

# **Comprendre et mettre en œuvre l'open data**

**Formation AAF**

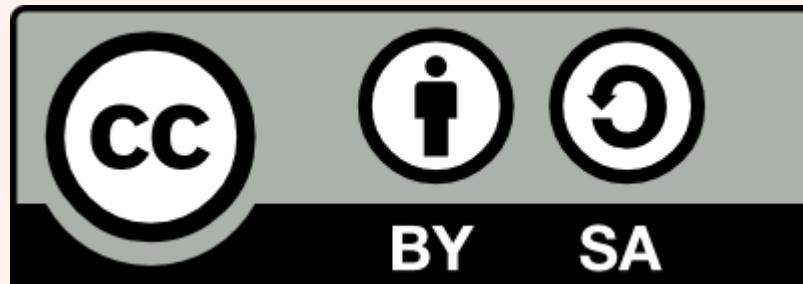
**Samuel Goëta, Anne-Laure Donzel,  
Dataactivist**

**02 mai 2022**

Ces slides en ligne : <http://dataactivist.coop/aaf>

Sources : <https://github.com/dataactivist/aaf>

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



We open data, we make them useful



# Qui sommes nous ?

- Dataactivist est un **pure player de l'open data** créé en 2016, par Samuel Goëta et Joël Gombin.
- Se positionnant sur **toutes les étapes du travail d'ouverture des données**, Dataactivist travaille tant avec les producteurs de données qu'avec les réutilisateurs et participe à l'appropriation des données par chacun.
- Nous appliquons nos propres **valeurs** : nous sommes une coopérative ; nos supports de formation et nos contenus sont librement réutilisables, publiés en licence Creative Commons.
- Une approche issue de la recherche : voir notamment <https://dataactivist.coop/these>.
- Nous animons également la communauté **#TeamOpenData**.

# Nos activités

- Conseil sur les **stratégies d'ouverture de données** : nous aidons les organisations dans la conception et la mise en œuvre de leur stratégie d'ouverture de données.
- **Accompagnement dans la réutilisation de données ouvertes** : nous aidons les organisations à utiliser les données au quotidien.
- **Sensibilisation et formation à la donnée** : nous formons à la culture générale des données, nous enseignons les grands principes et bonnes pratiques de l'open data et nous introduisons à la data science.
- **Médiation de données** : nous organisons des hackathons, des open data camps, des expéditions de données...

# Notre vision

## D'un open data de l'offre à un **open data de la demande**

Aujourd'hui, les administrations choisissent les données à ouvrir, elles décident du quand et du comment **sans consulter les usagers**.

Les rares données ouvertes sont souvent décevantes : pas à jour, trop agrégées, mal documentées, mal formatées, parfois inutilisables.

Il faut aujourd'hui passer à l'étape supérieure : **un open data piloté par la demande**.

Mais c'est difficile : demande peu visible, pas de possibilité de soumettre la communication des données à l'expression d'une demande. Leviers : demandes d'ouverture, plateformes, animations (ex : hackathon), gouvernance partagée...

# Rappel des objectifs de la formation

- Comprendre les grands principes de l'ouverture des données
- Appréhender le cadre juridique spécifique aux données publiques
- Maîtriser les enjeux liés à la qualité des données publiques pour faciliter la réutilisation
- Disposer des informations essentielles pour initier une démarche d'open data dans son institution

# Rappel du programme de la formation

## **lundi 02 mai**

- Éléments de l'histoire de l'ouverture des données
- Revue des grands principes de l'open data
- Le droit de la réutilisation des données publiques et les obligations issues de la loi pour une République numérique
- Panorama des données ouvertes et de leurs usages

## **Lundi 11 mai**

- Préparer l'ouverture de ses (premières) données : le registre des entrées !
- Importance de la standardisation
- Le schéma de donnée du registre
- Mise en pratique

# **Avant de commencer**

## **Présentons nous**



Public Domain Review

# **Origines et principes de l'open data**

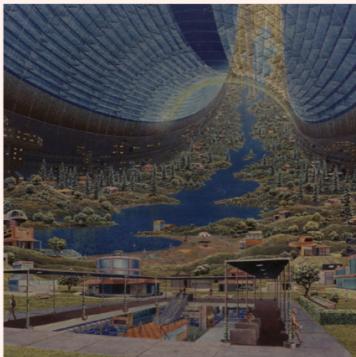
# Les multiples facettes de l'open data



La transparence



Le partage des données en science



La libre circulation de l'information



Les industries de la donnée



La transformation des administrations

Selon vous, quelle facette est-elle la plus porteuse actuellement ?

# **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.

**Explorons maintenant quelques grandes dates de l'apparition de l'open data.**

# 2005 : Open Definition

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

## The Open Definition

The [Open Definition](#) sets out principles that define "openness" in relation to data and content.

It makes precise the meaning of "open" in the terms "**open data**" and "**open content**" and thereby ensures **quality** and encourages **compatibility** between different pools of open material.

It can be summed up in the statement that:

*"Open means anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)."*

Put most succinctly:

*"Open data and content can be freely used, modified, and shared by anyone for any purpose"*

# 2007 : la rencontre de Sebastopol

**Quoi ?** Une rencontre de l'Open Government Group à Sebastopol (Californie), 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





# Revue des principes

# **1. Des données complètes**

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



# BY DEFAULT

## **2. Des données primaires**

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

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

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

## **4. Des données accessibles**

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

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

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

## **6. Des données non discriminatoires**

**Elles peuvent être utilisées par tous sans réclamer un enregistrement préalable**

## **7. Des données 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**

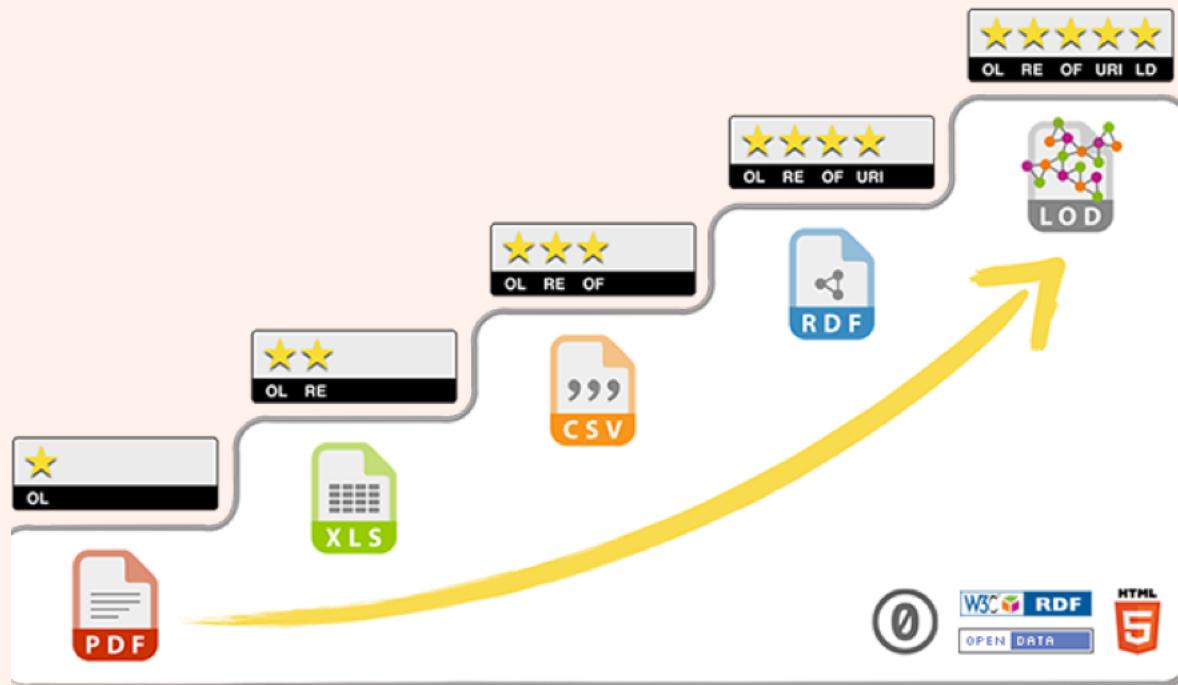
## **8. Des données dans une licence ouverte**

