

# **Comprendre les enjeux de l'ouverture des données publiques culturelles**

**Ministère de la Culture**

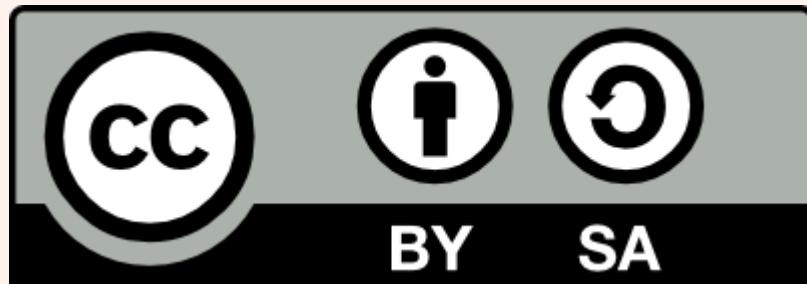
**Dataactivist**

**Octobre 2022**

Ces slides en ligne :<https://dataactivist.coop/ministere-culture/jour1.html>

Sources : <https://github.com/dataactivist/ministere-culture>

Les productions de Dataactivist sont librement réutilisables selon les termes de la licence [Creative Commons 4.0 BY-SA](#).



# Rappel des objectifs de la formation

## Open Data : comprendre les enjeux de l'ouverture des données publiques culturelles

- Appréhender le contexte et les opportunités d'une politique Open Data dans la culture.
- Découvrir des cas pratiques appliqués aux données culturelles.
- Comprendre les problématiques liées aux données du secteur culturel.

# **Programme de la formation**

## **Open Data : comprendre les enjeux de l'ouverture des données publiques culturelles**

1. Origines historiques de l'ouverture des données
2. Les grands principes de l'ouverture des données
3. Enjeux de l'ouverture des données
4. Cadre juridique de l'ouverture des données culturelles
5. Plusieurs exemples #OpenGLAM
6. Concevoir une stratégie d'ouverture de données : quelques grands principes.

# De quoi parle t-on ?

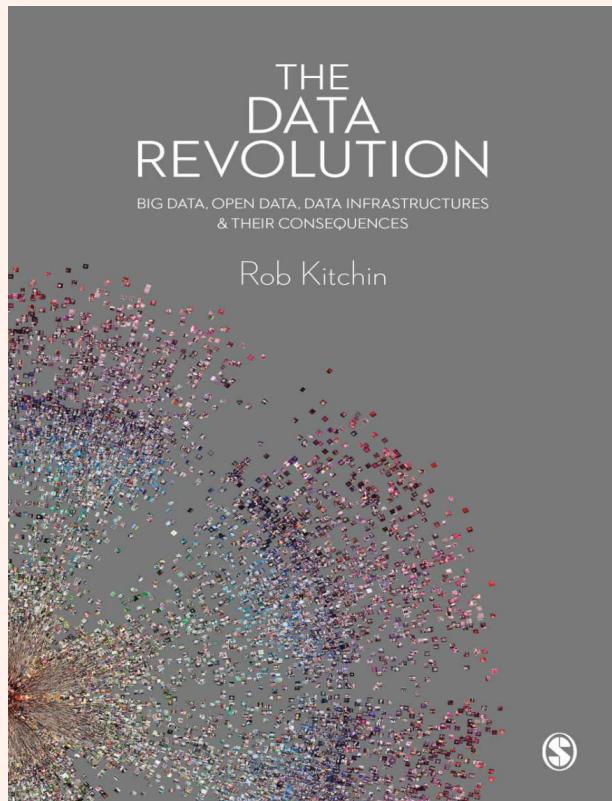
# Imaginons...

Vous êtes journaliste (ou étudiant·e en master ou simplement avec une grande curiosité) et vous travaillez sur les festivals en France. Vous souhaitez étudier : la fréquentation, l'historique des festivals...

- De quoi avez-vous besoin ?
- Comment allez-vous procéder ?

# 1 - Comprendre la donnée

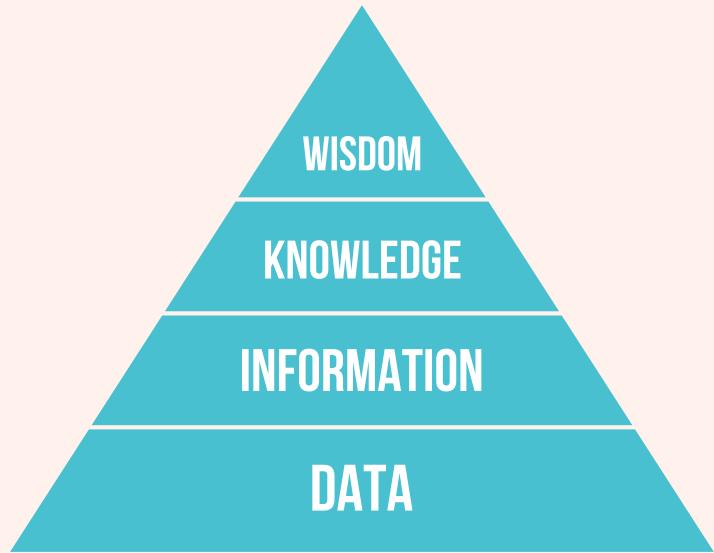
# Une définition générale des données



**Les données sont couramment comprises comme les matériaux bruts produits dans l'abstraction du monde en catégories, mesures et toute autre forme de représentation-nombres, caractères, symboles, images, sons, ondes électromagnétiques, bits qui constituent les fondations sur lesquelles l'information et le savoir sont créés.**

Cette définition est une des rares à s'extraire des discours entourant les données ou leurs infrastructures. L'auteur de cet ouvrage de référence en sociologie s'attache à définir précisément l'ontologie et l'épistémologie de la donnée.

# La pyramide Data-Information-Knowledge-Wisdom



Attribuée à **Russell Ackoff** en 1989, elle signifie que :

- Les **données** sont la matière "brute" de l'information conçues plutôt pour des machines.
- **L'information** pourrait être définie comme des données qui ont été interprétées pour dégager du sens pour des humains.
- En donnant du sens à de l'information, on obtient de la **connaissance**
- En donnant du sens à la connaissance on obtient de la **sagesse**.

# Une définition des "bases de données"

Source : [Source : Guide Open Data Cnil](#)

On entend par base de données un recueil d'œuvres, de données ou d'autres éléments indépendants, disposés de manière systématique ou méthodique, et individuellement accessibles par des moyens électroniques ou par tout autre moyen (art L112-3 du code de la propriété intellectuelle)

En d'autres termes il s'agit d'un ensemble de données organisé en vue de son utilisation par des programmes correspondant à des applications distinctes et de manière à faciliter l'évolution indépendante des données et des programmes (en anglais : data base)

Le document donne également une **définition de ce qu'est une "donnée"** :

Une donnée correspond à la représentation d'une information sous une forme conventionnelle destinée à faciliter son traitement (en anglais : data).

# Les données, c'est aussi tout ce qui circule dans un ordinateur

Les données ne sont pas seulement le fondement du savoir, elles sont aussi la base de l'informatique. Tout ce qui circule dans un ordinateur, ce sont des données.

Data inflation		
Unit	Size	What it means
Bit (b)	1 or 0	Short for "binary digit", after the binary code (1 or 0) computers use to store and process data
Byte (B)	8 bits	Enough information to create an English letter or number in computer code. It is the basic unit of computing
Kilobyte (KB)	1,000, or $2^{10}$ bytes	From "thousand" in Greek. One page of typed text is 2KB
Megabyte (MB)	1,000KB; $2^{20}$ bytes	From "large" in Greek. The complete works of Shakespeare total 5MB. A typical pop song is about 4MB
Gigabyte (GB)	1,000MB; $2^{30}$ bytes	From "giant" in Greek. A two-hour film can be compressed into 1-2GB
Terabyte (TB)	1,000GB; $2^{40}$ bytes	From "monster" in Greek. All the catalogued books in America's Library of Congress total 15TB
Petabyte (PB)	1,000TB; $2^{50}$ bytes	All letters delivered by America's postal service this year will amount to around 5PB. Google processes around 1PB every hour
Exabyte (EB)	1,000PB; $2^{60}$ bytes	Equivalent to 10 billion copies of <i>The Economist</i>
Zettabyte (ZB)	1,000EB; $2^{70}$ bytes	The total amount of information in existence this year is forecast to be around 1.2ZB
Yottabyte (YB)	1,000ZB; $2^{80}$ bytes	Currently too big to imagine

The prefixes are set by an intergovernmental group, the International Bureau of Weights and Measures.  
Yotta and Zetta were added in 1991; terms for larger amounts have yet to be established.

Source: *The Economist*

Pensez à votre abonnement téléphonique, chaque mois, vous payez pour consommer un certain volume de données quantifié en octet ou en bit.

Le volume des données créées et traitées ne cesse de croître en même temps que les capacités de calcul des ordinateurs.

# **2 - Histoire du mouvement de l'open data**

# Open data : un terme récent

Le terme est apparu pour la première fois dans les années 1970 dans les accords qu'a signés la NASA avec des pays partenaires en vue du partage de données satellitaires.

C'est en 1995 qu'on en voit le premier usage public aux Etats-Unis dans un rapport de la National Academy of Science intitulé "*On the Full and Open Exchange of Scientific Data*". Dans le langage courant, cela fait **près de 10 ans** que l'ouverture des données est devenue un élément majeur des politiques publiques numériques.

# Open data, les débuts

Au début des années 2010, c'est dans le cadre d'un **mouvement international** que l'ouverture des données publiques émerge en France. L'initiative vient du monde anglo-saxon, **les États-Unis et le Royaume-Uni mettant en place au même moment des plateformes gouvernementales de partage des données publiques, dans une démarche d'open data.**

Les objectifs de ce mouvement d'ouverture des données publiques sont à la fois

- **citoyens** : une plus grande transparence de l'action publique garantit un meilleur contrôle démocratique
- **économiques** : les données de l'administration peuvent être le support de nombreuses activités économiques et devraient être diffusées aussi largement que possible.

