



I DONNÉES OUVERTES

D'origine publique ou privée, les données ouvertes sont librement partagées, utilisées et mises à profit par quiconque, sans restriction. Ainsi, dans le contexte muséal, il peut s'agir d'une sélection de données descriptives d'objets de collection (date de production, origine, date d'acquisition, etc.).

Des jeux de données, ou ensembles de données, peuvent être publiés, c'est-à-dire mis à disposition sous forme de fichiers prêts à être téléchargés et utilisés librement.

Pour voir différentes utilisations de données ouvertes, visionnez les interventions d'Ana-Laura Baz, Emmanuel Château-Dutier, Jean-François Gauvin et Patrick M. Lozeau qui répondent aux questions de Bruno Guglielminetti dans le cadre de la journée professionnelle du 22 juin 2016.



VIDÉO

POURQUOI opter pour des données ouvertes?

- Pour se positionner favorablement dans un contexte où l'évolution technologique transforme les façons de créer du contenu numérique et pour être visible dans le Web des données.
- Pour contribuer à l'offre de contenu culturel québécois et augmenter sa visibilité sur le Web.
- Pour faciliter l'accessibilité et la réutilisation des données muséales (collections, localisation, heures d'ouverture, expositions en cours, activités, etc.).
- Pour être partie prenante du développement culturel et économique, résultant de la création de services autour de l'exploitation d'une masse critique de données ouvertes dans le secteur culturel.
- Parce que de nouvelles combinaisons de données provenant de différentes sources peuvent mener à des projets innovants (applications, services numériques, etc.).
- Parce que les données ouvertes peuvent être extraites et téléchargées dans un format exploitable par des systèmes informatiques, contrairement aux données accessibles en ligne dans une base de données comme celle d'Info-Muse de la SMQ.
- Pour jouer un rôle central dans la diffusion des connaissances à l'ère numérique.

EXEMPLES

Avec la disponibilité de jeux de données regroupant géolocalisation, calendriers d'activités de municipalités, de musées, de bibliothèques, des adresses, des œuvres et objets de collection, des personnages historiques, des fonds photographiques, etc., des applications proposeraient des visites culturelles, ou des parcours sur un territoire en faisant des liens avec des activités ponctuelles offertes, en plus d'informer les utilisateurs sur des personnages historiques ou des artistes ayant habité ce même territoire.

Le musée royal des Pays-Bas, **Rijksmuseum**, a libéré les données et les photographies des œuvres de grands maîtres de la peinture hollandaise, comme *La laitière* de Vermeer afin de se positionner en tant qu'autorité dans ce domaine, tout en accroissant son rayonnement culturel et touristique.



VIDÉO

Pour voir jusqu'où vont les propositions de la Société des musées de sciences et technologies du Canada, visionnez la conférence d'Alex Benay.

COMMENT publier des données ouvertes?