**Idéalement dans le domaine public  
sinon dans une licence conforme à  
l'Open Definition : Licence Ouverte  
(CC-BY) ou ODBL (CC-BY-SA)**

# **2008 : Tim Berners-Lee "Raw data now"**

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

# 2010 : le modèle en 5 étoiles de Tim Berners-Lee



# Des principes techniques sans portée politique

"Il peut être vrai en un sens qu'un régime devient plus ouvert chaque fois qu'il fournit des données ouvertes supplémentaires, même pour les sujets banals et apolitiques, mais il est facile d'imaginer qu'un régime fermé pourrait divulguer de grandes quantités de données conformes à ces huit exigences [de Sébastopol] **sans pour autant accroître sa responsabilité réelle en tant que gouvernement** .

Une diffusion électronique des déclarations de propagande faites par les dirigeants politiques de la Corée du Nord, par exemple, pourrait satisfaire à ces huit exigences et ne pas promouvoir une transparence ou une responsabilité supplémentaires de la part d'un régime notoirement fermé et opaque."

Yu & Robinson (2012) « [The New Ambiguity of Open Government](#) »

# 2011 : la déclaration pour un gouvernement ouvert



## OPEN GOVERNMENT DECLARATION

As members of the Open Government Partnership, committed to the principles enshrined in the Universal Declaration of Human Rights, the UN Convention against Corruption, and other applicable international instruments related to human rights and good governance:

**WE ACKNOWLEDGE** that people all around the world are demanding more openness in government. They are calling for greater civic participation in public affairs, and seeking ways to make their governments more transparent, responsive, accountable, and effective.

**WE RECOGNIZE** that countries are at different stages in their efforts to promote openness in government, and that each of us pursues an approach consistent with our national priorities and circumstances and the aspirations of our citizens.

**WE ACCEPT** responsibility for seizing this moment to strengthen our commitments to promote transparency, fight corruption, empower citizens, and harness the power of new technologies to make government more effective and accountable.

**WE UPHOLD** the value of openness in our engagement with citizens to improve services, manage public resources, promote innovation, and create safer communities. We embrace principles of transparency and open government with a view toward achieving greater prosperity, well-being, and human dignity in our own countries and in an increasingly interconnected world.

**TOGETHER, WE DECLARE OUR COMMITMENT TO:**

**Increase the availability of information about governmental activities.** Governments collect and hold information on behalf of people, and citizens have a right to seek information about governmental activities. We commit to promoting increased access to information and disclosure about governmental activities at every level of government. We commit to increasing our efforts to systematically collect and publish data on government spending and performance for essential public services and activities. We commit to proactively provide high-value information, including raw data, in a timely manner, in formats that the public can easily locate, understand and use, and in formats that facilitate reuse. We commit to providing access to effective remedies when information or the corresponding records are improperly withheld, including through effective oversight of the recourse process. We recognize the importance of open standards to promote civil society access to public data, as well as to facilitate the interoperability of government information systems. We commit to seeking feedback from the public to identify the information of greatest value to them, and pledge to take such feedback into account to the maximum extent possible.

**Support civic participation.** We value public participation of all people, equally and without discrimination, in decision making and policy formulation. Public engagement, including the full participation of citizens, is critical to the development of effective policies and programs.

mechanisms to enable greater collaboration between governments and civil society organizations and businesses.

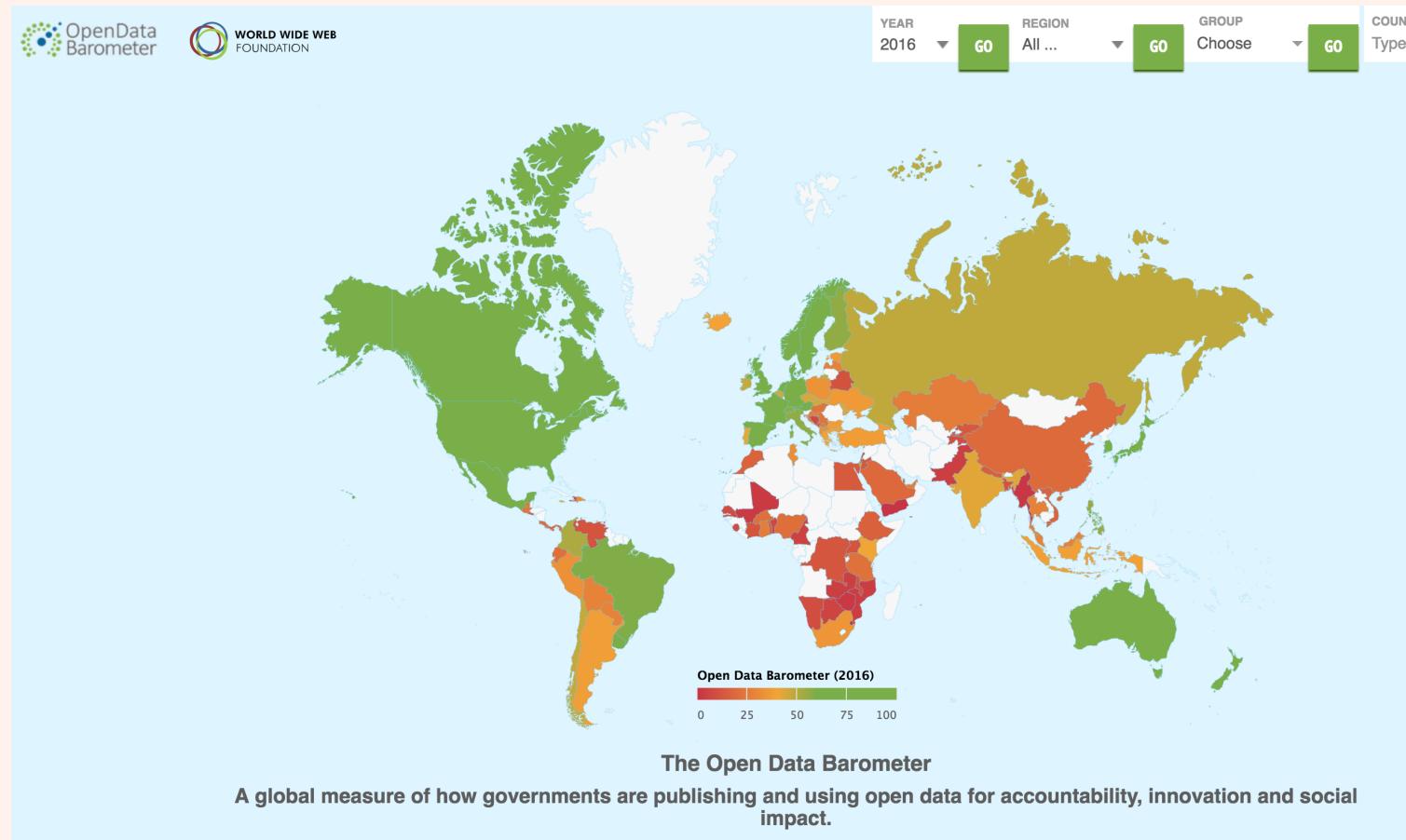
**Implement the highest standards of professional integrity throughout our administrations.** Accountable government requires high ethical standards and codes of conduct for public officials. We commit to having robust anti-corruption policies, mechanisms and practices, ensuring transparency in the management of public finances and government purchasing, and strengthening the rule of law. We commit to maintaining or establishing a legal framework to make public information on the income and assets of national, high ranking public officials. We commit to enacting and implementing rules that protect whistleblowers. We commit to making information regarding the activities and effectiveness of our anticorruption prevention and enforcement bodies, as well as the procedures for recourse to such bodies, available to the public, respecting the confidentiality of specific law enforcement information. We commit to increasing deterrents against bribery and other forms of corruption in the public and private sectors, as well as to sharing information and expertise.