Le contexte est alors marqué par l'essor considérable des technologies et de l'économie numérique.

**L'open data s'inscrit dans une série de continuités historiques qui en font un mouvement aux multiples facettes...**

Il est important de connaître ses fondements philosophiques pour mieux comprendre ses déclinaisons locales et verticales.

# L'influence de la culture du logiciel libre

Le mouvement de l'open data est largement **influencé par le mouvement du logiciel libre** qui le précède de quelques décennies, et dont les créateurs participèrent à en dessiner les contours, notamment en 2007, lors de la conférence de Sebastopol sur laquelle nous reviendrons.

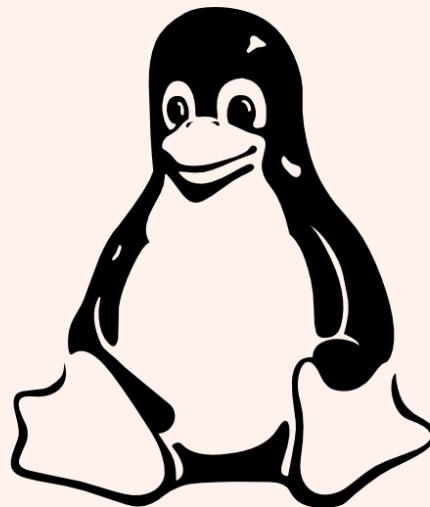
## Comment ?

- Historiquement venu des pratiques des chercheurs en informatique des années 70-80, le partage d'information, de données est un modèle de la recherche scientifique de l'époque.
- À partir de 1980, des licences d'utilisation des logiciels prévues pour limiter la liberté de l'utilisateur apparaissent.
- 1983 : Réaction d'un chercheur du MIT (Richard Stallman) avec le projet GNU (le logiciel, une forme de connaissance scientifique, qui doit être partagée) formalise le concept de Logiciel Libre.
- 1985: Création de la Fondation pour le Logiciel Libre et rédaction de la General Public License, GPL

# Les principes du logiciel Libre

## Les quatre libertés fondamentales

- La liberté d'utilisation d'un programme
- La liberté d'étudier le fonctionnement d'un programme
- La liberté de modifier un programme
- La liberté de distribuer un programme



Il faut penser à travers ces principes à la libre expression et liberté d'utilisation, non pas à la gratuité

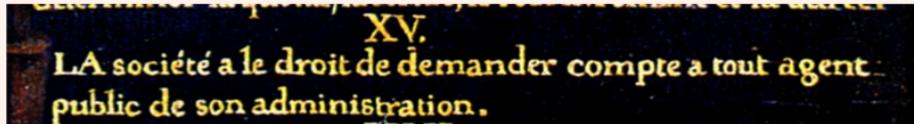
# **Open data : un mouvement aux multiples facettes**

**Brève introduction aux principales facettes :**

- La transparence
- La libre circulation de l'information
- La demande de données brutes
- L'industrie de la donnée
- La modernisation des administrations

# Facette n°1 : la transparence

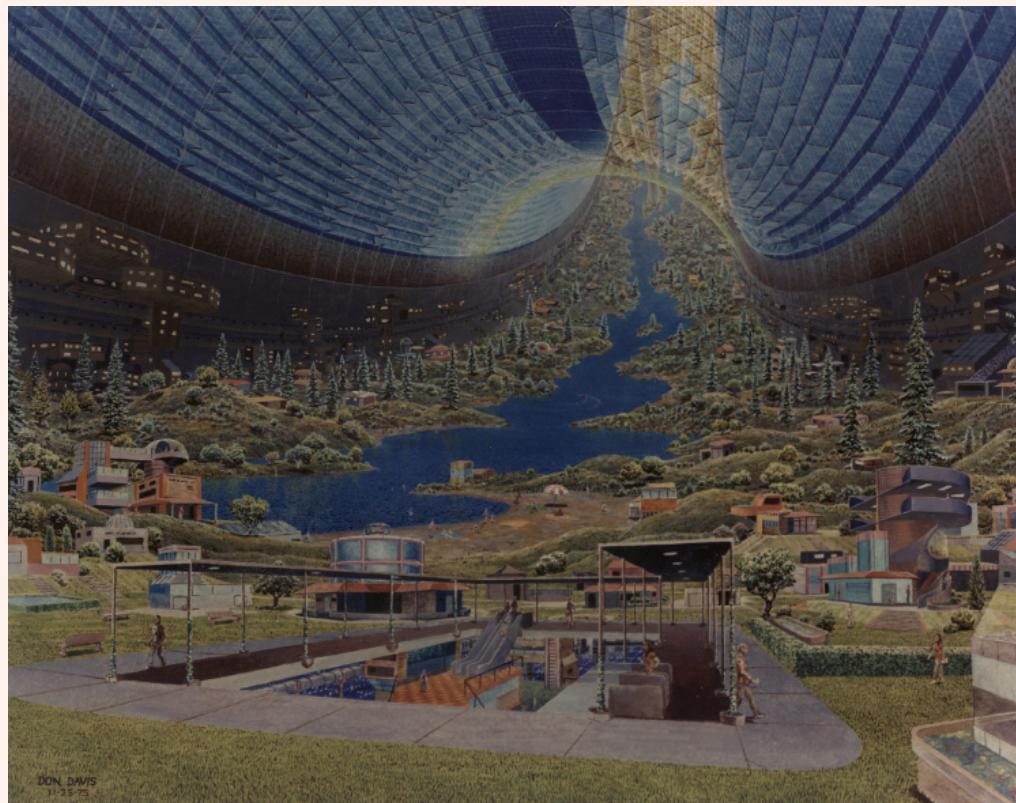
Aller au-delà du droit d'accès et de la réédition des comptes (*accountability*) : réduire les asymétries d'information en accédant au matériau brut de l'information et du savoir



AUX REPRESENTANS DU PEUPLE FRANCOIS ~

## Facette n°2 : la libre circulation de l'information

Dans la lignée de la cybernétique et du logiciel libre, la circulation des données comme valeur essentielle



# Facette n°3 : la demande de données brutes

Dans la lignée des premières initiatives de partage de données scientifiques en génétique ou en astronomie, obtenir les données sous leur forme pré-interprétable pour refaire les calculs

1	6650	623	8960	-588	3750	3648	489	926	678	37378
2	455	545	5511	552	5552	851	54	233	1888	4568
3	656	321	4598	564	654	321	381	3134	356	569
4	890	9023	8734	789	234	6789	2678	789	234	978
5	22	322	1112	322	322	322	322	322	322	322
6	388	3866	623	623	3888	3888	3888	3888	3888	3888
7	662	3222	422	422	6622	6622	6622	6622	6622	6622
8	20045	6222	3222	3222	20045	20045	20045	20045	20045	20045
9	656	47004	3222	3222	656	656	656	656	656	656

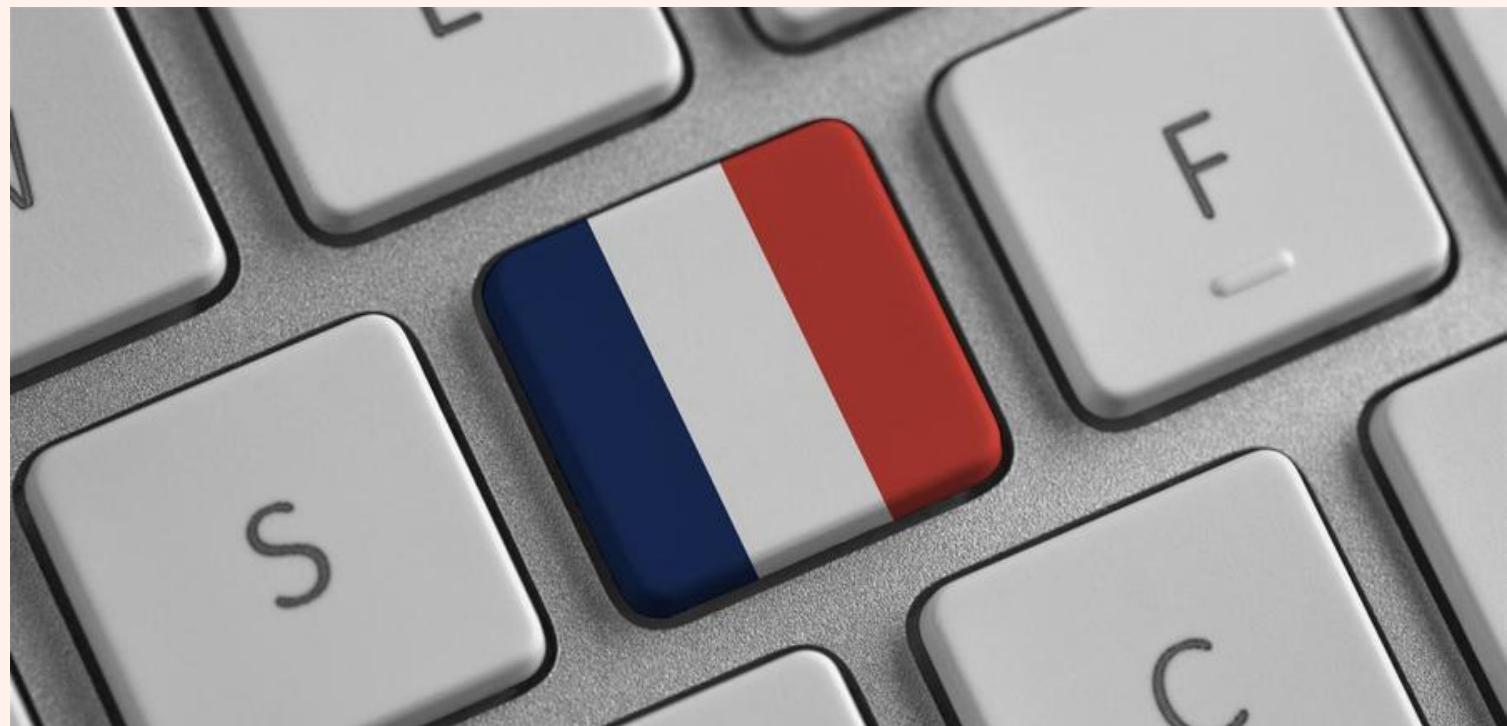
# Facette n°4 : l'industrie de la donnée

Le nouveau pétrole et les promesses démesurées de croissance



# **Facette n°5 : la modernisation des administrations**

Désiloter l'administration et ouvrir l'innovation, collaborer avec la société civile dans l'élaboration et la mise en œuvre des politiques publiques



# Que signifie “Open” dans un contexte culturel ?

- **L'Open Content** (les contenus ouverts) est la mise à disposition gratuite sur Internet des reproductions numériques des œuvres sous une licence garantissant leur libre accès et leur réutilisation par tous, sans restriction technique, juridique ou financière.
- **L'Open Access** (le libre accès) est la mise à disposition gratuite en ligne de contenus numériques, sous licences libres (Creative commons, etc.) ou non.
- **L'Open Knowledge** (la culture libre) est un mouvement social qui promeut la liberté de distribuer et de modifier des œuvres de l'esprit sous la forme d'œuvres libres. Il puise sa philosophie dans celle du logiciel libre en l'appliquant à la culture et à l'information.
- **L'Open Data** (les données ouvertes), est la mise à disposition gratuite sur Internet de données sous une licence libre.

