmes d'information a été recruté, afin d'améliorer les systèmes informatiques du musée Rodin et diriger la réinformatisation actuelle du musée, notamment par l'élaboration d'un cahier des charges en vue d'une nouvelle base de gestion des collections. La nouvelle base de gestion des collections est, en effet, une étape clé dans le processus de réinformatisation du musée Rodin.

1.1.3 Les collections sur le web

Le musée Rodin a procédé à la refonte complète et à la mise en ligne de son nouveau site internet en 2009. Ce site est un outil d'information et de documentation qui souhaite favoriser la découverte de la vie et de l'œuvre de Rodin.

Toutefois, la base de gestion des collections ne bénéficiant pas d'interface de restitution en ligne sur le site internet, le site web présente donc comme beaucoup d'autres sites web de musée, les œuvres qui sont considérées comme majeures, avec l'exposé de notices d'œuvres choisies, le développement d'une collection de fiches éducatives et la présentation d'œuvres en 3D.

Mais, première étape du projet de réinformatisation, le service de la recherche, de la documentation, de la bibliothèque et des archives vient de procéder à la mise en ligne de l'interface de recherche du catalogue de la bibliothèque et de la bibliothèque numérique[32].

[32] Bibliothèque numérique : par comparaison avec un catalogue de bibliothèque en ligne, dans une bibliothèque numérique, l'utilisateur accède au document numérique lui-même et non à sa seule notice.

Il est prévu, à terme, et cela montre la volonté d'ouverture du musée Rodin, de mettre en ligne un portail documentaire numérique proposant une recherche fédérée sur les différentes bases de données du musée : collections, fonds historiques, bibliothèque, archives, fonds photographique.

Cette présentation des collections du musée Rodin pose déjà quelques problématiques que nous allons développer maintenant en présentant l'analyse de l'existant et des besoins.

1.2 Analyse de l'existant et des besoins

Afin de procéder à cette analyse, nous allons faire un point sur les ressources documentaires liées aux collections du musée Rodin, sur les modes de diffusion de l'information et des connaissances ainsi que sur le projet de changement de logiciel de gestion des collections, ce qui nous permettra de présenter le besoin fondamental du musée Rodin concernant l'harmonisation de ses données.

1.2.1 Ressources documentaires

Documentation des collections

Le service des collections dispose d'une réserve où est classée la documentation papier des collections (les dossiers d'œuvres) par domaine (antiques, dessins et sculptures). Il n'existe pas de plan de classement commun, ni de dossier par œuvre.

Cette documentation était classée jusqu'en 2005 dans les bureaux des conservateurs. Chaque conservateur était chargé d'un domaine de collection, en quelque sorte « sa » collection, et constituait « sa » documentation pour laquelle il avait son propre classement.

En 2005, suite au déménagement des bureaux, tous les dossiers d'œuvre ont été rassemblés dans un seul espace, le classement restant toutefois par domaine de collections :

- Antique : par zone géographique ;

- Dessins, gravures, peintures : par auteur et par numéro d'inventaire ;

- Sculptures : par sujet et par titre.

Ce changement est symptomatique d'une évolution dans l'organisation du travail au sein du musée où les collections et la documentation « n'appartiennent » plus aux seuls conservateurs. Auparavant, un conservateur pouvait passer sa carrière dans un même musée et être en charge d'une collection pendant de longues années. Il avait ainsi la mémoire de l'information et était quasiment le seul référent sur le sujet. Actuellement, un

conservateur change de poste et de musée beaucoup plus souvent. L'information doit être transmise et diffusée de manière beaucoup plus rapide, pérenne, structurée et accessible à tous.

Le chargé d'études documentaires des collections travaille ainsi, depuis son arrivée en 2009, à une homogénéisation du classement de la documentation afin que les documents puissent être retrouvés plus facilement. Il répond ainsi à un besoin primordial pour la recherche d'information découlant d'une situation dont l'historique est encore très prégnant.

Une partie seulement des données contenues dans la documentation papier a été reprise puis enrichie dans la base de gestion des collections dès les années 1990. Cette base est constituée de modules distincts par domaine de collection, reprenant ainsi la séparation existante dans la documentation papier :

- Antiques
- Dessins
- Gravures
- Mobilier
- Peintures
- Photographies
- Sculpture

La recherche peut se faire soit par module, soit sur la collection complète, mais elle reste difficile à mener. Il est donc nécessaire de pouvoir harmoniser les données et de créer des outils communs de description et de recherche.

Fonds historique, fonds photographique, bibliothèque, archives

Le service de la recherche, de la documentation, de la bibliothèque et des archives dispose de réserves où sont classés le fonds historique du musée, le fonds photographique, les dossiers documentaires, les ouvrages de la bibliothèque et les archives institutionnelles.

L'informatisation de ces fonds est inégale selon les ressources, hormis le catalogue de la bibliothèque et de la bibliothèque numérique dont le processus de réinformatisation est abouti.

Le catalogue de la bibliothèque et de la bibliothèque numérique est géré par le logiciel PMB (système intégré de gestion des bibliothèques), développé et intégré par la société PMB Services. C'est un logiciel libre de gestion des bibliothèques et full web. La base

« bibliothèque », initialement gérée sur 4D, a été exportée en 2010 sur le logiciel PMB répondant aux standards en vigueur pour les bibliothèques.

Les données du fonds historique, des archives institutionnelles et du fonds historique sont gérées soit sur le logiciel 4D qui est obsolète, soit sur Microsoft Excel. La migration de ces données sur des logiciels adaptés à chacune des ressources constituera la prochaines étape du projet de réinformatisation qui interviendra après la mise ne place d'un nouveau logiciel de gestion des collections.

1.2.2 Diffusion de l'information

La documentation des collections est à disposition de tous les agents du musée Rodin en accès libre. La communication de ces dossiers d'œuvres peut se faire également au sein de la salle de lecture du musée Rodin, notamment pour des chercheurs extérieurs. C'est une documentation qui reste très utile pour la compréhension des œuvres, même si elle est parfois datée. Elle devrait surtout être reprise dans la base de gestion des collections pour une gestion efficace de l'information, ce qui n'est pas toujours le cas.

La base de gestion des collections est pour le moment uniquement accessible aux agents du musée Rodin. Toutes les personnes concernées devraient pouvoir accéder à cette base très régulièrement et y trouver les informations utiles à leur travail. En réalité, cette base étant compliquée d'utilisation et n'ayant pas d'interface de recherche, un certain nombre de personnes, même au sein du service des collections, ne va pas y chercher les informations, alors même qu'elles s'y trouvent. Les utilisateurs perdent donc du temps en cherchant l'information soit dans la documentation papier, soit dans la base de gestion des collections et ne la trouvent pas toujours. Un travail de rationalisation serait donc nécessaire afin de faciliter la recherche d'information sur les collections.

Les fonds historiques du musée, les dossiers documentaires, les ouvrages de la bibliothèque et les archives institutionnelles sont mis à la disposition des lecteurs au sein de la salle de lecture du musée Rodin, les fonds en réserve sur demande, et les ouvrages en accès libre. Le public visé, outre les agents du musée Rodin, est un public extérieur : lecteurs, étudiants, chercheurs, professionnels de l'art et amateurs.

Ceux-ci peuvent désormais consulter sur le web le catalogue de la bibliothèque et de la bibliothèque numérique, tout récemment mis en ligne et accessible via le site internet du musée, mais n'ont pas accès aux autres données numériques existantes sur les autres fonds.

L'accès aux données reste donc difficile : il faut prendre rendez-vous et justifier sa demande. Le lecteur est tributaire du responsable de chaque fonds pour la diffusion des documents et de l'information. Pourtant, la demande d'information venant d'un public extérieur augmente

chaque année et est de plus en plus internationale, Rodin, ses œuvres et ses collections ayant une réputation mondiale.

Le portail documentaire numérique prévu dans le cadre du projet de réinformatisation permettra de rendre mieux accessibles à tous, les collections, les ressources et les données du musée Rodin.

1.2.3 Logiciel de gestion des collections

Le musée Rodin est doté actuellement, pour la gestion de ses collections, du logiciel 4D (4e Dimension), développé et intégré par la société Graine de code. C'est un système de gestion de base de données relationnelle (SGBDR). Ce logiciel est utilisé par le musée Rodin, depuis 1989, en version monoposte, comme système d'information et de gestion des collections du musée, mais également pour le service courrier et la réservation des groupes.

Il permet de traiter les données documentaires sur l'œuvre du sculpteur et ses collections, gérer le courrier administratif et la réservation des visites. En 2002, le musée Rodin décide de refondre et de fusionner ses 43 bases de données sur le logiciel 4D. En 2006, le logiciel est divisé en 4 bases distinctes :

- Base 4D « collections » (environnement client-serveur) ;
- Base 4D « manuscrits » (stockée sur un monosposte) ;
- Base 4D « réservation » (partage du même serveur que celui des collections) ;
- Base 4D « courrier » (hébergé sur un serveur dédié).

La base 4D « collections » a donc été développée en propre par le musée Rodin, et notamment par le service des collections, composé d'historiens de l'art sans formation documentaire, bien avant l'arrivée d'un chargé d'études documentaires des collections et sans administrateur de base de données. Il en résulte une base tentaculaire composée de multiples champs non contrôlés. Ces champs, jusqu'à l'arrivée du chargé d'études documentaires des collections, étaient en texte libre sans référence à des listes de vocabulaire, sans vocabulaire contrôlé.

Depuis son arrivée en 2009, le chargé d'études documentaires des collections restructure la base, établit des listes de vocabulaire et met en place un contrôle de saisie sur les champs. L'information est ainsi mieux structurée. Mais le travail restant à faire est encore très important.

La base 4D « collections » est de plus obsolète. Elle ne permet aucune des fonctionnalités proposées par les logiciels de gestion des collections sur le marché actuellement :

impression, saisie intuitive, gestion des droits en lecture écriture, traçabilité des modifications apportées au contenu ou à la structure des notices, sécurité des données, interface de gestion, interface de recherche, recherche fédérée, travail collaboratif, *full web*, *open source*, évolutivité. C'est également un logiciel inadapté car il a été conçu pour le monde de la santé.

La sécurité des données n'est pas non plus assurée car il est arrivé plusieurs fois que des centaines de données disparaissent sans qu'aucune explication ne soit trouvée. La sauvegarde de la base n'étant pas assurée, ces données sont définitivement perdues.

Nouveau logiciel de gestion des collections

La base 4D « collections » ne permettant plus de répondre aux besoins du musée Rodin, le choix d'un nouveau logiciel est en cours parmi les standards du marché. Cette orientation est considérée par la direction du musée Rodin comme une urgence pour le quotidien comme pour le travail et la crédibilité scientifique de l'établissement. Ce projet a été confié au responsable des systèmes d'information, recruté par le musée Rodin en octobre 2012, ainsi qu'au chargé d'études documentaires des collections. L'outil devra remplir des critères d'ergonomie, de pérennité, d'ouverture vers d'autres systèmes, d'interface avec lnternet, et être un outil *full web*. Il devra être doté d'un moteur de recherche adapté et performant.

Le chargé d'étude documentaire des collections continue de travailler sur une meilleure structuration des données dans la base 4D « collections » afin de préparer au mieux la migration des données dans le nouveau logiciel qui sera choisi.

Un audit de la base a été réalisé, en juillet 2011, par le chargé d'études documentaires des collections assisté de la société APIMEC. Des ateliers de travail ont été instaurés au premier semestre 2013 avec les personnes qui utilisent la base, afin de réaliser une étude des besoins qui sert à l'élaboration du cahier des charges, actuellement en cours avec l'aide d'une assistance à maîtrise d'ouvrage.

1.2.4 Harmonisation des données

Historiquement, les ressources du musée Rodin sont gérées séparément par le service des collections et le service de la recherche, de la documentation, de la bibliothèque et des archives, et au sein même de ces services, de manière hétérogène. Il en résulte des données redondantes et il en découle des recherches souvent fastidieuses.

L'harmonisation des données est donc un besoin fondamental dans le cadre du projet de réinformatisation et du changement de l'outil de gestion des collections. Pour cela, il est nécessaire :

- dans un premier temps, de réfléchir à une meilleure structuration des données sur le logiciel de gestion des collections dans le cadre de l'élaboration d'un cahier des charges pour un nouveau logiciel ;
- dans un second temps, de s'appuyer sur la restructuration des données afin d'élaborer une liste de vocabulaire contrôlé, sous forme de thésaurus, propre aux collections du musée, permettant une description plus fine et plus homogène et une recherche plus efficace des données ;
- dans un troisième temps, d'utiliser ce vocabulaire de référence et de l'enrichir afin qu'il soit utilisé pour les autres ressources du musée Rodin, le but étant d'harmoniser les données ;
- dans un dernier temps, de mettre en place un portail documentaire numérique proposant une recherche fédérée sur toutes les bases de données.

Ma mission de stage est intervenue au niveau de la deuxième étape avec l'élaboration d'un thésaurus utilisable, dans un premier temps, sur le logiciel de gestion des collections. Cette mission s'est faite dans le cadre de la réflexion concernant la restructuration des données, déjà initiée depuis son arrivée par le chargé d'études documentaire des collections, et au moment de l'élaboration du cahier des charges fonctionnel pour un nouvel outil de gestion des collections.

1.3 Organisation de la mission

Afin de réaliser au mieux ma mission, j'ai réalisé un cahier des charges délimitant le cadre de la mission et constitué d'un rétroplanning mis à jour toutes les semaines.

1.3.1 Cahier des charges

La rédaction d'un cahier des charges s'avère, en effet, extrêmement utile, la réalisation d'un thésaurus étant complexe et fastidieuse [9, DÉGEZ]. Il est également utile pour cadrer ensuite la mission, même si elle n'est pas réalisée de bout en bout par la même personne. Les éléments à prendre en compte sont le type et le nombre d'œuvres sur lesquels portera le thésaurus, les spécialités concernées, les publics et les besoins, les objectifs et enfin les ressources disponibles.

La difficulté de ce cahier des charges était qu'il ne m'était pas possible, dans les délais qui m'étaient impartis, d'élaborer un thésaurus dans son entier. En effet, même si un thésaurus

n'est jamais terminé, il faut compter, dans le cas du musée Rodin, entre six mois et un an pour en élaborer un complet. Ce cahier des charges délimite donc ma mission à une première ébauche du thésaurus, ce qui permettra, s'il est jugé pertinent, de l'ajuster et de le compléter dans un deuxième temps. Le but en quelque sorte est de créer un thésaurus « test ».

Type et nombre d'œuvres

La base de gestion des collections comprend l'ensemble de la collection du musée Rodin, ce qui correspond à 23 000 notices, si l'on excepte les photographies. Comme je viens de l'expliquer, il n'était pas possible dans le court laps de temps de ma mission d'élaborer une liste de vocabulaire contrôlée référente pour l'ensemble de la collection. Il a donc été décidé que je travaillerai plus précisément sur les dessins et sur les sculptures :

- Dessins : 7 977 notices ;
- Sculptures : 6 774 notices.

J'ai toutefois pris en compte le fait que ce travail devra être rapidement élargi au reste de la collection :

- Antiques : 6 454 notices ;
- Gravures : 1 157 notices ;
- Mobilier : 256 notices ;
- Peintures : 193 notices.
- Photographies : 17 000 items (dont 10 000 étant considérées comme appartenant aux collections et l'ensemble étant actuellement géré dans la base de gestion des collections).

Publics et besoins

Le service de la conservation est le public principal visé par les premières étapes du projet et dont les besoins sont les plus importants :

- besoins pour l'indexation : avoir des termes et des concepts bien définis ; être aidé par la classification des termes ; pouvoir mettre en relation des œuvres par les concepts associés aux termes ;
- besoins pour la recherche : retrouver les œuvres ; éviter le bruit et le silence ; mettre en relation les œuvres ; aider au travail.

Les autres services du musée Rodin ont des besoins d'informations spécifiques et opérationnelles : aider à la recherche ; éviter le bruit et le silence.

Le public extérieur au musée Rodin sera visé par ce projet lors de la mise en ligne de la base de gestion des collections sur le portail documentaire. Cette mise en ligne est d'ores et déjà à prendre en compte pour une cohérence du projet dans son ensemble et la visibilité à moyen et long terme.

- besoins d'informations générales et scientifiques : aider à la recherche ; éviter le bruit et le silence ;
- besoin d'une valorisation des collections auprès du public : faire connaître les collections : avoir une crédibilité scientifique.

Objectifs

La conception d'un vocabulaire contrôlé s'inscrit dans la réflexion sur une restructuration des données dans le cadre du changement de logiciel de gestion des collections et va ainsi aider à préparer au mieux la migration des données dans le nouveau logiciel, améliorer la description et l'indexation des œuvres et faciliter la recherche en utilisant les nouvelles technologies d'aide à la recherche, par exemple les facettes.

Dans ce cadre, l'objectif principal est d'arriver à mettre en relation les œuvres et les concepts qui y sont associés. Actuellement, la recherche se fait uniquement par le titre des œuvres. Or il est important de pouvoir regrouper les œuvres par concepts avec une hiérarchie.

Pour cela, l'objectif est d'établir un vocabulaire contrôlé spécifique aux concepts propres aux collections du musée Rodin, dans un souci d'aide aux utilisateurs et d'une mise en valeur des œuvres. Ce vocabulaire contrôlé sera ainsi un outil pour l'indexation et pour la recherche par la mise en relation entre les termes, les concepts et les œuvres.

Ressources

Aucun budget spécifique n'a été affecté à ce projet, dirigé par le chargé d'études documentaire des collections dans le cadre de la réinformatisation, en coopération avec le bibliothécaire et l'archiviste.

1.3.2 Planning

Le planning réalisé m'a permis de tenir les délais qui m'étaient impartis et d'avoir une vision, dès le début de ma mission, des tâches et des objectifs à réaliser. Il a été fait en fonction du cahier des charges et a pris en compte le fait que la mission se déroulait sur une période courte. Chaque tâche a donc été « compressée » afin d'être réalisées dans le bref délai imparti.

Ce planning a bien sûr évolué tout au long de la mission et a été mis à jour chaque semaine. Il comprenait :

- les étapes de réalisation de la liste de vocabulaire contrôlée ;
- des points réguliers avec le chargé d'études documentaires des collections ;
- des entretiens avec les conservateurs et les attachés de conservation du service des collections, ainsi que le bibliothécaire et l'archiviste ;
- des contacts avec des responsables de gestion documentaire sur les collections dans d'autres musées ;
- une présentation de mon travail, à l'issue de ma mission, à l'ensemble du service des collections et du service de la recherche, de la documentation, de la bibliothèque et des archives.

L'élaboration du thésaurus se faisant dans le cadre d'une restructuration des données sur le logiciel de gestion des collections, il a été nécessaire que je participe à cette réflexion et que je me penche sur la manière dont un objet, appartenant à une collection muséale, peut être décrit et indexé afin de mieux cerner l'existant et les besoins.

2 Description et indexation d'un objet de collection muséale

La description et l'indexation d'un objet de collection muséale nécessitent en amont une réflexion approfondie, tout d'abord parce que le traitement documentaire d'une œuvre présente des particularités par rapport au traitement d'un document, mais aussi parce qu'il répond à des réglementations et à des normes bien précises. Dans ce cadre, quelles sont les spécificités des collections du musée Rodin ? Quelles sont les méthodes utilisées dans d'autres musées ? Quelles préconisations peut-on faire pour une meilleure structuration des données ?

2.1 Particularité d'une œuvre par rapport à un document

La description et l'indexation d'un objet appartenant à une collection muséale poursuit le même objectif que pour tout autre document, dans le sens où ce qui est décrit est un objet avec son titre, son auteur, son sujet, la date, etc. On représente ainsi, par les termes d'un langage documentaire, l'information résultant de l'analyse de l'objet afin d'en faciliter l'identification et la recherche.

Mais, contrairement au traitement documentaire qui peut être fait par un documentaliste, cette description est réalisée par un conservateur à l'issue 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.

Les musées ne sont également pas libres de se séparer d'une partie de leurs collections et d'en laisser une partie en friche, contrairement aux bibliothèques par exemple qui ont plus de liberté sur leurs fonds. Les musées ont, en effet, une obligation juridique de traiter l'intégralité de leurs œuvres, le traitement des collections étant encadré, comme nous l'avons vu dans la première partie, 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.

La représentation d'un objet de collection muséale est ainsi effectuée, de manière très normée, dans le cadre de l'inventaire des biens d'un musée de France défini par la loi et régi par le service des musées de France au MCC.

2.2 Réglementations et normes

2.2.1 Cadre juridique

L'inventaire des biens d'un musée de France est défini et rendu obligatoire par le titre premier du décret du 2 mai 2002 pris en application de la loi n°2002-4 du 4 janvier 2002 relative aux musées de France :

> « Art. 1er. - La personne morale propriétaire des collections d'un musée de France établit et tient régulièrement à jour un inventaire des biens affectés aux collections de ce musée […]
>
> Art. 3. - Est inventorié tout bien acquis à titre gratuit ou onéreux affecté aux collections du musée de France par un acte émanant de la personne morale propriétaire du bien […] La personne morale propriétaire des collections d'un musée de France fait procéder en permanence par les professionnels mentionnés à l'article 6 de la loi du 4 janvier 2002 susvisée aux opérations nécessaires au récolement des collections dont elle est propriétaire ou dépositaire et à la mise à jour de l'inventaire et du registre des dépôts ».

L' arrêté du 25 mai 2004 fixe les normes techniques relatives à la tenue de l'inventaire, du registre des biens déposés dans un musée de France et au récolement. L'annexe 1 de cet arrêté liste les trois rubriques de l'inventaire des musées de France ;

Rubrique relative au statut juridique des biens et aux conditions de leur acquisition :

- N° d'inventaire ;
- Mode d'acquisition ;
- Nom du donateur, testateur ou vendeur ;
- Date de l'acquisition et d'affectation du musée ;
- Avis des instances scientifiques ;
- Prix d'achat – subvention publique ;
- Date d'inscription au registre d'inventaire.

Rubrique portant description des biens :

- Désignation du bien ;
- Marques et inscriptions ;
- Matières ou matériaux ;
- Techniques de réalisation, préparation, fabrication ;
- Mesures ;
- Indications particulières sur l'état du bien au moment de l'acquisition.

Rubrique complémentaire :

- Auteur, collecteur, fabricant, commanditaire.
- Époque, datation ou date de récolte.
- Fonction d'usage.
- Provenance géographique.
- Observations.
- Première date de présence attestée dans le musée si origine inconnue.
- Utilisateur illustre, premier et dernier propriétaire (pour les collections scientifiques et techniques uniquement, à la demande du ministère chargé de la recherche ;
- Ancien ou autres numéros d'inventaire ;
- Mentions à porter en cas de radiation ;
- Date de vol ou de disparition et, le cas échéant, la date à laquelle le bien a été retrouvé ;
- Dans le cas d'un ensemble complexe : existence du sous inventaire nombre de registres de sous-inventaire annexés avec, pour chaque sous-inventaire et chaque registre, le nombre de biens inscrits.

Les articles annexes 1d et4c précisent les conditions de réalisation de l'édition informatisée des registres et de dépôt à partir de l'outil informatique du musée.

2.2.2 Méthode de description des objets

Respectant ce cadre juridique, le service des musées de France, en collaboration avec les équipes de nombreux musées, propose une méthode de rédaction informatisée des notices d'objets de musées (alliant les notions d'inventaire réglementaire et de base documentaire) [3, MCC].

Cette méthode est préconisée par la Direction des musées de France pour indexer un objet, quelle que soit sa nature (archéologie, arts décoratifs, beaux-arts, ethnologie, sciences et techniques, histoire). Elle regroupe les rubriques réglementaires fixées par l'arrêté du 25 mai 2004 ainsi que de nombreux champs documentaires. L'ensemble des champs est réparti en quatre catégories :

Catégorie relative à l'identification et à la description de l'objet :

- N° inventaire
- Ancien(s) numéro(s)
- Autre(s) numéro(s)
- N° de dépôt
- Domaine (catégorie du bien)
- Dénomination

- Appellation
- Titre
- Auteur / exécutant / collecteur
- Précisions sur l'auteur / exécutant / collecteur
- École
- Anciennes attributions
- Période de création / exécution
- Millésime de création / exécution
- Époque / style / mouvement
- Matériaux et techniques
- Mesures
- Inscriptions
- Précisions sur les inscriptions
- Onomastique
- Description
- État du bien au moment de l'acquisition ou du dépôt
- État actuel
- Représentation (décor porté)
- Précisions sur la représentation (décor porté)
- Date de la représentation
- Source de la représentation

Catégorie précisant le contexte historique de l'objet :

- Genèse (stade de création) ;
- Objet(s) associé(s) ;
- Historique ;
- Lieu(x) de création / d'exécution ;
- Précisions sur le(s) lieu(x) de création, d'exécution ;
- Géographie historique ;
- Utilisation, destination ;
- Précisions sur l'utilisation, destination ;
- Lieu(x) d'utilisation, destination ;
- Précisions sur les lieu(x) d'utilisation, destination ;
- Période d'utilisation, destination ;
- Millésime d'utilisation, destination ;

- Lieu de découverte, de collecte ou de récolte ;
- Type de site de découverte, de collecte ou de récolte ;
- Méthode de collecte / récolte ;
- Date de découverte, de collecte ou de récolte ;
- Découvreur ;
- Précisions sur la découverte / collecte ;
- Numéro de site SDA.

Indiquer le statut juridique de l'objet et les conditions de son acquisition ou de son dépôt :

- Date d'inscription au registre d'inventaire ;
- Date d'inscription au registre des biens reçus en dépôt ;
- Type de propriété ;
- Mode d'acquisition ;
- Prix d'achat (en euros)Mention des concours publics ;
- Institution propriétaire ;
- Établissement affectataire ;
- Date et références de l'acte d'acquisition ;
- Date d'affectation au musée ;
- Avis des instances scientifiques compétentes en matière d'acquisition ;
- Nom du donateur, testateur ou vendeur ;
- Anciennes appartenances ;
- Dépôt / Établissement dépositaire ;
- Date et références de l'acte de dépôt (ou date de prise en charge du bien ;
- Date et références de l'acte mettant fin au dépôt (et / ou date de restitution du bien) ;
- Anciens dépôts ;
- Localisation de l'objet dans l'établissement responsable.