**Increase access to new technologies for openness and accountability.** New technologies offer opportunities for information sharing, public participation, and collaboration. We intend to harness these technologies to make more information public in ways that enable people to both understand what their governments do and to influence decisions. We commit to developing accessible and secure online spaces as platforms for delivering services, engaging the public, and sharing information and ideas. We recognize that equitable and affordable access to technology is a challenge, and commit to seeking increased online and mobile connectivity, while also identifying and promoting the use of alternative mechanisms for civic engagement. We commit to engaging civil society and the business community to identify effective practices and innovative approaches for leveraging new technologies to empower people and promote transparency in government. We also recognize that increasing access to technology entails supporting the ability of governments and citizens to use it. We commit to supporting and developing the use of technological innovations by government employees and citizens alike. We also understand that technology is a complement, not a substitute, for clear, useable, and useful information.

We acknowledge that open government is a process that requires ongoing and sustained commitment. We commit to reporting publicly on actions undertaken to realize these principles, to consulting with the public on their implementation, and to updating our commitments in light of new challenges and opportunities.

We pledge to lead by example and contribute to advancing open government in other countries by sharing best practices and expertise and by undertaking the commitments expressed in this declaration on a non-binding, voluntary basis. Our goal is to foster innovation and spur progress, and not to define standards to be used as a precondition for cooperation or

# Depuis 2010 : l'ère des classements



# 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



# 2015 : la charte internationale de l'open data



## 1. Open By Default

This represents a real shift in how government operates and how it interacts with citizens. At the moment we often have to ask officials for the specific information we want. Open by default turns this on its head and says that there should be a presumption of publication for all. Governments need to justify data that's kept closed, for example for security or data protection reasons. To make this work, citizens must also feel confident that open data will not compromise their right to privacy.



## 2. Timely and Comprehensive

Open data is only valuable if it's still relevant. Getting information published quickly and in a comprehensive way is central to its potential for success. As much as possible governments should provide data in its original, unmodified form.



## 3. Accessible and Usable

Ensuring that data is machine readable and easy to find will make data go further. Portals are one way of achieving this. But it's also important to think about the user experience of those accessing data, including the file formats that information is provided. Data should be free of charge, under an open license, for example, those developed by Creative Commons.



## 4. Comparable and Interoperable

Data has a multiplier effect. The more quality datasets you have access to, and the easier it is for them to talk to each other, the more potential value you can get from them. Commonly-agreed data standards play a crucial role in making this happen.



## 5. For Improved Governance & Citizen Engagement

Open data has the capacity to let citizens (and others in government) have a better idea of what officials and politicians are doing. This transparency can improve public services and help hold governments to account.



## 6. For Inclusive Development and Innovation

Finally, open data can help spur inclusive economic development. For example, greater access to data can make [farming](#) more efficient, or it can be used to tackle [climate change](#). Finally, we often think of open data as just about improving government performance, but there's a whole universe out there of entrepreneurs making money off the back of open data.

# **Éléments du cadre juridique de l'open data**

# **Le guide CADA-CNIL : le document de référence**

Je n'indique ici que quelques éléments généraux sur le cadre juridique, ce guide constitue votre source de référence. Joël Gombin, co-fondateur de Dataactivist, a publié sa [fiche de lecture](#) sur TeamOpenData.

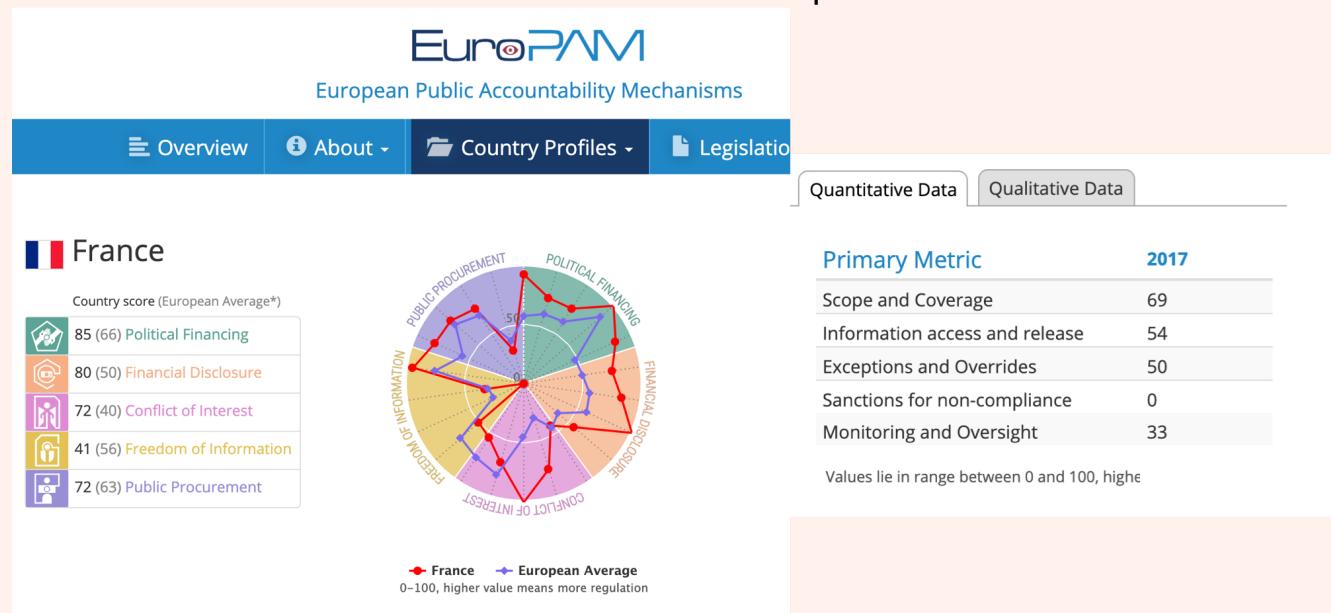
# 1978 : La loi CADA, vers le "droit de savoir"