# **3 - Les principaux enjeux de l'open data**

# **La transparence de la vie publique et la participation citoyenne à l'action publique**

- Exemple avec les **Données essentielles des conventions de subvention**,  
année 2018

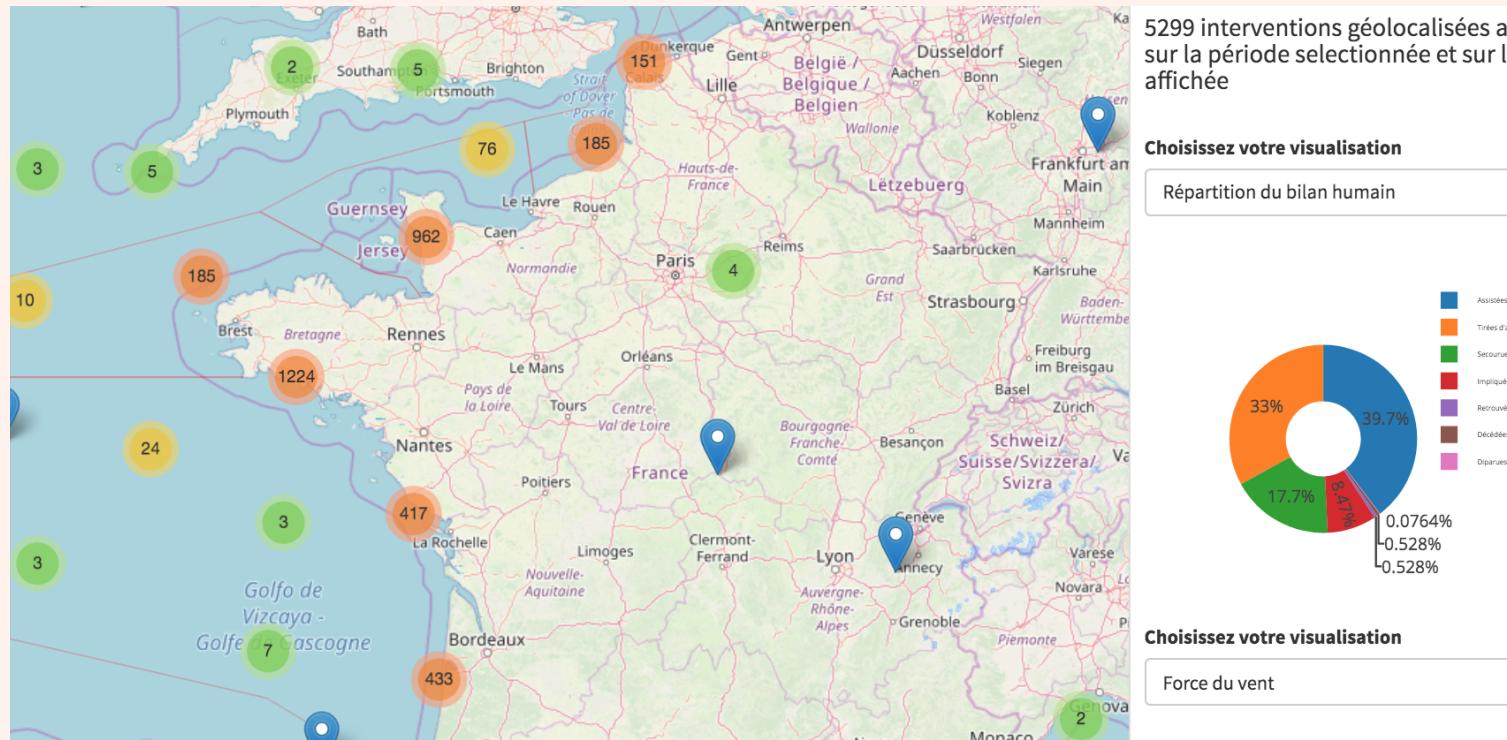
# La participation citoyenne à l'action publique

- Exemple : le **Grand Débat National**, mars 2019



# L'efficacité de l'action publique : amélioration de la qualité des services publics

- Exemple : sauver des vies en mer en prévenant les accidents maritimes



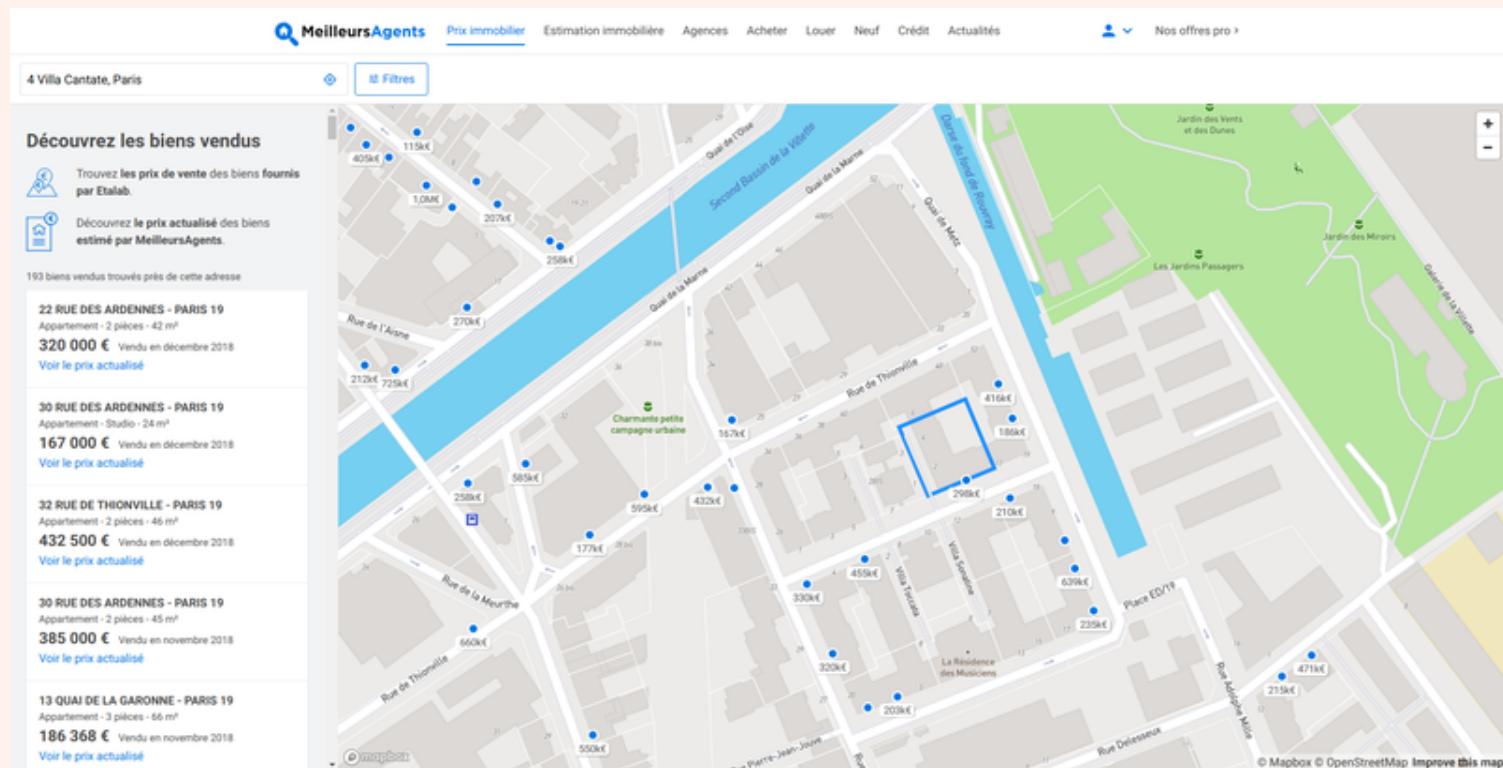
# Le développement de l'attractivité des territoires

- Le portail **DATAtourisme**



# Le soutien à l'innovation économique et sociale

- Les données des **valeurs foncières** reprises par le site meilleurs agents, avril 2019



# Une opportunité en matière de démocratisation culturelle et de transmission des savoirs

- Le projet **UrbanHist** à Toulouse



# **Origines historiques de l'open data dans le monde et en France**

# **Quelques dates jalonnent dans l'histoire mondiale de l'open data**

# 2005 : L'Open Definition

Une définition juridique des droits de l'usager d'un savoir ouvert

## L'Open Definition

L'[Open Definition](#) énonce les principes définissant l'“ouverture”, qu'il s'agisse de données ou de contenu.

Le définition précise le sens du terme “ouvert” dans les expressions “open data” et “open content” garantissant ainsi la qualité et encourageant la compatibilité entre différentes ensembles de matériaux ouverts.