Donner accès à de plus amples informations sur l'objet :

- Exposition ;
- Bibliographie ;
- Commentaires ;
- Observations ;
- Crédits photographiques ;
- Rédacteur.

Le service des musées de France répertorie également les logiciels de gestion des collections qui fonctionnent selon les prescriptions et la procédure de validation des modules d'inventaire réglementaire et de récolement décennal[33].

2.2.3 Vocabulaires

Le service des musées de France propose des vocabulaires scientifiques servant à l'interrogation et à l'alimentation des champs.

Ces listes de vocabulaire, qui sont des listes d'autorité ou des thésaurus, sont enrichies et mises à jour régulièrement :

- Liste des auteurs et exécutants : 35 370 termes dont 5034 synonymes :
    - tous les intervenants : auteurs ou exécutants (fabricants), ou commanditaires ou émetteurs pour les collections scientifiques et techniques. Un intervenant peut être une personne physique, un groupe de personnes physiques, une personne morale, une population, une civilisation, etc ;
- Liste des rôles des auteurs et exécutants : 205 termes dont 12 synonymes ;
- Liste des méthodes de collecte, types de sites, lieux géographiques de découverte : 4 186 termes dont 1 535 synonymes ;
- Liste des dénominations : 5 050 termes dont 331 synonymes ;
- Liste des domaines : 155 termes dont 5 synonymes ;
- Liste des écoles : 605 termes dont 156 synonymes ;
- Liste des époques et styles : 910 termes dont 106 synonymes ;
- Liste de la genèse (stades de création) : 171 termes dont 17 synonymes ;
- Liste des inscriptions : 275 termes dont 18 synonymes ;
- Liste des lieux géographiques de création, exécution, utilisation : 4 235 termes dont 1 150 synonymes ;
- Liste du statut juridique : 75 termes dont 5 synonymes ;
- Listes des périodes : 342 termes ;
- Liste des sujets représentés : 15 039 termes dont 4 209 synonymes (à partir du thésaurus Garnier) ;

[33] Espace professionnel du portail des collections des musées de France « Joconde ». http://www.culture.gouv.fr/documentation/joconde/fr/partenaires/AIDEMUSEES/societe-info.htm [consulté le 05/12/2013]

- Liste des sources de la représentation : 1 778 termes dont 77 synonymes ;
- Liste des techniques - matériaux : 3 238 termes dont 264 synonymes ;
- Liste des utilisations - destinations : 1 338 termes dont 110 synonymes.

Afin de décrire les représentations portées sur les objets, le service des musées de France préconise le *Thésaurus iconographique*, système descriptif des représentations établi par François Garnier[34].

2.3 Spécificité des collections du musée Rodin

La particularité des collections du musée Rodin et la difficulté de les décrire et de les indexer, selon les normes et méthodes que nous venons de voir, est liée au fait qu'une grande partie des collections est du même artiste ou en lien avec cet artiste : Rodin.

Seule une petite partie des sculptures, des dessins et des peintures a été réalisée par d'autres artistes que Rodin. Leur description doit donc être beaucoup plus précise que dans un autre musée contenant des collections plus éclectiques. En effet, si on décrit une œuvre créée par *Rodin*, *La porte de l'enfer*, *groupe de figures*, par exemple, le résultat ne sera pas pertinent car il existe de très nombreuses œuvres de Rodin liées à la porte de l'enfer et composées de groupe de figures. La datation n'est pas toujours d'une grande aide car elle n'est pas toujours connue et, de plus, Rodin, continue d'intervenir sur ses œuvres, les modifie et les complète. Il y a rarement une œuvre finale chez Rodin. Ses œuvres font partie d'un processus créatif lié à des sujets qui évoluent en permanence. La technique et le matériau sont très utiles, mais insuffisants lorsqu'il faut mettre en relation des œuvres ou chercher un sujet précis.

Il est donc important de définir des concepts reliés aux sujets de Rodin, pour une meilleure appréhension de ses œuvres et de son processus mental (reprises, œuvres fragmentées, assemblage). Les œuvres de Rodin sont complexes à appréhender en raison du processus créatif propre au sculpteur.

[34] GARNIER François, Thesaurus iconographique : système descriptif des représentations, Paris, Le Léopard d'or, 1984, 240 p.

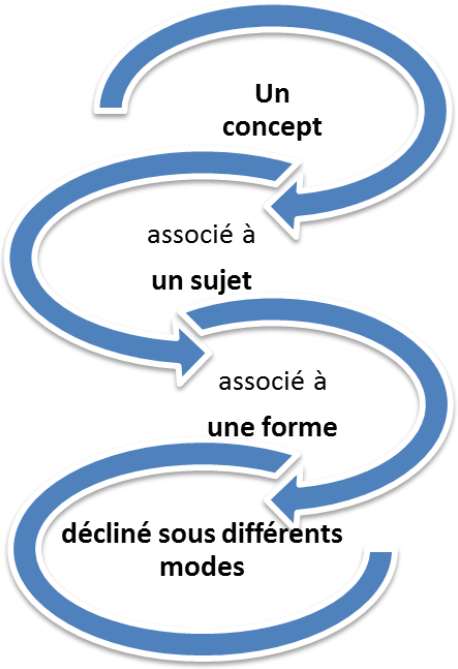


Figure 4 : Processus créatif de Rodin.

La collection personnelle de Rodin est également complexe à décrire, notamment sa collection d'antiques, dont Rodin s'inspirait pour ses œuvres. Il est donc nécessaire, à la fois de la traiter comme une collection « classique », à l'instar de la collection d'antiques conservée au musée du Louvre, mais également en relation avec les œuvres de Rodin. Ces œuvres doivent ainsi prendre tous leur sens au sein de la collection.

Mais ce trait d'union est complexe à mettre en œuvre car chaque œuvre est à la fois spécifique et reliée aux autres. D'autre part, jusqu'à présent, chaque domaine de collection a été traité séparément, que ce soit dans la documentation papier ou dans le logiciel de gestion des collections (un module par domaine de collection). Il n'existe pas actuellement de vocabulaire commun et propre à toute la collection du musée Rodin.

2.4 Ce qui se fait dans d'autres musées

Dans le cadre de ma mission, j'ai pris contact avec les services de documentation des collections d'autres musées afin de découvrir quelle est leur méthode de travail, et de pouvoir la comparer à celle du musée Rodin.

Je suis allée notamment au centre de documentation des collections du château de Versailles. Il m'a paru intéressant de voir comment une collection aussi importante (environ 100 000 objets) est traitée. De plus, comme au musée Rodin, la collection est variée et composée de dessins, de peintures, de sculptures, de gravures et de mobilier.

Je suis également allée au musée Bourdelle dont la collection est très proche de celle du musée Rodin : collection monographique (sculptures, dessins, photographies) et collection personnelle d'Antoine Bourdelle (antiques, moulages en plâtres, peintures, dessins et gravures). Bourdelle était également un ami et un collaborateur de Rodin.

La visite de ces musées a confirmé les observations que j'ai pu effectuer au musée Rodin :

- la difficulté pour les conservateurs de suivre les contraintes documentaires et de se restreindre aux vocabulaires préconisés. Ils ne sont en effet pas documentalistes et leur rôle est plus scientifique que documentaire ;
- la particularité de chaque œuvre, de chaque collection et de chaque musée, et la difficulté d'utiliser les vocabulaires préconisées par le service des musées de France qui ne permettent pas toujours de décrire les collections avec la précision voulue et de pouvoir exploiter ensuite les données ;
- la redondance des informations entre la base de gestion des collections et les bases de données gérant les fonds bibliographiques, les archives, les fonds iconographiques et multimédia.

Je suis allée également consulter la toute nouvelle base de gestion des collections du Centre des monuments nationaux (CMN). Cet établissement ne gère pas des musées mais des monuments et ne dépend donc pas du service des musées de France mais du service du patrimoine[35]. La collection du CMN (environ 100 000 objets) est donc décrite selon la méthode de description préconisée par le service du patrimoine, différente de celle du service des musées de France pour les collections muséales, et s'appuie sur le *Thésaurus des objets mobiliers*. Il est intéressant d'observer qu'un nombre important de termes se retrouve dans les différents vocabulaires préconisés à la fois par le service des musées de France et celui du patrimoine. D'autre part, même si la collection n'est pas une collection muséale et traitée en tant que telle, il m'a paru utile de voir de quelle manière les données étaient structurées et d'observer les mêmes problématiques que dans les musées.

Ces difficultés récurrentes dans les musées mettent en évidence la nécessité d'avoir des référentiels communs et de les décliner selon les spécificités qui vont permettre à leur tour d'enrichir ces référentiels communs, afin que tous les musées puissent les utiliser. Le MCC travaille actuellement sur cette problématique et sur ces enjeux avec le projet GINCO (cf. partie 3 section 3.2.1). Concernant la récurrence des termes dans différents thésaurus et vocabulaires, le MCC aborde également cette problématique à travers le *tagging* sémantique

[35] Le service des musées de France et le service du patrimoine ont été réunis dans un souci d'homogénéisation de gestion, au sein de la nouvelle Direction générale des patrimoines du Ministère de la culture et de la communication, créée par le décret du 11 novembre 2009, dans le cadre de la révision générale des politiques publiques.

et notamment le projet HDALab (cf. partie 3 section 3.2.1). Nous aborderons plus en détails ces projets dans la troisième partie.

Mais avant même de travailler sur les vocabulaires et d'utiliser des référentiels communs, il est nécessaire de bien structurer les données au préalable et ma mission au musée Rodin m'a permis de pouvoir faire quelques préconisations à ce sujet.

2.5 Préconisations pour une structuration des données

Dans un premier temps, il est nécessaire de répondre aux normes fixées par la loi sur l'inventaire des biens d'un musée de France et par la méthode préconisée par le service des musées de France.

Même si les collections du musée Rodin ont leur spécificité, le vocabulaire utilisé pour les décrire est en grande partie celui fourni par la méthode de rédaction informatisée des notices d'objets de musées. Il est donc tout à fait profitable de s'appuyer sur le vocabulaire existant, puis de l'enrichir lorsque celui-ci n'est pas suffisant pour décrire un objet.

Comme nous l'avons vu précédemment, la recherche sur le logiciel des collections du musée Rodin se fait actuellement en grande partie par le titre des œuvres. Ces titres fournissent, en effet, beaucoup d'informations qui, si elles étaient reprises et structurées, pourraient être très utiles pour la recherche. J'ai ainsi listé les types de données que l'on peut trouver dans les titres et qui auraient leur place dans des champs d'indexation préconisés par la méthode de rédaction informatisée des notices d'objets de musées :

- Dénomination
- Appellation
- Matériau
- Technique
- Genèse

Ces champs existent mais ils sont peu ou mal utilisés, en raison d'un passif que nous avons déjà décrit. Il en résulte des listes de vocabulaire qui ne sont pas adaptées et un travail d'indexation et de correction important à réaliser. Une réflexion sur le vocabulaire a déjà été initiée par le chargé d'études documentaire des collections, avec un vaste chantier d'indexation et de réflexion, sur les dénominations, les appellations, les matériaux et les techniques.

Les données concernant la genèse de l'œuvre auraient également toute leur place dans un champ indexé. Cela améliorerait de manière significative la recherche. Nous avons ainsi listé et hiérarchisé, avec le chargé d'études documentaire des collections, les termes utiles pour

les collections du musée (25 termes). Cette liste de termes est largement inspirée de la liste de vocabulaire préconisée par le service des musées de France :

- Œuvres en rapport
- Stade de la création
    - Agrandissement
    - Réduction
    - Réplique
    - Petit modèle
    - Moyen modèle
    - Grand modèle
- Œuvre préparatoire
    - Ébauche
    - Esquisse
    - Étude
    - Maquette
    - Modèle original
    - Projet
    - Découpage
    - Fragment
    - Modèle de fonderie
    - Modèle de mise au point
    - Moulage
    - Sur moulage
- Épreuve
    - Épreuve (sculpture)
    - Épreuve d'état (gravure)
- Marcottage

Nous avons aussi mis en avant l'utilité de créer un index des titres principaux des œuvres de Rodin et de relier chaque œuvre à un ou plusieurs de ces titres. Nous pouvons prendre en exemple le titre de *la Porte de l'Enfer*. C'est un titre qui est repris pour un nombre important d'œuvres de Rodin et qui est relié à un certain nombre d'autres œuvres. Pouvoir indexer une œuvre avec ce titre « générique » tout en gardant le titre original de l'œuvre est une vraie valeur ajoutée pour la recherche.

Cette structuration des données constitue un premier pas vers une recherche plus efficace, mais elle ne suffit pas pour mettre en relation les œuvres, les sujets et les concepts, d'où la nécessité de compléter la description des œuvres avec l'aide d'une liste de vocabulaire contrôlé, sous forme de thésaurus, propre aux collections du musée.

Ce thésaurus s'appuie sur le thésaurus préconisé par le service des musées de France mais celui-ci pourtant très complet reste très général et ne permet pas de retrouver toutes les sensibilités particulières de l'œuvre de Rodin et de quoi décrire distinctement la collection du musée.

3 Élaboration d'un thésaurus

Nous avons, dans la première partie, expliqué ce qu'est un thésaurus, comment il est constitué et à quoi il sert (cf. partie 1 section 2.4). Nous avons ensuite, dans le cahier des charges, établi les besoins et les objectifs (cf. section 1.3.1). Nous allons maintenant montrer quelle a été la méthodologie utilisée pour construire le thésaurus, puis nous ferons un point sur le complément et la mise à jour à mettre en œuvre. Nous terminerons en mettant en lumière les difficultés rencontrées et en proposant quelques préconisations

3.1 Méthode d'élaboration d'un thésaurus monolingue

Afin d'élaborer un thésaurus selon la bonne méthode, il existe un certain nombre d'étapes qui varient en fonction du type de thésaurus (monolingue, bilingue ou multilingue), à suivre dans un ordre précis par souci de cohérence et de rigueur.

3.1.1 Collecte du vocabulaire ou lexique

Je me suis tout d'abord munie des documents déjà existants pouvant m'aider dans la réalisation du thésaurus :

- les listes de vocabulaire existant dans le domaine d'expertise : les vocabulaires scientifiques préconisés par le service des musées de France et le *Thésaurus iconographique Garnier* ;
- la norme internationale ISO 25964 « *Thésaurus et interopérabilité avec d'autres vocabulaires* » (cf. partie 1 section 2.4.4).

Afin de choisir une première liste de termes et de concepts à traiter, j'ai travaillé à partir d'une sélection d'œuvres issue d'une liste d'œuvres reliées à l'exposition « Rodin, les origines d'un génie », présentée par le musée Rodin en 2010 au Palazzo Leone da Perego de Legnano en Italie, et constituée de dessins et de sculptures (une centaine d'œuvres). Je me suis attachée, pour chaque œuvre, à définir les termes et les concepts qui pourraient être utiles pour leur indexation et leur recherche. Ce travail m'a à la fois permis d'élaborer une première liste de vocabulaire et de concepts, mais aussi de réfléchir à une meilleure structuration des données tel que nous l'avons vu précédemment (cf. tableau en annexe 1).

Puis, afin d'enrichir cette première liste de vocabulaire et de concepts, je me suis appuyée sur :

- des données contenues dans le logiciel de gestion des collections. Comme nous l'avons vu, ce logiciel est obsolète et inadapté et contient des données qui ne sont pas bien

structurées, mais il possède néanmoins des informations utiles et nécessaires à ce travail, notamment des informations d'identité sur les œuvres (n° d'inventaire, domaine, dénomination, titre, auteur, matériau, technique) ;

- des données issues de la documentation papier, particulièrement l'index thématique des dessins ;
- des données contenues dans des ouvrages de référence sur les dessins et les sculptures de Rodin : *l'Inventaire des dessins de Rodin*[36] et *Rodin et le bronze, Catalogue des œuvres conservées au musée Rodin*[37].

J'ai ensuite tenté de définir le périmètre exact du thésaurus, les besoins spécifiques des utilisateurs et le degré de précisions. Pour cela, j'ai régulièrement fait des points avec mon maître de stage et rencontré les conservateurs, et les attachés de conservation, afin de valider la pertinence des termes et des concepts choisis et connaître de manière plus précise leurs besoins d'indexation et de recherche.

La difficulté a été pour moi de comprendre à la fois les besoins exacts des utilisateurs mais aussi de tenter d'avoir une connaissance même partielle sur les œuvres et sur la collection en général, afin de savoir comment les décrire et de quelle manière pour que la description soit pertinente. Il s'en est suivi des échanges nombreux avec les personnes ayant une bonne connaissance des œuvres et pour moi des remises en question permanentes jusqu'au dernier jour de ma mission.

3.1.2 Listage des catégories et plan de classement

Une fois les termes et les concepts choisis (environ 600 termes), je les ai rassemblés en micro-disciplines regroupées autour de cinq catégories, elles-mêmes regroupées en deux groupes :

- Description du processus de création :
    - Thématiques
    - Phases de conception
- Description iconographique :
    - Description par mots clés
    - Personnages représentés
    - Lieux représentés

[36] JUDRIN Claudie. Inventaire des dessins. Paris, Éditions du musée Rodin, tome I à V, 1984-1992.

Le choix de ces catégories s'appuie à la fois sur le *Thésaurus iconographique Garnier* pour la description iconographique, et sur les particularités des collections du musée Rodin pour la description du processus de création.

3.1.3 Établissement des règles d'écriture et relation entre les termes

Une fois le plan de classement effectué, il est nécessaire d'établir des règles d'écriture et de relation entre les termes, en créant, comme nous l'avons vu dans la première partie (cf. partie 1 section 2.4.2), des relations d'équivalence, des relations hiérarchiques et des relations associatives.

- Relations d'équivalence sémantique

Il a fallu attribuer à chaque concept, une liste de synonymes et de quasi-synonymes. J'ai sélectionné les termes préférentiels et les termes non préférentiels et je les ai liés par un rapport d'équivalence. Cette étape est la plus délicate car le choix d'un terme préférentiel pose certaines difficultés. Il doit être le plus « neutre » possible, donc le moins ambigu.

- Relation hiérarchique

Le choix du nombre de niveaux hiérarchiques pose souvent problème, et dans le cas de ce thésaurus, a été délimité à trois niveaux maximum, car le nombre de termes n'est pas très élevé et les besoins ne vont pas au-delà.

- Relation associative

Les relations associatives sont nombreuses entre les termes des catégories « description par mots-clés » et « personnages représentés » et la catégorie « thématique ».

3.1.4 Test et corrections

Le test sert à vérifier la cohérence du thésaurus et la pertinence des termes choisis. Pour cela j'ai partiellement (par manque de temps) indexé un corpus d'œuvres, issu d'une liste d'œuvres reliées à l'exposition *Rodin, la lumière de l'antique*, présentée à Arles pendant la durée de mon stage et au musée Rodin à Paris de novembre 2013 à janvier 2014. Ce test m'a permis de consigner les termes préférentiels les plus souvent utilisés, de rajouter des termes préférentiels et des termes non préférentiels et de modifier certaines relations d'équivalence, hiérarchiques et associatives.

---

[37] LE NORMAND-ROMAIN Antoinette. Rodin et le bronze. Catalogue des œuvres conservées au musée Rodin. Éditions du musée Rodin / Éditions de la Réunion des musées nationaux, Paris, 2007, 2 vol.

De plus, ce corpus d'œuvres comportait tous les domaines de collections du musée Rodin, ce qui a permis d'entamer une réflexion sur la manière dont ce thésaurus peut être complété par la suite pour être utilisable de manière cohérente et pertinente sur l'ensemble de la collection.

3.2 Complément et mise à jour

Le principal objectif du complément et de la mise à jour de ce thésaurus est donc de l'étendre dans un premier temps à l'ensemble de la collection. Il n'est pour le moment utilisable que pour les peintures et les sculptures, et même dans ce cadre, il nécessite d'être largement enrichi.

Un thésaurus est un outil qui n'est jamais véritablement achevé. Il nécessite un enrichissement permanent, en particulier s'il est très spécifique à une discipline ou à un artiste comme au musée Rodin. Généralement, on observe qu'un thésaurus devient dysfonctionnel parce que d'une part le choix et l'intégration des termes sont effectués avec moins de rigueur et que d'autre part les velléités pour le mettre à jour s'affaiblissent au fil du temps.

3.3 Difficultés et préconisations

Une des principales difficultés a été le choix des termes, puis des catégories et des groupes. Pour cela il m'a fallu comprendre précisément la particularité des collections du musée Rodin et les besoins exacts des utilisateurs. D'autre part, il était nécessaire de combiner l'utilisation des vocabulaires déjà existants dans le domaine, comme le *Thésaurus iconographique Garnier*, puis les adapter et les enrichir avec des termes spécifiques et utiles pour la collection du musée Rodin, tout en respectant les normes en vigueur.

Une autre difficulté était que le corpus d'œuvres sur lequel portait l'élaboration du thésaurus était limité aux dessins et aux sculptures, mais il me fallait toutefois prendre en compte que ce thésaurus devait rapidement s'étendre à l'ensemble des collections et des ressources du musée Rodin. J'ai donc choisi des termes utiles à la description des dessins et des sculptures tout en présentant un thésaurus qui puisse facilement être étendu dans la suite du projet avec la création d'un vocabulaire commun et l'harmonisation des données. Lors de la présentation de mon travail au service des collections et au service de la recherche, les conclusions ont porté sur le fait qu'il apportait un véritable éclairage et des éléments concrets permettant de mettre en valeur les collections du musée Rodin. Mais les discussions ont également porté sur la difficulté de prolonger ce travail tout en mettant en avant sa nécessité. Mon travail n'était pas de rendre un projet abouti mais a surtout consisté à créer une première ébauche, à poser les jalons et le cadre pour la suite et a été une valeur ajoutée pour la réflexion menée au sein du musée Rodin.

Enfin, alors que l'objectif de ma mission était l'élaboration d'un vocabulaire contrôlé sous forme de thésaurus afin de faciliter l'indexation et la recherche, je me suis souvent heurtée à une méconnaissance de l'utilité des langages documentaires et de la structuration des données. Jusqu'à l'arrivée du chargé d'études documentaires des collections, aucun langage documentaire n'était utilisé et tous les champs de la base de gestion des collections étaient des champs libres et non indexés. Un certain nombre de champs sont maintenant indexés mais le logiciel de gestion des collections étant obsolète et mal utilisable, les utilisateurs n'ont pas conscience de la manière dont peut être valorisée l'utilisation des langages documentaires, notamment pour la recherche. Il est également vrai qu'au début de la période informatique, on voyait clairement comment était effectué le travail, mais aujourd'hui beaucoup de choses sont invisibles : la recherche se fait sur des liens et des champs qui peuvent être cachés ; le résultat est affiché mais pas forcément les étapes qui ont permis d'y arriver ; l'information peut être concaténée (mise côte à côte), donnant l'impression de l'existence d'un seul champ alors qu'il y en a plusieurs. Pour des personnes qui ne sont pas formées à la structuration des bases de données, il n'est pas toujours facile d'en comprendre le fonctionnement.

J'ai donc pris le temps, tout en m'attachant à comprendre les besoins des utilisateurs et les spécificités des collections du musée Rodin, d'expliquer ce qu'une meilleure structuration des données et l'utilisation d'un vocabulaire contrôlé pouvait apporter, et montrer les profits que l'on peut retirer d'un logiciel plus performant et bien paramétré.

J'ai pu également tenter de montrer l'utilité d'un langage documentaire tel que le thésaurus, qui permet justement de décrire correctement les données de façon homogène. Ce langage documentaire est, en effet, un outil sémantique performant qui permet de créer des liens et de donner du sens.

Or, justement, le web de données, dans lequel les données liées aux collections muséales doivent être intégrées, nécessite des données très structurées et porteuses de sens. On peut se demander si ce n'est pas là un des enjeux de la valorisation des collections muséales. Nous allons essayer, dans la troisième partie, de définir plus précisément ces enjeux. Quelles sont les politiques culturelles et comment les musées doivent y répondre ? Comment les musées peuvent-ils ouvrir leurs données, garantir leurs échanges et leur interopérabilité dans le cadre du web de données ?

Troisième partie :
Quels enjeux pour la valorisation des collections muséales

Dans cette dernière partie, et afin de répondre à ces questions, nous replaçons les collections muséales dans le contexte plus large des données culturelles et des grandes missions du MCC : garder trace de la connaissance et la diffuser. De ces deux grandes missions vont découler naturellement les enjeux de la valorisation des collections muséales.

En premier lieu, et afin de répondre à ces missions, le MCC impulse trois grandes politiques culturelles auxquelles les musées doivent répondre : la collecte des données avec le premier récolement décennal des collections, la numérisation des ressources avec la numérisation du patrimoine culturel et l'ouverture des données avec l'*open data*.

En répondant à ces politiques culturelles, les musées peuvent favoriser ainsi le partage et la diffusion des connaissances qui doivent se faire là où sont les utilisateurs. Or les utilisateurs, à l'heure actuelle, sont sur le web et ne sont pas seulement français mais du monde entier. En utilisant tous les canaux de diffusion de leurs données et de leurs ressources sur le web, et en mettant à contribution les utilisateurs, les musées peuvent ainsi acquérir une visibilité nationale et internationale d'envergure, qu'ils amélioreront encore en étant capables de comprendre les enjeux et les fonctionnalités du web de données.

En effet, il ne suffit pas de mettre ces données sur le web, il faut les rendre exploitables et ainsi mieux visibles. Et pour cela, et afin de tirer le meilleur profit du web de données, les musées doivent pouvoir garantir l'interopérabilité de leurs données. Cela nécessite des données très structurées qui reposent sur l'utilisation des langages documentaires comme les thésaurus et les vocabulaires contrôlés utilisés selon les principes du web sémantique. Les données vont pouvoir être porteuses de sens, favoriser le multilinguisme et ainsi permettre à l'utilisateur d'accéder à une recherche facilitée et unifiée, non seulement sur les collections muséales mais également sur l'ensemble des données culturelles.