- 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.

Il faudrait maintenant désigner la loi comme le Code des relations entre le public et l'administration (CRPA) qui, dans son **livre 3**, codifie le droit d'accès et de réutilisation mais il est encore moins connu que la loi CADA...

# La loi CADA : un point noir dans la transparence de la vie publique

Même si la France est le 3e pays à avoir adopté un droit d'accès, la loi CADA n'est pas un outil fort de la transparence de la vie publique. Ses principaux défauts : l'absence de sanctions et une procédure longue qui décourage les demandeurs et favorise le contournement par les administrations.

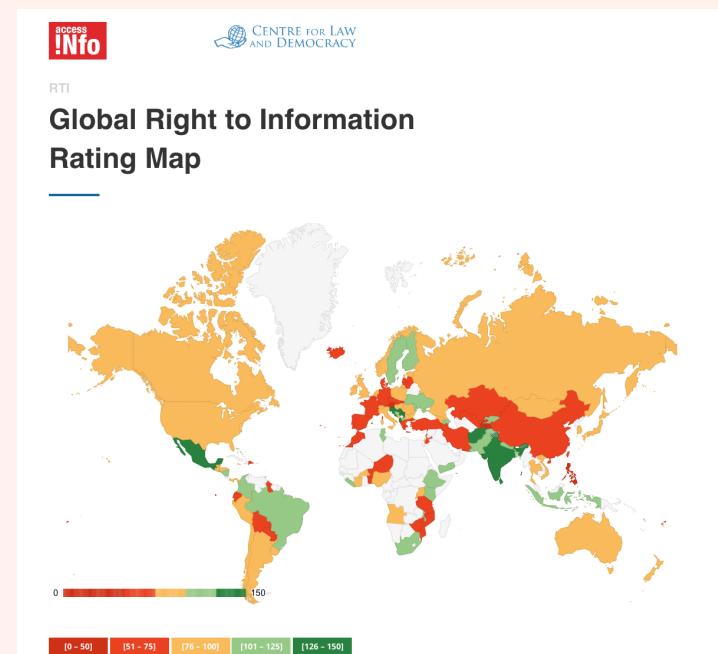


# Un droit d'accès très faible au regard des standards internationaux

Paradoxe : la France a une législation parmi les plus ambitieuses au monde en matière d'open data mais le droit d'accès y est un des plus mauvais au monde...

La France est classée **105e sur 123** du classement mondial des lois d'accès à l'information publique RTI Rating.

L'open data est comme la Tour de Pise : un très bel édifice posé sur des fondations instables.



# Les conditions d'exercice du droit d'accès

## 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 aux documents administratifs produits 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".

# Le cas des données personnelles

La prohibition de la communication à des tiers ne porte pas sur les données personnelles en général, mais **sur les données à caractère personnel relevant de la vie privée** des personnes concernées (âge, coordonnées personnelles, situation patrimoniale et financière, sympathies politiques, les croyances religieuses...)

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 : **Transparence Santé**)
2. si les personnes intéressées ont donné leur accord (consentement au sens de l'article 4 du RGPD)
3. les documents relevant de **l'article D312-3-1 du CRPA**

Hors de ces cas, il faut **anonymiser les données** : elles ne doivent plus se référer à une personne réelle (exit les noms, identifiants...) et ne doivent plus être spécifiques à un individu mais communes à un ensemble de personnes. L'indexation de ces données par un moteur de recherche externe est proscrite.

# Du droit d'accès au droit de réutilisation

Alors qu'aux Etats-Unis les données publiques sont placées dans le domaine public depuis les années 70, la question de la réutilisation des données publiques a émergé en France et en Europe dans les années 1990 avec le développement d'Internet facilitant la diffusion des données publiques et le développement de nouveaux services au public. Quelques jalons :

```
## Warning: package 'timevis' was built under R version 4.0.5
```

```
<div id="htmlwidget-f4ff52ce42939897c157" class="timevis html-widget">
<div class="btn-group zoom-menu">
<button type="button" class="btn btn-default btn-lg zoom-in" title="+">
<button type="button" class="btn btn-default btn-lg zoom-out" title="-">
</div>
</div>
<script type="application/json" data-for="htmlwidget-f4ff52ce42939897c157">
```

# Le droit de réutilisation en archives

La **réutilisation des informations publiques** désigne l'utilisation de documents ou de données par un tiers, pour d'autres fins que celles pour lesquelles ils ont été produits ou reçus initialement.

Les **informations publiques** sont tous les documents sauf ceux couverts par un délai de communicabilité et ceux sur lesquels des tiers détiennent des droits de propriété intellectuelle.

Un principe : la **gratuité de la réutilisation** mais avec des exceptions, notamment lorsque la réutilisation porte sur :

*Des informations issues des opérations de numérisation des fonds et des collections [...] des archives et, le cas échéant, sur des informations qui y sont associées lorsque ces dernières sont commercialisées conjointement.*

Les licences :

- Si gratuité : licence facultative (à choisir dans une liste fixée par l'Etat)
- Si redevance : licence obligatoire

# Loi pour une République Numérique : l'ouverture des données par défaut

Promulguée le 7 octobre 2016, après 4 ans d'annonces et de consultation, le titre 1 de la [loi pour une République Numérique](#) porte sur l'ouverture des données publiques.

La loi impose un principe d'**ouverture des données par principe** qui ne fait pas l'objet de sanctions à toutes les administrations, les entreprises délégataires d'une mission de service public et les **collectivités locales de plus de 3500 habitants et 50 agents**.



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.

# La définition des données

Le guide CADA-CNIL rappelle la **définition des "bases de données"**

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).