La définition peut se résumer dans la déclaration suivante :

*Est considéré comme "ouvert" des données ou du contenu pouvant être librement utilisé, modifié et diffusé par quiconque et à quelque fin que ce soit"*

[Lire l'Open Definition en détails »](#)

# 2007 : la rencontre de Sebastopol aux États-Unis

**Quoi ?** Une rencontre de l'Open Government Group à Sebastopol (Californie), le siège des éditions O'Reilly

**Pourquoi ?** Influencer le futur président des Etats-Unis pour faire avancer l'open data

**Comment ?** En adoptant une déclaration définissant les grands principes de l'Open Government Data

Ces principes sont fondamentaux dans le futur développement de politiques nationales d'ouverture des données. Ils sont issus de la culture du mouvement du Libre dont les protagonistes sont à l'origine de cette rencontre et réflexion.

# **2008 : Tim Berners-Lee "Des données brutes maintenant!"**

**La demande de données brutes sous les projecteurs**

# **Depuis 2010 : l'ère des classements**

**Le baromètre international de l'Open Data** est un classement de référence, tout comme l'Open Data Index de l'Open Knowledge Foundation, organisation internationale ayant pour objet de promouvoir la transparence des données publiques.

# Depuis 2010 : l'ère des classements

Rank	Place	Government Budget	National Statistics	Procurement	National Laws	Administrative Boundaries	Draft Legislation	Air Quality	National Maps	Weather Forecast	Company Register	Election Results	Locations	Water Quality	Government Spending	Land Ownership	Score
1	Taiwan	██████	██████	██████	████	██████	██████	██████	██████	██████	██████	██████	██████	██████	██████	██████	90%
2	Australia	██████	██████	██████	████	██████	████	████	████	████	██████	██████	██████	██████	██████	██████	79%
2	Great Britain	██████	██████	████	████	██████	████	██████	██████	████	██████	██████	████	████	████	████	79%
4	France	██████	██████	████	████	██████	████	████	████	████	██████	██████	████	████	████	██████	70%
5	Finland	████	██████	████	████	██████	████	████	██████	████	██████	██████	████	████	██████	██████	69%
5	Canada	████	██████	██████	██████	██████	████	████	██████	████	██████	██████	████	████	████	██████	69%
5	Norway	██████	██████	████	████	████	████	████	██████	██████	██████	██████	████	████	████	████	69%
8	New Zealand	██████	████	████	████	██████	████	████	████	████	██████	██████	████	████	████	██████	68%
8	Brazil	██████	██████	████	████	████	████	████	██████	██████	██████	██████	████	████	████	████	68%
10	Northern Ireland	████	████	████	████	██████	████	████	████	████	██████	██████	████	████	████	██████	67%
11	Denmark	██████	██████	████	████	████	████	████	████	████	██████	██████	████	████	████	████	65%
11	Mexico	██████	██████	████	████	████	████	████	████	████	████	████	████	████	██████	██████	65%

# 2013 : la charte du G8 sur l'open data



## Charte du G8 pour l'Ouverture des Données Publiques

### Préambule

1. Le monde assiste à la montée en puissance d'un mouvement planétaire favorisé par la technologie et les médias sociaux et stimulé par l'information — un mouvement au potentiel extraordinaire pour encourager l'émergence d'entreprises et de gouvernements plus responsables, efficaces, proactifs et efficents, et pour stimuler la croissance économique.
2. **L'ouverture des données publiques est au cœur de ce mouvement mondial.**

# 2015 : la charte internationale de l'open data

## Principes

[Home](#) / [Charter](#) / [Principles](#)

[ENG](#) | [ESP](#) | [FRA](#)

### CHARTE INTERNATIONALE SUR LES DONNÉES OUVERTES

*Les données ouvertes sont des données numériques accessibles dont les caractéristiques techniques et juridiques permettent la libre utilisation, réutilisation et redistribution par quiconque, en tout temps, en tout lieu.*

# Conclusion : des principes et une définition qui se normalisent

- **Une première définition du savoir ouvert avec l'Open Definition :**  
"Ouvert signifie que tout le monde peut librement accéder, utiliser, modifier et partager (*des données ou informations*) dans n'importe quel but, sous réserve d'exigences qui en préservent la provenance et l'ouverture."
- **Des principes appropriables par les gouvernements :** le travail du G8 en 2013 et plus tard de la Charte Internationale de l'Open Data rend les principes intégrables par les gouvernements.

Les données ouvertes deviennent des données auxquelles l'accès est totalement public et libre de droit, au même titre que l'exploitation et la réutilisation.

Les données ouvertes culturelles présentent des particularités juridiques que nous allons détailler par la suite

**Sur ces bases, la France déploie sa politique nationale de données ouvertes**

# **Les origines de l'open data en France**

# Les acteurs publics de l'open data

- **DITP**
- **DINUM**
- **DINUM - Etalab** : Crée en 2011, la mission coordonne la politique d'ouverture et de partage des données publiques, et anime la plateforme nationale data.gouv.fr
- **Mission Société Numérique de l'ANCT** : Structure les acteurs de l'accompagnement au numérique et les stratégies numériques des territoires, dont l'open data fait partie.

Au ministère de la Culture :

- **Depnum** (Département du numérique pour la transformation des politiques culturelles et de l'administration des données) au sein du Service du numérique est chargé d'identifier au ministère et dans les établissements culturels, les informations produites dans le cadre des missions de service public, de coordonner leur recensement, leur qualification, et la mise à disposition en données ouvertes qui ont vocation à être librement accessibles et gratuitement réutilisables. (*n'hésitez pas à le contacter pour toute question sur l'Open data*).

# Gouvernement Ouvert et Open Data

Rappel : l'open data n'est pas qu'une pratique d'ouverture et de publication de données, c'est également un **mouvement et une philosophie relativ à un meilleur accès à l'information publique**.

Plus largement, l'open data s'inscrit dans le mouvement du **gouvernement ouvert**, aussi appelé **OpenGov** en anglais. Celui-ci promeut :

- La plus grande **transparence de l'information** (notamment à travers l'open data), pour permettre entre autre le suivi des politiques publiques
- Une plus forte **participation des citoyens** dans l'élaboration des politiques publiques : à la fois en étant davantage consulté mais aussi en étant acteur dans la construction de celles-ci
- Une plus forte **collaboration** au sein des administrations publiques (casser les silos) et à l'extérieur (travailler avec des acteurs associatifs, privés) pour concevoir et mettre en oeuvre les politiques publiques

# **OGP/PGO : Partenariat pour un Gouvernement Ouvert**



En septembre 2011, un partenariat multilatéral s'est créé afin de promouvoir le gouvernement ouvert et traduire ses valeurs en des actions gouvernementales concrètes : **le partenariat pour un gouvernement ouvert** (PGO), aussi appelé **Open Government Partnership** (OGP)

Initialement fondé par 8 pays (Brésil, Indonésie, Mexique, Norvège, Philippines, Afrique du Sud, Royaume-Uni et États-Unis), le PGO en compte désormais 70, dont la France.

# **Le plan d'action national pour la France**

Le 17 décembre 2021, le 3e plan d'action national du Partenariat pour un gouvernement ouvert a été présenté par la ministre de la transformation et la fonction publiques.

[Lire le plan d'action 2021-2023](#)

# Data.gouv.fr

Data.gouv est la plateforme nationale ouverte des données publiques françaises. Elle a été conçue et inaugurée en 2011, par la mission Etalab, puis refondue en 2013.

On y trouve des milliers de jeux de données, provenant principalement de producteurs publics (Ministère de l'intérieur, Agence française de développement, Insee, Ville de Paris, Région Bretagne, CNIL...) mais aussi de producteurs tiers (OpenStreetMap, OpenFoodFacts...).

# Au-delà de data.gouv.fr

Data.gouv.fr est la plateforme **nationale** d'open data mais ce n'est pas la seule plateforme publique qui met à disposition des données ouvertes. Il existe :

- Des plateformes open data gérées par les **collectivités territoriales**. Par exemple, [Paris Data](#), [la Région Occitanie](#), l'agglomération de Saint-Malo...
- Des plateformes open data gérées par des **ministères**. Exemple : [data.culture.gouv.fr](#)
- Des plateformes open data gérées par des **entreprises**. Exemple [DataNova](#) (La Poste)
- Des jeux de données directement publiés sur le site des producteurs de données, sans plateforme dédiée.

# Les grandes fonctionnalités d'un portail open data

Un portail open data est un espace de présentation des données publiques ouvertes comprenant idéalement les grandes fonctions suivantes :

- recherche des jeux de données
- présentation détaillées des jeux de données
- accès aux données, sous différents formats, y compris API
- présentation graphique et cartographique des données quand cela est possible
- outils d'extraction ou de transfert des données
- dispositifs éditoriaux pour animer le projet open data

## **Et au Ministère de la Culture**

Le ministère de la culture s'est doté en avril 2013 d'une « feuille de route open data », comportant dix actions pour une stratégie numérique de diffusion et de réutilisation des données publiques numériques du secteur culturel.

# Les dates importantes à retenir

**2011** : Le ministère s'engage dans la politique gouvernementale en faveur de l'ouverture et du partage des données publiques. Les premiers jeux de données sont publiés par le Ministère sur [data.gouv.fr](http://data.gouv.fr)

**2013** : Adoption de la première feuille de route « open data culturel ». Elle est mise à jour en **2019**. Elle comporte 6 axes pour une stratégie numérique de diffusion et de réutilisation des données publiques numériques du secteur culturel.

**2014** : La BnF ouvre ses métadonnées descriptives en "Licence Ouverte"

**2014/2015** : Publication de la feuille de route «Métadonnées culturelles et transition Web 3.0»

**2016** : Ouverture du portail [data.culture.gouv.fr](http://data.culture.gouv.fr))

**2017** : Les Archives nationales optent pour l'ouverture et la gratuité des données

**2021** : Une nouvelle feuille de route "Données et contenus culturels" et nomination d'un AMDAC

# **La feuille de route Données et contenus culturels**

Cette **feuille de route**. Elle concerne les données, les algorithmes et les codes sources.