1 Répondre aux politiques culturelles

Comme nous venons de le voir, les deux grandes missions du MCC sont la conservation et le partage des connaissances. Ces deux grandes missions sont actuellement déclinées avec des politiques de grande envergure auxquelles les musées sont tenus de répondre : le premier récolement décennal des collections qui s'achève en 2014, et qui a permis aux musées de collecter et d'informatiser des informations primordiales sur les œuvres ; la numérisation du patrimoine culturel qui concerne également l'ensemble des ressources et des données culturelles et qui incite à la mise en relation et en commun des données ; le mouvement de l'*open data*, dont le principe est la publication, sur le web, des données publiques librement accessibles et exploitables, et dans lequel les données culturelles, et notamment les collections muséales, ont toute leur place.

1.1 Premier récolement décennal des collections

Le premier récolement décennal des collections est pour les musées une obligation juridique qui a été un formidable levier pour l'informatisation des données sur les collections muséales.

1.1.1 Cadre juridique

Le récolement est une obligation prévue par le Code du patrimoine qui consiste à vérifier la présence des œuvres à partir des inventaires. Cette obligation a été inscrite récemment dans la loi 2002 sur les musées de France et codifiée à l'article L.451-2 du Code du patrimoine qui précise que « les collections des musées de France font l'objet d'une inscription sur un inventaire. Il est procédé à leur récolement tous les dix ans. »

Le premier des récolements décennaux obligatoires s'achèvera au plus tard en juin 2014[38] et devra être entériné par la commission de récolement des dépôts d'œuvres d'art (CRDOA) du MCC. Le récolement consiste donc à vérifier, sur pièce et sur place, à partir d'un bien ou de son numéro d'inventaire[39] :

- la présence du bien dans les collections ;
- sa localisation ;

[38] Circulaire n° 2006/006 du 27 juillet 2006 relative aux opérations de récolement des collections du musée de France.
http://www.culture.gouv.fr/documentation/joconde/fr/partenaires/AIDEMUSEES/circulaire_recolement2006-006.pdf [consulté le 05/12/2013]

- l'état du bien ;
- son marquage ;
- la conformité de l'inscription à l'inventaire avec le bien ainsi que, le cas échéant, avec les différentes sources documentaires, archives, dossiers d'œuvres, catalogues.

La loi prévoit que chaque campagne de récolement doit faire l'objet d'un procès-verbal rédigé par le professionnel responsable des collections au sens de l'article L442-8 du code du patrimoine, et peut également bénéficier d'un traitement informatisé. Dans ce sens, le service des musées de France du MCC incite fortement les musées à informatiser les données relatives au récolement et explique l'intérêt d'un support informatisé par rapport au papier : « *l'intérêt d'une solution informatisée est la possibilité d'accès, de traitements multiples et de centralisation de l'information, ceci à tout moment. L'outil informatique constitue en cela une aide considérable pour faire aboutir la procédure du récolement* ».

Le service des musées de France du MCC répertorie également les logiciels de gestion des collections qui fonctionnent selon les prescriptions et la procédure de validation des modules de récolement décennal (cf. partie 2 section 2.2.2)[40].

1.1.2 Traitement des données

Ce premier récolement décennal obligatoire, qui s'achève très bientôt, a donc été une formidable opportunité pour les musées d'informatiser ou de réinformatiser les données sur leurs collections, voire de numériser les œuvres et les ressources, d'utiliser, selon les normes des musées de France, les langages documentaires et les vocabulaires préconisés, et ainsi de mettre en valeur leurs collections.

Les musées ont dû mettre en place :

- soit une informatisation complète (aucune informatisation au départ) ;
- soit une informatisation à partir de plusieurs sources informatisées documentaires ;
- soit la restructuration d'une base de données documentaire existante.

Par exemple, le musée Rodin, qui avait déjà une base de données documentaire lors de la mise en place de l'organisation du récolement, a été dans l'obligation de refondre la structure d'une partie des données de la base.

---

[39] Arrêté du 25 mai 2004 fixant les normes techniques relatives à la tenue de l'inventaire, du registre des biens déposés dans un musée de France et au récolement.
[40] Espace professionnel du Portail des collections des musées de France « Joconde ». http://www.culture.gouv.fr/documentation/joconde/fr/partenaires/AIDEMUSEES/societe-info.htm [consulté le 05/12/2013]

Ce récolement nécessite pour les musées un grand investissement humain, financier et leur prend beaucoup de temps du fait d'un gros travail de recherche et d'analyse devant être fait par des conservateurs, des attachés de conservation, ou des personnes compétentes dans le domaine d'expertise des collections. Et lorsque les collections sont importantes comme au château de Versailles, au musée Rodin ou au musée Bourdelle, le travail fourni est d'autant plus important.

Mais le récolement a permis de récolter un grand nombre de données, de retrouver parfois des œuvres perdues, de compléter l'inventaire et les données sur les collections, d'analyser, de mettre à jour et d'enrichir les informations, d'améliorer la gestion des collections, etc.

En revanche, j'ai pu observer, en allant voir les bases de gestion des collections dans différents musées, que le temps imparti pour terminer le récolement étant court, il a été jugé plus nécessaire de rentrer le maximum d'informations sans qu'il y ait toujours de véritable réflexion sur la structuration des données et sur la manière dont elles peuvent être exploitées. Il y a eu généralement une réflexion au départ, puis dans l'urgence et avec la masse de travail à réaliser, l'importance de rentrer les données a pris le dessus sur la structuration. Il est vrai, comme nous l'avons vu dans la deuxième partie (cf. partie 2 section 2.1), que les métiers documentaires ne sont pas du ressort des conservateurs ou des attachés de conservation, à qui il est davantage demandé une analyse scientifique sur les œuvres qu'un travail documentaire.

Dans la politique culturelle, le récolement est une demande forte et les musées ont à cœur d'y répondre. Mais actuellement, le récolement arrivant à son terme, et beaucoup de musées se trouvant en possession d'un nombre de données très importantes dans leur base documentaire, il est nécessaire que les musées puissent, en s'appuyant sur les langages documentaires, les traiter et les exploiter afin de mettre en valeur les collections, améliorer la recherche, l'analyse et la connaissance des œuvres, les rendre accessibles au plus grand nombre, en diffusant cette connaissance et en la structurant sur le web de données.

Cette problématique est identique et même démultipliée dans le cadre de la numérisation du patrimoine culturel.

1.2 Numérisation du patrimoine culturel

Alors que le récolement décennal obligatoire ne concerne que les collections publiques, la numérisation du patrimoine culturel va bien au-delà et concerne toutes les données culturelles du patrimoine (archives, bibliothèques, musées, services patrimoniaux, audio-visuel, etc).

La Commission européenne a adopté une recommandation invitant les États membres à intensifier leurs efforts, à mettre en commun leurs ressources et à faire appel au secteur

privé pour mener à bien la numérisation du matériel culturel. Ceci est indispensable pour rendre le patrimoine culturel européen accessible à un plus grand nombre et stimuler la croissance des industries de la création en Europe. Le matériel numérisé devrait être accessible via Europeana[41] (cf. section 2.1.3) point d'accès multilingue à tous les contenus culturels du patrimoine et de la création contemporaine répartis en Europe. Nous parlerons plus longuement de ce projet dans la section suivante.

Dans sa recommandation, la Commission engage les États membres à élaborer des programmes fiables et à nouer des partenariats en vue de rassembler sur Europeana 30 millions d'objets d'ici 2015 (contre 19 millions à l'heure actuelle).

En France, la numérisation du patrimoine culturel est donc l'une des priorités définies par le MCC afin de valoriser les collections et les fonds documentaires pour les diffuser plus largement sur le web.

Cette volonté de rendre accessibles à tous les données culturelles rentre dans le cadre de l'ouverture des données publiques, l'*open data*.

1.3 *Open data*

Nous allons dans un premier temps expliquer ce qu'est l'*open data*, puis dans un second temps montrer les préconisations et l'enjeu de cette politique pour le domaine culturel dans lequel les musées ont toute leur place.

1.3.1 Définition de l'*Open data*

L'o*pen data* (en français, ouverture des données) est une pratique de publication de données librement accessibles et exploitables. Une donnée ouverte est :

- une donnée numérique d'origine publique : produite par une collectivité ou un service public ;
- une donnée structurée selon les besoins et les principes du web de données ;
- une donnée diffusée selon une licence ouverte ;
- et ainsi, une donnée librement accessible et réutilisable par tous, sans restriction technique, juridique et financière.

En France, l'ouverture des données publiques est fondée sur la loi n° 78-753 du 17 juillet 1978 relative au droit d'accès aux documents administratifs : *« (…) quels que soient leur*

[41] Europeana : bibliothèque numérique européenne.
http://www.europeana.eu/ [consulté le 05/12/2013]

date, leur lieu de conservation, leur forme et leur support, les documents produits ou reçus, dans le cadre de leur mission de service public, par l'État, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission. (…) ». (article 1er).

Le droit d'accès aux documents administratifs a été reconnu comme une « liberté publique » par le Conseil d'État[42]. En 1997, le Gouvernement en a élargi le principe en décidant la mise en ligne gratuite des *« données publiques essentielles »*. En 2011, le Gouvernement pose le principe de la gratuité du droit à réutilisation des informations publiques détenues par l'État et ses établissements publics administratifs[43].

Le Gouvernement créé en 2011 la mission « Etalab » chargée d'assurer le pilotage de la politique d'ouverture en ligne des données publiques, sous l'autorité du Premier ministre[44]. Elle coordonne l'action des services de l'État et de ses établissements publics afin de faciliter la réutilisation la plus large possible de leurs informations publiques, notamment par la création d'un portail unique d'hébergement et de partage de données publiques ouvertes, gratuites et réutilisables : www.data.gouv.fr.

Ces données ouvertes, présentées sous forme de jeux de données, sont par exemple, si on se réfère aux données les plus recherchées sur le site data.gouv.fr au 4 décembre 2013 : les résultats des élections, les effectifs de la fonction publique, la sécurité et la criminalité, les indicateurs de résultats des lycées, l'emploi et le chômage, la fiscalité locale, les établissements de santé, les émissions industrielles de polluants , les aides de la politique agricole commune, la géolocalisation, le catalogue de la BnF en format sémantique.

Aujourd'hui, la majorité des données utilisent soit l'*Open Database License* (ODbL), développée en Grande Bretagne par l'*Open Knowledge Foundation*, puis traduite en français avec le concours de la mairie de Paris, soit la « Licence Ouverte/*Open Licence* », développée en France par « Etalab ».

Un projet comme Europeana que nous avons rapidement évoqué dans la section précédente et que nous verrons de manière plus approfondie dans la section suivante (cf. section 2.1.3) montre toute l'importance des données culturelles au sein de l'*open data* et l'on peut se demander quelles sont les préconisations et l'enjeu de cette politique pour le domaine culturel et les musées.

---

[42] CE, 29 avril 2002, U., n°228830.
[43] Décret n°2011-577 du 26 mai 2011.
[44] Site internet de la mission « Etalab ».
http://www.etalab.gouv.fr/ [consulté le 05/12/2013]

1.3.2 *Open data* culturel

Nous allons donc essayer de dresser rapidement les contours des données culturelles, de montrer comment le MCC souhaite standardiser ces données, soumises à un statut juridique spécifique.

Périmètre d'ouverture des données culturelles

Le domaine culturel offre de nombreuses données qui peuvent être utilisables. Cela concerne, non seulement les données conservées par les musées, mais aussi par toutes les institutions culturelles comme les archives, les bibliothèques et tous les lieux de culture. Ces données sont très diverses : images ou textes numérisés, contenus audio-visuels, métadonnées associées à des œuvres patrimoniales, statistiques de fréquentation de lieux culturels, données de programmation d'un cinéma ou d'une manifestation artistique, registres généalogiques, etc. On observe donc que même au sein des musées, les données ne concernent pas seulement les collections mais aussi toute la vie des musées.

Pour donner quelques exemples d'ouverture des données, la BnF a regroupé sur un même portail, data.bnf.fr, toutes les informations issues de ses différents catalogues, ainsi que de sa bibliothèque numérique Gallica[45] (cf. section 3.1.2) ; le Centre National de la Cinématographie et de l'image animée (CNC) met à disposition des données de fréquentation de salles, des statistiques d'audience, des statistiques d'œuvres aidées, etc[46].

Préconisation du Ministère de la culture et de la communication

Le groupe de travail sur le patrimoine culturel numérisé a rendu en 2009 un rapport « Partager notre patrimoine culturel – Propositions pour une charte de la diffusion et de la réutilisation des données publiques culturelles numériques ». Ce rapport préconise une standardisation et une centralisation des données qui sont très nombreuses et hétérogènes dans le monde culturel.

Ce rapport met également en avant « le cercle vertueux de la réutilisation des données » à travers les quatre étapes fondamentales de l'*open data* culturel :

- *« des données bien exposées et référencées donnent envie aux tiers de les réutiliser ;*
- *les réutilisations de toute nature améliorent leur remontée par les moteurs de recherche et accroissent la notoriété de la "marque" culture.fr et de celle de l'institution culturelle productrice ;*

[45] Site internet de data.bnf.fr.
http://data.bnf.fr/ [consulté le 05/12/2013]

- *cela stimule la fréquentation en ligne de culture.fr et des sites des établissements, ainsi que la fréquentation physique des lieux, des œuvres et des documents publics ;*
- *cette fréquentation accroît le désir et les opportunités de réutilisation des œuvres numérisées pour des usages innovants et créatifs, individuels comme commerciaux. »*

Spécificité juridique des données culturelles

Il faut bien distinguer les données (dans le secteur culturel : les œuvres elles-mêmes, numériques ou numérisées, et mises à la disposition de réutilisateurs) dont la réutilisation est liée à un statut juridique spécifique, et les métadonnées dont la diffusion et l'appropriation par des acteurs publics et privés ne rencontrent habituellement pas d'obstacle [15, Net : Lab].

Les données culturelles ont un statut spécifique car elles sont soumises à un régime dérogatoire. La loi n° 78-753 du 17 juillet 1978 relative au droit d'accès aux documents administratifs précise que « *ne sont pas considérées comme informations publiques (…), les informations contenues dans les documents sur lesquels des tiers détiennent des droits de propriété intellectuelle* ». Cette loi permet ainsi aux établissements culturels de déterminer les conditions de réutilisation des données.

Enjeux de l'*open data* culturel

Ce statut juridique freine donc le développement de l'*open data* culturel, à l'inverse du mouvement général de l'*open data* et des enjeux de la libre réutilisation des données culturelles. Un certain nombre d'institutions culturelles, surtout les bibliothèques, ont malgré tout mis en place ces dernières années une politique d'ouverture des données avec des solutions techniques et juridiques, en accomplissant un travail de numérisation, d'indexation, de documentation et de mise en ligne.

D'autre part, il est important de souligner que la politique d'ouverture des données peut se faire hors du cadre de l'*open data*, qui ne peut être une solution unique, mais également dans le cadre de la réflexion sur les stratégies de diffusion, de mise à disposition et de valorisation des données culturelles par la mise en place des services à valeur ajoutée économique. Cette réflexion est menée depuis 2007 par l'Agence du patrimoine immatériel de l'État et concerne les données juridiques, économiques, géographiques, sociales mais aussi culturelles.

---

[46] Site internet du Centre national du cinéma et de l'image animée. http://www.cnc.fr/web/fr/statistiques [consulté le 05/12/2013]

Les défis de l'ouverture des données pourront donc être relevés avec :

- un éclairage sur le statut juridique des données culturelles ;
- une réflexion sur l'impact économique de la gestion du patrimoine immatériel culturel sur l'économie ;
- le développement des projets initiés par le MCC ;
- la mise en place de projets et d'actions à mener par les institutions culturelles et notamment les musées.

Nous ne traiterons pas plus en détail les deux premiers points qui ne concernent pas directement notre sujet, mais nous montrerons justement, dans notre dernière section sur l'interopérabilité des données, quels projets sont actuellement menés par le MCC permettant une meilleure accessibilité de la culture sur le web grâce à une meilleure structuration des données. Nous allons montrer maintenant quelles actions peuvent mener les musées afin de partager et diffuser les connaissances en utilisant tous les canaux de diffusion sur le web.

2 Diffuser les données sur le web et comprendre les enjeux du web de données

En répondant aux politiques culturelles impulsées par le MCC et dont la réalisation leur est confiée, les musées peuvent favoriser ainsi le partage et la diffusion des connaissances sur le web. Nous allons présenter dans cette partie toutes les actions que peuvent mettre en place les musées, tous les projets auxquels ils peuvent participer. Les modes de mise en ligne des données sur le web et les points d'accès pour les utilisateurs sont multiples : propres à chaque musée, fédérateurs au niveau national et au niveau international. D'autre part, la diffusion des données sur le web favorise un vrai partage des connaissances dans le sens où, non seulement les données sont reliées et ainsi porteuses de sens mais aussi où chacun de nous peut devenir contributeur et enrichir ainsi la connaissance. C'est une petite « révolution » dans le monde fermé des collections muséales et les musées ont tout intérêt à profiter de cette opportunité qui leur est donnée. Enfin nous allons voir que la multiplicité des points d'accès et des données sur le web ne favorise pas toujours une recherche efficace de l'information pour l'utilisateur, et que pour retirer tout le profit du web, rendre visibles leurs données et leurs ressources, et valoriser ainsi leur collection, les musées doivent être en mesure de comprendre les enjeux et les fonctionnalités du web de données.

2.1 Mise en ligne des données

Nous allons présenter ici quelques exemples de projets, soit initiés par les musées eux-mêmes (musée du quai Branly et Centre Pompidou), soit auxquels les musées sont associés au niveau national (Portail national des collections des musées de France Joconde), et également au niveau international (projet Europeana et Google Art Project).

2.1.1 Sur les sites web des musées

Le premier prolongement dans le cadre du web, pour les musées, de la diffusion traditionnelle de l'information sur leurs collections, est de mettre en ligne leurs données sur leur propre plate-forme de diffusion de contenus numériques accessibles via leur site internet.

Comme nous l'avons vu dans la première partie (cf. partie 1 section 1.4), les musées ont privilégié ces dernières années le choix d'outils dédiés à chaque type de ressources. Les données sont ainsi mieux structurées et répondent aux normes spécifiques de chaque métier.

Afin de donner accès à l'ensemble des données à travers un point d'entrée unique et de créer une recherche unifiée entre ces données traitées ainsi en silos[47], certains musées se sont dotés d'interface comme les OPACs (*online public access catalogues*[48]), et ont mis en place une recherche unique sur plusieurs catalogues. C'est le cas du musée du quai Branly qui propose sur son site internet une recherche fédérée sur ses différentes bases de données (objets, images, médias, archives et documentation). L'utilisateur peut ainsi sélectionner plusieurs catalogues afin d'effectuer une recherche. La recherche valorise les éléments communs aux catalogues : titres et appellations, auteurs, artistes et personnes liés à une collection, sujets thématiques, etc. Les résultats sont affichés catalogue par catalogue[49].

Mais dans cet exemple, les résultats renvoient vers les différentes bases de données sans fédération des données, sans harmonisation et sans relation entre les données obligeant l'utilisateur à naviguer entre les bases de données et à revenir à sa recherche initiale. Ce n'est pas une véritable contrainte mais les technologie du web sémantique permettent une véritable convergence des données facilitant la recherche et présentant une véritable recherche unifiée pour l'utilisateur.

Souhaitant ainsi fédérer de manière plus forte ses données, le Centre Pompidou s'appuie, pour le *Centre Pompidou virtuel*, créé en 2007, et redéveloppé en 2012, sur le principe du web de données qui permet de relier entre elles les ressources dont le format est hétérogène, en utilisant des identifiants web, les URI [18, BERMÈS]. Le but de cette technologie est de fournir à l'utilisateur final un espace global d'informations dans lequel il peut naviguer intuitivement. Le *Centre Pompidou virtuel* donne ainsi accès à l'ensemble de sa production numérique et à certaines ressources de ses établissements associés (BPI, Ircam et Centre Pompidou Metz) : œuvres d'art numérisées, documents sur l'art et l'histoire de l'art, enregistrements sonores, vidéos, archives, notices bibliographiques, etc. Plusieurs bases de données existantes sont ainsi agrégées : la base de gestion des collections du musée, l'agenda de la programmation culturelle, les catalogues des bibliothèques, les inventaires d'archives, les produits de la boutique, mais aussi des bases de données contenant des ressources audio-visuelles, des articles, des biographies, des dossiers pédagogiques, etc.

---

[47] En silos : données traitées dans des bases de données distinctes avec un cloisonnement de l'information empêchant l'échange d'information et impliquant des redondances d'information.
[48] OPAC : interface publique d'une base de données, par opposition à l'interface dédiée à l'administrateur.
[49] Site internet du musée du Quai Branly.
http://www.quaibranly.fr/ [consulté le 15/12/2013]

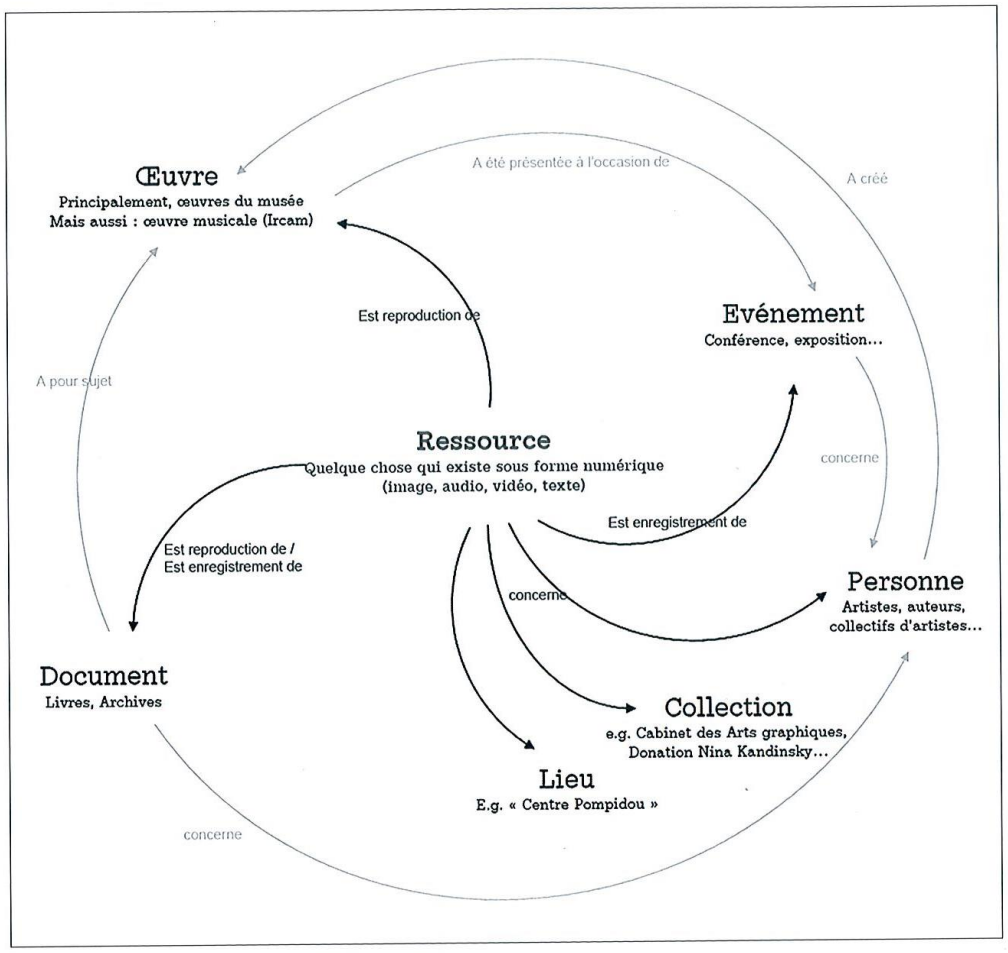


Figure 5 : Vue d'ensemble du modèle de données du Centre Pompidou virtuel [18, BERMÈS].

Le *Centre Pompidou virtuel* est totalement novateur au sein des musées dans sa volonté de valoriser ses collections en s'appuyant sur les principes de structuration du web de données. Seul le MuCEM[50] a également mis en place un projet de consolidation de ses données provenant de ressources différentes et traitées dans quatre outils différents (collections, multimédia, bibliothèque, archives). Ce projet s'est appuyé sur un modèle en UML d'interconnexion entre les données.

Ces deux musées ouvrent la voie à tous les autres musées qui n'ont pas forcément conscience des possibilités offertes par les nouvelles technologies ou n'ont pas toujours les moyens de mener de tels projets à terme. De plus, les outils proposés sur le marché actuellement, s'ils répondent aux normes spécifiques à chaque métier et à chaque ressource, ce qui est déjà indispensable, obligent toutefois les musées à traiter les données en silo. Les informations sont souvent répétées et non reliées entre elles.

[50] Site internet du musée des civilisations de l'Europe et des Méditerranées (MuCEM). www.mucem.org/, [consulté le 11/1é/2013]

L'enjeu à l'heure actuelle, pour les musées, est de pouvoir à la fois traiter les données selon les normes et les outils spécifiques à chaque ressource, tout en harmonisant ces données afin de pouvoir les consolider et fédérer les informations.

Un autre enjeu pour les musées est de normaliser et d'harmoniser les informations documentaires sur leurs collections afin de pouvoir les rassembler sur un même portail numérique avec l'ensemble des données de collections muséales de tous les musées de France.

2.1.2 Sur le Portail national des collections des musées de France : Joconde

Les musées ont également la possibilité d'exporter les données se rapportant à leurs collections sur le Portail des collections des musées de France Joconde[51]. Ce catalogue collectif a été créé en 1975 et est en ligne depuis 1995. Il est administré, au sein du MCC, par le bureau de la diffusion numérique des collections du service des musées de France, en partenariat avec les musées participants.