# Loi pour une République numérique : 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.



# Loi pour une République numérique : 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 DINSIC 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.

# Loi pour une République numérique : une limitation des licences



## Panorama des principaux types de licences

### 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

### 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 : SGAR Occitanie <http://sgar.asso.fr/licence-ouverte-open-licence>

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

# Où en est l'open data ?

# **Des politiques d'open data "de l'offre"**

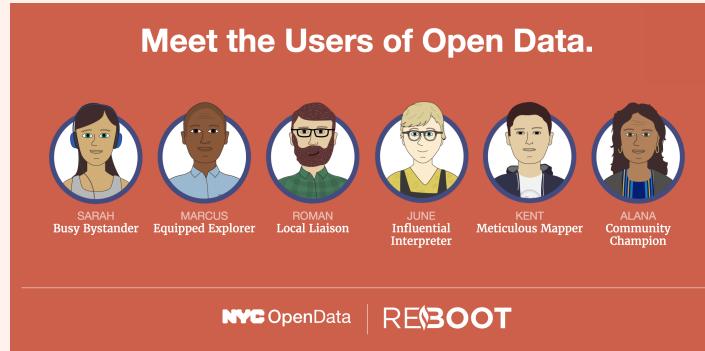
## **En théorie**

**Toutes les données sont ouvertes sous leur forme brute dès leur production**

## **En pratique**

**L'administration choisit quelles données ouvrir, comment les ouvrir, quand et ce qu'elles contiennent**

# La réutilisation : un parcours fait de frictions et de déceptions



**Des données difficiles à trouver,  
pas ouvertes partout**

**Granularité temporelle et spatiale  
souvent trop faible**

**Des données souvent inutilisables :  
manque de documentation**

**Des données pas mises à jour**

**Des portails qui ciblent les  
développeurs**

# Le problème de la qualité des données

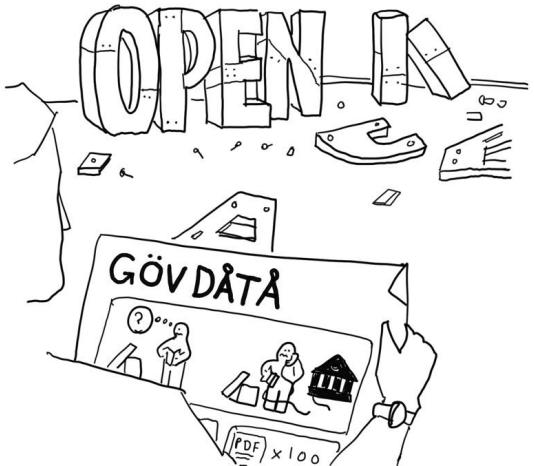
**Les données des gouvernements sont souvent incomplètes, pas actualisées, de mauvaise qualité et fragmentaires.**

Dans la plupart des cas, les catalogues ou portails de données ouvertes sont alimentés manuellement, conséquence d'une approche informelle de la gestion des données.

De plus, les procédures, les calendriers et les responsabilités des institutions publiques chargées de leur gestion manquent souvent de clarté. Ainsi, de manière générale, la gestion et la publication des données ouvertes sont fragiles et sujettes à des erreurs multiples.

Source : [Rapport mondial du 4e Open Data Barometer](#)

# Des données bien cachées



"La découverabilité des données est un défi majeur. Nous avons des portails et des registres de données, mais les organismes gouvernementaux qui relèvent d'un seul gouvernement national publient toujours les données de différentes façons et à différents endroits (...).

**La découverabilité des données est une condition préalable pour que les données ouvertes réalisent leur potentiel et, actuellement, la plupart des données sont très difficiles à trouver."**

<https://index.okfn.org/insights/>

# Un exemple de données introuvables

Ce fichier contient tous les signes officiels de la qualité et de l'origine (SIQO) : les AOC et AOP, IGP, Label rouge, spécialité traditionnelle garantie, Agriculture biologique... Très utile mais qui peut le trouver sur data.gouv.fr ?

The screenshot shows the data.gouv.fr homepage with the following elements:

- Header:** French flag, REPUBLIQUE FRANCAISE, data.gouv.fr, Se connecter, S'enregistrer.
- Search Bar:** Recherche with a magnifying glass icon.
- Navigation:** Données (selected), Réutilisations, Organisations, Actualités, À propos, Nous contacter.
- Breadcrumbs:** Accueil > Jeux de données > SIQO Publiéés.
- Title:** SIQO Publiéés (with a star icon indicating 4 favorites).
- Update Info:** Mis à jour le 4 avril 2022 — Licence Ouverte / Open Licence.
- Tabs:** Producteur (selected), Métadonnées, Actions.
- Producteur Section:** Institut national de l'origine et de la qualité (INAO). It features the INAO logo, a blue circular badge with a white question mark, and a brief description: "L'Institut national de l'origine et de la qualité est un établissement public administratif, doté de la personnalité civile, sous tutelle du Ministère de l'agriculture et de la pêche. Par la Loi d'orientation agricole du 5 janvier 2006, l'INAO est chargé de la mise en œuvre de la politique...". Below this is a link to "6 jeux de données".
- SOMMAIRE Section:** A sidebar with numbered links: 1. Description, 2. Fichiers (1), 3. Ressources communautaires (0), 4. Réutilisations (0).

# Un autre exemple de données vraiment introuvables

Ce jeu de données contient plus de 6000 marchés publics attribués par l'Etat. Mais, malgré plusieurs **signalements**, tout est fait pour le rendre introuvable !

The screenshot shows a dataset page on the data.gouv.fr platform. At the top, there's a navigation bar with the French Republic logo, the site name "data.gouv.fr", and links for "Se connecter" and "S'enregistrer". Below the navigation is a search bar. The main menu includes "Données", "Réutilisations", "Organisations", "Actualités", "À propos", and "Nous contacter". The breadcrumb navigation shows the path: Accueil > Jeux de données > AIFE\_DE\_44785462100045. The title of the dataset is "AIFE\_DE\_44785462100045", last updated on June 13, 2019, with a "License Not Specified". There are tabs for "Producteur", "Métadonnées", and "Actions". The "Actions" tab is active. Below the tabs, there's a logo for "aife." and a badge indicating "4 jeux de données". The "Métadonnées" section contains a summary of the dataset, mentioning the Agence pour l'Informatique Financière de l'Etat (AIFE) and its role in managing the Système d'Information Chorus. The "Actions" section lists a "SOMMAIRE" with five items: 1. Description, 2. Fichiers (84), 3. Ressources communautaires (0), 4. Réutilisations (0), and 5. Discussions (1).