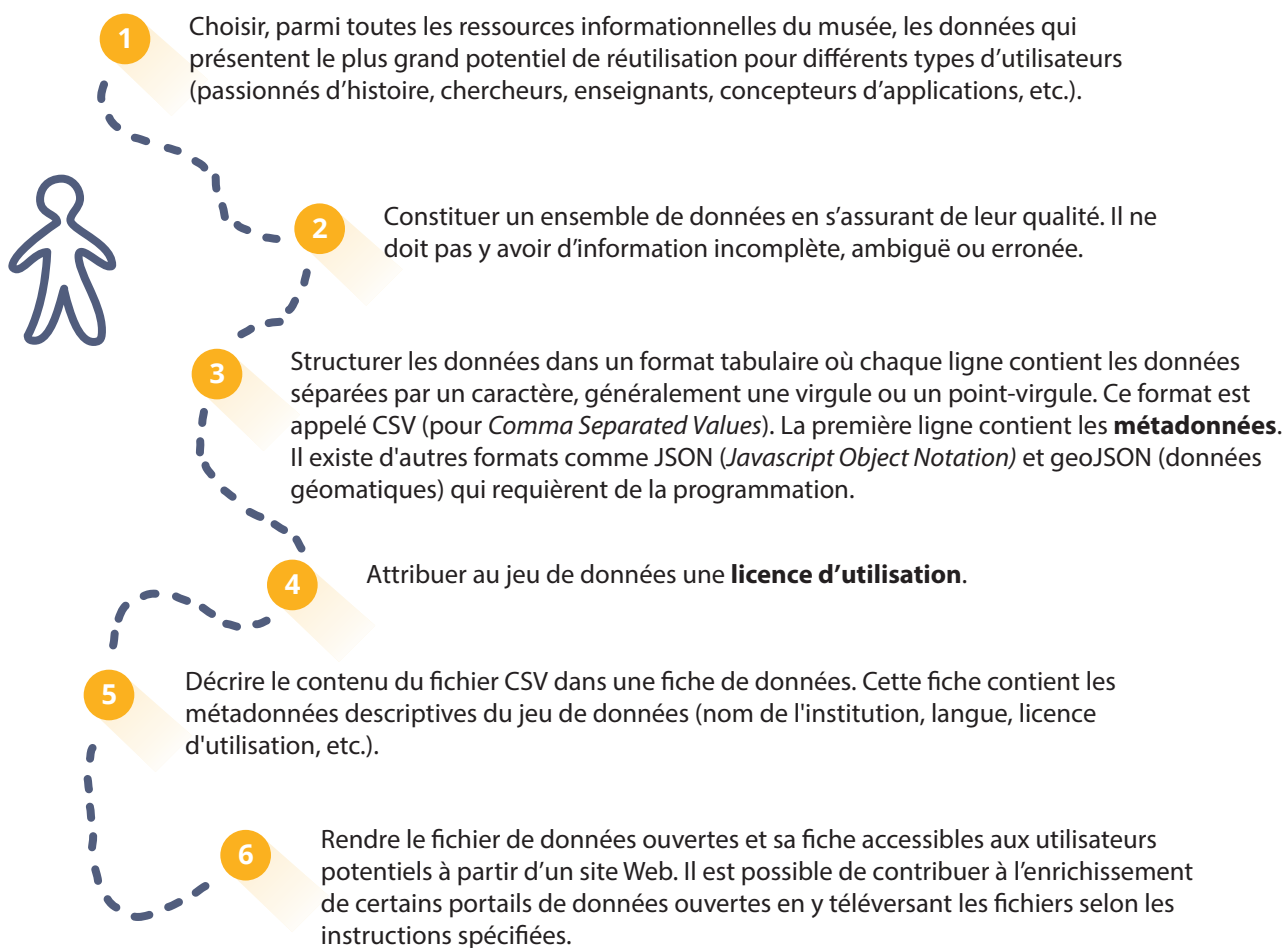
Pour la plupart des musées du Québec, une partie du travail sur les collections a déjà été réalisée avec l'implantation de Réseau Info-Muse. En effet, les données sont structurées selon les normes du système documentaire Info-Muse et stockées dans des bases de données.



Comment documenter vos collections?

Le guide de documentation du Réseau Info-Muse

Alors, que reste-t-il à faire?

- 
- 1 Choisir, parmi toutes les ressources informationnelles du musée, les données qui présentent le plus grand potentiel de réutilisation pour différents types d'utilisateurs (passionnés d'histoire, chercheurs, enseignants, concepteurs d'applications, etc.).
 - 2 Constituer un ensemble de données en s'assurant de leur qualité. Il ne doit pas y avoir d'information incomplète, ambiguë ou erronée.
 - 3 Structurer les données dans un format tabulaire où chaque ligne contient les données séparées par un caractère, généralement une virgule ou un point-virgule. Ce format est appelé CSV (pour *Comma Separated Values*). La première ligne contient les **métadonnées**. Il existe d'autres formats comme JSON (*Javascript Object Notation*) et geoJSON (données géomatiques) qui requièrent de la programmation.
 - 4 Attribuer au jeu de données une **licence d'utilisation**.
 - 5 Décrire le contenu du fichier CSV dans une fiche de données. Cette fiche contient les métadonnées descriptives du jeu de données (nom de l'institution, langue, licence d'utilisation, etc.).
 - 6 Rendre le fichier de données ouvertes et sa fiche accessibles aux utilisateurs potentiels à partir d'un site Web. Il est possible de contribuer à l'enrichissement de certains portails de données ouvertes en y téléversant les fichiers selon les instructions spécifiées.