Il regroupe au 1er juillet 2013 :

- près de 500 000 notices d'objets de toute nature (archéologie, beaux-arts, ethnologie, histoire, sciences et techniques, etc) ;
- dont environ 300 000 illustrés ;
- issus de 370 collections sur les 1315 musées de France classés ;
- incluant des collections très diverses : 211 199 dessins, 64 692 peintures, 56 965 pièces archéologiques, 44 895 pièces ethnologiques, 37 245 estampes, 34 693 sculptures, 25 501 photographies, 22 111 céramiques ou 16 738 costumes et leurs accessoires.

Mais tout cela n'est rien à côté des millions d'œuvres conservées dans les musées. L'effort à faire est donc encore considérable. Des perspectives nouvelles doivent être explorées : modes de collaboration entre institutions, modèles de données ou encore modes de navigation dans les contenus.

D'autant plus que la demande des utilisateurs est forte : en 2010, la base avait été consultée 15 millions de fois. Au cours de l'année 2012, le catalogue des collections des musées de France a reçu plus de 1 096 000 visites du monde entier (chercheurs, étudiants ou particuliers), lesquelles ont généré 62 millions de recherches, la consultation de 7,7 millions d'images et la lecture de plus d'un million de pages.

---

[51] Portail des collections des musées de France Joconde.
http://www.culture.gouv.fr/documentation/joconde/fr/pres.htm/ [consulté le 15/12/2013]

Le versement des données dans Joconde est fait sur une base volontaire pour les musées et le défi pour ceux-ci est donc d'amplifier l'export des données vers Joconde, afin de donner aux utilisateurs la possibilité d'avoir sur un portail unique l'opportunité de pouvoir faire des recherches sur l'ensemble des collections muséales.

Et pour que la recherche soit la plus efficace et la plus intéressante possible, l'enjeu est cette fois-ci d'harmoniser les données sur des collections très diverses conservées dans des dizaines de musées différents. Joconde s'appuie ainsi sur des vocabulaires contrôlés communs à tous les domaines de collections. Mais les musées ont parfois du mal à s'approprier les normes et les référentiels communs. En effet, comme nous l'avons dit, les conservateurs ne sont pas documentalistes et leur mission est un travail scientifique et d'analyse sur les collections. Il parait donc parfois difficile et réducteur de combiner l'analyse de l'œuvre et son aspect purement documentaire.

Mais il est dommage, avec les possibilités qu'offre le web de données, de ne pas chercher à mettre en relation les données et de fédérer les informations. D'autant plus qu'un concept est le même, par exemple, en ethnologie et en arts décoratifs : un lion sera toujours un lion quelle que soit la manière dont il est traité.

Et la mise en relation entre les données ne doit pas se limiter aux collections et aux ressources muséales, mais aussi au patrimoine de l'architecture et du patrimoine mobilier (recensés sur la base mobilier Palissy et sur la base architecture Mérimée[52]) et à tout l'ensemble des ressources culturelles (archives, bibliothèques, etc).

Pour reprendre l'exemple du lion, le concept reste toujours le même qu'il soit traité pour une œuvre, un élément architectural ou mobilier, un ouvrage, etc. Nous présenterons dans notre dernière section, les projets du MCC qui permettent de traiter cette problématique.

Mais le partage des données ne se limite pas aux données culturelles françaises, le web de données est international. Les utilisateurs sont originaires du monde entier et les projets européens et internationaux ont toute leur place et présentent un intérêt non négligeable dans la diffusion des connaissances.

2.1.3 Sur des « bibliothèques » numériques internationales

Des projets visant à mettre en commun les ressources des données culturelles au niveau international voient, en effet, le jour depuis plusieurs années. Nous allons présenter ici deux projets : Eurepeana et Google Art Project.

---

[52] Bases de données documentaires de la direction de l'architecture et du patrimoine du MCC.
http://www.culture.gouv.fr/culture/inventai/patrimoine// [consulté le 15/12/2013]

Europeana est un projet de grande envergure lancé par la Commission européenne en 2008 qui vise à créer une bibliothèque numérique européenne rassemblant les ressources des bibliothèques nationales et européennes, des archives et des musées.

Le projet concerne 77 millions de livres, 24 millions d'heures de documents multimédias, 358 millions de photographies, 75 millions d'œuvres d'art et 10 milliards de pages d'archives[53]. A ce jour, 20 millions d'objets numériques sont sur Europeana sous forme d'images, de textes, de sons et de vidéos. Plus de 1500 institutions participent à ce projet, telles que la British Library au Royaume Uni, le Rijksmuseum à Amsterdam ou le Louvre en France.

Si l'on fait une recherche sur Rodin par exemple, le résultat affiche 7 539 images, 263 textes, 24 sons et 18 vidéos, pouvant être triés par langue, année, pays, *copyright* et contributeurs. En revanche, ce n'est pas une plate-forme d'archivage des œuvres mais un catalogue de recherche, les œuvres étant accessibles par des liens pointant vers les institutions les ayant numérisées.

Un autre projet de grande ampleur, privé et non public cette fois-ci, a vu le jour récemment : *Google Art Project*, accessible sur le portail de l'Institut culturel de Google[54]. Ce service, mis en ligne par Google en février 2011, permet de visiter virtuellement différents musées. Grâce à la technologie *Street View*, Google a permis la numérisation en 3D, de 32 000 œuvres de 151 musées ou lieux différents à travers 40 pays.

L'utilisateur peut rechercher par le nom du musée, le nom d'un artiste, ou le titre d'une œuvre. Lorsqu'on effectue une recherche sur « Rodin » par exemple, le résultat affiche 155 éléments pouvant être triés par musée, auteur, support, lieu, personne, type de support et date. Les 155 œuvres affichées proviennent de 38 musées différents à travers le monde.

Ces projets montrent que la problématique de structuration et la mise en relation des données ne se limitent pas non plus aux données françaises mais doivent être aussi internationales. La question du multilinguisme se pose d'ailleurs de manière importante. Nous aborderons plus longuement cette question dans la dernière section.

Tous ces dispositifs offrent aux musées de vrais moyens de diffusion en ligne et de partage de leurs ressources. Le déploiement de la publication des données culturelles multiplie les points d'accès à l'information et la rend disponible à tous. Les utilisateurs, de plus en plus variés (étudiants, chercheurs, visiteurs) peuvent devenir acteurs en venant réutiliser et

---

[53] Wikipédia.
http://fr.wikipedia.org/wiki/Europeana/ [consulté le 15/12/2013]

contribuer à enrichir les données. Les musées doivent ainsi s'adapter à une nouvelle approche des collections muséales largement ouverte sur les réseaux.

2.2 Réutilisation et contribution des utilisateurs

En effet, l'intégration des collections muséales sur le web permet d'envisager des perspectives inédites d'interaction avec les utilisateurs et d'enrichissement de la description des collections. Les internautes ont des habitudes de consommation en réutilisant les données, et de participation en contribuant à les enrichir. Ils sont habitués à être partie prenante : visiter, télécharger, acheter, s'abonner, contribuer, promouvoir[55]. Cette demande forte des utilisateurs s'oppose au mode de fonctionnement habituel des musées, mais ceux-ci peuvent tout à fait tirer profit des véritables compétences et connaissances des internautes qui peuvent être des spécialistes, des chercheurs, des étudiants ou tout simplement des amoureux de l'art. Ceux-ci, en provenance du monde entier, peuvent ainsi pallier les manques de moyens humains et financiers des musées.

Le terme généralement employé pour désigner ce type de projets est celui de *crowdsourcing*, c'est-à-dire des contenus ou des informations (*source*) produits par la foule (*crowd*) des usagers[56]. Cette activité participative peut se faire par l'indexation collaborative (folksonomie[57]), par l'identification de documents iconographiques ou audio-visuels, etc. On parle alors de « métadonnées » sociales car elles viennent améliorer l'accès des utilisateurs aux collections dont la description est enrichie et améliorée. Mais pour le moment cette pratique est encore jeune et les projets sont innovants et expérimentaux.

En Angleterre, un site ouvre la voie dans ce domaine *: Yourpaintings tagger*[58]*.* Ce projet a été initié par la Public Catalogue Fondation et la BBC. 200 000 peintures à l'huile publiques ont été mises en ligne sur le site web.

L'objectif du projet est de compléter les informations existantes (titre, artiste, date d'exécution) par des informations sur le type de peinture, les sujets décrits dans les

---

[54] Portail de l'institut culturel de Google. http://www.google.com/culturalinstitute/project/art-project [consulté le 15/12/2013]

[55] Ministère de la culture et de la communication. Service des musées de France. Bureau de la diffusion numérique. Portail des collections des musées de France Joconde. Journée professionnelle. "Les musées à l'heure du numérique : travailler en réseau, réutiliser et contribuer". Paris, 7 juin 2013. http://www.culture.gouv.fr/documentation/joconde/fr/partenaires/AIDEMUSEES/journee_BDNC_2013/journee-pres.htm [consulté le 15/12/2013]

[56] MOIREZ Pauline. Bibliothèques, crowdsourcing, métadonnées sociales. Bulletin des bibliothèques de France, 2013, n°5, p32-36.

[57] Folksonomie : système de classification collaborative décentralisée spontanée, basé sur une indexation effectuée par des non-spécialistes.

peintures, les styles et les mouvements représentés. L'indexation des œuvres par ces informations complémentaires permet aux utilisateurs d'effectuer ainsi des recherches bien plus fines et efficaces. La qualité des contributions s'appuie sur une validation par un programme (une indexation est entérinée au bout d'un certain nombre de validation par les utilisateurs ; tant qu'elle n'est pas entérinée, elle est annotée « regard humain » par des experts).

En France, un site est pionnier dans le monde des musées français : *Les Herbonautes, l'herbier numérique collaboratif citoyen*[59]. Ce site propose de contribuer à la création d'une base de données scientifique, à partir des millions de photos des plantes de l'herbier de Paris situé au Muséum national d'Histoire naturelle.

Le projet s'appuie sur le fait que nous sommes tous différents, que nous avons chacun des compétences et des intérêts extrêmement variés : une expertise sur un groupe de plantes, sur une région, sur l'identification d'écritures anciennes, et que nous pouvons tous apporter notre pierre à l'édifice. Des procédures de validation de ces données ont été développées par le Conservatoire National des Arts et Métiers afin d'assurer la qualité et la fiabilité des contributions : formations, et assistance des contributeurs, évaluation de leurs compétences et distribution de rôles différenciés, corrections multiples des mêmes données, vérifications systématiques ou échantillonnées par des professionnels.

[58] Site internet Yourpaintings tagger.
http://tagger.thepcf.org.uk/ [consulté le 15/12/2013]
[59] Site internet Les Herbonautes, l'herbier numérique collaboratif citoyen.
http://lesherbonautes.mnhn.fr/ [consulté le 15/12/2013]

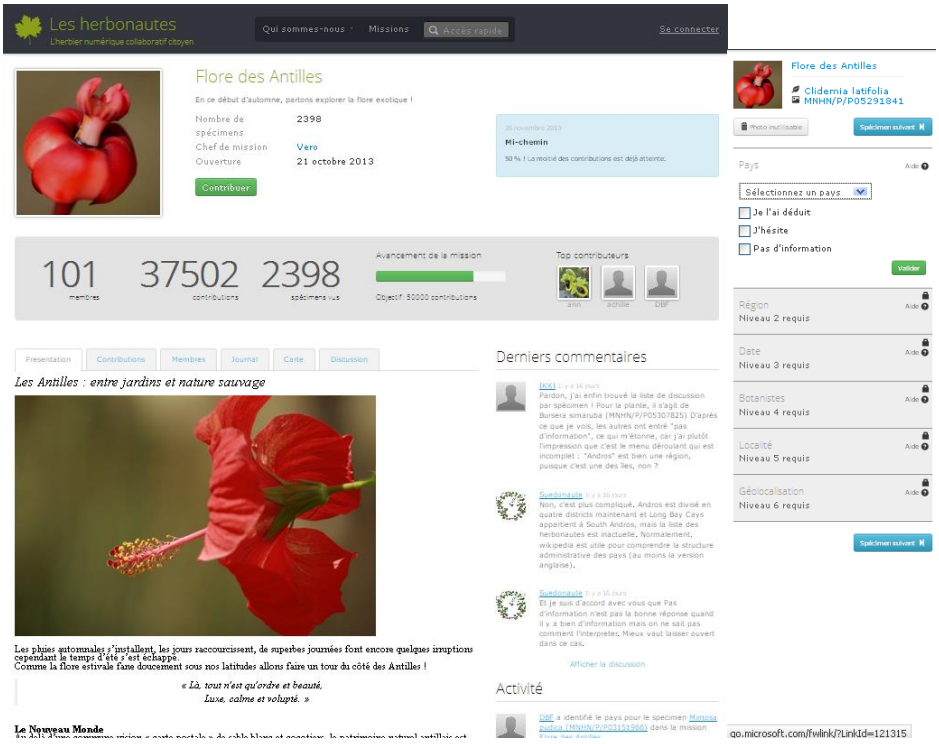


Figure 6 : Extraits du site *Les Herbonautes, l'herbier numérique collaboratif citoyen.*

Cet enrichissement des données par les utilisateurs demande également en amont une réflexion sur la structuration des données avec un travail sur les vocabulaires et les référentiels. L'utilisation des langages documentaires tels que les thésaurus, en lien avec les technologies du web sémantique, permettent ainsi aux données d'être plus accessibles sur le web et aux musées d'accroître leur visibilité au niveau mondial, de partager et d'enrichir leur données avec d'autres musées et d'autres institutions culturelles, valorisant ainsi leurs collections. Ce sont là tous les enjeux du web de données que nous allons détailler maintenant.

2.3 Enjeux du web de données

Nous avons donné quelques repères dans la première partie sur ce qu'est le web de données (cf. partie 1 section 3.3:). Nous allons maintenant expliquer plus en détail quels en sont les enjeux pour les données culturelles et notamment pour les collections muséales, puis comment fonctionne ce web de données.

2.3.1 Quel apport pour les musées ?

Comme nous venons de le voir, les musées ont l'opportunité d'aller au plus près des utilisateurs en répondant aux politiques culturelles du MCC et en diffusant les collections et les connaissances sur le web.

Mais les données sur les collections muséales, mises en ligne sur le web, sont peu indexées par les moteurs de recherche. Or les utilisateurs sont habitués à chercher les informations sur Google et ne cherchent pas à savoir quel musée possède telle œuvre, telle ressource ou telle donnée. Les utilisateurs retrouvent donc difficilement les données sur le web.

La mise en place de projets de valorisation des données culturelles dans le cadre du web de données permet de faciliter l'accès à l'information, de relier les ressources entre elles et de les rendre visibles sur le web par les moteurs de recherche. Les données peuvent ainsi « sortir » des silos documentaires, être échangées plus facilement entre les musées et sont rendues plus accessibles aux utilisateurs.

Comme nous l'avons vu, le web de données s'appuie sur de nouveaux langages technologiques permettant de structurer les données et de les lier, mais cela ne fonctionne qu'avec un travail important sur le vocabulaire et sur la mise en place de référentiels communs. Les langages documentaires ont donc là encore toute leur place, et notamment les thésaurus. Nous verrons, dans notre dernière section, comment permettre d'aligner les vocabulaires avec les thésaurus, permettent à l'utilisateur de trouver la réponse à sa question quelle que soit sa demande. Mais pour en comprendre tout l'intérêt, expliquons tout d'abord comment fonctionne exactement le web de données.

2.3.2 Comment cela fonctionne ?

Comme nous l'avons vu dans la première partie, le web de données consiste à lier et structurer la multitude d'informations présente sur internet afin de la rendre accessible, exploitable et réutilisable. L'objectif est en quelque sorte de créer une immense base de données fonctionnant avec les technologies du web sémantique. Le web de données repose sur les notions fondamentales du web, l'identifiant URI, qui permet de nommer les ressources, et le protocole HTTP, et sur les nouveaux langages du web sémantique permettant de décrire les données et de les relier entre elles, les langages XML et RDF et le standard SparQL [11, GANDON].

Tim Berners-Lee, inventeur du web, a défini une notation en cinq étoiles permettant d'évaluer la qualité de l'intégration des données sur le web de données, à savoir des données accessibles, structurées, dans un format non propriétaire, identifiées et liées entre elles sémantiquement :

- niveau * : mise en ligne des données sur le web sous licence libre (quel que soit le format) ;
- niveau ** : structurer et expliciter les données (par exemple, utiliser Excel plutôt qu'une image scannée) ;
- niveau *** : utiliser un format non propriétaire (par exemple, utiliser CSV plutôt qu'Excel) ;
- niveau **** : utiliser un identifiant URI pour identifier les sujets et les objets ;
- niveau ***** : lier les données entre elles[60].

Afin de mettre en œuvre ces recommandations et pouvoir lier les données entre elles, il faut tout d'abord comprendre comment fonctionne le langage RDF, qui est l'unité de base du web sémantique [13, SAJUS]. Bertrand Sajus (département numérique du MCC) explique très bien cela et montre de quelle manière fonctionne le triplet RDF (sujet, prédicat, objet).

Nous nous inspirons de sa présentation pour tenter de l'expliquer ici avec un exemple sur Rodin et la Porte de l'Enfer :

- en langage naturel :

  « Rodin est l'auteur de La Porte de l'Enfer »

- selon la méthodologie du web sémantique, avec le triplet RDF :

  Sujet « Rodin »

  Prédicat « est l'auteur de »

  Objet « La Porte de l'Enfer »

Chaque élément :

- peut être identifiée par un URI ;
- peut provenir d'une source de référentiels sur le web : par exemple : « Rodin » est référencé dans VIAF (*Virtual International Authority File*) ;

[60] 5 star *open data.*
http://5stardata.info/ [consulté le 15/12/2013]

- peut être connecté avec d'autres triplets (exemple : « La Porte de l'Enfer est inspiré de la Divine Comédie », « Dante est l'auteur de la Divine Comédie », « La Porte de l'Enfer est conservé au musée Rodin », etc) ;

Plusieurs triplets interconnectés :

- composent un graphe ;
- permettent d'enrichir les données par liage (par exemple : lier « la Porte de l'Enfer », « le musée Rodin » et « Dante ») ;
- de créer des liens entre les corpus de données (par exemple entre le corpus du portail des collections des musées de France pour les données propres au musée Rodin et le catalogue de la BnF pour les données sur l'œuvre de Dante).

Enfin le langage de requête SparQL permet :

- d'interroger les graphes RDF ;
- d'en extraire les données.

Pour que cela fonctionne, et que les données soient complétement interopérables il faut ainsi :

- avoir des référentiels communs (par exemple, afin d'éviter que Rodin soit écrit une fois « A. Rodin » et une autre fois « Auguste Rodin », ce qui empêche les données d'être interopérables, il est nécessaire de créer des référentiels normalisant les données, comme le référentiel auteur VIAF).
- utiliser des vocabulaires structurés tels que les thésaurus et les classifications ou autres langages documentaires, qui pour être publiés dans le web sémantiques doivent être traduits en SKOS. Les vocabulaires sont représentés de manière conceptuelle avec SKOS selon les principes du web sémantique, un objet n'étant pas décrit par un terme mais par un concept. Un « concept SKOS » est ainsi défini comme une ressource RDF et donc identifié par une URI. Des correspondances peuvent être réalisées entre des concepts provenant de différents systèmes d'organisation de connaissances permettant ainsi un alignement des vocabulaires, de façon à rendre ces différents langages davantage interopérables.

Un certain nombre de projets reprenant les principes du web de données et du web sémantique sont initiés et soutenus actuellement par le MCC afin de valoriser l'ensemble des données culturelles en garantissant l'interopérabilité des données.

3 Garantir l'interopérabilité des données

Le MCC, en effet, met en place et soutient des projets et des applications du web de données sur l'ensemble des données culturelles. Ces projets visent en garantissant l'échange et l'interopérabilité des données sur le web, à valoriser les données culturelles en France et à l'étranger et à permettre à l'utilisateur d'y accéder plus facilement en favorisant leur visibilité sur le web. Ces applications portent essentiellement sur une meilleure structuration et une harmonisation des données et sur la gestion des vocabulaires et des référentiels.

3.1 Structuration et harmonisation des données culturelles

Les projets concernant la structuration et l'harmonisation des données culturelles sont de grande ampleur comme Dbpedia en français (structurer les données de Wikipédia et les partager sur le web), comme data.bnf.fr (exploiter et rassembler les données de différents catalogues et les rendre plus visibles sur le web), comme HADOC (harmoniser l'ensemble des données culturelles en définissant un nouveau modèle conceptuel de production des données).Data.bnf.fr n'est pas un projet initié directement par le MCC mais son ampleur et son application d'envergure en font un exemple pour l'ensemble des données culturelles.

3.1.1 Dbpedia en français dans le cadre de Semanticpdia

Sémanticpédia est une plate-forme de collaboration entre le MCC, l'Inria et Wikimédia France. Son objectif est la mise en place de projets de recherche et de développement à partir des techniques du web sémantique sur les ressources francophones de Wikimédia. Le but de ces projets est de développer un écosystème de données culturelles liées et publiquement accessibles [26, Semanticpedia].

Ce projet répond à un vrai besoin de gestion des données culturelles, car sur les 1 300 000 entrées de Wikipédia, 45 % concernent le domaine de la culture. C'est aussi une plate-forme de promotion de la langue française permettant une présence forte de la langue française dans le web de données. Son projet est d'ailleurs suivi par la Délégation générale à la langue française et aux langues de France.

Dbpedia en français est le premier projet de Sémanticpédia. Il est le chapitre francophone de DBpedia qui a pour but d'extraire des informations structurées à partir des ressources offertes en libre accès par Wikipédia, afin de rendre ces données disponibles sur le Web. L'objectif étant de relier Wikipédia à d'autres ensembles de données ouvertes provenant du

Web des données. Nous verrons un peu plus loin deux projets allant dans ce sens, HdALab et Joconde Lab (cf. section 3.2.1).

Dbpedia repose donc sur les standards du web de données [27, DBpedia]. Les ressources sont livrées sous forme de données normalisées au format du web sémantique (format RDF et identifiant URIs) et sont regroupées au sein de documents dérivés de l'encyclopédie Wikipédia. Ainsi, pour chaque document encyclopédique, existe une page de ressources, contenant toutes les données sous formes de triplets RDF.

DBpedia est initialement un projet anglophone lancé par l'université libre de Berlin et l'université de Leipzig. Le premier ensemble de données accessibles au public a été publié en 2007. En novembre 2010, la base de données décrivait 3,4 millions d'entités. Mais la création de données d'après une encyclopédie anglophone a rapidement montré des limites de couverture et le projet a été internationalisé. Différentes versions locales de DBpedia, dont la version francophone, ont donc pour but de donner accès aux données publiées dans les versions respectives de Wikipédia. La version francophone de DBpedia offre ainsi la possibilité d'interroger ces données mais aussi d'intégrer différentes ressources francophones au web de données.

Ce dispositif offre aux établissements culturels et donc aux musées des perspectives inédites de diffusion et de partage de leurs ressources en ligne. Les possibilités de recherches (chronologiques, thématiques, géographiques, etc) sont démultipliées. Le projet JocondeLab que nous allons décrire un peu plus loin montre toute l'opportunité pour les musées de cette démarche.

3.1.2 data.bnf.fr

Le projet data.bnf.fr lancé en 2011 a pour but de valoriser la richesse des fonds de la BnF (plusieurs millions d'ouvrages, de manuscrits, de documents divers) sur le web en exploitant les données de différents catalogues (catalogue général de la BnF, bibliothèque numérique Gallica, Europeana, Wikipédia, Sudoc) et de les rendre davantage présentes sur le web [19, BnF].

Le projet permet ainsi, en rendant les données compréhensibles par les machines, de les sortir des silos en reliant les différents catalogues, de créer un point d'entrée unique améliorant la recherche et la navigation pour l'utilisateur, à la fois au sein du site de la BnF et des catalogues mais aussi à partir des moteurs de recherche sur le web.

Pour cela, le projet s'appuie sur les principes du web de données et du web sémantique, en travaillant sur l'alignement des données :

- à partir des millions de notices bibliographiques présentes dans les catalogues, structurées et reliées ensemble ;
- reposant sur les listes d'autorité fiables et exactes, sur les auteurs, les œuvres et les sujets (noms communs, géographiques et subdivisions chronologiques RAMEAU), maintenus par des URISs.

La BnF propose :

- une nouvelle interface de navigation avec des pages HTML, réunissant sur chaque page toutes les informations relatives aux œuvres, aux auteurs et aux sujets, destinées au grand public. L'utilisateur accède ainsi par des liens à toutes les ressources de la BnF et aux ressources extérieures.
- Des données structurées et interopérables selon les techniques du web sémantique (RDF) exposées sur le web de données. Data.bnf.fr expose notamment les notices de sujets de RAMEAU, converties dans le langage RDF SKOS.

Ainsi, en structurant les données existantes, en les enrichissant et en liant son jeu de données avec des ressources externes et internes, en publiant des pages HTML sur lesquelles les utilisateurs et les moteurs de recherche puissent naviguer, et en exposant des données brutes en RDF selon les principes du *linked open data*, la BnF ne remplace pas les catalogues existants mais propose un outil destiné à aider les internautes à trouver dans les ressources ce dont ils ont besoin et à les partager plus facilement.

La BnF souhaite étendre ce projet au-delà des bibliothèques, à toutes les institutions culturelles et notamment aux musées. Mais se posent des difficultés liées aux différences de types de ressources et de données. Le projet HADOC que nous allons décrire maintenant travaille justement sur l'harmonisation des données culturelles afin de pouvoir relier des données hétérogènes et les structurer dans le web de données.