La publication des feuilles de route est le fruit d'une volonté politique renouvelée et portée par le Premier ministre qui s'est vu remettre le 23 décembre 2020 le rapport de la mission parlementaire présidée par Eric Bothorel et par la publication de la circulaire d'avril 2021 du Premier Ministre.

Elle s'articule autour de trois axes prioritaires :

- Améliorer la capacité du ministère à diffuser et rendre visible les contenus culturels ;
- Contribuer à l'efficacité des actions du ministère et à la simplification des relations entre les usagers et le service public culturel ;
- Faire bénéficier un secteur économique de la culture fragilisé par la crise COVID-19 des opportunités de reprise et de développement présentées par l'usage des données.

# Pour les services centraux, les DRAC et les SCN : des données hébergées sur [data.culture.gouv.fr](http://data.culture.gouv.fr)

**data.culture.gouv.fr**

la plate-forme de données ouvertes du ministère de la Culture

Pour les données de l'administration centrale, d'une DRAC ou d'un SCN, la mise à disposition des données est à effectuer sur le site [data.culture.gouv.fr](http://data.culture.gouv.fr).

Afin d'améliorer la visibilité des données culturelles, **le portail data.culture.gouv.fr est moissonné automatiquement par data.gouv.fr**, pour pouvoir les exposer sur les deux plateformes à la fois.

*Pour toute information sur la mise à disposition de nouveaux jeux de données, contacter le département du numérique pour la transformation des politiques culturelles et de l'administration des données (Service du numérique) via l'onglet « Contact » de la plateforme data.culture.gouv.fr.*

# **Pour les établissements publics : des données hébergées sur [data.gouv.fr](#) et visibles (aussi) sur [data.culture.fr](#)**

Pour les établissements publics, la mise à disposition des données est à effectuer sur le site [data.gouv.fr](#).

Afin d'améliorer la visibilité des données culturelles, le Ministère référence également les données des établissements sur la plateforme ministérielle [data.culture.gouv.fr](#).

Cela implique la création d'une page dédiée à l'établissement, puis la publication des jeux de données conformément [aux modalités précisées par la DINUM](#)

**Le ministère mettra en place les mécanismes nécessaires pour que toutes les données publiées de cette façon soient ensuite automatiquement visibles sur [data.culture.gouv.fr](#).**

# Culture : principales catégories de données concernées par l'ouverture

- Les **données statistiques** des institutions culturelles : fréquentation, entrées, équipements
- Les **bases de données des collections** : métadonnées et plus largement les données associées aux œuvres (localisation, description, auteur, œuvres associées, prix, propriétaire) avec images ou sans image
- Les **listes de structures** : musées, monuments, cinéma, etc.
- Les référentiels ou nomenclatures
- Les données de type **agenda**, événements, programmations
- Les **données financières** : financement, ressources, budgets annuels, subventions

Important : les œuvres protégées par le droit d'auteur ou un droit voisin ne peuvent faire l'objet d'un droit de réutilisation, mais les œuvres entrées dans le domaine public relèvent pleinement de l'ouverture des données publiques.

# **4 - Éléments juridiques relatifs à l'ouverture des données**

# Du droit d'accès au droit de réutilisation des données publiques en France

Trois grandes étapes caractérisent en France la **construction d'un droit de l'ouverture des données publiques** :

- 1 - La création du droit d'accès aux documents administratifs
- 2 - La création du droit de réutilisation des informations publiques
- 3 - Le développement d'une politique active d'ouverture des données, assorti d'un essor de la gratuité.

# Un appareil législatif étoffé

1. **Loi Informatique et liberté** du 6 janvier 1978
2. **Loi CADA** du 17 juillet 1978
3. **Directive INSPIRE** de 2007
4. **Directive PSI** de 2013
5. **Loi Valter** de 2015
6. **Loi pour une République Numérique** (loi Lemaire) du 7 octobre 2016.

2 éléments cohabitent dans ces textes : le droit d'accès et le droit de réutilisation

# 1978 : La loi CADA, vers le "droit de savoir"

- Le fondement : la **Déclaration des Droits de l'Homme et du Citoyen de 1789** dans son article 15, "la Société a le droit de demander compte à tout Agent public de son administration."
- Le droit d'accès des citoyens à l'information publique émerge en **1978 avec la loi dite CADA** du nom de la Commission d'Accès aux Documents Administratifs.
- La France était le **3e pays au monde** après la Suède en 1766 et les Etats-Unis en 1966 avec le Freedom of Information Act (FOIA) en 1966 à accorder un "droit de savoir" avec pour but d'améliorer les relations entre le public et l'administration.

La loi CADA est intégrée au Code des relations entre le public et l'administration (CRPA) qui, dans son **livre 3** codifie le droit d'accès et de réutilisation

# Les conditions d'exercice du droit d'accès aux documents administratifs

Une définition **très large** des documents administratifs

Les documents administratifs peuvent revêtir de nombreuses formes (dossiers, rapports, études, comptes rendus, procès-verbaux, statistiques, directives, instructions, circulaires, codes sources, etc.) et adopter tout support (écrit, enregistrement sonore ou visuel, forme numérique ou informatique).

Le droit d'accès s'applique dans le cadre d'une **mission de service public** ie :

Une mission d'intérêt général sous le contrôle de l'administration et qui est dotée à cette fin de prérogatives de puissance publique est chargée de l'exécution d'un service public.

Il ne s'exerce que si l'administration a effectivement en sa possession le document demandé, que si le document est formellement **achevé**, qu'il existe et qu'il n'est pas **préparatoire à une décision administrative** en cours.

# Les exclusions au droit d'accès

Le droit d'accès **ne doit pas porter atteinte aux secrets protégés** :

- le fonctionnement de l'État : délibérations du Gouvernement, secret de la défense nationale, la conduite de la politique extérieure de la France, la sûreté de l'État...
- la protection de la vie privée
- le secret médical
- le secret des affaires (secret des procédés, des informations économiques et financières et des stratégies commerciales ou industrielle)
- les mentions portant une appréciation ou un jugement de valeur sur une personne physique ou faisant apparaître un comportement d'une personne pouvant lui porter préjudice...

Le document ne peut être publié qu'après avoir fait l'objet d'un traitement permettant d'**occulter les mentions protégées**. L'administration n'est pas tenue de publier un document lorsque "les travaux d'occultation dénatureraient ou videraient de sens le document" ou si "le document est indivisible ou l'occultation est trop complexe".

# **La création du droit de réutilisation des informations publiques avec la Directive ISP/PSI**

- Partant du constat que le secteur public produit et diffuse un grand nombre d'informations et qu'elles « constituent une matière première importante pour les produits et services de contenu numérique », la directive PSI de 2003 concernant la réutilisation des informations du secteur public se fonde sur des considérations économiques. **Elle introduit la notion de réutilisation commerciale et non-commerciale.**

En France, elle est transposée par l'ordonnance du 6 juin 2005, qui insère dans la loi CADA de 1978 un nouveau chapitre relatif à la réutilisation des informations publiques.

# Le droit de réutilisation des informations publiques

Le droit de réutilisation issu de la directive PSI se combine donc en France avec le droit d'accès aux documents administratifs reconnu depuis 1978 : les informations contenues dans tout document administratif communicable sont, en principe, réutilisables « par toute personne qui le souhaite à d'autres fins que celles de la mission de service public pour les besoins de laquelle les documents ont été élaborés ou sont détenus ».

Par cette combinaison, le droit français garantit ainsi une liberté effective de réutilisation bien plus grande que celle imposée par la directive ISP. Le principal apport de la transposition au droit français est d'avoir mis fin à l'interdiction de réutilisation des documents administratifs à des fins commerciales.

# **La loi pour une République Numérique : une obligation de diffusion et un accès actif aux données**

- La loi est promulguée le **7 octobre 2016**

Elle a pour objectifs de :

- Encourager l'innovation et l'économie numérique
- Promouvoir une société numérique protectrice
- Garantir l'accès de tous au numérique (Haut débit..)

Source

# L'ouverture des données par défaut, un premier principe fondamental de la Loi Lemaire

Les administrations devront publier en ligne dans un standard ouvert leurs principaux documents, y compris leurs codes sources, ainsi que leurs bases de données qui présentent un intérêt économique, social, sanitaire ou environnemental.

Cette obligation concerne toutes les *administrations*, établissements publics, collectivités et organismes privés chargés d'un service public, à deux conditions : être doté d'au moins **50 agents** et, pour les collectivités locales, concerner au moins **3500 habitants**.

# Le périmètre de l'open data par défaut

L'ouverture concerne un **périmètre très large** :

- les documents communiqués suite à des demandes CADA
- les "bases de données"
- les données "dont la publication présente un intérêt économique, social, sanitaire ou environnemental".

Rares sont donc les données publiques qui ne sont pas concernées par cette obligation d'ouverture.

# **Création d'un service public de la donnée**

- La loi souhaite faciliter la réutilisation des principales bases de données de l'État par les acteurs privés ou publics (entreprises, associations, chercheurs, etc.), en leur garantissant un niveau élevé de qualité de service.
- La loi propose de mettre à disposition comme "service public" les grandes bases de données de référence de l'État , comme par exemple le répertoire SIRENE des entreprises ou le cadastre.

# Les 9 jeux de données du Service Public de la Donnée

- Le répertoire des entreprises et des établissements (base Sirene)
- Le répertoire national des associations (RNA),
- Le plan cadastral informatisé,
- Le registre parcellaire graphique (RPG)
- Le référentiel à grande échelle (RGE),
- La base adresse nationale (BAN),
- Le répertoire opérationnel des métiers et des emplois (ROME),
- Le référentiel de l'organisation administrative de l'Etat,
- Le code officiel géographique (COG).

Le service public de la donnée sur [data.gouv.fr](http://data.gouv.fr)

# La LRN encourage l'utilisation de standards ouverts

Afin de faciliter la réutilisation des données par les machines, la loi pour une République Numérique consacre **l'utilisation de standards ouverts** pour la communication des documents administratifs.

L'administration est dorénavant tenue, lorsqu'elle communique un document administratif au format électronique, de le mettre à disposition du citoyen **"dans un standard ouvert, aisément réutilisable et exploitable par un système de traitement automatisé"**.