Pour voir comment les Musées de la civilisation et le Musée national des beaux-arts du Québec ont procédé, visionnez les présentations d'Isa Mailloux et de Nathalie Thibault.



I DONNÉES OUVERTES LIÉES

Des données liées sont publiées ou mises à disposition selon les principes du Web sémantique, c'est-à-dire, structurées et normalisées pour être lues par des systèmes informatiques et reliées les unes aux autres par des liens qui ont un sens. Ces données sont rendues disponibles dans le Web de façon similaire à la publication de pages Web.

POURQUOI publier des données liées?

- Pour que les données muséales soient repérables par des moteurs de recherche et par des applications qui collectent des données.
- Pour multiplier les possibilités d'orienter les internautes et les utilisateurs d'applications vers des ressources muséales et des contenus culturels québécois.
- Parce qu'elles favorisent le rayonnement et la **découvrabilité** des musées et des collections en favorisant l'interconnexion des données qui les décrivent avec d'autres données provenant de sources diverses ailleurs dans le monde.
- Pour assurer une présence numérique forte des musées du Québec sans dépendre des grands générateurs de contenus de type Google.

AVANTAGES

- Mise en valeur de collections éloignées des circuits touristiques ou ne disposant pas de grandes possibilités de diffusion.
- Contribution au décroisement des données culturelles et connexion à d'autres données publiées par les institutions muséales ou d'autres organismes de conservation de la mémoire à travers le monde.
- Développement de l'expertise des musées québécois pour le traitement et la publication de données dans le Web.

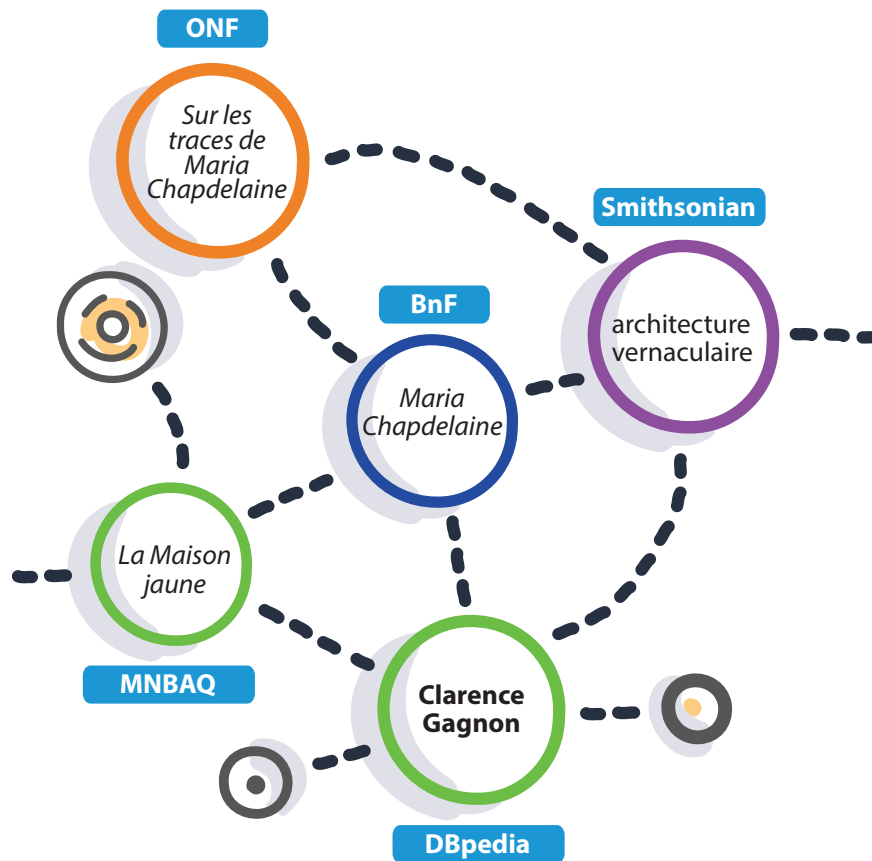




EXEMPLES

Différentes relations peuvent être établies entre le tableau de Clarence Gagnon intitulé *La Maison jaune*, appartenant au Musée national des beaux-arts du Québec (MNBAQ), et d'autres œuvres ou ressources informationnelles selon le sujet, le thème, le type de support, l'auteur, etc.