3.1.3 HADOC

Le projet HADOC (*Projet d'harmonisation des données culturelles*) a été lancé dans le cadre du schéma directeur des systèmes d'information 2005-2008 du MCC et répond à une constatation concernant le traitement des données culturelles à l'heure actuelle [21, BRIATTE] :

- une multiplicité des pratiques et des outils ;
- une obsolescence de la chaîne de production ;
- un mode de fonctionnement en silo ;
- des problèmes de redondance et d'incohérence informationnelle des données.

Les équipes travaillant sur le projet ont mis en avant le besoin de travailler autant sur les données elles-mêmes, afin d'accroître leur qualité, que sur la production de ces données, afin d'optimiser leur processus, en agissant sur trois leviers :

- l'harmonisation des données ;
- l'harmonisation des processus liés à la production de ces données ;
- la généralisation des normes et des standards de production.

Le projet souhaite pour cela changer les pratiques et proposer de nouveaux modes de production, concevoir des modèles de données communs garantissant plus d'interopérabilité entre les données, mettre en œuvre des référentiels communs et partagés, décloisonner les métiers culturels et permettre ainsi la création de portails numériques qui puissent fédérer toutes les ressources malgré leur hétérogénéité. Afin de mettre en œuvre tout cela, le but du projet est de modéliser les données d'identification des biens culturels, en concevant un méta-modèle dans lequel chaque métier puisse se retrouver.

Le projet s'applique plus ou moins bien selon les domaines culturels. En effet, au sein des métiers culturels, comme nous l'avons déjà montré, certains domaines sont très normalisés, comme les archives et les bibliothèques, et mettent en place plus facilement des projets sémantiques tel que data.bnf.fr, ou s'intègrent assez bien dans des projets d'envergure comme HADOC, mais ce n'est pas le cas pour les collections, pour lesquelles il n y a pas de culture normalisée. Or justement, et c'est là tout l'intérêt de ce projet, le champ d'application de ce nouveau modèle doit être applicable à l'ensemble des données culturelles permettant une véritable harmonisation des données entre les bibliothèques, les archives et les musées.

Pour les musées, le projet s'appuie tout de même sur les normes édictées par le MCC (cf. partie 2 section 2.2) et sur les normes internationales élaborées afin de faciliter l'échange des données des musées et des données sur le patrimoine culturel au niveau international mais relativement peu utilisées en France :

- La norme SPECTRUM[61], créée par la Collection Trust au Royaume-Uni, est un guide des pratiques en matière de documentation muséologique, élaboré avec le milieu muséal. Cette norme comprend des procédures de documentation des objets et des processus qui s'y rapportent. Elle indique aussi comment identifier l'information qu'il faut enregistrer pour étayer ces procédures. C'est une norme procédurale relative aux métadonnées pour la documentation des collections des musées, employée pour les échanges de données muséologiques à l'international, utilisée par plus de 8 000 musées dans 40 pays différents.

[61] Site internet de la Collection Trust.
http://www.collectionslink.org.uk/programmes/spectrum [consulté le 15/12/2013]

- La norme des CDWA Lite[62] (*Categories for the Description of Works of Art*) a été créée par the Art Information Task Force (AITF) avec l'appui financier du J. Paul Getty Trust. Ces catégories permettent de décrire, sur la base d'un schéma en langage XML, le contenu de bases de données sur les œuvres d'art en proposant un cadre conceptuel pour décrire l'information sur des objets et des images et y donner accès.

- Le modèle CRM (*Conceptual Reference Model*), élaboré depuis 1994, au sein de l'ICOM, par le Documentation Standards Working Group du CIDOC (Comité international pour la documentation)[63], est un modèle sémantique de référence qui constitue une « ontologie » de l'information relative au patrimoine culturel, sur la base d'un schéma en langage XML. Patrick Leboeuf (BnF, référent français du CIDOC CRM) explique que c'est un modèle *« qui vise à expliciter sans ambiguïté le sens des informations relatives à un « objet de musée » dans l'acception la plus large du terme (œuvre d'art, vestige archéologique tel qu'un clou ou un four à pain, corps momifié, fleur séchée ou animal empaillé, mais aussi quelque chose qu'il est matériellement impossible de conserver à l'intérieur d'un musée mais qu'une institution du patrimoine culturel peut gérer, comme un monument ou un site tel qu'une grotte préhistorique), et ce quelle que soit la structure des données qui véhiculent ces informations. Le CRM peut en outre se définir comme : un modèle orienté objet ; un modèle sémantique ; une ontologie de domaine »*[64]. Il permet, en effet, de décrire, les concepts et les relations utilisés pour la documentation muséale et de promouvoir un partage des données en créant un cadre sémantique commun, un langage commun pour tous les domaines d'expertise et en proposant un guide des bonnes pratiques de la modélisation des concepts. En 2003, un rapprochement a été créé entre le modèle CIDOC CRM des musées et le modèle FRBR[65] des bibliothèques, conçu selon le formalisme dit « entité relation ». Ce rapprochement a entraîné une reformulation du modèle FRBR dans le formalisme dit « orienté objet » dans lequel est exprimé le modèle CIDOC CRM (il existe donc aujourd'hui deux versions

---

[62] Site internet du Getty.
http://www.getty.edu/research/publications/electronic_publications/cdwa/introduction.html [consulté le 15/12/2013]

[63] Site internet de l'ICOM.
http://icom.museum/the-committees/international-committees/international-committee/international-committee-for-documentation/ [consulté le 15/12/2013]

[64] LE BŒUF Patrick. Le modèle CRM pour la documentation muséographique : s'attacher au sens pour n'être pas piégé par la forme. Journée d'étude organisée par l'ADBS : « La modélisation : pourquoi l'intégrer dans les systèmes d'information documentaire ? », Paris-La Défense, 20 mai 2003.
http://www.cidoc-crm.org/docs/adbs_crm.pdf [consulté en ligne le 15/12/2013]

[65] Modèle FRBR (Functional Requirements for Bibliographic Records, en français Fonctionnalités requises des notices bibliographiques) : modèle conceptuel de données bibliographiques élaboré par l'International Federation of Library Associations and Institutions (IFLA).

distinctes du modèle FRBR : $FRBR_{ER}$ et $FRBR_{OO}$[66]). La combinaison du CIDOC CRM et de FRBRoo constitue désormais un modèle conceptuel commun à l'information muséographique et à l'information bibliographique. En 2006, le CRM est devenu une norme ISO, ce qui ne le rend pas obligatoire mais lui donne une légitimité au-delà même de la documentation muséographique. Et en 2010, un projet a été mis en place afin d'étendre également ce modèle conceptuel commun aux archives afin de couvrir la totalité du spectre de l'information relative au patrimoine culturel.

- La norme museumdat[67], permet également de décrire les objets du patrimoine culturel sur la base d'un langage en schéma XML. C'est une version de la norme CDWA, à la fois simplifiée et compatible avec le CRM (avec une approche multidisciplinaire du CRM).
- Le projet LIDO (*Lightweight Information Describing Objects*, en français Informations légères pour la description des données) fédère ces différentes normes et a été lancé en 2010, par l'ICOM, au sein du groupe de travail sur la récolte et l'échange des données du CIDOC [4, ICOM-CIDOC]. Il est fondé sur les schémas en XML CDWA Lite et museumdat, est conforme au modèle CIDOC CRM et est en phase avec SPECTRUM. Il permet de manière puissante de transférer les données des bases de données de collections muséales aux agrégations de données et de partager les données muséales sur le Web, quelle que soit l'hétérogénéité des données.

Le modèle de données HADOC s'appuie sur le développement de ces normes et procède à un alignement systématique du modèle CIDOC CRM sur HADOC, afin de fédérer les informations sur les collections françaises traitées en silos, même au niveau national avec la base Joconde (collections muséales), la base Palissy (collections mobilières), la base Mérimée (collections architecturales). Le but n'est pas de fusionner ces bases mais de consolider les informations afin de traiter de manière commune les concepts identiques dans ces différents domaines de connaissances et de ressources, de les partager avec les autres données culturelles et ainsi de modéliser l'ensemble des données d'identification des biens culturels.

La première version du modèle est sortie en juillet 2013, avec la mise à disposition d'un modèle d'implémentation pour les outils informatiques qui peut être utilisé par les éditeurs de logiciels.

---

[66] Site internet de la BnF.
http://www.bnf.fr/fr/professionnels/modelisation_ontologies/a.modele_FRBR.html [consulté le 15/12/2013]
[67] Site internet de museumdat.
http://www.museumdat.org/index.php?ln=en [consulté le 15/12/2013]

Ce projet montre également toute l'importance pour mener de tels programmes sémantiques, du travail à mener sur l'alignement et la gestion des vocabulaires et l'utilisation des référentiels.

3.2 Gestion des vocabulaires et utilisation des référentiels

En effet, un projet permettant une meilleure structuration des données comme le modèle HADOC ne peut être totalement mis à profit sans un travail sur les vocabulaires. Nous allons voir comment la mise en place de *tagging* sémantique comme HdALab et JocondeLab, le développement d'un outil de gestion des vocabulaires, tel que GINCO et l'utilisation de référentiels peuvent être un formidable levier pour la valorisation des données sur le web.

3.2.1 *Tagging* sémantique

3.2.1.1 HdALab

Le projet HdALab a été lancé par le MCC et l'Institut de recherche et d'innotation (IRI) afin de mettre en valeur les données grâce au *tagging* sémantique. Le corpus utilisé pour ce projet est issu du portail *Histoiredesarts* dont le but est de proposer aux utilisateurs des documents fiables sur l'histoire de l'art [23, HdALab].

Le portail *Histoiredesarts* propose ainsi un accès à une sélection de 5 000 ressources qualifiées et accessibles, classées selon le programme d'enseignement d'histoire des arts, et alimenté par un réseau de 350 institutions culturelles. Chaque ressource est décrite par des notices qui renvoient par des liens vers des ressources en ligne et qui contiennent des mots-clés qui sont au départ de simples tags.

Dans ce projet, le *tagging* sémantique, qu'on devrait appeler plutôt « liage sémantique » consiste à relier chaque tag (mot-clé), en appliquant des liens conformes aux technologies du Web sémantique, aux termes de DBpedia en français.

Le fait d'utiliser un référentiel aussi complet que DBpedia permet de donner du sens aux tags, de partager des compléments d'information et facilite ainsi la recherche des utilisateurs. Ce qui est intéressant dans le retour d'expérience de l'équipe du projet HdALab, c'est que 80% des mots-clés du portail *Histoiredesarts* ont pu être liés au dictionnaire Wikipédia en français et que par exemple, au mois d'octobre 2012, 100 % des 100 mots-clés les plus fréquemment utilisés par les visiteurs du portail *Histoiredesarts* sont liés à ce dictionnaire.

D'autre part, les termes de DBpedia étant traduits dans des dizaines de langues, ce liage sémantique permet par ricochet un accès multilingue aux mots-clés du portail *Histoiredesarts* et donc une recherche en plusieurs langues.

3.2.1.2 JocondeLab

Le projet JocondeLab se situe dans le prolongement du projet HdALab. Il est développé par l'IRI et est dirigé, au sein du MCC, par la délégation générale à la langue française, en collaboration avec le service des musées de France et le département des programmes numériques [24, MCC].

Ce projet s'applique au corpus de la base Joconde du MCC, constitué de plus de 400 000 notices d'œuvres des musées de France, et consiste, comme pour le projet HdALab à relier les mots-clés documentaires de la base Joconde et les termes de DBpedia en français.

Le projet permet ainsi d'explorer une nouvelle navigation dans les contenus et les images de Joconde et d'expérimenter l'indexation collaborative.

Mais surtout l'un des principaux objectifs de ce liage consiste à enrichir les données et surtout à internationaliser l'accès à Joconde par ces mots-clés. Les liens avec Dbpédia permettent en effet, d'obtenir automatiquement des équivalents de ces mots-clés dans diverses langues et favorise ainsi le multilinguisme. L'accès des collections des musées de France en ligne sur le catalogue national peut ainsi s'étendre à un public non francophone et favorise la promotion de la culture française à l'étranger.

Ces deux projets, HdALab et JocondeLab montrent toute la nécessité et tout l'intérêt dans le web de données, du travail sur les vocabulaires et sur leur alignement, et sur l'utilisation de référentiels communs.

3.2.2 Outil de gestion des vocabulaires : GINCO

L'application GINCO *(application de gestion informatisée de nomenclatures collaboratives et ouvertes)* est un projet du MCC, actuellement en cours de développement, et constitue une brique du projet HADOC. Il part du principe que les langages scientifiques et techniques garantissent la qualité de la production documentaire et sont des outils indispensables pour faciliter la navigation dans le contexte du web sémantique.

L'objectif du projet est de donner accès, de manière homogène et centralisée, à l'ensemble des vocabulaires scientifiques et techniques produits par le MCC et ses partenaires.

Cet outil permet de concevoir, de produire et de diffuser des vocabulaires tels que les listes d'autorité, les taxonomies et les thésaurus et s'appuie sur les normes et les standards en

cours, la norme ISO 25964-1:2011, dont la méthodologie met en particulier en œuvre la différenciation entre concept et terme, et les langages SKOS.

Il offre les fonctionnalités classiques d'un outil de gestion de thésaurus :

- création de termes et de concepts ;
- relations d'équivalence entre termes ;
- relations de hiérarchie (mono- ou polyhiérarchie) et relations 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 ;
- éditions alphabétiques et hiérarchiques ;
- import et export de vocabulaires au format SKOS/RDF.

L'application, en cours de développement, est accessible sur la plate-forme de diffusion des vocabulaires du MCC accessible à l'adresse data.culture.fr/thesaurus, et s'enrichit progressivement de nouveaux vocabulaires. Les vocabulaires peuvent être consultés en ligne, téléchargés au format SKOS et interrogés par les machines au moyen du langage de requête SparQL.

3.2.3 Projet de référentiel des acteurs historiques

Une autre « brique » du projet HADOC est l'identification et le partage de référentiels de données, c'est-à-dire des regroupements de données de référence qui sont des informations partagées et fréquemment utilisées par l'ensemble des acteurs culturels (archives, musées, bibliothèques, etc). Cela permet ainsi de résoudre des problématiques de diffusion et de produire des données solides et consolidables.

Le projet part du principe que toute donnée partagée et utilisée de manière répétitive doit être sortie des applications et intégrée dans une « zone référentielle » qui peut ensuite être utilisée par les différentes applications.

Le projet prévoit ainsi de créer un référentiel des acteurs historiques dans l'écosystème du web sémantique, sous la forme d'un modèle RDF ou plutôt d'une ontologie décrivant les acteurs historiques, et qui sera une source d'informations de référence sur les autorités parmi d'autres référentiels (cela implique d'interconnecter les jeux de données diffusés par le MCC avec les gisements de données existants) [25, BELKOFF].

Plusieurs référentiels d'autorité majeurs participent au web de données selon les technologies du web sémantique et ont potentiellement de nombreuses intersections possibles avec le futur référentiel d'acteurs historiques :

- Le VIAF (Virtual International Authority File, en Français Fichier d'autorité international virtuel)[68] a été fondé en 2003 par un consortium constitué par la Bibliothèque du Congrès et la Deutsche Nationalbibliothek, que la Bibliothèque nationale de France a rejoint en 2007. Il s'agit d'un fichier fédérant des données d'autorités provenant de vingt-sept établissements, parmi lesquels une majorité de bibliothèques, ainsi qu'un musée le Getty et de Wikipédia. Le VIAF regroupe la plus importante quantité de données d'autorités présentes sur le web, soit trente-quatre millions. Le VIAF a commencé par traiter les données d'autorité sur les noms de personne et intègre progressivement les données d'autorité sur les collectivités, les noms géographiques, les œuvres et les expressions[69].
- L'ISNI[70] (*International Standard Name Identifier*) est un code international normalisé servant à identifier de manière univoque, sur le long terme et à l'échelle internationale, les personnes et les organismes ayant été impliqués dans la création, la production, la gestion ou la distribution de contenus intellectuels ou artistiques ou faisant l'objet de ces contenus (auteurs, comédiens, les producteurs, les éditeurs). Aujourd'hui, sa base de données contient une quantité d'autorités équivalente à un cinquième de celle du VIAF, soit 6 636 916 identités publiques. Mais la part d'acteurs associés à des biens patrimoniaux y est probablement assez réduite, l'attribution d'un identifiant à un acteur historique n'étant pas nécessaire pour inventorier et gérer les biens culturels[71].
- Nous pouvons citer également l'ULAN (*Union List of Artist Names*) qui est une liste anglophone du site internet de l'institut de recherche Getty. Elle compte actuellement plus de 293 000 noms d'artistes en tout genre et de toute époque, dont une partie est reversée sur le VIAF. Elle doit être également entièrement publiée sur le web de données.
- DBpedia en Français, dont l'objectif est de relier Wikipédia à d'autres ensembles de données ouvertes provenant du Web de données (et inversement), est également un

[68] Site internet du VIAF.
http://viaf.org/ [consulté le 20/12/13]
[69] Présentation du VIAF sur le site de la BnF.
http://www.bnf.fr/fr/professionnels/donnees_autorites/a.viaf.html [consulté le 15/12/2013]
[70] Site internet de l'ISNI.
http://www.isni.org/ [consulté le 20/12/13]
[71] Présentation de l'ISNI sur le site de la BnF.
http://www.bnf.fr/fr/professionnels/s_informer_autres_numeros/a.isni_autres_numeros.html
[consulté le 15/12/2013]

référentiel sémantique particulièrement intéressant pouvant présenter des interactions avec le futur référentiel d'acteurs historiques.

Il a été décidé en mars 2013 de porter un premier jeu de données du MCC dans le VIAF en 2014 : la partie orfèvres de la base de données Poinçons constituée par la Mission de l'Inventaire à partir d'informations issue de registres et de plaques d'insculpation de la Garantie de Paris (1875-1924). Ce jeu de données sur les orfèvres dans le VIAF va être versé dans le fichier d'autorité de la Bibliothèque nationale de France où il sera converti dans le format requis par le VIAF, le format MARC[72].

L'intégration du référentiel des acteurs historiques dans le fichier d'autorité international permettra aux données culturelles d'être accessibles à un public plus large qui pourra avoir accès aux informations que contiennent les bases de données nationales sans qu'ils aient à connaître leur existence ni à faire une recherche dans ces outils. L'intégration du VIAF permettra également la conversion des données du référentiel en triplets ce qui devrait permettre de faciliter leur réutilisation et leur assurer une meilleure accessibilité.

Mais ce projet de référentiel des acteurs historiques est encore expérimental et il faudra encore beaucoup de temps pour qu'il soit véritablement opérationnel pour les institutions culturelles et notamment les musées.

Les autres projets que nous avons décrits dans cette section sont également très récents même s'ils commencent à faire leurs preuves dans le web de données. Il n'est pas facile d'en faire un bilan et d'en tirer des préconisations concrètes pour les musées.

3.3 Bilan sur ces projets

En effet, nous n'avons pas encore le recul nécessaire pour effectuer une analyse poussée de ces projets et de leurs applications dans les musées. C'est un peu la limite de ce travail qui fait un état des lieux sur le traitement de l'information sur les collections muséales et sur l'ensemble des projets et des normes en cours de développement dans le cadre du web sémantique mais qui ne peut en tirer des conclusions définitives sur l'ensemble des préconisations à mettre en œuvre. Au sein des musées, seuls le Centre Pompidou Virtuel et le MuCEM ont développé des projets dans le cadre du web sémantique, mais ces projets sont encore en construction. Ce mémoire constitue donc une étape dans une réflexion qu'il faut continuer à mener dans les années à venir.

[72] Formats MARC (*MAchine-Readable Cataloging*) : formats d'échanges de données bibliographiques permettant d'informatiser les catalogues de bibliothèques.

Mais ces projets montrent d'ores et déjà toute l'importance d'une meilleure visibilité des données culturelles sur le web, réalisable par l'harmonisation de l'ensemble des données culturelles s'appuyant sur une structuration des données selon les technologies du web sémantique et sur les vocabulaires contrôlés et les référentiels.

D'autre part, la mise en place concrète des projets que nous venons de décrire et leur intégration dans la gestion et le traitement au quotidien des données sur les collections muséales, présentent des problématiques de mutualisation et de gestion centralisée des informations. Cette gouvernance de l'information paraît compliquée à mettre en œuvre et il paraît plus intéressant de travailler à l'instauration d'une collaboration étroite avec les producteurs de données dans les musées qui sont au plus près des œuvres et qui, par conséquent, en ont une meilleure connaissance.

Cela implique une bonne communication entre le MCC et les musées, et entre les différents corps de métiers, car ces projets présentent des notions qui paraissent compliquées à saisir pour les musées (ils traitent d'informatique, de modélisation et de web sémantique qui reste encore une notion émergente et dont les technologies peuvent être difficiles à appréhender même en comprenant les enjeux du web de données), alors qu'au contraire ces projets visent à simplifier et à rationaliser le traitement de l'information, à donner du sens aux données et à les rendre compréhensibles par les machines facilitant ainsi le travail humain.

Le MCC porte donc la responsabilité de ces projets et doit continuer à les développer et les coordonner au niveau national, tout en nouant des partenariats au niveau international, et mettre à disposition des musées et des éditeurs de logiciels, tous les éléments techniques, et toutes les indications précises sur l'évolution des modèles et des normes.

Dans le même temps, les musées, afin de pouvoir s'intégrer dans les projets initiés par le MCC dans le cadre du web de données et du web sémantique, doivent être capables de structurer leurs données sur les collections muséales en s'appuyant sur les vocabulaires contrôlés et les référentiels existants dans leurs domaines de connaissances. L'utilisation de langages et de référentiels communs permet ainsi aux musées de partager respectivement leurs données et leurs connaissances sur les œuvres, ce qui est un premier pas pour garantir l'échange et l'interopérabilité des données sur le web.

Mais le traitement documentaire des données est un métier qui ne s'improvise pas et les musées ont tout intérêt à s'appuyer sur les professionnels de l'information et de la documentation, qui peuvent faire le lien avec les projets du MCC. Ces professionnels de l'information et de la documentation peuvent faire remonter les besoins concrets des musées d'autant plus que les vocabulaires existants ne sont pas toujours assez précis et doivent continuellement être enrichis ou adaptés. Ce travail ne se fait pas sans une réflexion importante demandant une bonne compréhension à la fois des besoins des conservateurs et

des projets à mettre en œuvre, en quelque sorte « le pont » nécessaire à une bonne collaboration entre le MCC, les musées et l'ensemble des institutions culturelles.

La mise en place de ces projets pourra ainsi être profitable, non seulement aux musées qui pourront valoriser de manière tout à fait novatrice leurs collections, à l'ensemble des données culturelles qui pourront être diffusées de manière harmonisée et bien mieux visibles, au niveau national et international, et surtout aux utilisateurs qui auront un meilleur accès à l'information, grâce à une recherche unifiée et facilitée.

Conclusion

En travaillant tout au long de ce mémoire sur la problématique posée en introduction, nous avons pu constater que les musées doivent définitivement ajouter un troisième et nouveau mode de valorisation de leurs collections en diffusant, sur le web, les données et les ressources dont ils disposent. Ce nouveau mode de communication est impulsé par les politiques culturelles du Ministère de la culture et de la communication (MCC) et une forte demande des utilisateurs, qui souhaitent non seulement accéder aux données pour les consulter sans limite mais aussi pouvoir contribuer à les enrichir. Le web de données constitue un véritable outil d'interaction entre les musées et le public. Ce nouvel outil de communication ne viendra pas se substituer aux deux modes de diffusion traditionnels des musées que sont la présentation sur place des collections au public et la diffusion écrite (ouvrages et documentation papier) dont l'essor a été très important durant tout le 20ème siècle. Mais, à l'aube du 21ème siècle, l'informatisation et la diffusion des collections muséales sur le web sont devenues incontournables et constituent une nécessité pour les musées, s'ils souhaitent atteindre un public plus large y compris au niveau international. Cela leur permettra d'accroître leur visibilité et participer à la valorisation et l'essor de la culture française à l'étranger.

Ce nouveau mode de diffusion révolutionne l'organisation du travail au sein des musées et surtout l'approche documentaire sur les œuvres. Les langages documentaires, associés aux technologies du web sémantique, sont ainsi devenus des outils indispensables pour la valorisation des collections. L'utilisation des thésaurus, des vocabulaires contrôlés, des listes d'autorités et des référentiels permettent, en effet, une meilleure description des objets et favorisent une recherche facilitée pour les utilisateurs. Associés aux technologies du web sémantique, les domaines de connaissances sont mis en relation et enrichis. Les données deviennent porteuses de sens, peuvent être interprétées automatiquement par des machines et traduites dans des dizaines de langues favorisant ainsi l'utilisation de catalogues multilingues. Un accès unifié pour la recherche peut ainsi être proposé à l'utilisateur qui obtiendra ainsi plus facilement les réponses à ses questions sur les collections muséales mieux référencées.

Pour cela les musées doivent, afin de traiter les données sur les collections, suivre les préconisations du MCC, s'attacher à extraire des informations structurées à partir de l'analyse scientifique de l'œuvre, utiliser pour cela les thésaurus et l'ensemble des vocabulaires contrôlés et des référentiels qui leur sont proposés dans leur domaine de connaissance. Ils ne doivent pas hésiter à utiliser des thésaurus différents, les outils de gestion actuels et les technologies du web sémantique permettent de les mettre en relation et de les utiliser simultanément. Et si les thésaurus ne sont pas assez précis et ne répondent pas à la spécificité des collections, les musées doivent les enrichir, les compléter et ensuite en faire profiter les autres musées et l'ensemble des donnés culturelles, ce qui améliorera

encore l'accès des informations aux utilisateurs. Les musées peuvent ainsi répondre de manière efficace aux politiques culturelles du MCC qui découlent de ses deux grandes missions : conserver et diffuser la connaissance. Cette utilisation des langages documentaires est une première étape qui, associée aux technologies du web sémantique, permet de garantir les échanges et l'interopérabilité des données avec l'ensemble des données culturelles quelle que soit leur hétérogénéité. Voilà tout l'enjeu de la valorisation des collections muséales et de l'utilisation des langages documentaires dans le cadre du web de données.