# Le développement de l'open data en France



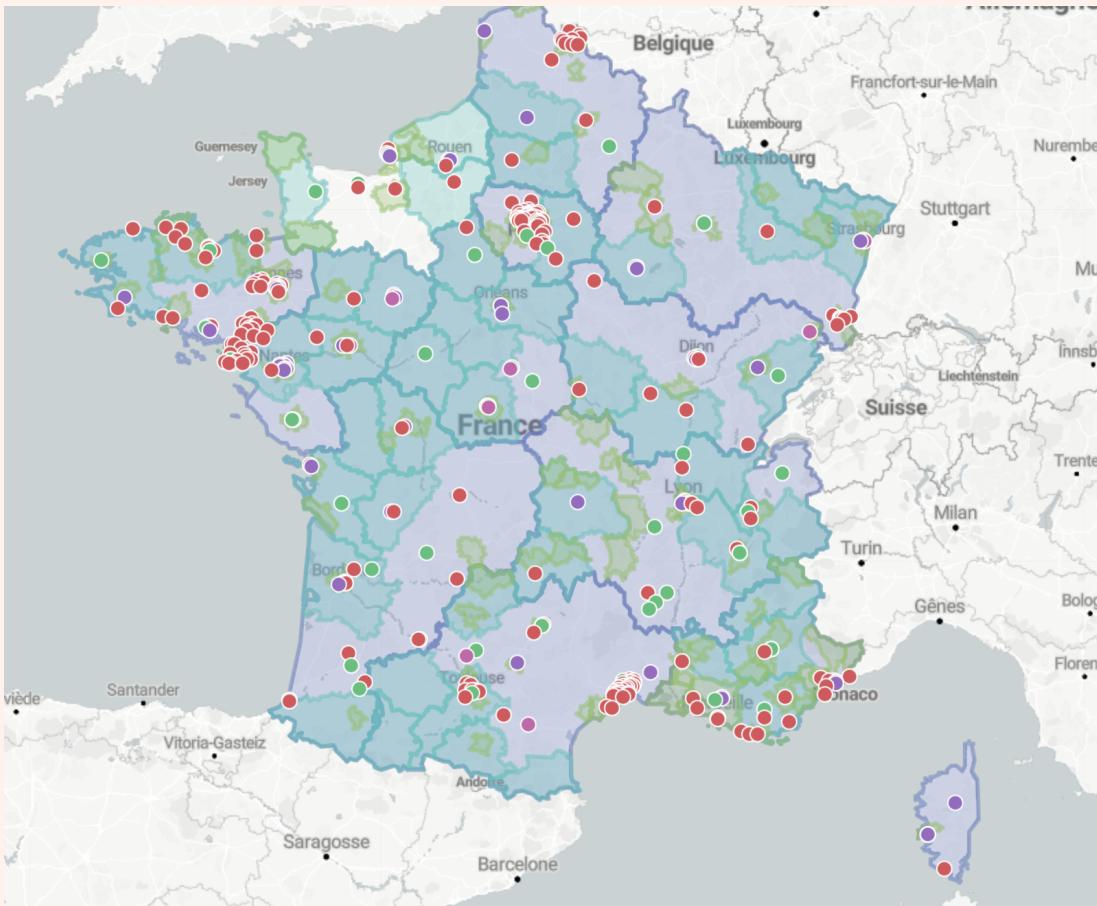
## Observatoire open data des territoires

Prendre la mesure de l'ouverture des données publiques dans les territoires pour évaluer, améliorer et faciliter son développement



(<http://slides.com/loichay/resultats-odt-oct18#/>)

# Le développement de l'open data en France



# Une multitude de portails territoriaux

Airtable

Plateformes territoriales open data

Sign up

Filter Sort ...

http://departement-ain.opendata.arcgis.com

TECHNO ArcGis

THÈME Généraliste

COUVERTURE Départementale

RÉGION Auvergne-Rhône-Alpes

http://opendata.auvergnerhonealpes.eu

TECHNO Spécifique

THÈME Généraliste

COUVERTURE Régionale

RÉGION Auvergne-Rhône-Alpes

http://data.metropolegrenoble.fr

TECHNO CKAN

THÈME Généraliste

COUVERTURE Métropolitaine

RÉGION Auvergne-Rhône-Alpes

https://data-atmoaura.opendata.arcgis.com

TECHNO ArcGis

THÈME Environnement

COUVERTURE Régionale

RÉGION Auvergne-Rhône-Alpes



Data Grand Lyon

TYPE Géoportal

URL <https://data.grandlyon.com>



Opendata Puy de Dôme

TYPE Site web

URL <https://puy-de-dome.fr/opendata/>



Data Grand Chambéry

TYPE Portail

URL <https://donnees.grandchambery.fr>



CRAIG

TYPE Géoportal

URL <https://ids.craig.fr/geocat/>

airtable.com/shrmZOih5lf14Q1PVN.../recqTC09fxU8b9E6R?blocks=hidelink

# Une multitude de portails nationaux

Plateformes nationales open data

Sign up

Filter Sort ...

Search

<p><b>data.gouv.fr</b></p> <p><b>TYPE</b> Portail</p> <p><b>URL</b> <a href="https://www.data.gouv.fr">https://www.data.gouv.fr</a></p> <p><b>PORTEUR</b> Etalab</p> <p><b>TECHNO</b> uData</p> <p><b>THÈME</b> Magasin général</p>	<p><b>adresse.data.gouv.fr</b></p> <p><b>TYPE</b> Concentrateur</p> <p><b>URL</b> <a href="https://adresse.data.gouv.fr">https://adresse.data.gouv.fr</a></p> <p><b>PORTEUR</b> Etalab</p> <p><b>TECHNO</b> Spécifique</p> <p><b>THÈME</b> Adresse</p>	<p><b>CeremaData</b></p> <p><b>TYPE</b> Géoportail</p> <p><b>URL</b> <a href="https://www.cdata.cerema.fr">https://www.cdata.cerema.fr</a></p> <p><b>PORTEUR</b> Cerema</p> <p><b>TECHNO</b> GeoNetwork</p> <p><b>THÈME</b> Référentiels géographiques</p>	<p><b>geo.data.gouv.fr</b></p> <p><b>TYPE</b> Concentrateur</p> <p><b>URL</b> <a href="https://geo.data.gouv.fr">https://geo.data.gouv.fr</a></p> <p><b>PORTEUR</b> Etalab</p> <p><b>TECHNO</b> Spécifique</p> <p><b>THÈME</b> Référentiels géographiques</p>

[airtable.com/shr3lYfHZW1cVTS3s.../recJiaEHQ83HrSPWc?blocks=hide](http://airtable.com/shr3lYfHZW1cVTS3s.../recJiaEHQ83HrSPWc?blocks=hide)

# Une multitude de solutions de portails sur le marché

	ARCGis Hub	CKAN	Geo Network	Junar	Opendata soft	Socrata	Udata	Koumou ul
<b>PERSONNALISATION</b>	✓	✓	✓	✓	✓	✓	✓	✓
<b>OPEN SOURCE ?</b>	✗	✓	✓	✗	✗	✗	✓	✓
<b>TYPE D'API</b>								
	Jeux de données	Catalogue	Catalogue	Jeux de données	Jeux de données	Jeux de données	Catalogue	Jeux de données
<b>RÔLES ET PERMISSIONS</b>	✓	✗	✗	✓	✓	✓	✓	?
<b>INTÉGRATION DANS SITES EXTERNES (EMBED)</b>	✓	✗	✗	✓	✓	✓	✗	✓