Ainsi, au fil des liens, nous pourrions découvrir, à la Bibliothèque nationale de France (BnF), le roman *Maria Chapdelaine* de Louis Hémon, illustré par Clarence Gagnon. Ce roman est lié à plusieurs films ayant ce personnage pour sujet. De la BnF, nous pourrions être dirigés vers le site de l'Office national du film du Canada (ONF) pour visionner quelques images du film de Jean-Claude Labrecque, *Sur les traces de Maria Chapdelaine*. Nous serions également tentés de suivre les liens pointant une étude sur l'architecture vernaculaire du XIX^e siècle dans les collections du Smithsonian Institute. Les données de ce document seraient ainsi liées aux données décrivant des éléments patrimoniaux appartenant aux collections des musées du Québec.



JocondeLab propose une interface multilingue pour découvrir près de 300 000 œuvres des musées de France.

<http://jocondelab.iri-research.org/jocondelab>



VIDÉO



DOCUMENT

Pour mieux comprendre la différence entre les données ouvertes et les données liées, visionnez la conférence en anglais de Sara Snyder du Smithsonian American Art Museum, ou consultez la version française de sa présentation.



PAR OÙ COMMENCER?



Se familiariser avec les concepts relatifs aux données ouvertes en les expérimentant avec un projet facile à circonscrire : nouvelle exposition, petite collection, nouvelles acquisitions...



Mutualiser les ressources et les expertises avec des partenaires ayant des objectifs communs.



Constituer une équipe pluridisciplinaire pour tirer parti d'une diversité de points de vue (communication et marketing, conservation, médiation, technologies de l'information, documentation, etc.).



Cibler des utilisateurs potentiels (à l'externe et à l'interne) afin que le jeu de données représente une valeur pour le plus d'utilisations possibles.

COMMENT publier des données ouvertes liées dans le Web?

1

Définir chaque donnée à partir du **format RDF** (*Resource Description Format*). Il s'agit d'une syntaxe de représentation, soit un modèle de description des ressources appelé triplet et composé de trois parties : sujet – prédicat – objet.



Marcelle Ferron, peintre et sculpteur, a créé une verrière pour la station de métro Champ-de-Mars.
Cet énoncé se traduit par les triplets suivants :

verrière	a pour auteure	Marcelle Ferron
Marcelle Ferron	a pour métier	peintre
Marcelle Ferron	a pour métier	sculpteur
verrière	est localisée	métro Champ-de-Mars
donnée	lien	donnée

2

Attribuer à chacune de ces données une **adresse URI** (*Uniform Resource Identifier*). Il s'agit d'un identifiant de ressource unique dont la syntaxe est similaire à celle de l'URL : nom de domaine, suivi de segments qui permettent d'accéder à la ressource (de manière comparable à la structure hiérarchique de répertoires et de dossiers de fichiers), puis du nom de la ressource.

Voici la représentation de l'identification de la ressource Clarence Gagnon :
<http://dbpedia.org/resource/Clarence_Gagnon>

3

Publier les données exprimées sous forme d'URI dans le Web à l'aide d'un serveur, grâce au protocole standard d'échange de données HTTP.

Voici les triplets de données ouvertes et liées où Clarence Gagnon a pour sujet la catégorie Peintres canadiens du XX^e siècle, dans DBpedia (la version Web sémantique de Wikipedia) :

Clarence Gagnon – fait partie – Peintres canadiens du XX^e siècle
<http://dbpedia.org/resource/Clarence_Gagnon>
<<http://purl.org/dc/terms/subject>><http://dbpedia.org/resource/Category:20th-century_Canadian_painters>



VIDÉO

Pour voir un exemple de données ouvertes liées sur des instruments de musique conservés dans différents musées, visionnez la présentation de Rodolphe Bailly, Cité de la musique - Philharmonie de Paris.