Concrètement, cela devrait empêcher la publication de fichiers PDF là où un fichier CSV pourrait être diffusé.



# Ainsi qu'une limitation des licences

La loi **limite par décret les licences** qui peuvent être utilisées par les administrations. Le [décret du 27 avril 2017](#) autorise deux licences :

- la Licence Ouverte d'Etalab
- la licence ODbL (Open Database License).

Ces deux licences offrent **un choix aux administrations** : la licence ouverte étant plus permissive que la licence ODBL qui a été conçue dans une logique de "pot commun" permettant de garantir que les données resteront ouvertes après leur réutilisation.

Utiliser une autre licence impose une **homologation** auprès de la DINUM avec un exposé des motifs ayant conduit à sortir de la liste fixée dans le décret et une consultation des usagers affectés par la licence proposée. L'homologation doit être faite pour chaque jeu de données même si la licence reste la même.

# Zoom sur la Licence Ouverte et l'ODBL



## Panorama des principaux types de licences

### 1 La licence ouverte

L'utilisateur est libre de :

- Partager (copier, distribuer et utiliser) la base de données
- Produire des créations à partir de la base de données
- Adapter (modifier, transformer, construire) la base de données
- Faire un usage commercial

Tant qu'il :

- Mentionne la paternité et la date de dernière mise à jour

### 2 La licence ODBL

L'utilisateur est libre de :

- Partager (copier, distribuer et utiliser) la base de données
- Produire des créations à partir de la base de données
- Adapter (modifier, transformer, construire) la base de données
- Faire un usage commercial

Tant qu'il :

- Mentionne la paternité
- Partage aux conditions identiques
- Garde ouverte la base de données

Source : <https://www.etalab.gouv.fr/licence-ouverte-open-licence>

Source : <http://vvlibri.org/fr/licence/odbl/10/fr>

# **Les données à exclure du périmètre de l'open data**

**Les informations sur lesquelles des tiers détiennent des droits de propriété intellectuelle**

**Les informations contenant des données à caractère personnel ou pouvant affecter la vie privée des citoyens**

**Les données couvertes par un secret**

- Exemple : Vous voulez publier des photographies ? Il faut vérifier que :
  - Soit, les images appartiennent au domaine public
  - Soit, vérifier si un contrat de cession de droits existe entre l'établissement et le/la photographe
  - Sinon, l'auteur (ou ayant-droit) dispose de droits sur l'image

# Le cas des données personnelles

La réutilisation de données à caractère personnel doit être faite dans le respect de la règlementation sur les données à caractère personnel (loi relative à l'informatique, aux fichiers et aux libertés et au RGPD).

La mise en ligne de données à caractère personnel sans anonymisation n'est permise que dans **trois cas** :

1. une disposition législative contraire le prévoit (exemple : [l'article D312-3-1 du CRPA](#))
2. si les personnes intéressées ont donné leur accord (consentement au sens de l'article 4 et 7 du RGPD)
3. Les données ont été [anonymisées](#) : elles ne doivent plus se référer à une personne réelle (exit les noms, identifiants...).

# La réutilisation des données et le régime dérogatoire pour le secteur culturel

Un principe : **la gratuité** et un régime dérogatoire pour le secteur culturel.

- Lorsque la réutilisation porte sur *des informations issues des opérations de numérisation des fonds et des collections des bibliothèques, y compris des bibliothèques universitaires, des musées et des archives et, le cas échéant, sur des informations qui y sont associées lorsque ces dernières sont commercialisées conjointement*. Dans ce cas, une **redevance** peut être mise en place, en respectant une règle de calcul prévue par la réglementation.

Les données publiques nativement numériques ne peuvent pas faire l'objet d'une redevance.

# Réutiliser des informations publiques dans le domaine de la culture

Pour résumer :

- La licence ce n'est pas obligatoire
- La gratuité est obligatoire **sauf** exception

Mais, il faut garder à l'esprit que cela va dépendre du type de réutilisation :

- Dans le cas de réutilisations non commerciales : La licence ce n'est pas obligatoire, mais si c'est le cas : LO ou OdBL (ou une licence homologuée sur autorisation)
- Dans le cas de réutilisations commerciales : **si** une redevance est appliquée, l'adoption d'une licence spécifique est obligatoire

En juillet 2017, les Archives nationales se sont engagées dans l'Open data en consacrant le principe de gratuité pour toute réutilisation aux ressources libres de droits qu'elles conservent.

# **Prenons maintenant connaissance des principes généraux d'ouverture des données et regardons leur application concrète**

Tels que définis dans les "huit principes des données publiques ouvertes", définis les 7-8 décembre 2007 lors de la rencontre Sébastopol, citée plus haut. Source: [opengovdata.org](http://opengovdata.org)

## **1. Des données complètes**

**Toutes les données publiques doivent être rendues disponibles dans les limites légales liées à la vie privée ou la sécurité**

C'est le **principe d'open data par défaut**

## **2. Des données primaires**

**Les données ouvertes sont telles que collectées à la source, non-agrégées avec le plus haut niveau de granularité**

# Trouver des données granulaires

Le portail Data Culture basé sur la solution OpenDataSoft permet de trouver rapidement les données plus granulaires en triant les données par le nombre d'enregistrements dans le catalogue de données.

The screenshot shows the homepage of the data.culture.gouv.fr portal. At the top, there's a navigation bar with icons for home, 'Les données', 'La démarche', and 'Créer vos cartes'. Below this, a section titled '67 jeux de données' is displayed. A dropdown menu is open over this section, showing filter options: 'Alphabétique', 'Anti alphabétique', 'Récemment modifiés', 'Anciennement modifiés', and 'Le + d'enregistrements' (which is selected, indicated by a blue border). Other visible filters include 'Modifié', 'Populaires', 'De A à Z', 'Le - d'enregistrements', 'Le + de téléchargements', 'Le - de téléchargements', 'Les + populaires', 'Les - populaires', and 'Thèmes'. On the right side of the page, there's a sidebar with links for 'Tableau', 'Carte', 'Analyse', 'Export', and 'API'. The main content area features a large image of a document titled 'Collections des musées de France : extrait de la base Joconde' with the subtitle 'La base Joconde décrit les œuvres des musées de France (600 000 notices environ)'. Below the image, there's a table showing statistics for different domains: dessin (248 059), archéologie (73 624), peinture (70 378), ethnologie (53 156), estampe (50 113), and sculpture (40 535). A note at the bottom states: 'La géolocalisation est effectuée via la BAN sur le nom de la ville.'

On trouvera ainsi facilement la [base Joconde](#) qui décrit les œuvres des musées de France et contient environ 600 000 notices.

The screenshot shows the 'Collections des musées de France : extrait de la base Joconde' interface. At the top, it displays '592 279 enregistrements' and a note 'Aucun filtre actif'. Below this is a 'Filtres' section with a search bar and a 'Domaine' section listing categories with their counts: dessin (248 059), archéologie (73 624), peinture (70 378), ethnologie (53 156), estampe (50 113), and sculpture (40 535). To the right, there are links for 'Informations', 'Tableau', 'Carte', 'Analyse', 'Export', and 'API'. A note below the filters says: 'La base Joconde décrit les œuvres des musées de France (600 000 notices environ). La base Joconde est le catalogue collectif des collections des musées de France : <http://www2.culture.gouv.fr/documentation/joconde/fr/pres.htm>. Depuis mars 2019, elle est accessible sur la plateforme ouverte du patrimoine, POP.' A 'Liste des champs:' section follows, listing various metadata fields. At the bottom, a note states: 'La géolocalisation est effectuée via la BAN sur le nom de la ville.'

### **3. Des données *fraîches* (*timely*)**

**Les données doivent être disponibles dès qu'elles sont produites**

# Un exemple de données temps réel

La ville d'Issy les Moulineaux ouvre les **données temps réel de l'affluence** dans ses structures dont la médiathèque. Ces données sont exploitées notamment par l'application **Affluences**.

## **4. Des données accessibles**

**Les données doivent être utilisables par le plus grand nombre d'usagers potentiels**

## **5. Des données exploitable par les machines**

**Les données peuvent être traitées automatiquement par les machines**

## Exemple : le nombre de flashes par radars en France

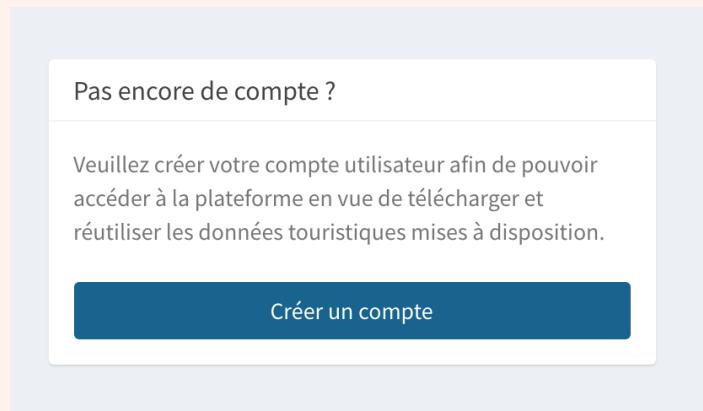
Le ministère de l'Intérieur publie les statistiques du nombre de flashes sur les radars en France mais sous la forme de fichiers PDF. Ces tableaux de données sont inexploitables par les machines, on ne peut pas les importer simplement dans un outil d'analyse de données.

A la suite d'une conversation sur [TeamOpenData](#), un citoyen a extrait ces données avec l'outil [Tabula](#) et les a republié sur [data.gouv.fr](#) dans un format lisible par les machines. Plusieurs cartes ont été produites à la suite.

## 6. Des données non discriminatoires

Elles peuvent être utilisées par tous sans réclamer un enregistrement préalable

Pour consulter ou télécharger les données du portail DataTourisme, il faut s'inscrire sur le site et créer un compte obligatoire.