Mais les musées ne pourront répondre à ces enjeux sans l'impulsion et l'appui du MCC. Celui-ci a une vision large des enjeux de la visibilité des données culturelles au niveau national et international, de l'état des lieux, des projets, des normes et des partenariats en cours en France et à l'étranger, des incidences juridiques, économiques et technologiques de la diffusion des données culturelles. Il peut ainsi coordonner et faire aboutir l'ensemble des projets qui sont encore à l'heure actuelle expérimentaux, proposer aux musées des applications concrètes, présenter aux éditeurs de logiciels des solutions techniques et préconiser à l'ensemble des institutions culturelles un ensemble de modèles et de normes.

Toutefois, une véritable coordination est à mettre en place afin de pouvoir faire le lien entre les projets extrêmement novateurs du MCC, les besoins de l'ensemble des institutions culturelles et les besoins particuliers des musées. Au MCC, donc, de bien communiquer, de proposer des solutions concrètes aux musées, et aux musées de faire remonter leurs besoins et leurs difficultés pour s'adapter à une réalité nouvelle, ouverte sur l'ensemble du monde culturel en France comme à l'international.

Cela ne pourra se faire sans une réflexion approfondie sur les moyens humains et financiers indispensables à la réalisation de ces projets. Le web de données implique une rationalisation et une simplification du traitement des données, les machines viennent remplacer le travail humain sur certains aspects. Mais, les projets, encore à l'état expérimental, doivent être développés et mis en application. Cela demande du temps et des moyens, une réflexion approfondie sur la gestion mutualisée de l'information, des collaborations et des partenariats à mettre en place, une nouvelle organisation du travail au sein des musées. D'autre part, les technologies du web de données ne sont pas faciles à appréhender, à utiliser et à mettre en place. Le personnel dans les musées a déjà énormément de tâches à réaliser et n'a pas forcément de temps à consacrer à ces problématiques qui peuvent lui sembler loin du terrain. Les musées doivent ainsi s'appuyer sur le personnel qualifié et compétent, pouvant faire le lien entre les différents corps de métiers et leurs besoins respectifs, entre les préconisations du MCC et les besoins des musées et comprendre l'enjeu de la valorisation des données culturelles.

Enfin, on ne peut conclure ce travail sans évoquer le problème de la pérennité des données, de la transmission de la connaissance aux générations futures. Il y a environ 15 000 ans, l'Homo sapiens nous a légué un patrimoine culturel sans précédent : qui d'entre nous n'a pas le cœur qui bat en admirant les fresques polychromes conservées dans la grotte préhistorique de Font-de-Gaume[73] ? Plus récemment, le roi Soleil nous a légué un patrimoine culturel extraordinaire : qui d'entre nous n'est pas ému en lisant un document annoté et signé de la main de Louis XIV présentant le parcours à suivre dans les jardins du château de Versailles afin d'en admirer toutes les merveilles ? Nos ancêtres ont utilisé les techniques de leur temps. À nous de nous interroger sur ce que nous léguerons aux générations futures. Nos données seront-elles encore accessibles dans 50 ans, dans 10 ans voire dans 5 ans ? C'est le défi que nous devons relever, charge à nous de nous assurer que nous mettons à profit toutes les techniques modernes et que celles-ci assurent de transmettre aux générations futures les richesses culturelles, dont nous avons hérité, en ajoutant nos contributions et notre connaissance à cette chaîne intergénérationnelle du savoir et de la création artistique.

[73] Site internet du Centre des monuments nationaux. Stes préhistoriques des Eyzies-de-Tayac. Grotte de Font-de-Gaume.
http://eyzies.monuments-nationaux.fr/fr/#details [consulté le 31/12/2013]

Bibliographie

Cette bibliographie analytique est arrêtée au 20 décembre 2013. Elle se compose de 27 références. Ce sont les références qui ont été les plus utilisées afin de préparer et de réaliser ce travail de mémoire. Cette liste n'est en rien exhaustive des références nécessaires pour couvrir tout le domaine et la problématique abordée.

Les références bibliographiques ci-dessous sont organisées par thème puis par ordre d'auteur.

Les thèmes sont organisés de la façon suivante :

Collections muséales

- État des lieux
- Normes

Langages documentaires

- Repères
- Élaboration d'un thésaurus

Web de données

- Éléments de compréhension
- Politiques culturelles
- Projets et utilisations

    Centre Pompidou virtuel

    Data.bnf.fr

    HADOC

    HdALab et Joconde Lab

    GINCO

    Sémanticpédia et Dbpedia

Collections muséales

État des lieux

[1] Ministère de la culture et de la communication. Mini-site des musées de France.
http://www.culturecommunication.gouv.fr/Disciplines-et-secteurs/Musees
[consulté le 02/12/2013]

Le mini-site des musées de France du MCC donne, sur cette page, des éléments de compréhension sur les collections, l'appellation « musée de France » et sur les missions d'un musée de France. Il propose également une rubrique actualité, une rubrique documentation regroupant les documents indispensables à connaître sur le sujet et une liste exhaustive des organismes liés aux musées de France.

[2] SALLOIS Jacques. Les musées de France. 4e éd., Paris, P.U.F. « Que sais-je ? », 2008, 128 pages. ISBN : 9782130567677.

L'auteur, ancien directeur des musées de France et directeur de la Commission de récolement des dépôts d'œuvres d'art (CRDOA), décrit, dans cet ouvrage, comment les musées de France sont un élément essentiel de notre système culturel, scientifique et pédagogique. Les musées de France ont connu depuis une vingtaine d'années une mutation sans précédent. L'auteur présente les nouvelles lois qui ont modifié leur statut juridique, les méthodes de conservation et d'enrichissement des collections, les initiatives pédagogiques et les évolutions de la muséographie qui séduisent et attirent les visiteurs.

Normes

[3] Ministère de la culture et de la communication. Service des musées de France. Bureau de la diffusion numérique. Portail des collections des musées de France Joconde. Méthode de rédaction informatisée des notices d'objets de musées (inventaire réglementaire et base documentaire). Mise en ligne : 23 juin 2005 ; mise à jour : 19 mars 2013.
http://www.culture.gouv.fr/documentation/joconde/fr/partenaires/AIDEMUSEES/inventaire-informatise.htm
[consulté le 02/12/2013]

Sur cette page du Portail des collections des musées de France, mis en ligne par le bureau de la diffusion numérique du service des musées de France du MCC, sont présentés la méthode d'inventaire informatisé des notices d'objets de musées, ainsi qu'un ensemble de

documents administratifs, méthodologiques et complémentaires, très utiles pour les professionnels.

[4] The International Council of Museums (ICOM). Mini site du CIDOC. Working groups. Data Harvesting and Interchange. LIDO overwiew.
http://network.icom.museum/cidoc/working-groups/data-harvesting-and-interchange/what-is-lido/
[consulté le 15/12/2013]

Le mini-site du CIDOC, mis en ligne par l'ICOM, expose, sur cette page, le projet LIDO et présente une vue d'ensemble des formats préalablement existants, la norme des CDWA Lite et la norme museumdat qu'il vient remplacer, ainsi que la norme SPECTRUM et la norme CIDOC-CRM.

Langages documentaires

Repères

[5] DALBIN Sylvie. Blog. Descripteurs. Site dédié aux thésaurus et autres vocabulaires contrôlés pour l'accès à l'information.
http://dossierdoc.typepad.com/descripteurs/
[consulté le 04/12/2013]

L'auteure, consultante en organisation et ingénierie documentaires au sein d'Assistance & Techniques Documentaires, présente, sur ce blog, tout ce qu'il faut savoir sur l'actualité liée aux thésaurus et autres vocabulaires contrôlés pour l'accès à l'information.

[6] MENON Bruno. « Les langages documentaires » Un panorama, quelques remarques critiques et un essai de bilan. Revue Documentaliste-Sciences de l'Information, 2007/1, vol. 44, p. 18-28. ISSN 0012-4508.

L'auteur, maître de conférences associé en science de l'information-documentation à l'Université Paris-8, revient, dans cet article, sur le développement et l'usage des langages documentaires : il examine les taxonomies et les ontologies, nouveaux outils de recherche d'information apparus avec l'explosion des technologies Internet. Il montre également que les langages documentaires peuvent s'imposer dans l'économie numérique.

[7] SERRES Alexandre. Introduction à l'Indexation. URFIST Bretagne-Pays de Loire, Janvier 2002. Support de cours. Dernière mise à jour : Septembre 2003.
http://www.sites.univ-rennes2.fr/urfist/Supports/Indexation/PlanCoursIndexation.html#INTRODUCTION%20A
[onsulté le 07/12/2013]

L'auteur, maître de conférences en sciences de l'information et de la communication et co-responsable de l'URFIST Bretagne-Pays de Loire, a élaboré et mis en ligne ce support de cours qui peut être librement exploité sous réserve de citer son origine. Ce support méthodologique introduit les problématiques de l'indexation (repères historiques, définition, fonction, complexité, évolutions actuelles) et présente les langages documentaires (définition, fonction, utilisation, typologie.

Élaboration d'un thésaurus

[8] DALBIN Sylvie, YAKOVLEFF Nathalie, ZYSMAN Hélène. Livre blanc. Norme ISO 25964-1 - Thésaurus pour la recherche documentaire. Avec la collaboration de François FEYLER, Roselyne BLOCH, Michèle LÉNARD, KATELL Briatte. Éditions Afnor, janvier 2013.
http://www.bivi.fonctions-documentaires.afnor.org/livres-blancs/livre-blanc-sur-la-norme-iso-25964-1-thesaurus-pour-la-recherche-documentaire-parue-en-version-anglaise-en-aout-2011
[téléchargeable en PDF, consulté le 06/12/2013]

Ce livre blanc présente de manière synthétique la dernière norme parue sur les thésaurus en version anglaise en août 2011. Ce guide expose les grandes lignes et les nouveautés d'un guide métier à l'attention de professionnels ayant déjà une connaissance des outils d'organisation que sont les thésaurus, les taxonomies, les classifications et les vocabulaires contrôlés en général. Il identifie les points importants de la norme et précise, quand c'est important pour la communauté francophone, leur caractère obligatoire.

[9] DEGEZ Danièle. Construire un thésaurus. Revue Archimag, Mars 2009, n°222, p 44.

L'auteure, ancienne présidente de l'ADBS et co-auteure de "Thésauroglossaire des langages documentaires" (ADBS, 2001), participe aux travaux de mise à jour des normes Iso sur le thesaurus. Elle présente, dans cet article, un certain nombre de règles et de bonnes pratiques pour la création d'un thésaurus.

[10] HUDON Michèle. Guide pratique pour l'élaboration d'un thésaurus documentaire. Avec la collaboration de Danièle Dégez et Dominique Ménillet. Revue Documentaliste, Dossier sur le web sémantique (Cairn), Québec, Les éditions ASTED, 2009, ISBN : 978-2-923563-17-6.

L'auteure, professeure à l'École de bibliothéconomie et des sciences de l'information (EBSI) de l'Université de Montréal, présente, dans cet ouvrage, une approche méthodologique de l'élaboration d'un thésaurus. Après avoir redéfini les fondamentaux du thésaurus comme lexique de descripteurs associé à une structure relationnelle, elle situe cet outil par rapport aux référentiels terminologiques comme les listes de vedettes matières ou à ceux qui sont utilisés sur le web : taxonomies et ontologies. L'auteure analyse la démarche nécessaire pour rédiger un cahier des charges avant de créer un thésaurus et propose une méthodologie de construction d'un thésaurus par étapes successives.

Web de données

Éléments de compréhension

[11] GANDON Fabien. Technologies et architecture du web de données. p. 27-30. In Dossier : Web sémantique web de données, quelle nouvelle donne ?. Dossier réalisé sous la direction de Muriel Amar (BnF) et Bruno Menon (Université Paris 8). Revue Documentaliste – Sciences de l'information, décembre 2011, vol. 48 : n°4, p. 20-61.
http://www.adbs.fr/revue-docsi-volume-48-n-4-decembre-2011-dossier-web-semantique-web-de-donnees-quelle-nouvelle-donne--111766.htm
[articles téléchargeables en PDF, consulté le 15/12/2013]

L'auteur, vice-responsable de l'équipe de recherche Edelweiss (Inria Sophia-Antipolis) et membre du W3C, présente, dans cet article, l'ensemble de standards et leurs fonctions dans un web de données en perpétuelle évolution depuis le début du web. L'ensemble du 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.

[12] RABAUT Hélène, ZYSMAN Hélène. Les schémas de concepts et le Web sémantique : la norme sur les thésaurus ISO 25964 et le Web sémantique. Journée d'étude AFNOR CG 46 / BnF. « Référentiels et données d'autorité à l'heure du Web sémantique ». 27 mai 2011.
http://www.bnf.fr/documents/afnor2011_norme_thesaurus.pdf
[consulté en ligne le 05/12/2013]

Hélène Rabault est chef de projet chez Lexis-Nexis (édition professionnelle) et Hélène Zysman est consultante fonctionnelle GED/ECM, Smile (Open Source Solutions). Les auteures exposent, dans cette présentation réalisée lors d'une journée d'étude organisée conjointement par l'AFNOR et la BnF, la norme sur les thésaurus ISO 25964 dans le cadre du web sémantique, en donnent quelques repères, insistent sur les principes d'interopérabilité et replace l'utilisation du langage SKOS dans la publication des thésaurus sur le web.

[13] SAJUS Bertrand. Repères sur le web sémantique / Web de données. INTD, 5 avril 2013.
http://fr.slideshare.net/Be_Sa/reperes-web-semantiqueintd20130405
[consulté le 15/12/2013]

L'auteur, du département des programmes numériques du MCC, expose, dans cette présentation réalisée lors d'une journée organisée par le CNAM-INTD à l'attention des professionnels, les enjeux, les fonctionnalités et les technologies liées au web sémantique et aux web de données.

[14] W3C Incubator Group Report. Library Linked Data Incubator Group Final Report. 25 October 2011.
http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/
[consulté le 15/12/2013]

Le site officiel de l'incubateur du W3C propose, sur cette page, leur rapport final mis en ligne en 2011, et présente les bénéfices de l'approche du web de données, fait un état des lieux sur la situation et propose quelques recommandations notamment pour les collections.

Politiques culturelles

[15] Net : Lab. Rapport sur l'état des lieux de l'*Open Data*. Secteur de la culture.
http://opendata-culture.net-lab.fr/
[consulté le 05/12/2013]

Ce rapport en ligne explique ce qu'est l'open data, ce que change l'open data pour la culture et les enjeux, ainsi que le cadre juridique particulier des données culturelles. Net:lab est une plate-forme ouverte de débat et de collaboration sur le thème Internet et Société rassemblant des experts de la société civile, du secteur académique et du secteur privé.

[16] 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.
http://www.culture.gouv.fr/culture/editions/r-cr/cr118-119.pdf
[consulté en ligne le 15/12/2013]

Ce numéro, mis en ligne par le MCC, est consacré à la politique de numérisation du patrimoine culturel et présente les enjeux de la numérisation ainsi que de la valorisation des collections numériques pour les institutions culturelles. Il s'appuie sur de nombreux exemples de projets en cours en 2008 en France et au niveau européen.

[17] 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.
http://cblog.culture.fr/wp-content/uploads/2013/07/Feuille-de-route-open-data-MCC.pdf
[consulté le 17/12/2013]

Ce document, mis en ligne par le MCC, présente la feuille de route ministérielle concernant l'ouverture et le partage des données culturelles publiques.

Projets

Centre Pompidou Virtuel

[18] BERMÈS Emmanuelle. Les parcours de sens dans le Centre Pompidou virtuel. Bulletin des bibliothèques de France, 2013, n°5, p52-57.

L'auteure, chef du service multimédia au Centre Pompidou et membre active du W3C, explique ,dans cet article, comment, en 2012, le Centre Pompidou a renouvelé son site web sous la forme d'une plate-forme de diffusion de contenus numériques en ligne, donnant accès à travers un point d'entrée unique, à l'ensemble de la production numérique de l'institution : œuvres d'art numérisées, documents sur l'art et l'histoire de l'art, enregistrements sonores, vidéos, archives, notices bibliographiques, etc. L'auteure explique les enjeux du projet, la mise en place et les contraintes et pose un premier bilan.

Data.bnf.fr

[19] BnF. data.bnf.fr. Juillet 2011.
http://data.bnf.fr/about
[consulté le 09/12/2013]

Le site data.bnf.fr, présente, sur cette page, le projet data.bnf.fr, les chiffres clés et des informations sur les technologies utilisées dans le but d'informer le grand public. Cette page présente également des liens vers des documents complémentaires et permet de suivre les communications de la BNF sur le projet.

GINCO

[20] Ministère de la culture et de la communication. Le projet GINCO. Paris.
www.culturecommunication.gouv.fr/Etudes-et-documentation/HADOC/Referentiels2/Les-vocabulaires-scientifiques-et-techniques/Le-projet-GINCO
[consulté le 10/12/2013]

Cette page, sur le site de MCC, présente l'ensemble du projet HADOC et expliquent notamment en détail le projet GINCO.

HADOC

[21] 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.
http://blogpeda.ac-poitiers.fr/vocabnomen/files/2010/05/1-Harmoniser-les-donnees-culturelles_K.-BRIATTE_20100504.pdf
[consulté le 02/12/2013]

Dans ce document, Katell Briatte, chef de projet maîtrise d'ouvrage au département des systèmes d'information patrimoniaux du MCC, présente le point de départ du projet HADOC, ses problématiques ainsi que les solutions proposées.

[22] Ministère de la culture et de la communication. Présentation du programme [HADOC]. Paris : 2 juillet 2013.
http://www.culturecommunication.gouv.fr/index.php/Etudes-et-documentation/HADOC/Presentation-du-Programme/Presentation-du-programme
[consulté le 02/12/2013]

Cette page, sur le site de MCC, présente l'ensemble du projet HADOC et détaille les différentes briques du projet : le modèle de données, les référentiels, les normes et l'ouverture des données.

HdaLab et Joconde Lab

[23] Portail HdALab.
http://hdalab.iri-research.org/hdalab/
[consulté le 15/12/2013]

Le portail HdALab, mis en ligne par le MCC et l'IRI, présente de manière officielle le projet HdALab.

[24] Ministère de la culture et de la communication. Service des musées de France. Bureau de la diffusion numérique. Portail des collections des musées de France.
http://portail-joconde.over-blog.com/article-jocondelab-premiers-pas-de-joconde-sur-le-web-semantique-119255498.html
[consulté le 15/12/2013]

Le blog du Portail des collections des musées de France présente, dans l'article du 26 juillet 2013, tous les enjeux du projet JocondeLab.

Projet de référentiel des Acteurs historiques

[25] BELKOFF Elodie. Contribution à l'élaboration d'un référentiel d'acteurs historiques dans l'écosystème du web sémantique. Mémoire de Master 2. École du Louvre. Septembre 2013.

L'auteure, étudiante à l'École du Louvre, a effectué un stage auprès de Katell Briatte, chef de projet maîtrise d'ouvrage au département des systèmes d'information patrimoniaux du MCC. Elle présente, dans son mémoire de stage, le projet de référentiel des acteurs historiques, la problématique de sa mise en place et les difficultés rencontrées.

Semanticpedia et Dbpedia

[26] Blog de Sémanticpédia.
http://www.semanticpedia.org/
[consulté le 15/12/2013]

Le blog officiel de Sémanticpédia présente la plate-forme de collaboration Sémanticpédia, les partenaires, les projets en cours et à venir.

[27] Site internet de DBpedia.
http://dbpedia.org/Internationalization
[consulté le 15/12/2013]

Le blog officiel de Dbpedia international fournit des informations sur le projet DBpedia et sur les problématiques rencontrées par les versions locales de DBpedia. Il présente les problématiques du projet, propose un guide de développement, des contacts dans différents pays, les différents langages et un ensemble de documentations sur le sujet.

Annexes

Annexe 1 - Liste d'œuvres de l'exposition « Rodin, les origines d'un génie », Legnano, 2010

Fiche d'identité de l'œuvre											Description du processus de création		Description iconographie		
N° d'inv.	Domaine	Dénomination	Appellation	Titre indexé (sujets)	Titre	Auteur	Nature des matériaux	Technique	Date	Genèse	Thématique	Phase de conception	Description par mots-clés	Personnage représenté	Lieu représenté
S.06411	Sculpture	sculpture, ronde bosse	statue		Les Trois ombres GM	Rodin	bronze	fonte à la cire perdue	1881	grand modèle	Ombre		groupe de figures (3) ; homme ; debout ; en pied ; nu ; avec main		
S.00161	Sculpture	sculpture, ronde bosse	statue		Le Penseur GM	Rodin	plâtre		1903	grand modèle	Penseur		figure ; homme ; assis ; en pied ; monumental		
S.01302	Sculpture	sculpture, ronde bosse	statue		Eve	Rodin	bronze	fonte au sable	1881	grand modèle	Eve		figure ; femme ; debout ; en pied ; bras replié ; tête penchée	*Eve*	
S.00962	Sculpture	sculpture, ronde bosse			Adam	Rodin	bronze	fonte au sable	1880		Adam		homme ; homme ; debout ; en pied ; avec main	*Adam*	
D.00052	Dessin	dessin			Porteuses d'offrandes d'après la frise du Parthénon	Rodin	papier calque crème	crayon au graphique ; plume encre brune	Avt 1870?		Antique	Jeunesse-Formation	scène antique ; groupe de figuress (7) ; debout ; offrande ; frise		Parthénon
P.07214	Peinture	peinture			Chaland au bord du canal de Willebroeck	Rodin	huile sur carton		1871-1877		Paysage	En Belgique	paysage ; canal ; chaland		
P.07215	Peinture	peintre			Une péniche au bord du canal de Willebroek	Rodin	huile sur carton		1871-1877		Paysage	En Belgique	paysage ; péniche ; canal		Belgique ; Willebroeck ;
D.00063	Dessin	dessin			Scène orientale	Rodin	papier vélin	aquarelle ; gouache	Avt 1870?	étude	Orient	Jeunesse-Formation	paysage proche d'une ville ; groupe de figures		Orient
D.00070	Dessin	dessin			Personnage, étude d'après l'antique	Rodin	papier vélin ; papier vergé	aquarelle ; crayon au graphite	Avt 1870?		Antique	Jeunesse-Formation	d'après l'antique ; figure ; personnage ; de profil		

Fiche d'identité de l'œuvre											Description du processus de création		Description iconographie		
N° d'inv.	Domaine	Dénomination	Appellation	Titre indexé (sujets)	Titre	Auteur	Nature des matériaux	Technique	Date	Genèse	Thématique	Phase de conception	Description par mots-clés	Personnage représenté	Lieu représenté
D.00071	Dessin	dessin			Paysage de Belgique	Rodin	papier vélin ; carton	crayon ; crayon au graphite	1871-1877?		Paysage	En Belgique	paysage ; nature ; arbre		Belgique
D.00088	Dessin	dessin			Cortège d'après la frise du Parthénon	Rodin	papier crème		Avt 1870?		Antique	Jeunesse-Formation	scène mythologique ; d'après l'antique ; chevaux ; personnages ; cortège ; frise		Parthénon
P.07217	Peinture	peinture			Ferme dans son verger vers Groenendael à l'entrée de la forêt de Soignes	Rodin	papier collé sur carton	huile	1871-1877		Paysage	En Belgique	paysage ; ferme ; forêt		Belgique ; Groenendael ; Soignes
P.07224	Peinture	peinture			Le coin du balai en forêt de Soignes	Rodin	papier collé sur carton		1871-1877		Paysage	En Belgique	paysage		Belgique ; Soignes
S.00971	Sculpture	sculpture, ronde bosse			Jean-Baptiste Rodin	Rodin	bronze	fonte au sable	1860		Portrait de proches	Jeunesse-Formation		Jean-Baptiste Rodin	
P.07226	Peinture	peinture			Coin de hameau vers Boitsfort en forêt de Soignes	Rodin	papier collé sur carton	huile	1871-1877		Paysage	En Belgique	hameau ; forêt ; paysage		Belgique ; Boitsfort ; Soignes
S.01720	Sculpture	sculpture, ronde bosse	statue		Le Père Eymard	Rodin	plâtre teinté		1863	grand modèle	Portrait de proches	Jeunesse-Formation	figure	Pierre-Julien Eymard	
S.00212	Sculpture	sculpture, ronde bosse	buste		Jeune fille au chapeau fleuri de marguerites	Rodin	terre cuite		1865-1870			Jeunesse-Formation	portrait ; jeune fille ; en buste ; fleur ; chapeau		
S.00496	Sculpture	sculpture, ronde bosse	tête		L'Homme au nez cassé	Rodin	bronze	fonte au sable	1864	type II, 2ème modèle	Homme au nez cassé	Jeunesse-Formation	tête ; homme ; nez cassé		
S.00974	Sculpture	sculpture, ronde bosse	buste avec épaule		L'Homme au nez cassé	Rodin	marbre		1874-75		Homme au nez cassé	Jeunesse-Formation	en buste ; homme ; nez cassé		