Tour d'horizon de ces solutions

# Exemple de données ouvertes : la base SIRENE

REPUBLIC FRANÇAISE  
Liberté Egalité Fraternité

[data.gouv.fr](#)

Se connecter | S'enregistrer

Recherche

Données Réutilisations Organisations Actualités À propos Nous contacter

Accueil > Jeux de données > Base Sirene des entreprises et de leurs établissements (SIREN, SIRET)

Mis à jour le 29 avril 2022 — Licence Ouverte / Open Licence version 2.0

Producteur Métadonnées Actions

 Institut National de la Statistique et des Etudes Economiques (Insee)

L'Institut national de la statistique et des études économiques (Insee) collecte, produit, analyse et diffuse des informations sur l'économie et la société françaises. Ces informations intéressent les pouvoirs publics, les administrations, les entreprises, les chercheurs, les médias, les...

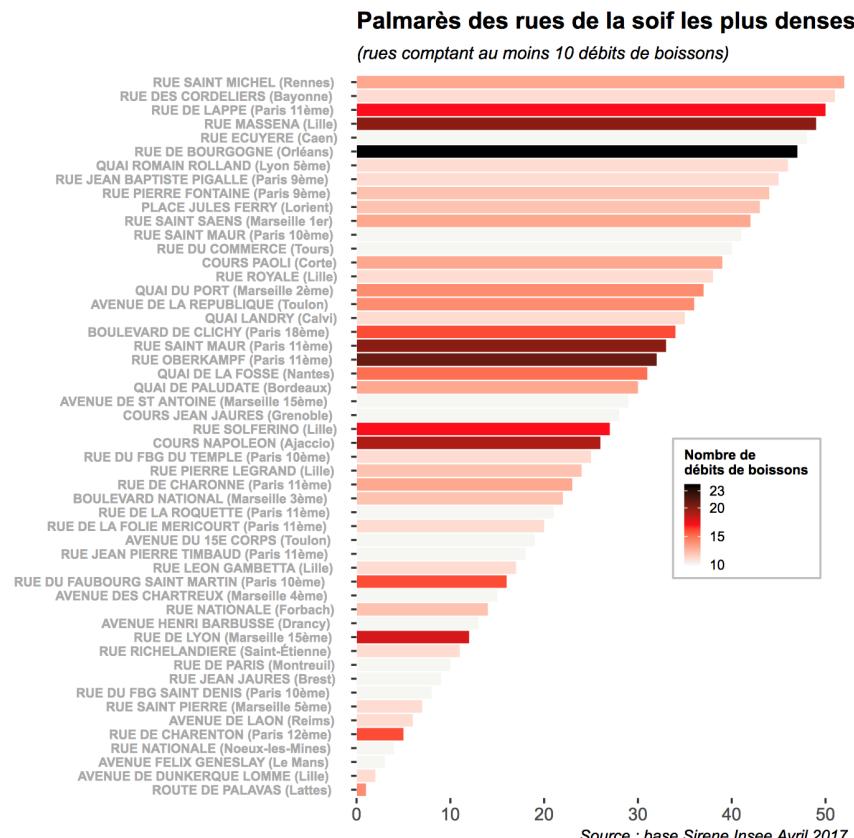
43 jeux de données 2 réutilisations

59 favoris

SOMMAIRE

1. Description
2. Fichiers (15)
3. Ressources communautaires (4)
4. Réutilisations (24)
5. Discussions (396)

# La base SIRENE : exemple de cas d'utilisation



# Exemple de données ouvertes : la base adresse nationale

The screenshot shows the data.gouv.fr website interface. At the top, there is a navigation bar with the French Republic logo, the URL "data.gouv.fr", and links for "Se connecter" and "S'enregistrer". A search bar with a magnifying glass icon is also present. Below the navigation bar, there are menu links: "Données" (highlighted in blue), "Réutilisations", "Organisations", "Actualités", "À propos", and "Nous contacter". The main content area shows the "Base Adresse Nationale" dataset. The title "Base Adresse Nationale" is displayed with a small "BAN" badge next to it. To the right, a button indicates "46 favoris". Below the title, a note states "Mis à jour le 9 novembre 2020 — Licence Ouverte / Open Licence version 2.0". A callout box provides information about consulting the dataset on the Point d'Accès National. The dataset card includes tabs for "Producteur", "Métadonnées" (selected), and "Actions". The "Producteur" tab shows the "BAN" logo and a gold medal icon. The "Métadonnées" tab contains a detailed description of the dataset: "Base Adresse Nationale. Elle est constituée par la collaboration entre: des acteurs nationaux tels que l'IGN, La Poste et la mission Etalab des acteurs locaux tels que les collectivités, les communes, les SDIS, des citoyens par exemple à travers le projet OpenStreetMap et l'association OpenStreetMap France." It also mentions "1 jeux de données". The "Actions" tab is currently inactive. On the right side, a sidebar titled "SOMMAIRE" lists five sections: "1. Description", "2. Fichiers (1)", "3. Ressources communautaires (1)", "4. Réutilisations (5)", and "5. Discussions (67)".