Suite au signalement de cette obligation et à une [discussion sur TeamOpenData](#), ces données ont finalement été rendues accessibles à tous sur le portail [data.gouv.fr](http://data.gouv.fr)

A screenshot of the data.gouv.fr website. The top navigation bar includes the French national emblem and the text "data.gouv.fr". Below the navigation is a banner with the text "Plateforme ouverte des données publiques françaises". A search bar is visible. The main content area features a large blue box with the text "DATAtourisme, la base nationale des données du tourisme en Open Data".

## **7. Des données publiées dans un format ouvert**

**Ce format ne doit pas être la propriété d'une organisation en particulier (.xls) et doit être gouverné par ses usagers**

# Exemples courants de fichiers excels publiés

A éviter : police ou colonnes en couleur pour signifier une légende particulière pour les données.

725320	MUSEE DEPARTEMENTAL STEPHANE MALLARMÉ	VULAINES-SUR-SEINE	
701720	MUSEE DE LA BATELLERIE	CONFLANS-ST-GERMAIN	
703200	MUSEE DE LA TOILE DE JONVY	JONVY EN JOSSELAIN	
703200	MUSEE DES TISSUS ET VILLA MEDICIS	LUCIEN-PIERRE-SAINTE-ADRESSE	
503700	MUSEE EMMILE JEAN	VILLERS-SUR-MARNE	
504800	MUSEE D'ART CONTEMPORAIN DU VAL-DE-MARNE	VITRY-SUR-SEINE	
502000	MUSEE D'ART ET D'HISTOIRE	ARGENTEUIL	
502000	MUSEE DES BEAUX-ARTS DE LA REMPARTANCE	COURBEVOIE	
902900	MUSEE ARCHEOLOGIQUE DU VAL D'ORNE	GUERRY EN YVELINES	
903100	MUSEE LOUIS SEMECQ	ISLE ADAM	
902500	MUSEE DE L'ESPRESSO, L'IMAGE, L'INFO ET D'ARCHIVES DE LA PRESSE	KOMMUNIQUE	
902500	MUSEE JEAN-JACQUES ROUSSEAU	KOMMUNIQUE	

<http://www.data.gouv.fr/DataSet/0303823877/mcu-frequentation-des-musees-de-france/excel>

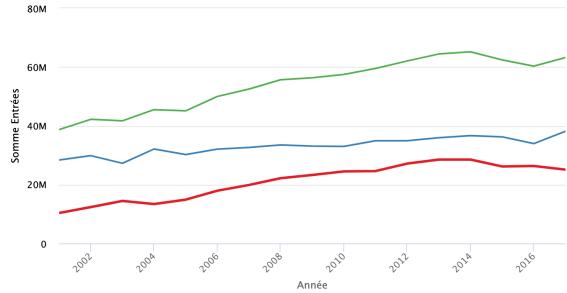
Ce fichier de la fréquentation des musées en France a depuis été mis à jour sur Data-Culture avec une infographie.

## Fréquentation des Musées de France

Informations Tableau Analyse Infographie Export API

### Tendances clés sur la fréquentation des musées

#### Evolution de la fréquentation des musées



## **8. Des données publiées sous une licence ouverte**

| Licences prévues par la réglementation

# 6 - Des exemples #OpenGLAM

# **Open GLAM : une communauté engagée pour l'ouverture des données et contenus culturels**

**OpenGLAM** est un réseau mondial de personnes et d'organisations engagées à ouvrir du contenu et des données détenues par des galeries d'art, des bibliothèques, des archives et des musées.



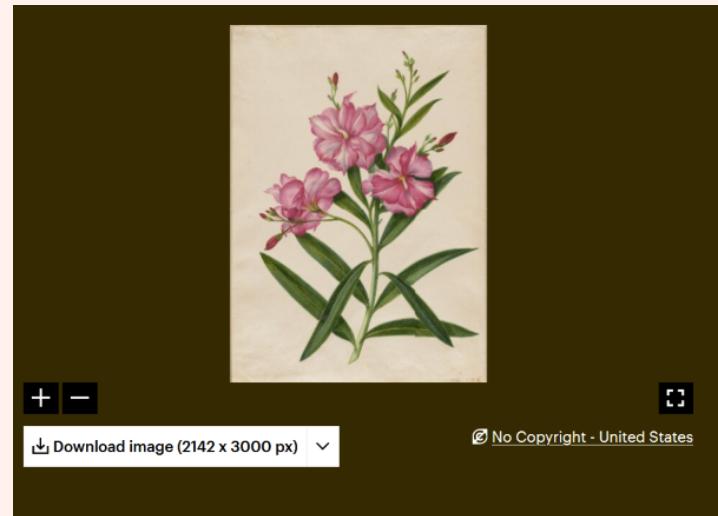
**GLAM** : Galleries, Libraries, Archives and Museums.

## **Des exemples internationaux**

# Le Getty Museum

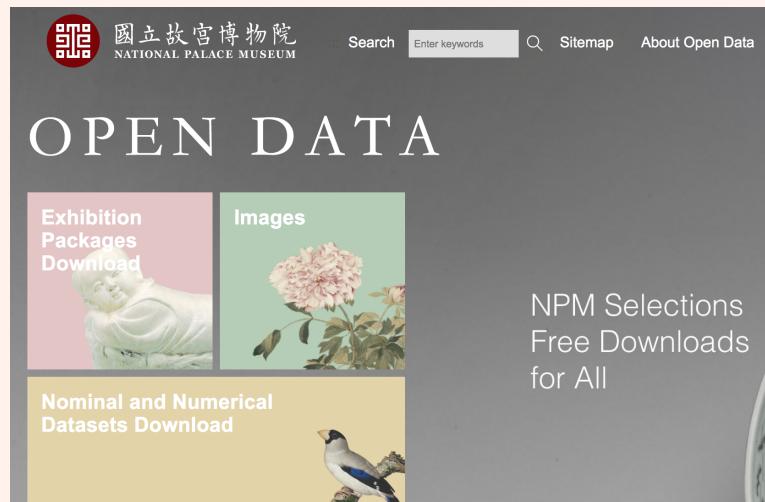
- En 2013, le Getty Museum place les reproductions de 4 600 de ses œuvres sous le régime de l'Open Content.
- Aujourd'hui, c'est près de 200 000 œuvres qui sont concernées.
- Les images, téléchargeables en HD, sont réutilisables pour tout usage, commercial ou pas.
- Licence : aucune licence (Domaine Public)

[En savoir plus](#)



# Le Palais national de Taiwan

Entre 2015 et 2017, le Palais national de Taiwan, a ouvert 70 000 images sur une plate-forme dédiée National Palace Museum Open Data, qui permet de télécharger sans restriction des images d'œuvres et des jeux de données sur ses expositions. Ceci a été réalisé au terme d'un grand projet de numérisation de ses collections et en lien avec la politique de gouvernement ouvert du pays.



Licence : Open government data licence ou CC-BY-4.0

[En savoir plus](#)

# Rijksstudio

Lancé en 2012 pendant la période de travaux du **Rijksmuseum d'Amsterdam**, le Rijksstudio est d'abord le catalogue en ligne des collections de l'institution.

The screenshot shows the Rijks Studio website. At the top, a large painting of a windmill by Johannes Vermeer serves as the background. Overlaid on it is the text "RIJKS STUDIO" in large white letters. In the bottom right corner of the main image, there is a small "Made possible by" logo. Below the main image is a black bar containing the text "Discover the possibilities of the masterpieces" and a red button labeled "Create your own Rijksstudio >". The main content area features a grid of nine smaller images of historical paintings. On the left, there are two images: one by Rembrandt van Rijn showing a group of people in a dark setting, and another by Johannes Vermeer showing a man and a woman in a room. In the center, there are three images: one by Johannes Vermeer showing a street scene with a brick building, one by Vermeer showing a woman in a blue dress, and one by Rembrandt van Rijn showing a woman in a green dress. On the right, there are four images: one by Rembrandt van Rijn showing a group of people in a room, one by Vermeer showing a group of people in a room, and two other smaller images of figures. Below each row of images are captions: "Rembrandt van Rijn" and "RIJKS MUSEUM" under the first row; "Johannes Vermeer" and "RIJKS MUSEUM" under the second row; and "Paintings" and "RIJKS MUSEUM" under the third row.

Licence : ?

[En savoir plus](#)

Open Data, comprendre les enjeux de l'ouverture des données publiques culturelles

# Cleveland Museum of Arts

Le 23 janvier 2019, le Cleveland Museum of Art a annoncé qu'il devenait une institution en libre accès, utilisant la désignation Creative Commons Zero (CC0) pour les images et les données haute résolution liées à sa collection.

Les + de la démarche : tableau de bord des effets de la [stratégie open content](#) et contenus 3D ouverts.

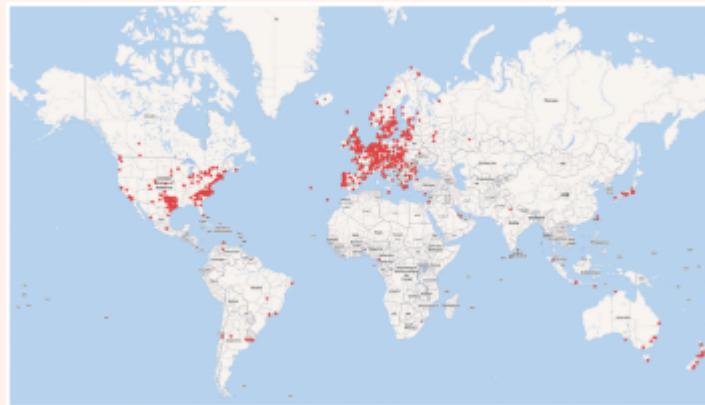
Licence : CC0

[En savoir plus](#)

# Mais aussi...

- Le MET, CC0
- Le Smithsonian, CC0

Carte de Douglas McCarthy, *Global instances of Open GLAM, March 2022*



## **Des exemples en France**

# **POP, la plate-forme ouverte du patrimoine du Ministère de la Culture**

La plateforme POP regroupe les contenus numériques du patrimoine français afin de les rendre accessibles et consultables au plus grand nombre sur : [pop.culture.gouv.fr](http://pop.culture.gouv.fr)

Plus de 3,3 millions de fiches y sont actuellement présentes, issues des bases historiques du ministère.