Fiche d'identité de l'œuvre											Description du processus de création		Description iconographie		
N° d'inv.	Domaine	Dénomination	Appellation	Titre indexé (sujets)	Titre	Auteur	Nature des matériaux	Technique	Date	Genèse	Thématique	Phase de conception	Description par mots-clés	Personnage représenté	Lieu représenté
D.00152	Dessin	dessin			Cavalier au galop	Rodin	papier vélin	crayon au graphite ; lavis d'encre	Avt 1870?		Cheval	Jeunesse-Formation	cavalier ; cheval ; galop		
D.00160 à 00175	Dessin	dessin			Feuille de croquis d'après MA & monuments italiens	Rodin	papier	crayon au graphite ; lavis d'encre	1875-1876		Regarder les maîtres	Voyage en Italie			
D.00176	Dessin	dessin			Scène orientale au cheval cabré	Rodin	papier vélin ; papier vergé	plume-encre	avant 1870?		Cheval	Jeunesse-Formation	cavalier ; cheval		Orient
D.00177	Dessin	dessin			Scène orientale avec des chevaux et un âne	Rodin	papier vergé	crayon au graphite ; lavis d'encre	avt 1870?		Cheval	Jeunesse-Formation	cavalier ; cheval ; âne		Orient
D.00200 à 00204	Dessin	dessin			Projets de monuments/feuille de croquis	Rodin	plume et encre sur papier				Regarder les maîtres	Voyage en Italie			
D.00223	Dessin	dessin			Scène orientale	Rodin	papier	gouache ; lavis d'encre	Avt 1870?		Orient	Jeunesse-Formation	groupe de figuress ; homme ; femme ; enfant		Orient
D.00224	Dessin	dessin			Paysage oriental	Rodin	papier vélin	aquarelle ; pierre noire	Avt 1870		Orient	Jeunesse-Formation			Orient
D.00296	Dessin	dessin			Copie d'après l'Enfance de Bacchus de Poussin	Rodin	papier vélin	fusain ; gouache		copie		Jeunesse-Formation	enfance ; groupe de figures	Poussin ; *Bacchus*	
D.00311 à 00323	Dessin	dessin			Feuille de croquis d'après Donatello & monuments italiens	Rodin	papier découpé ; papier cartonné	crayon au graphite ; gouache	1875-1876	feuille de croquis	Regarder les maîtres	Voyage en Italie			
D.00400	Dessin	dessin			Cavalier	Rodin	papier vélin/réglé ; papier filigrané	crayon ; gouache	ca. 1880?		Cheval	Jeunesse-Formation	cavalier ; cheval ; homme		

Fiche d'identité de l'œuvre											Description du processus de création		Description iconographie		
N° d'inv.	Domaine	Dénomination	Appellation	Titre indexé (sujets)	Titre	Auteur	Nature des matériaux	Technique	Date	Genèse	Thématique	Phase de conception	Description par mots-clés	Personnage représenté	Lieu représenté
D.00405	Dessin	dessin			Homme retenant un cheval cabré	Rodin	papier vélin	crayon	vers 1880		Cheval	Jeunesse-Formation	cheval ; homme		
D.00414	Dessin	dessin			Cavalier de profil	Rodin	papier vélin/réglé	lavis d'encre ; plume-encre	vers 1880		Cheval	Jeunesse-Formation	cavalier ; cheval ; homme ; de profil		
D.00416	Dessin	dessin			Courses de chevaux	Rodin	papier calque ; papier vergé	plume-encre	Avt 1880?		Cheval	Jeunesse-Formation	course ; cavalier		
P.07227	Peinture	peinture			L'étang de la patte d'oie à Groenendae en forêt de Soignes	Rodin	papier collé sur carton	huile	1871-1877		Paysage	En Belgique	paysage ; arbre ; forêt ; étang		Belgique ; Groenendae ; Soignes
P.07228	Peinture	peinture			Vieille chaumière à flanc de coteau à Boitsfort en forêt de Soignes	Rodin	papier collé sur carton	huile	1871-1877		Paysage	En Belgique	paysage ; chaumière ; forêt		Boitsfort ; Soignes
P.07230	Peinture	peinture			Académie d'homme vu de dos	Rodin	toile	huile	1855-60 ?		Académie	Jeunesse-Formation	homme ; de dos		
P.07231	Peinture	peinture			Académie d'homme vu de face	Rodin	toile	huile	1855-60 ?		Académie	Jeunesse-Formation	homme ; de face		
P.07234	Peinture	peinture			Fûtaie sur taillis de chênes et de bouleaux dans la forêt de Soignes	Rodin	carton	huile	1871-1877		Paysage	En Belgique	paysage ; chêne ; bouleau ; forêt		Belgique ; Soignes
P.07237	Peinture	peinture			Portrait de Rose Beuret	Rodin	toile	huile	1865		Portrait de proches	Jeunesse-Formation	portrait ; femme	Rose Beuret	
D.00424	Dessin	dessin			Projet de monument	Rodin	papier vélin	lavis d'encre ; plume-encre	v. 1880 ?		Concours	Retour à Paris	monument ; architecture		

Fiche d'identité de l'œuvre											Description du processus de création		Description iconographie		
N° d'inv.	Domaine	Dénomination	Appellation	Titre indexé (sujets)	Titre	Auteur	Nature des matériaux	Technique	Date	Genèse	Thématique	Phase de conception	Description par mots-clés	Personnage représenté	Lieu représenté
P.07238	Peinture	peinture			Les maisons rouges au bord de l'eau vers Boisfort en forêt de Soignes	Rodin	papier collé sur toile	huile	1871-1877		Paysage	En Belgique	paysage ; maison		Belgique ; Boisfort ; Soignes ;
P.07240	Peinture	peinture			Chemin de terre à Watermael en forêt de Soignes	Rodin	papier collé sur carton	huile	1871-1877		Paysage	En Belgique	paysage ; chemin		Belgique ; Watermael ; Soignes
P.07241	Peinture	peinture			Cheval	Rodin	carton	huile	1864		Cheval	Jeunesse-Formation	cheval ; de profil		
P.07242	Peinture	peinture			Sortie de la forêt de Soignes vers Auderghem	Rodin	papier collé sur carton	huile	1871-1877		Paysage	En Belgique	paysage ; forêt ; village		Belgique ; Soignes ; Auderghem ;
P.07243	Peinture	peinture			Le vieux Boitsfort en forêt de Soignes	Rodin	papier collé sur carton	huile	1871-1877		Paysage	En Belgique	paysage ; forêt ; arbre ; nuage		Boitsfort ; Soignes ; Belgique ;
P.07244	Peinture	peinture			Copie d'après le Coup de lance de Rubens	Rodin	papier collé sur toile	fusain	1871-1877	copie	Regarder les maîtres	Voyage en Italie	crucifixion ; lance ; croix ; homme	Rubens	
S.00294	Sculpture	sculpture, ronde bosse			Vénus et l'amour	Rodin	terre cuite		1871		Jeune femme et enfant	En Belgique	amour		
S.00295	Sculpture	sculpture, ronde bosse			Vénus	Rodin	terre cuite		1871			En Belgique	nymphe ; assise ; se peignant		
S.00961	Sculpture	sculpture, ronde bosse	buste		Suzon	Rodin August e	bronze	fonte au sable	1875		Suzon	En Belgique			
S.00355	Sculpture	sculpture, ronde bosse			Secret d'amour	Rodin	terre cuite ; bois		1871		Petite sculpture décorative ; Jeune femme et enfant	En Belgique	femme ; assise ; enfant ; amour		
S.06672	Sculpture	sculpture, ronde bosse	buste		Orpheline alsacienne	Rodin	marbre		1871		Petite sculpture décorative	En Belgique	jeune fille ; en buste		

Fiche d'identité de l'œuvre											Description du processus de création		Description iconographie		
N° d'inv.	Domaine	Dénomination	Appellation	Titre indexé (sujets)	Titre	Auteur	Nature des matériaux	Technique	Date	Genèse	Thématique	Phase de conception	Description par mots-clés	Personnage représenté	Lieu représenté
S.01193	Sculpture	sculpture, ronde bosse			La Toilette de Vénus	Rodin	terre cuite		1871		Petite sculpture décorative ; Jeune femme et enfant	En Belgique	femme ; assise ; enfant ; amour		
S.01665	Sculpture	sculpture, ronde bosse			La Lorraine	Rodin	plâtre		1874		Petite sculpture décorative	En Belgique			Lorraine
S.02491	Sculpture	sculpture, ronde bosse	buste		Jeune fille aux écoutes	Rodin	plâtre		V. 1875 ?		Petite sculpture décorative	En Belgique	jeune fille ; en buste		
S.02042	Sculpture	sculpture, ronde bosse			Jeune femme et enfant	Rodin	plâtre		v. 1875 ?		Petite sculpture décorative ; Jeune femme et enfant	En Belgique	femme ; assise ; enfant ; amour		
S.01441	Sculpture	sculpture, ronde bosse	buste		Flora	Rodin	plâtre		V. 1870-75 ?		Petite sculpture décorative	En Belgique	femme ; en buste	*Flora*	
S.02321	Sculpture	sculpture, ronde bosse			Enfants s'embrassant (variante)	Rodin	plâtre patiné		v. 1875-80 ?		Petite sculpture décorative	En Belgique	enfant ; amour		
S.01047	Sculpture	sculpture, ronde bosse			Dosia	Rodin	terre cuite		1874		Petite sculpture décorative	En Belgique		*Dosia*	
S.05822	Sculpture	sculpture, ronde bosse	buste		Diane	Rodin	marbre		1875		Petite sculpture décorative	En Belgique	femme ; en buste	*Diane*	
S.00283	Sculpture	sculpture, ronde bosse			Amour debout	Rodin	terre cuite		1871		Petite sculpture décorative	En Belgique	amour		
S.00429	Sculpture	sculpture, ronde bosse	buste		Mme Cruchet	Rodin	terre cuite		1878 £		Portrait	En Belgique	femme ; en buste	Mme Cruchet	
S.06671	Sculpture	sculpture, ronde bosse	buste		Mme Almire Huguet	Rodin	terre cuite		1874 £		Portrait	En Belgique	femme ; en buste	Mme Almire Huguet	
S.02130	Sculpture	sculpture, ronde bosse	buste		Henry Thorion	Rodin	terre cuite		v. 1875-1876 ?		Portrait	En Belgique	homme ; en buste	Henry Thorion	
S.01455	Sculpture	sculpture, ronde bosse			Docteur Jules-Adrien Thiriar	Rodin	plâtre		av 1874		Portrait	En Belgique	homme	Docteur Jules-Adrien Thiriar	

Fiche d'identité de l'œuvre											Description du processus de création		Description iconographie		
N° d'inv.	Domaine	Dénomination	Appellation	Titre indexé (sujets)	Titre	Auteur	Nature des matériaux	Technique	Date	Genèse	Thématique	Phase de conception	Description par mots-clés	Personnage représenté	Lieu représenté
S.00447	Sculpture	sculpture, ronde bosse			Alexandre Van Berckelaer	Rodin	plâtre		v. 1875 ?		Portrait	En Belgique	homme	Alexandre Van Berckelaer	
S.05707	Sculpture	sculpture, ronde bosse			Cariatide	Rodin	stuc peint		1874		Décoration monumentale ; Cariatide	En Belgique			bd d'Anspach ; Bruxelles
S.05714	Sculpture	sculpture, ronde bosse			Atlante, coude gauche tourné vers l'extérieur	Rodin	Stuc peint		1874		Décoration monumentale ; Cariatide	En Belgique	Atlante		bd d'Anspach ; Bruxelles
S.05711	Sculpture	sculpture, ronde bosse			Atlante, coude droit vers l'extérieur	Rodin	stuc peint		1874		Décoration monumentale ; Cariatide	En Belgique	Atlante		bd d'Anspach ; Bruxelles
S.01220	Sculpture	sculpture, ronde bosse			Ugolin assis	Rodin	plâtre		V. 1875		Regarder les maîtres		homme ; assis	*Ugolin*	
S.02682 Co.3498	Sculpture	sculpture, ronde bosse	piédestal ; vasque		Jardinière aux Titans	Rodin	terre cuite		1878		Regarder les maîtres		homme ; assis ; Titans		
P.07246	Peinture	peinture			Copie d'après le Mariage par procuration	Rodin	huile sur toile		1871-1877	copie	Regarder les maîtres		mariage ; groupe de figuress	Marie de Médicis ; Rubens	
D.00449	Dessin	dessin			Homme nu de dos contre un piédestal et dressant une couronne	Rodin	papier vélin	lavis d'encre ; plume-encre	v. 1880 ?		Concours	Retour à Paris	homme ; nu		
D.05104	Dessin	dessin			Académie d'homme, un bâton à la main	Rodin	papier vergé ; filigrané	fusain-estompe	1854-57 ?		Académies	Jeunesse-Formation	homme ; debout ; nu ; bâton		
D.05105	Dessin	dessin			Académie d'homme barbu	Rodin	papier vergé ; filigrané	fusain-estompe	1854-57 ?		Académies	Jeunesse-Formation	homme ; debout ; nu		
D.05106	Dessin	dessin			Académie d'homme barbu	Rodin	papier vergé ; filigrané	fusain-estompe	1854-57 ?		Académies	Jeunesse-Formation	homme ; debout ; nu		

Fiche d'identité de l'œuvre											Description du processus de création		Description iconographie		
N° d'inv.	Domaine	Dénomination	Appellation	Titre indexé (sujets)	Titre	Auteur	Nature des matériaux	Technique	Date	Genèse	Thématique	Phase de conception	Description par mots-clés	Personnage représenté	Lieu représenté
D.05114	Dessin	dessin			Académie d'homme barbu	Rodin	papier vergé ; filigrané	fusain-estompe	1854-57 ?		Académies	Jeunesse-Formation	homme ; debout ; nu ; bâton		
P.07247	Peinture	peinture			Jean-Baptiste Rodin	Rodin	toile	huile	1860		Portrait de proches	Jeunesse-Formation	homme ; en buste	*Jean-Baptiste Rodin*	
S.00175	Sculpture	sculpture, ronde bosse			Saint Jean-Baptiste	Rodin	plâtre patiné		1880		Grandes figures ; St Jean-Baptiste	Retour à Paris	homme ; nu ; bras levé	*St Jean-Baptiste*	
S.00468	Sculpture	sculpture, ronde bosse			L'Age d'airain	Rodin	bronze	fonte au sable	1877	grand modèle	Grandes figures ; Âge d'airain	Retour à Paris			
S.02205	Sculpture	sculpture, ronde bosse			Etude pour St Jean-Baptiste	Rodin	plâtre		1878	étude	Grandes figures ; St Jean-Baptiste	Retour à Paris	homme ; nu ; debout	*St Jean-Baptiste*	
S.06670	Sculpture	sculpture, ronde bosse	buste		Buste de saint Jean Baptiste	Rodin	bronze	fonte au sable	1879		Grandes figures ; St Jean-Baptiste	Retour à Paris			
Ph.00270	Photographie	photographie			Auguste Neyt	G. Marconi	papier albuminé		1877	modèle âge d'airain	Grandes figures ; Âge d'airain	Retour à Paris	homme ; nu ; debout ; de face	Auguste Neyt	
Ph.00272	Photographie	photographie			Auguste Neyt	G. Marconi	papier albuminé		1877	modèle âge d'airain	Grandes figures ; Âge d'airain	Retour à Paris	homme ; nu ; debout ; de dos	Auguste Neyt	
Ph.00278	Photographie	photographie			St Jean-Baptiste dans l'atelier	Braun ?	papier albuminé		v. 1881 ?	maquette porte de l'enfer	Concours ; St Jean-Baptiste ; Porte de l'enfer	Retour à Paris	modèle ; homme ; nu ; debout ; de dos		
Ph.00281	Photographie	photographie			La Défense devant la Porte de l'enfer	Druet	épreuve gélatinoargentique		v. 1898		Concours ; Porte de l'enfer	Retour à Paris			
D.05115	Dessin	dessin			Etude d'après le Narcisse de Valerio Cioli	Rodin	fusain sur papier crème		v. 1877 ?	étude	Regarder les maîtres	Voyage en Italie		ValérioCioli	
S.02279	Sculpture	sculpture, ronde bosse	buste		Projet de monument	Rodin	terre cuite		v. 1880 ??	projet	Concours	Retour à Paris	homme ; en buste		
S.02523	Sculpture	sculpture, ronde bosse	torse et tête		La Défense, torse du génie avec tête	Rodin	plâtre		1879		Concours ; Génie ; Défense	Retour à Paris			

Fiche d'identité de l'œuvre											Description du processus de création		Description iconographie		
N° d'inv.	Domaine	Dénomination	Appellation	Titre indexé (sujets)	Titre	Auteur	Nature des matériaux	Technique	Date	Genèse	Thématique	Phase de conception	Description par mots-clés	Personnage représenté	Lieu représenté
S.03078	Sculpture	sculpture, ronde bosse	tête		La Défense, tête du génie ailé, petit modèle	Rodin	plâtre		1879	petit modèle	Concours ; Génie	Retour à Paris			
S.00702	Sculpture	sculpture, ronde bosse	tête		La Défense, tête du guerrier	Rodin	plâtre		1879		Concours ; Défense ; Guerrier	Retour à Paris			
Ph.00314	Photographie	photographie			Buste de Mme Roll dans l'atelier	V. Panneli er	tirage argentique		1883		Relations amicales	Retour à Paris			
Ph.00667	Photographie	photographie			Bellone dans l'atelier, avec St Jean-Baptiste et l'Homme au nez cassé	Michele z	Papier albuminé - retouché		v. 1881 ?		Concours	Retour à Paris			
S.00469	Sculpture	sculpture, ronde bosse			La Défense	Rodin	bronze		1879		Concours	Retour à Paris			
D.05117	Dessin	dessin			Etude d'après L'Aurore de Michel Ange	Rodin	fusain sur papier crème		v. 1877 ?	étude	Regarder les maîtres, Aurore	Voyage en Italie		Michel Ange	
Ph.02153	Photographie	photographie			L'Age d'airain	G. Marcon i			1877		Grandes figures ; Âge d'airain ; Génie	Retour à Paris			
Ph.02264	Photographie	photographie			La Défense, génie ailé seul	Fiorillo	papier albuminé		1879 ?		Concours ; Défense ; Génie	Retour à Paris			
S.01836	Sculpture	sculpture, ronde bosse	buste		Bellone (sans piédouche)	Rodin	plâtre patiné		1881		Concours ; Bellone	Retour à Paris	figure ; en buste	*Bellone*	
S.02836	Sculpture	sculpture, ronde bosse	buste		Bellone, version au casque avec putti	Rodin	plâtre				Concours ; Bellone	Retour à Paris	figure ; en buste ; casque	*Bellone*	
S.06750	Sculpture	céramique	vase		Vase Shanghai "Femme et enfant"	Sèvres et Rodin	porcelaine de Sèvres	émail	1881-82		Sèvres	Retour à Paris	femme ; enfant		

Fiche d'identité de l'œuvre											Description du processus de création		Description iconographie		
N° d'inv.	Domaine	Dénomination	Appellation	Titre indexé (sujets)	Titre	Auteur	Nature des matériaux	Technique	Date	Genèse	Thématique	Phase de conception	Description par mots-clés	Personnage représenté	Lieu représenté
S.06751	Sculpture	céramique	vase		Vase Shanghai "Faune et enfant"	Rodin et Sèvres	porcelaine de Sèvres	émail	1881-82		Sèvres	Retour à Paris	faune ; enfant		
S.02416	Sculpture	sculpture, ronde bosse	vase		Seau de Pompéï "La Nuit"	Sèvres et Rodin	porcelaine de Sèvres		1881-82		Sèvres	Retour à Paris			Pompéï
S.02417	Sculpture	sculpture, ronde bosse	plaquette		Plaquette "Fillette et enfant"	Sèvres et Rodin	porcelaine de Sèvres		1881-82		Sèvres	Retour à Paris	fillette ; enfant ; frère ; sœur		
S.02490	Sculpture	sculpture, ronde bosse			Le Lion qui pleure	Rodin	plâtre patiné		1881		Relations amicales	Retour à Paris	lion		
Ph.05708	Photographie	photographie			Bellone dans l'atelier, avec St Jean-Baptiste et l'Homme au nez cassé	Michelez	papier albuminé		v. 1881 ?		Concours ; Bellone ; St Jean-Baptiste ; Homme au nez cassé	Retour à Paris	homme ; nu	*Bellone ; Saint Jean-Baptiste*	
S.00985	Sculpture	sculpture, ronde bosse			Buste de Mme Roll	Rodin	marbre		1882-83		Relations amicales	Retour à Paris	femme ; en buste	Mme Roll	
S.00158	Sculpture	sculpture, ronde bosse			Buste de Jean-Paul Laurens	Rodin	pâte à modeler ; plâtre	gomme laqué ; patine	1882		Relations amicales	Retour à Paris	homme ; en buste	Jean-Paul Laurens	
S.01981	Sculpture	sculpture, ronde bosse			Buste de Carrier-Belleuse	Rodin	terre cuite		1882		Relations amicales	Retour à Paris	homme ; en buste	Albert-Ernest Carrier-Belleuse	
S.01189	Sculpture	sculpture, relief			La Porte de l'Enfer (3ème maquette)	Rodin	plâtre		1880	3ème maquette	Porte de l'enfer				
S.00457	Sculpture	sculpture, relief			La Porte de l'Enfer (2ème maquette)	Rodin	plâtre		1880	2ème maquette	Porte de l'enfer				
S.02392	Sculpture	sculpture, ronde bosse	groupe sculpté		Ugolin	Rodin	plâtre			élément porte de l'enfer	Porte de l'enfer ; Ugolin dévorant ses enfants		homme ; enfant ; nu ; à genoux ; grimaçant	*Ugolin* ; Dante	

Fiche d'identité de l'œuvre											Description du processus de création		Description iconographie		
N° d'inv.	Domaine	Dénomination	Appellation	Titre indexé (sujets)	Titre	Auteur	Nature des matériaux	Technique	Date	Genèse	Thématique	Phase de conception	Description par mots-clés	Personnage représenté	Lieu représenté
S.01223	Sculpture	sculpture, ronde bosse	torse		Torse d'Adèle	Rodin	plâtre		1878	élément porte de l'enfer	Porte de l'enfer ; Torse d'Adèle ; Eternel Printemps		femme ; nu	Adèle Abruzzesi	
S.02521	Sculpture	sculpture, ronde bosse			Penseur sur élément de châpiteau	Rodin	plâtre		1881-1882	élément Porte de l'Enfer	Porte de l'enfer ; Penseur ; Poète		homme ; nu ; assis ; tête appuyée sur la main ; méditation		
S.00963	Sculpture	sculpture, ronde bosse			L'Homme qui tombe	Rodin	bronze			élément porte de l'enfer	Porte de l'enfer ; Homme qui tombe		homme ; nu		
S.00989	Sculpture	sculpture, ronde bosse			L'Eternel Printemps	Rodin	bronze	fonte au sable	1884	élément porte de l'enfer	Porte de l'enfer ; Eternel Printemps		couple ; amants		
S.02834	Sculpture	sculpture, ronde bosse			Le Baiser	Rodin	plâtre patiné			élément porte de l'enfer	Porte de l'enfer ; Baiser		couple ; amants	*Paolo et Francesca* ; Dante	
S.02396	Sculpture	sculpture, ronde bosse			La Femme accroupie	Rodin	plâtre		1882	élément porte de l'enfer ; petit modèle	Porte de l'enfer ; Femme accroupie		femme ; nue ; accroupie		
S.00598.(LUX)	Sculpture	sculpture, ronde bosse			FugitAmor	Rodin	bronze	fonte au sable	Av. 1885	élément porte de l'enfer	Porte de l'enfer ; Fugitamor		couple		
S.00467	Sculpture	sculpture, ronde bosse			Adolescent désespéré	Rodin	bronze	fonte au sable		élément porte de l'enfer	Porte de l'enfer ; Adolescent desespéré		homme ; nu ; à genoux		

Annexe 2 – Présentation du thésaurus

Les termes sont regroupés autour de cinq domaines, eux-mêmes regroupés en deux groupes :

Description du processus de création :

- Thématiques
- Phases de conception

Description iconographique :

- Description par mots clés
- Personnages représentés
- Lieux représentés

La description du processus de création est propre et spécifique aux collections du musée Rodin.

La description iconographique s'appuie sur le *Thésaurus iconographique Garnier*.