I STANDARDS DE PUBLICATION

Sur le Web

Web des documents



Dans le Web

Web des données



PAGES WEB

HTML

Langage de représentation

+

URL

Identification et adressage

+

HTTP

Protocole de communication



DONNÉES OUVERTES LIÉES

RDF

Modèle de description

+

URI

Identification et adressage

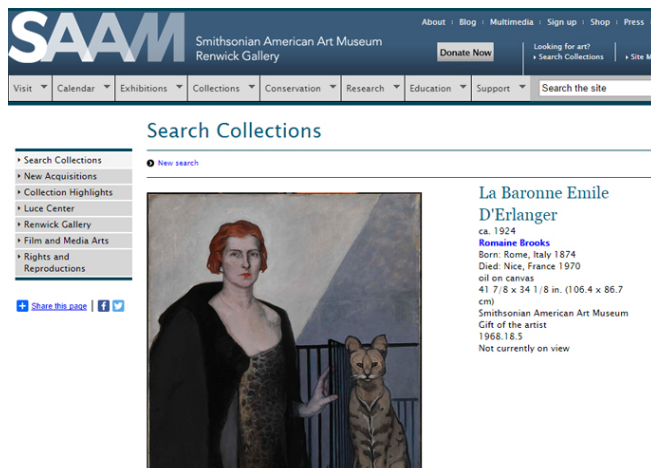
+

HTTP

Protocole de communication

EXEMPLE

EXEMPLE



SAAM Smithsonian American Art Museum
Renwick Gallery

Search Collections

La Baronne Emile D'Erlanger
ca. 1924
Romaine Brooks
Born: Rome, Italy 1874
Died: Nice, France 1970
oil on canvas
41 7/8 x 34 1/8 in. (106.4 x 86.7 cm)
Smithsonian American Art Museum
Gift of the artist
1968.18.5
Not currently on view

Référence : <http://americanart.si.edu/collections/search/artwork/?id=2884>



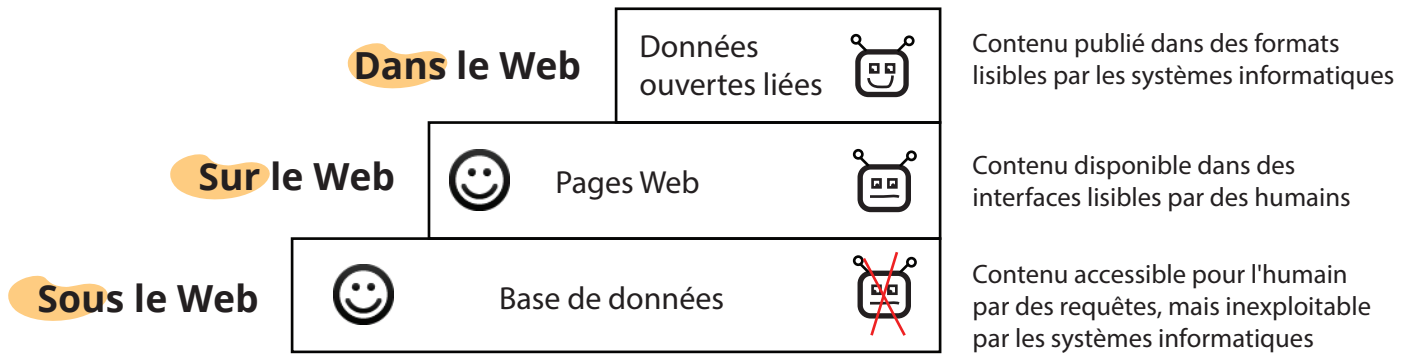
Smithsonian American Art Museum
Renwick Gallery