- POP propose plusieurs types de recherche : une recherche par texte libre, par facettes, une recherche cartographique ainsi qu'une recherche avancée par base de données.
- Conçue et développée dans le cadre d'une start-up d'État soutenue par les services du Premier ministre, POP a vocation à remplacer l'ancien système de bases de données du ministère de la Culture, utilisé depuis les années 1970.
- La plate-forme permet de verser l'essentiel des métadonnées (à ce stade pas les images) sur les plateformes [data.culture.gouv.fr](http://data.culture.gouv.fr) et [data.gouv.fr](http://data.gouv.fr)

# Paris Musées

Le 8 janvier 2020, Paris Musées lançait sa politique d'Open Content en proposant des reproductions numériques des œuvres des collections des musées de la Ville de Paris. Un an après ils annoncent : 2,5 millions d'images téléchargées en 2020 et 250 000 images disponibles.

Licence : CC0

[En savoir plus](#)



ZOOM  
+



TÉLÉCHARGER



# Archives Nationales

En 2017, les Archives Nationales ont fait le choix de l'open content en soumettant les documents qu'elles détiennent au principe de gratuité totale.

Licence : aucune

[En savoir plus](#)



AN, CP/F/12/4055/D/C, Pièce 25  
*"Exposition universelle de 1889".*  
*"Histoire de l'habitation Phénicie :*  
*façade principale"*

# Les collections en partage du musée de Bretagne

En 2017, la plateforme [collections.musee-bretagne.fr](http://collections.musee-bretagne.fr) propose ses collections en téléchargement gratuit. Une volonté de la métropole de Rennes de rendre accessibles les collections, participant ainsi à la consolidation de sa politique d'ouverture des données de la collectivité. La licence appliquée autorise la réutilisation des documents (images et notices), y compris à des fins commerciales.

Licence : de différents types



*Fileuse (Morbihan) – Négatif sur verre –  
Jeanne-Marie Barbey (1876-1960),  
début 20ème siècle – CC-BY-SA*

# **DoReMus : DOing REusable MUsic Data**

Un projet de recherche en collaboration **avec de multiples partenaires culturels**

Le projet s'appuie sur **l'expertise de catalogage des œuvres musicales** de :

- Radio France
- la BnF
- la Cité de la musique
- du projet collaboratif Discotheka de la société Meaning Engines
- et sur un corpus de modèles préexistants, en particulier FRBRoo.

**Un moteur de recherche liant les données sur les dates des performances, les interprètes, les enregistrements, les instruments, les partitions, etc..**

- **Site du projet** : <http://data.doremus.org/>
- **Github du projet** : <https://github.com/DOREMUS-ANR>

Licence : CC-BY 4.0

# SocFace

Projet de mise en données des recensements de population détenus par les services d'archives départementales de 1836 à 1936. Le projet est de développer des technologies de reconnaissance automatique d'écriture manuscrite pour analyser l'ensemble des listes nominatives et produire une base de données des individus ayant vécu en France durant cette période.

Projets multipartenarial et réalisé dans le cadre d'une ANR.

L'open access et la mise à disposition des données est prévue dans le cadre du projet.

[En savoir plus](#)

# **Quels sont les éléments important de ces démarches ?**

# Nouer des partenariats

Afin de donner de la visibilité ses collections, le département de Haute-Saône, a **réalisé un partenariat avec Wikimédia**, pour diffuser et valoriser **les données d'images haute définition des collections départementales** sous licence libre.

Le château de Champlitte, qui abrite le musée départemental d'Arts et Traditions populaires.

*Jérôme BARDIAU, CC BY-SA 4.0, via Wikimedia Commons*



# Verser une collection sur une plateforme collaborative

- Possibilités pour la recherche, l'innovation et la créativité de cette rencontre entre des collections de qualité bien documentées et l'encyclopédie collaborative en ligne.
- **Rendre accessible une collection tandis que le bâtiment est fermé** ou en l'absence (c'est le cas de petits musées) de portail des collections
- **Donner une visibilité aux œuvres** par leur remploi dans l'écriture des articles de Wikipédia, ou les inscrire dans le référencement sémantique multilingue que permet le développement du projet Wikidata.

En effet, les articles Wikipédia jouent un rôle important de “porte d'entrée” vers les œuvres utilisées pour les illustrer, les métadonnées peuvent être traduites en anglais et des catégories peuvent être créées pour une recherche plus facile.

**Les collections de musées versées sur Wikimédia bénéficient d'un potentiel de reconnaissance et d'étude inconnu jusque-là.**

# Mettre à disposition des données de références

Data.bnfr.fr extrait, transforme et regroupe dans une **base commune** des métadonnées issues de ressources internes et extérieures, produites dans des formats différents, afin de les lier entre elles et de les rendre interopérables. Ses pages sont indexées par les moteurs de recherche.

Le site permet de rassembler autour de ses pages auteur, œuvre, thème, date et lieu, des ressources de la Bibliothèque nationale de France, ainsi que des ressources extérieures.

# Profiter du potentiel du web semantique

[Crotos](#), est un moteur de recherche qui s'appuie sur Wikidata et donne ainsi accès à des œuvres en s'appuyant sur la force du web semantique.

[Paladia](#), est un moteur de recherche et d'affichage d'une partie des collections du Musée Saint-Raymond.

# 7 - Concevoir une stratégie open data

# **Comment ouvrir ses données : les différentes étapes pratiques à suivre**

- **Étape 1** : Sélectionner les données
- **Étape 2** : Préparer les données
- **Étape 3** : Diffuser et valoriser les données

# Réussir sa stratégie d'ouverture des données

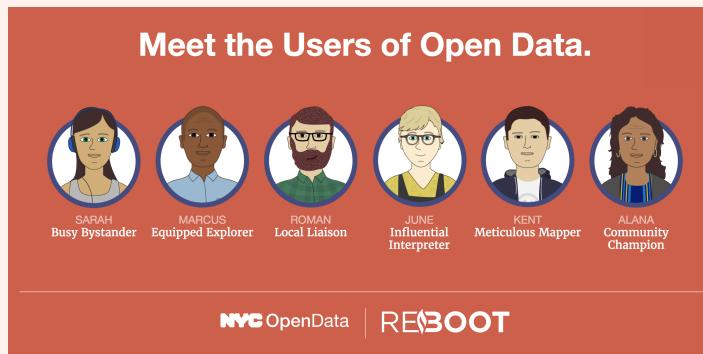
- Au-delà de l'obligation légale : **poser les objectifs** poursuivis par la stratégie
- Se donner les moyens de connaître la **demande de données** pour y répondre
- Travailler à la **qualité des données**, garante d'usages réels
- Mener le projet d'ouverture des données comme un projet de modernisation de l'administration (transversalité, agilité, lean...)

# Les freins à l'ouverture des données

- **Les ressources nécessaires (temps, RH et finances)** : extraction, anonymisation, hébergement des données, numérisation des œuvres ...  
*Pour y palier : accompagner les opérateurs et évaluer les risques/bénéfices, fixer une feuille de route open data pour prioriser/rationnaliser.*
- **Une mise en qualité insuffisante** : une qualité jugée suffisante pour des usages internes pour lesquels elles sont produites peut être jugée insuffisante pour le grand public : fichiers incomplets, de mauvaise qualité ou avec des erreurs.  
*Pour y palier : compter sur l'intelligence collective et l'interaction avec les usagers pour permettre l'amélioration des données (avec le crowdsourcing par exemple).*
- **La crainte d'un mauvais usage** : quelques rares cas de mauvaises réutilisations avec des données détournées de leur contexte ou utilisées à de mauvaises fins.  
*Pour y palier : renseigner les métadonnées qui accompagnent le jeu de données.*

# Bien penser sa stratégie open data, c'est s'assurer l'enthousiasme des gestionnaires et réutilisateurs

- De nombreux profils et équipes doivent s'emparer de l'open data



Une bonne stratégie consiste à réduire les frictions :

- Rendre les données moins difficiles à trouver (*enjeu de trouvabilité*)
- Publier les données avec une bonne granularité temporelle et spatiale
- Documenter les jeux de données pour en comprendre la structure
- Veiller à leur mise à jour la plus fréquente possible

# Résumé : les facteurs clés de succès d'une stratégie open data

## Obtenir un soutien au plus haut niveau

- Pour mettre en œuvre les moyens et les ressources à déployer

## Configurer l'organisation pour l'ouverture

- Organiser la gouvernance, les circuits de validation

## Faciliter la découverte et l'utilisation des données

- Faire attention au référencement, à la qualité des métadonnées

## Rester à l'écoute et interagir avec les usagers

- Prévoir des boucles de participation

# Bibliographie 1/3

- Politique du Ministère de la Culture pour les codes sources, algorithmes et codes sources, juin 2022
- **Feuille de route – Données et contenus culturels**, Publiée par le ministère de la Culture, septembre 2021
- **Guide pratique dédié à l'ouverture des données publiques culturelles**  
Publié par le ministère de la Culture, juillet 2019

# Bibliographie 2/3

- **Outil pédagogique facilitant la compréhension des droits de propriété littéraire et artistique applicables aux contenus culturels dont la diffusion est souhaitée**
- **Numérisation et valorisation des contenus culturels : guide ouverture et réutilisation des informations publiques numériques du secteur culturel**, publié par le ministère de la Culture, 2017
- **Guide de bonnes pratiques en matière de propriété littéraire et artistique**, publié par le ministère de la Culture, 2017
- **Rapport "Ouverture et partage des données publiques culturelles"**, publié par le ministère de la Culture, décembre 2013

# Bibliographie 3/3

- **Rapport open content dans les institutions culturelles en France**, Wikimédia, 2022
- **Retour d'expérience sur la mise en ligne des collections du musée de Bretagne**
- **Penser demain, L'Open Innovation dans les Musées & Institutions Culturelles**, l'ICART
- **Guide pratique pour la recherche et la réutilisation des images d'œuvres d'art**, INHA
- **Droits des images, histoire de l'art et société**

# Des questions ?

# Merci !