Description du processus de création

Thématique

Termes génériques (TG)	Termes spécifiques (TS)	Termes non employés	Termes associés (TA) *en italique : autre domaine*	Terme générique dans un autre domaine
Abattis				
Académie				
Adam			Eve	
Adolescent désespéré (l')		Désespéré		
Âge d'airain (l')		Airain		
Aile				
Amants				
Amour		Putti		
Ange				
Architecture			Monument ; Cariatide ; *Éléments d'architecture*	
	Église			
	Cathédrale			
	Château			

Termes génériques (TG)	Termes spécifiques (TS)	Termes non employés	Termes associés (TA) *en italique : autre domaine*	Terme générique dans un autre domaine
	Élément		*atlante ; cheminée ; colonne ; édifice entier ; fenêtre ; fragment ; frise ; lanterne ; moulure*	
Assemblée d'hommes			Funérailles ; Tombeau ; *homme*	
Bacchantes				*mythologie gréco-romaine (domaine "Personnages")*
Bacchanale				
Bacchus				*mythologie gréco-romaine (domaine "Personnages")*
Baigneuse				
Balzac				*écrivain (domaine "Personnages)*
	Têtes			
	Masques			
	Études de nu			
	Études drapées			
	Monument			
Baudelaire				*écrivain (domaine "Personnages")*
Bourgeois de Calais (les)			*Calais ; Aire, Jean d' ; Andres, Andrieu d' ; Fiennes, Jean de ; Saint-Pierre, Eustache de ; Wissant, Jacques de ; Wissant, Pierre de*	

Termes génériques (TG)	Termes spécifiques (TS)	Termes non employés	Termes associés (TA) *en italique : autre domaine*	Terme générique dans un autre domaine
	Maquettes			
	Figures			
	Têtes			
	Mains			
	Monument			
Cariatide			*Atlante*	
Cavalier				
Centaure				*mythologie gréco-romaine (domaine "Personnages")*
Centauresse				*mythologie gréco-romaine (domaine "Personnages")*
Cheval				*animal (domaine "Description par mots-clés")*
	Course de chevaux			
Croix				*personnage du Nouveau Testament (domaine "Personnages")*
Cortège			Ronde	
Couple enlacé				
Couple saphique				
Damné			Femmes damnées	

Termes génériques (TG)	Termes spécifiques (TS)	Termes non employés	Termes associés (TA) *en italique : autre domaine*	Terme générique dans un autre domaine
Danse			Mouvement de danse	
	Danseuse			
	Danseur			
	Danseuse cambodgienne			
	Danseuse javanaise			
Dante (dessins qui se rapportent à l'une des séquences de l'Enfer)				*écrivain (domaine "Personnages)*
Enguerrande				
Enfant			Femme et enfant	
Enlèvement		Rapt		
Ève			Adam ; Adam et Ève	
Faune				*mythologie gréco-romaine (domaine "Personnages")*
Faunesse				*mythologie gréco-romaine (domaine "Personnages")*
Femme et enfant			Enfant ; Jeune femme	
Femme en tunique bleu ?			*femme ; tunique*	
Femme moulure			*femme ; moulure*	
Femme pyjama (avant 1901)			*femme ; pyjama*	
Femme vase (après 1859)			*femme ; vase*	

Termes génériques (TG)	Termes spécifiques (TS)	Termes non employés	Termes associés (TA) *en italique : autre domaine*	Terme générique dans un autre domaine
Femmes damnées			*femme* ; Damné	
Funérailles			Assemblée d'homme ; Tombeau	
Furie				
Génie				
Guerrier				
Héros				
Idylle				
Lion				*animal (domaine "Description par mots-clés)*
Lutte		Combat		
Main			*bras*	
Masque				
Mercure				*mythologie gréco-romaine (domaine "Personnages)*
Métamorphoses d'Ovide (extraire les thèmes)			*Ovide*	
Minerve				*mythologie gréco-romaine (domaine "Personnages)*
Minotaure				*mythologie gréco-romaine (domaine "Personnages)*
Monument		Décoration monumentale	Architecture ; Cariatide	
Mort				*corps humain (domaine "Description par mots-clés)*

Termes génériques (TG)	Termes spécifiques (TS)	Termes non employés	Termes associés (TA) *en italique : autre domaine*	Terme générique dans un autre domaine
Mouvement de danse			Danse	
Nymphe				*mythologie gréco-romaine (domaine "Personnages)*
Ombre				
Paolo et Francesca		Francesca	Le Baiser	*personnage imaginaire (domaine "Personnages)*
Paysage				
Poète (le)				
Porte de l'Enfer		Enfer		
	Porte de l'Enfer : maquettes			
	Porte de l'Enfer sans les figures (1900)			
	Porte de l'Enfer avec les figures			
	Porte de l'Enfer : éléments (tympan, pilastres, vantaux)			
Portrait				*représentation humaine (domaine "Description par mots-clés)*
Printemps				
Psyché				*mythologie gréco-romaine (domaine "Personnages)*
Purgatoire			*figures du purgatoire*	
Ronde			Cortège	

Termes génériques (TG)	Termes spécifiques (TS)	Termes non employés	Termes associés (TA) *en italique : autre domaine*	Terme générique dans un autre domaine
Satyre				*mythologie gréco-romaine (domaine "Personnages)*
Serpent				*animal (domaine "Description par mots-clés)*
Sirène				*mythologie gréco-romaine (domaine "Personnages)*
Sommeil (le)				
Songe (le)				
Source (la)				
Sphinge				
Tentation				
Tête de la femme slave		Femme ; slave		
Tombeau		sarcophage	Assemblée d'hommes ; Funérailles	
Ugolin				*personnage imaginaire (domaine "Personnages)*
Vénus				*mythologie gréco-romaine (domaine "Personnages)*
Victor Hugo			*Hugo, Victor*	
	Bustes, têtes et masques de Victor Hugo 1er monument : Victor Hugo assis 2ème monument : Victor Hugo debout			
Vieillard				

Termes génériques (TG)	Termes spécifiques (TS)	Termes non employés	Termes associés (TA) *en italique : autre domaine*	Terme générique dans un autre domaine
	Vieillard assis		assis	
	Vieillard suppliant	Suppliant		
Assemblage avec vase				
Assemblage de figures				

Description du processus de création

Phases de création

Domaine de collection	Phase de création	Terme rejeté
dessin, sculpture, peinture	Jeunesse - Formation (1854-1859)	
dessin, sculpture	Pour Sèvres (1859-1862)	
dessin, sculpture, peinture	En Belgique (1871-1877)	
dessin	Dessins d'inspiration du voyage en Italie (1875-1876)	
dessin	Dessins noirs (vers 1880)	
sculpture	Années 1880	
dessin	Dessins de transition - D’après nature (1890-1896)	
sculpture	Années 1890	
dessin, sculpture	Exposition à l'Alma (1898-1900)	
dessin	Dessins musée secret	*érotique*
sculpture	Après 1900	
dessin	Dessins à l'estompe (1909-1913)	

Description iconographique

Description par mots-clés

Catégorie	Termes génériques (TG)	Termes spécifiques 1 (TS 1)	Termes spécifiques 2 (TS 2)	Termes non employés	Termes associés (TA)
Genre iconographique					
	représentation humaine				
		figures			
			figure		
			figure biblique		
			figure mythologique		
		groupe de figures			
			groupe de figures		
			groupe de figures bibliques		
			groupe de figures mythologiques		
		portrait			
			portrait		
			portrait de proche		
			portrait d'anonyme		
			autoportrait		
		scène			
			scène		
			scène antique		
			scène d'après l'antique		
			scène biblique		
			scène historique		
			scène mythologique		
Figuration partielle					
	à mi-corps				

Catégorie	Termes génériques (TG)	Termes spécifiques 1 (TS 1)	Termes spécifiques 2 (TS 2)	Termes non employés	Termes associés (TA)
	en buste				
	sans membre				
		sans bras			
			sans bras gauche		
			sans bras droit		
			sans avant-bras		
		sans main			
			sans main gauche		
			sans main droite		
		sans jambe			
			sans jambe gauche		
			sans jambe droite		
		sans pied			
			sans pied droit		
			sans pied gauche		
		sans tête		*sans visage*	
	avec membre				
		avec bras			
			avec bras gauche		
			avec bras droit		
			avec avant-bras		
		avec main			
			avec main gauche		
			avec main droite		
		avec jambe			
			avec jambe gauche		
			avec jambe droite		
		avec pied			
			avec pied droit		
			avec pied gauche		
		avec tête		*avec visage*	
Angle de vue					

Catégorie	Termes génériques (TG)	Termes spécifiques 1 (TS 1)	Termes spécifiques 2 (TS 2)	Termes non employés	Termes associés (TA)
	vu d'en bas				
	vu d'en haut			*de haut ; vu de haut*	
	de face			*vu de face*	
	de dos			*vu de dos*	
	de profil			*de profil à gauche ; de profil à droite ; sur le côté*	
	de trois-quarts				
Position					
	debout			*en pied*	
	accroupi				
	à genoux			*genoux en terre ; agenouillé*	
	à quatre pattes				
	assis				
	à califourchon				
	allongé			*couché ; étendu*	
		allongé sur le côté			
		allongé sur le ventre			
		allongé sur le dos			
		à demi-allongé			
	corps penché				
		corps penché en avant			
		corps penché vers la gauche		*vers la gauche*	
		corps penché vers la droite		*vers la droite*	
		corps penché vers l'arrière		*à la renverse ; corps renversé en arrière*	
	corps tourné				
		corps tourné vers la gauche		*vers la gauche*	
		corps tourné vers la droite		*vers la droite*	
	corps dressé			*se redressant*	

Catégorie	Termes génériques (TG)	Termes spécifiques 1 (TS 1)	Termes spécifiques 2 (TS 2)	Termes non employés	Termes associés (TA)
	sur le flanc				
		sur le flanc gauche			
		sur le flanc droit			
	sur les mains				
	sur les talons				
Être humain					
	enfance				
		enfant			
			fillette		
			garçon		
	jeunesse				
		jeunes gens			
			jeune fille		
			jeune homme		
	maturité				
		femme			
			femme vêtue	*vêtue*	
			femme nue	*nue*	
			femme à demi-nue	*demi-nue ; demi-vêtue ; femme à demi-vêtue*	
			femme s'habillant	*s'habillant*	
			femme se déshabillant	*se déshabillant*	
			femme drapée	*drapée*	
			femme enceinte	*enceinte*	
			femme et enfant		
		homme			
			homme vêtu	*vêtu*	
			homme nu	*nu*	
			homme à demi-nu	*demi-nu ; demi-vêtu ; homme à demi-vêtu*	
	vieillesse				
		vieille femme			

Catégorie	Termes génériques (TG)	Termes spécifiques 1 (TS 1)	Termes spécifiques 2 (TS 2)	Termes non employés	Termes associés (TA)
		vieil homme			
	personnage				
	couple				
		couple saphique			
		couple enlacé			
	frère				
	sœur				
Corps humain					
	dos				
	fesses				
	membre				
		bras			
			main		
		jambe			
			pied		
	peau				
	sexe				
		sexe féminin			
		sexe masculin			
	tête				
	partie de tête				
		barbe			
			barbu		
		moustache			
		bouche			
		dent			
		langue			
		cheveux			
			chignon		
		nez			
			nez cassé		
			nez retroussé		

Catégorie	Termes génériques (TG)	Termes spécifiques 1 (TS 1)	Termes spécifiques 2 (TS 2)	Termes non employés	Termes associés (TA)
		oreille			
	torse dénudé				
		sein			
	ventre				
	mort				
	repos				
	soin du corps				
		coiffure			
			chignon		
		bain			
Habillement	chapeau				
	chemise				
		chemise ouverte			
	dans ses voiles				
	demi-drapé				
	drapé			*dans des draperies*	
	drapé sous les hanches				
	jambes dénudées				
	pantalon				
	pyjama				
	sandale				
	tête couverte				
	torse dénudé				
	tunique				
	vêtement				
		vêtement autour des hanches		*vêtement sur les hanches*	
		vêtement autour des reins			
		vêtement entrouvert			
		vêtement ouvert			

Catégorie	Termes génériques (TG)	Termes spécifiques 1 (TS 1)	Termes spécifiques 2 (TS 2)	Termes non employés	Termes associés (TA)
				vêtement relevé jusqu'à la taille ; robe retroussée jusqu'à la taille ; vêtement relevé sur les reins ; vêtement retroussé sur les hanches ; vêtement relevé au dessus du visage ; vêtement retroussé ; vêtement relevé jusqu'au sein	
		vêtement relevé			
		vêtement sur l'épaule			
Posture					
	accoudé			*s'accoudant ; appuyé sur un coude*	
	en appui			*appuyé*	
	bras croisés				
	bras levé			*bras soulevé*	
		bras gauche levé			
		bras droit levé			
		bras levés			
	bras replié			*levant un avant-bras*	
		bras replié au-dessus de la tête			
	bras tendu				
		bras droit tendu			
		bras gauche tendu			
		bras tendus			
		bras tendu au-dessus de la tête		*bras repliés au-dessus du visage*	
	genou en terre				
	jambe repliée				
		jambes repliées			
	jambes croisées				
	jambes écartées				
	jambe relevée				
		jambe gauche relevée			
		jambe droite relevée			

Catégorie	Termes génériques (TG)	Termes spécifiques 1 (TS 1)	Termes spécifiques 2 (TS 2)	Termes non employés	Termes associés (TA)
		jambes relevées			
	main à la chevelure			*main aux cheveux*	
	main jointe	mains à la chevelure			
	main tendue			*imploration*	
	main serrée			*poing serré*	
	paume sur la joue				
	paume sur le menton				
	sur le dos d'une femme				
	sur les jambes d'une femme				
	tête appuyée sur la main				
	tête contre un bras				
	tête en bas				
	tête penchée				
	tête renversée				
	tête baissée				
	étreinte				
Action					
	aidant une autre femme à passer un vêtement				
	dansant				
	écrivant				
	enlaçant un enfant				
	faisant le poirier				
	lisant			*étreignant un enfant*	
	marchant				
	passant un vêtement				
	relevant un vêtement				
	se déshabillant				
	se drapant				
	se peignant				

Catégorie	Termes génériques (TG)	Termes spécifiques 1 (TS 1)	Termes spécifiques 2 (TS 2)	Termes non employés	Termes associés (TA)
	s'habillant				
	tenant un vêtement				
	tenant une draperie				
Architecture					Cariatide
	éléments d'architecture				
		atlante			
		cheminée			
		colonne			
		édifice entier			
		fenêtre			
		fragment			
		frise			
		lanterne			
		moulure			
		vase			
Accessoires					
	bâton				
	draperie				
	rideau				
	casque				
Nature					
	matière inanimée				
		brouillard			
		coquille			
		mer			
		nuée			
		nuage			
		rocher			
	végétal				
		arbre			

Catégorie	Termes génériques (TG)	Termes spécifiques 1 (TS 1)	Termes spécifiques 2 (TS 2)	Termes non employés	Termes associés (TA)
		branchage			
		caillou			
		fleur			
		fruit			
		plante			
	animal				
		âne			
		lion			
		serpent			
		aile			
		abeille			
		cheval	chevaux		
			cheval au galop		
				chevaux au galop	
			cheval se cabrant		
				chevaux se cabrant	

Description iconographique

Personnages représentés

Catégorie	Terme générique (TG)	Terme spécifique (TS)	Profession	Lien avec Rodin	Lien avec les œuvres	Nationalité
Sujet biblique						
	personnage de l'Ancien Testament					
		Putiphar (femme de)				
	personnage du Nouveau Testament					
		Christ				
		Enfant prodigue				
		Figure du purgatoire				
Sujet mythologique						
	mythologie gréco-romaine					
		Actéon				
		Andromède				
		Aphrodite				
		Apollon				
		Ariane				
		Bacchante				
		Bacchus				
		Bellone				
		Castor et pollux				
		Centaure				
		Centauresse				
		Cerbère				
		Chloé				

Catégorie	Terme générique (TG)	Terme spécifique (TS)	Profession	Lien avec Rodin	Lien avec les œuvres	Nationalité
		Cybèle				
		Diane				
		Enée				
		Daphnis				
		Dosia				
		Esculape				
		Euménides				
		Eurydice				
		Faune				
		Faunesse				
		Flora				
		Hamadryade				
		Icare				
		Iris				
		Isis				
		Mercure				
		Minerve				
		Minotaure				
		Muse				
		Naïade				
		Narcisse				
		Neptune				
		Nymphe				
		Niobé				
		Orphee				
		Psychée				
		Saturne				
		Satyre				
		Sirène				
		Thétis				
		Titan				
		Triton				
		Ulysse				

Catégorie	Terme générique (TG)	Terme spécifique (TS)	Profession	Lien avec Rodin	Lien avec les œuvres	Nationalité
		Vénus				
Personnage imaginaire						
		Ophelie				
		Paolo et Francesca				
		Ugolin			*figure ; sujet*	
		Comte Guidon				
Personnage historique						
	saint					
		Saint Antoine				
		Saint Francois d'Assise				
		Saint Pierre-Julien Eymard (1811-1868)				*français*
		Saint Jean-Baptiste				
		Saint Sébastien				
	artiste					
		Cigoli, Valério (1529-1599)	*sculpteur*			*italien*
		Lorrain, Claude (1600-1682)	*peintre*		*figure*	*lorrain*
		Michel-Ange (1475-1564)	*sculpteur ; peintre ; architecte ; poète*			*italien*
		Poussin, Nicolas (1594-1665)	*peintre*			*français*
		Rubens, Petrus Paulus (1577-1640)	*peintre ; graveur*			*flamand*
	écrivain					*français*
		Balzac, Honoré de (1799-1850)	*romancier ; auteur dramatique*			*français*
		Baudelaire, Charles (1821-1867)	*poète ; critique d'art*		*figure*	*italien*
		Dante Alighieri (1265-1321)	*poète*			
		Homère	*poète épique*			
		Ovide (0043 av. J.-C.-0017	*poète*			*français*
		Stendhal (1783-1842)	*écrivain*			

Catégorie	Terme générique (TG)	Terme spécifique (TS)	Profession	Lien avec Rodin	Lien avec les œuvres	Nationalité
	personnage politique					
		Cléopâtre VII (reine d'Égypte ; 0069-0030 av. J.-C.)				
		Marie de Médicis (reine de France ; 1573-1642)				
		Néron (empereur romain ; 0037-0068)	*empereur de Rome de 54 à 68*			
	autre					
		Aire, Jean d'		*commande officielle*	*figure*	
		Alembert, d'		*commande officielle*	*figure*	
		Andres, Andrieu d'		*commande officielle*	*figure*	
		Fiennes, Jean de		*commande officielle*	*figure*	
		Gherardesca, Ugolin della (1220-1289)	*modèle du héros damné de la Divine Comédie*	*commande officielle*	*figure ; lié à la porte de l'enfer*	
		Lynch, Général		*commande officielle*	*figure*	
		Saint-Pierre, Eustache de		*commande officielle*	*figure*	
		Wissant, Jacques de		*commande officielle*	*figure*	
		Wissant, Pierre de		*commande officielle*	*figure*	*italien*
Contemporain de Rodin						
	proche de Rodin					
		Beuret, Rose (1844-1917)		*maîtresse ; épouse*	*figure*	*française*
		Choiseul, duchesse de		*maîtresse*	*figure*	*d'origine*

Catégorie	Terme générique (TG)	Terme spécifique (TS)	Profession	Lien avec Rodin	Lien avec les œuvres	Nationalité
						américaine
		Claudel, Camille (1864-1843)	*sculptrice*	*élève ; modèle ; maitresse*	*figure*	*française*
		Cruchet, Mme			*figure*	
		Elisseieff, Madame			*figure*	
		Goloubeff, Madame de (1879-1941)				
		Haquette Maurice	*secrétaire d'administration - manufacture de Sèvres*			
		Hunter, Madame			*figure*	
		Morla-Vicuna, Madame				
		Rodin, Jean-Baptiste		*père de Rodin*		
		Rohozinska, Madame				
		Roll, Mme				
		Simpson Kate				
	modèle de Rodin					
		Abruzzesi, Adèle		*modèle*		
		Abruzzesi, Anna		*modèle*		
		Abruzzesi, François		*modèle*		
		Pignatelli	*modèle des Beaux-Arts*	*modèle*		
		Poirée	*modèle des Beaux-Arts*	*modèle*		
		Fanelli	*modèle des Beaux-Arts*	*modèle*		
		Valentin	*modèle des Beaux-Arts*	*modèle*		
		Corsi	*modèle des Beaux-Arts*	*modèle*		
		Bianci	*modèle des Beaux-Arts*	*modèle*		
		Neyt, Auguste		*modèle*		
		Damedoz, Carmen		*modèle*		
		Samuel S. White	*étudiant*	*modèle*		*américain*
	riche clientèle					
		Fairfax, Eve		*commande privée*	*figure*	
		Sackville		*commande*	*figure*	

Catégorie	Terme générique (TG)	Terme spécifique (TS)	Profession	Lien avec Rodin	Lien avec les œuvres	Nationalité
				privée		
		Fenaille, Madame		*commande privée*	*figure*	
		Nostitz, Hélène von		*commande privée*	*figure*	
		Harriman, Edward Henry	*homme d'affaires*	*commande privée*	*figure*	*américain*
		de Walden, Lord Howard (1880-1946)		*commande privée*	*figure*	anglais
	collectionneur				*figure*	
		Eddy, Arthur-Jérôme	*juriste ; collectionneur*	*commande privée*	*figure*	*américain*
		Ryan, Thomas Fortune		*commande privée*	*figure*	
		Fenaille, Maurice (1855-1937)		*commande privée*	*figure*	
	artiste					
		Besnard, Albert (1849-1934)				*français*
		Bastien-Lepage, Jules (1848-1884)			*figure*	*français*
		Boutet, Henri (1851-1919)	*peintre*			
		Chauvel, Théophile (1831-1910)	*peintre ; dessinateur ; graveur ; lithographe,*			*français*
		Carrier-Belleuse, Albert-Ernest (1824-1887)	*sculpteur ; peintre*	*professeur*	*figure*	
		Carrière, Eugène (1849-1906)	*peintre, litographe*		*figure*	
		Cesbron, Achille (1849-1913)	*peintre*			*français*
		Dalou, Jules (1838-1902)	*sculpteur*		*figure*	*français*
		Falguière, Alexandre (1831-1900)	*sculpteur ; peintre*		*figure*	*français*
		Guillaume, Eugène (1822-1905)	*sculpteur ; professeur ; critique d'art*			*français*
		Huguet, Almire	*scultpteur*	*ami*		

Catégorie	Terme générique (TG)	Terme spécifique (TS)	Profession	Lien avec Rodin	Lien avec les œuvres	Nationalité
		John, Gwen (1876-1939)				*galloise*
		Laurens, Jean-Paul (1838-1921)	*sculpteur ; peintre*			*français ; naturalisé britannique*
		Legros, Alphonse (1837-1911)	*sculpteur ; peintre ; graveur*			*français*
		Nadelman, Elie (1882-1946)	*sculpteur ; dessinateur ; collectionneur*			*polonais ; naturalisé américain*
		Puvis de Chavannes, Pierre (1824-1898)	*peintre*			*français*
		Silsby, Wilson (1883-1952)	*graveur*			*américain*
	compositeur					
		Franck, César (1822-1890)	*professeur ; organiste ; compositeur*		*figure*	*belge ; naturalisé français*
		Mahler, Gustav(1860-1911)	*compositeur ; pianiste ; chef d'orchestre*			*autrichien*
	danseur					
		Hanako (1868-1945)	*danseuse ; comédienne*	*modèle*	*figure ; sujet*	*japonaise*
		Loie Fuller (1862-1928)	*danseuse*			*américaine*
		Moreno, Alda	*danseuse ; acrobate*			
		Nijinsky, VatslavFomitch (1889-1950)	*danseur ; chorégraphe*			*russe ; origine polonaise*
		Saint Denis, Ruth (1878-1968)	*danseuse*			*américaine*
		Isadora duncan (1877-1927)	*danseuse*			*américaine*
	critique d'art					
		Coquiot, Gustave (1865-1926)	*écrivain ; critique d'art*	*secrétaire*		
		Geffroy, Gustave (1855-1926)	*journaliste ; critique d'art ; historien ; romancier*			*français*
		Marx, Roger (1859-1913)	*critique d'art*			*français*
		Mirbeau Octave (1848-1917)	*écrivain ; journaliste*			*français*
		Selz, Jean				
	journaliste					
		Thorion, Henry				

Catégorie	Terme générique (TG)	Terme spécifique (TS)	Profession	Lien avec Rodin	Lien avec les œuvres	Nationalité
	écrivain					
		Becque, Henry (1837-1899)	*dramaturge*		*figure*	*français*
		Bergerat, Emile (1845-1923)				
		Cladel, Judith				
		Clerget, Fernand				
		Hugo, Victor	*poète, dramaturge*	*commande officielle*	*figure, sujet*	
		La Varende, Jean de (1887-1959)	*écrivain*			*français*
		Meier-Graefe, Julius (1867-1935)	*écrivain ; critique d'art*			*allemand*
		Michelet, Victor-Emile (1861-1938)	*poète*			*français*
		Montesquiou, Robert de (1855-1921)	*écrivain ; poète ; essayiste*			*français*
		Morice, Charles (1860-1919)	*écrivain ; poète ; essayiste*			*français*
		Shaw, Bernard (1856-1950)	*critique musical ; dramaturge ; essayiste ; scénariste*			*irlandais*
		Rilke, Rainer Maria (1875-1926)	*écrivain*			
		Séverine,				
	personnage politique					
		Barres, Maurice (1862-1923)	*écrivain ; homme politique*			*français*
		Clemenceau, George (1841-1929)	*homme politique ; homme d'Etat*		*figure*	*français*
		Clémentel, Etienne (1864-1936)	*homme politique ; ministre*		*figure*	*français*
		Dewavrin, Omer	*maire de Calais*			
		Estournelles de Constant, Paul-Henri Benjamin, Baron d' (1852-1924)	*diplomate ; homme politique*		*figure*	*français*
		Herriot, Edouard (1872-1957)	*homme politique ; écrivain*			

Catégorie	Terme générique (TG)	Terme spécifique (TS)	Profession	Lien avec Rodin	Lien avec les œuvres	Nationalité
		Leygues, Georges (1857-1933)	*homme politique*			
		Proust, Antonin (1832-14905)	*journaliste ; homme politique ; critique d'art, collectionneur*			*français*
		Sarmiento, Domingo Faustino (1811-1888)	*écrivain,homme d'Etat*			*argentin*
		Sisowath (1840-1927)	*roi*			*cambodgien*
		Wyndham, George (1863-1913)	*écrivain ; homme politique*			*anglais*
	religieux					
		Benoit XV (1854-1922)	*pape*		*figure ; sujet*	*italien*
	scientifique					
		Berthelot, Marcelin (1827-1907)	*chimiste ; essayiste ; historien des sciences ; homme politique*		*figure*	
		Thiriar, Jules-Adrien (1846-1913)	*physicien*			*belge*
		Curie, Marie (1867-1934)	*physicienne ; chimiste*			*polonaise ; naturalisée française*
		Poincaré, Henri (1854-1912)	*mathématicien ; physicien ; philosophe ; ingénieur*			
	photographe					
		Eugène Druet (1867-1916)	*galeriste parisien ; photogaphe*			*français*
	fournisseur					
		Mommen, Félix				
	travaux pour Rodin					
		Danielli, J.	*Durcissement, Métallisation et décorations artistiques du plâtre*		*figure*	

Catégorie	Terme générique (TG)	Terme spécifique (TS)	Profession	Lien avec Rodin	Lien avec les œuvres	Nationalité
A classer	Basler, Adolphe					
	Bonnet, Anne-Marie					
	Brophy, John					
	Ciolkowska, Muriel					
	Dircks					
	Duchâtel, F.					
	Kieffer					
	Hunt, Lynn					
	Murat					
	Rohan, duc de					
	Van Berckelaer, Alexandre					
	Petit, Victor					
	Gsell, Paul					

Description iconographique

Lieux représentés

Lieux géographiques
Auderghem
Bd d'Anspach
Belgique
Blois
Boitsfort
Bruxelles
Chartres
Courbevoie
Dijon
Etampes
Groenendael
Houdan
La Seine
Lorraine
Orient
Parthénon
Pologne
Pompéï
Quimperle
Soignes
Watermael
Willebroeck