URI: <http://edan.si.edu/saam/id/object/1968.18.5>

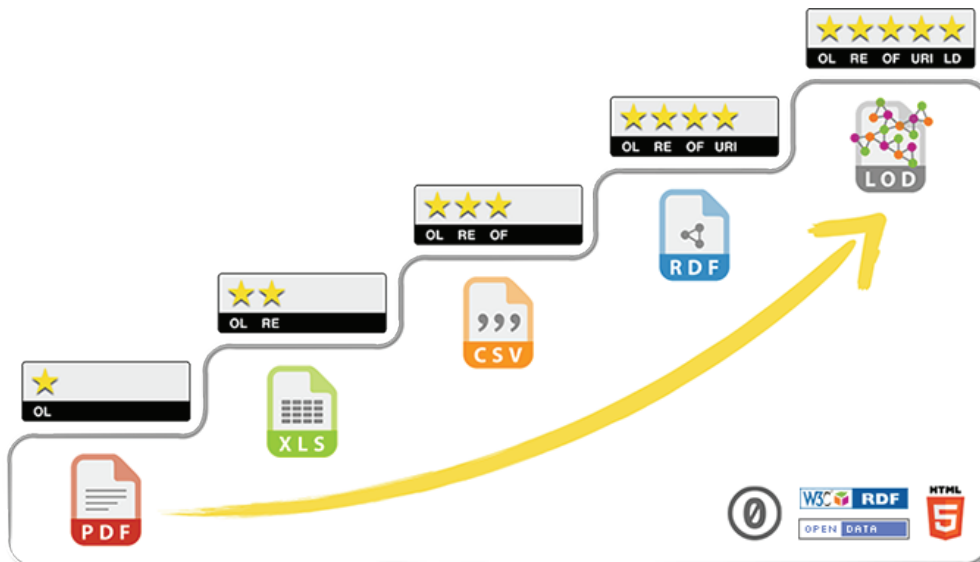
Property	Value
http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://www.cidoc-crm.org/cidoc-crm/P138i
http://www.cidoc-crm.org/cidoc-crm/P138i_has_representation	
http://www.cidoc-crm.org/cidoc-crm/P1_is_identified_by	http://edan.si.edu/saam/id/object/1968.18.5
http://www.cidoc-crm.org/cidoc-crm/P102_has_title	http://edan.si.edu/saam/id/object/1968.18.5
http://www.cidoc-crm.org/cidoc-crm/P30i_custody_transferred_through	http://edan.si.edu/saam/id/object/1968.18.5
http://www.cidoc-crm.org/cidoc-crm/P48_has_preferred_identifier	http://edan.si.edu/saam/id/object/1968.18.5

Référence : <http://edan.si.edu/saam/id/object/1968.18.5>

■ Présence numérique de contenus culturels



■ Vers la publication de données ouvertes liées



Tim Burners-Lee, World Wide Web Consortium

<http://5stardata.info/en>



I LEXIQUE ET RÉFÉRENCES

— Adresse URI (*Uniform Resource Identifier*)

« Chaîne de caractères à syntaxe normalisée qui sert à désigner de façon unique et permanente une ressource par sa localisation ou par son nom. »
(Grand dictionnaire terminologique de l'Office québécois de la langue française)

— Découvrabilité

« Potentiel pour un contenu, un produit ou un service de capter l'attention d'un internaute de manière à lui faire découvrir des contenus autres. »
(Grand dictionnaire terminologique de l'Office québécois de la langue française)

Découvrabilité : nos contenus culturels sont-ils visibles?

Par Josée Plamondon

— Données liées

Données publiées selon les principes du Web sémantique, c'est-à-dire, structurées et normalisées pour être lues par des systèmes informatiques et reliées les unes aux autres par des liens qui ont un sens.

CIDOC-ICOM, Déclaration sur les identifiants des objets muséaux pour les données liées

(en anglais)

— Données ouvertes

Données d'origine publique ou privée rendues accessibles et réutilisables sous une forme libre de droits. Les données ouvertes ne sont pas nécessairement des données liées (Web sémantique) et les données liées ne sont pas nécessairement ouvertes.

Le Manuel de l'Open Data

La notion des « données ouvertes »

Le cycle de vie des données et métadonnées publiques liées et ouvertes

Ce qu'il faut savoir avant de lancer un projet de données ouvertes.

OpenGLAM – Documentation

Ressources (documentation, projets, collections, lab culturels) partagées par les participants d'OpenGLAM, un collectif soutenu par Open Knowledge Foundation Network (en anglais)

— Données ouvertes liées

Données mises à la disposition des utilisateurs en vertu d'une licence d'utilisation. Ces données sont publiées en format RDF et selon des normes pouvant être repérées et lues par des systèmes informatiques. Dans ce format où les relations entre les données sont explicites, des requêtes faites par des utilisateurs ou des interfaces de programmation permettent de relier les données les unes aux autres pour générer une information plus riche et détaillée ou pour découvrir d'autres ressources.

Les 5 étoiles de l'Open Data

Les 5 étapes vers les données ouvertes et liées expliquées par Emmanuelle Bermès.

Le Web des données ouvertes et liées. Qu'est-ce que c'est?

Vidéo (3 min.) produite par Europeana et expliquant le concept.

Si les musées ne choisissent pas l'Open Content, ils deviendront invisibles et inutiles

Par Emmanuelle Delmas-Glass, Yale Center for British Art

Le Museo del Prado lance un nouveau site personnalisé et sémantique pour « transformer complètement l'expérience numérique du musée »

Par Clic France

— Format de données ouvertes

Pour être publiées, les données ouvertes doivent être représentées dans un format prédéfini, comme un tableau simple (colonnes et lignes) ou sous forme tabulaire avec des valeurs délimitées par des séparateurs de champs. Bien que les ensembles de données ouvertes puissent être mis à la disposition des utilisateurs sous différents formats de fichier, le format CSV (*Comma Separated Values*) demeure le plus recommandé parce qu'il est exploitable quel que soit le logiciel et l'environnement technologique.

— Format RDF (ou modèle RDF)

« Format universel de description de données qui a pour but de créer un système d'indexation adapté autant à la création de serveurs de recherche dans Internet qu'à l'indexation de fichiers stockés sur le disque dur de l'utilisateur. »

(Grand dictionnaire terminologique de l'Office québécois de la langue française)

— Licences d'utilisation

Licences définissant ce qu'il est permis de faire avec les données. Il existe différentes licences d'utilisation pour les données ouvertes. La licence Creative Commons 4.0 a été privilégiée par le gouvernement du Québec parce qu'elle présente peu de restrictions tout en protégeant les droits d'auteur. Pour en apprendre plus au sujet des types de licences ouvertes :

Licence Creative Commons

— Métadonnée

Donnée qui définit une autre donnée. Dans l'exemple ci-dessous, les métadonnées précèdent les données qui décrivent un objet.

Numéro d'accèsion : b.2011.57
Nom de l'objet : écritoire
Matériaux : verre
Établissement : Les Sœurs du Bon-Pasteur de Québec

— Norme

Ensemble de règles établies qui fait consensus dans un domaine. Par exemple, pour décrire leurs collections, la majorité des musées québécois utilisent les normes du système documentaire du Réseau Info-Muse de la SMQ.

— Web sémantique

« Web intelligent dans lequel les informations, auxquelles on donne une signification bien définie, sont reliées entre elles de façon à ce qu'elles soient comprises par les ordinateurs, dans le but de transformer la masse des pages Web en un index hiérarchisé et de permettre de trouver rapidement les informations recherchées. »
(Grand dictionnaire terminologique de l'Office québécois de la langue française)

Web sémantique, Web de données : définitions

Vidéo (10 min.) Emmanuelle Bermès,
Bibliothèque nationale de France

Web sémantique et Web de données

Par Inria

Équipe de réalisation

Société des musées du Québec

Supervision : Katia Macias-Valadez, directrice – Formation et développement professionnel
Recherche et rédaction des contenus : Judith Houde, spécialiste en gestion des collections
et Françoise Simard, directrice du Réseau Info-Muse

Consultants

Rédaction et validation des contenus : Josée Plamondon, cyberbibliothécaire, consultante
en exploitation de contenu numérique
Édition et mise en forme : Geoffroi Garon-Épaule, conseiller, Pygmalion numérique

Décembre 2016

Ce projet s'inscrit dans le cadre de la mise en œuvre du Plan culturel numérique du Québec.

Québec 

**COMPÉTENCE
CULTURE**
COMITÉ SECTORIEL DE
MAIN-D'ŒUVRE EN CULTURE



Ce document est mis à disposition selon les termes de la
Licence Creative Commons Attribution 4.0 International.