# Exemple de données ouvertes : les accidents de la route

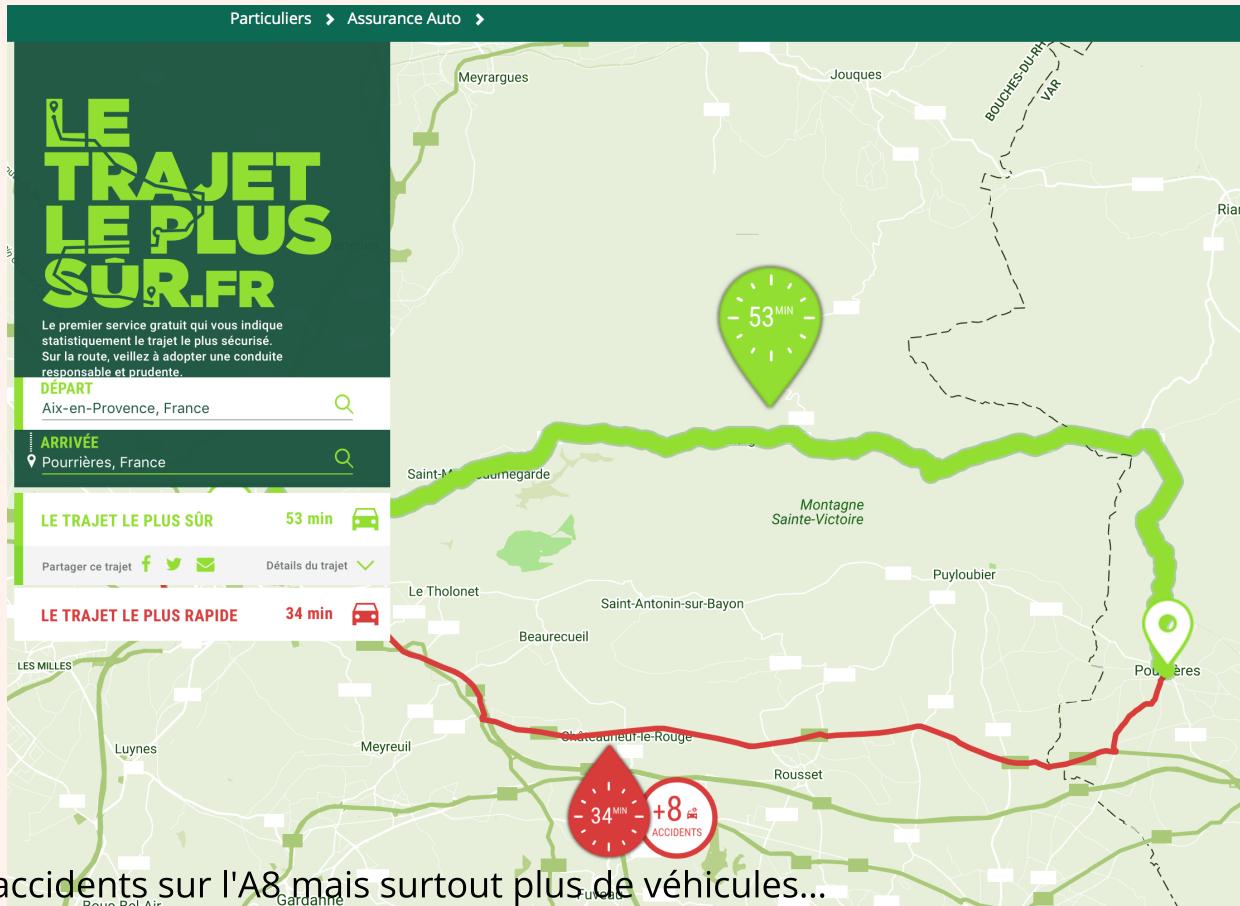
The screenshot shows the data.gouv.fr website interface. At the top, there is a navigation bar with the French flag, the text "RÉPUBLIQUE FRANÇAISE", and "data.gouv.fr". On the right side of the navigation bar are links for "Se connecter" and "S'enregistrer", and a search bar with a magnifying glass icon. Below the navigation bar, there is a horizontal menu with links: "Données" (which is underlined in blue), "Réutilisations", "Organisations", "Actualités", "À propos", and "Nous contacter".

In the main content area, the breadcrumb navigation shows "Accueil > Jeux de données > Bases de données annuelles des acci...". The title of the dataset is "Bases de données annuelles des accidents corporels de la circulation routière - Années de 2005 à 2020". To the right of the title is a button with "40 favoris" and a star icon. Below the title, it says "Mis à jour le 26 novembre 2021 — Licence Ouverte / Open Licence".

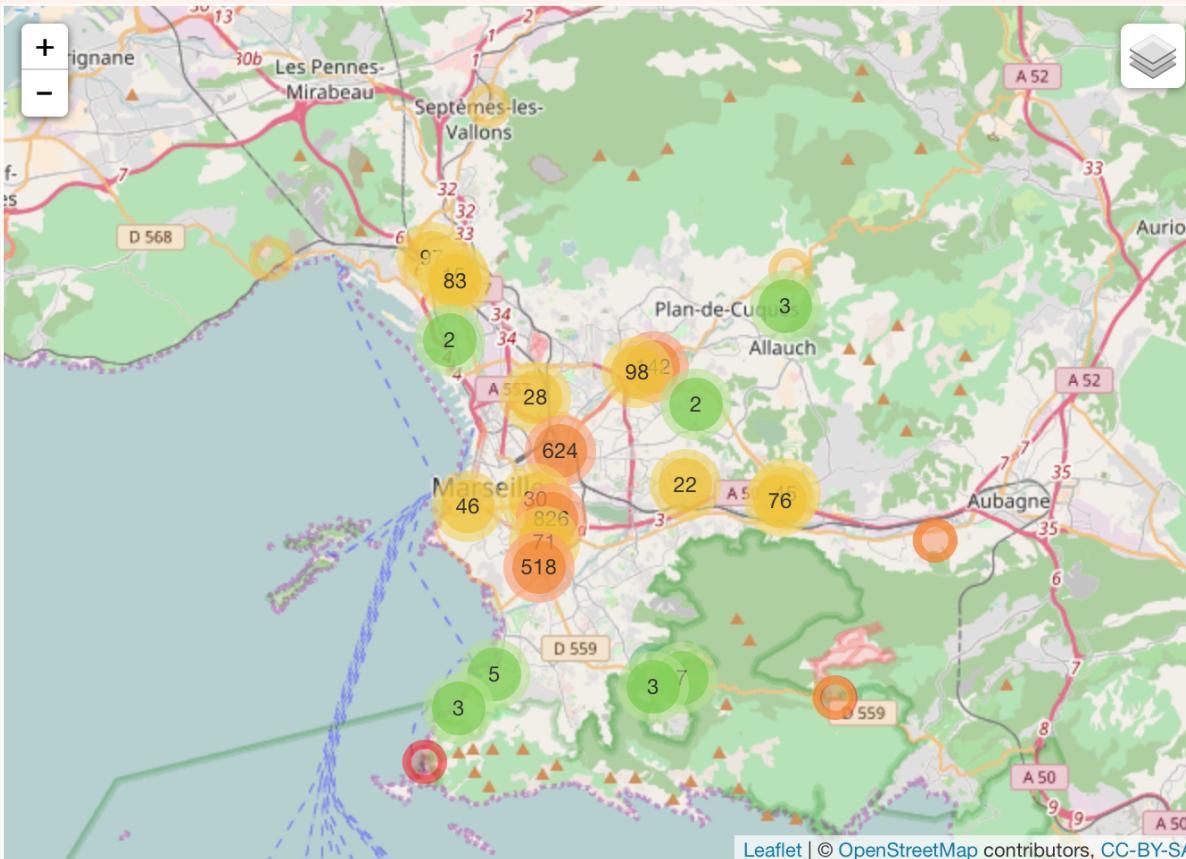
The page features three tabs: "Producteur" (highlighted in blue), "Métadonnées", and "Actions". Under the "Producteur" tab, there is a box containing the logo of the "Ministère de l'Intérieur" (Ministry of the Interior) and a brief description: "Placé au cœur de l'État, le ministère de l'Intérieur assure la permanence et la continuité de l'État. Cette fonction régaliennne se concrétise par le rôle majeur et les services rendus par le réseau des préfectures et des sous-préfectures aux citoyens sur tout le territoire national. Au quotidien,...". It also shows statistics: "630 jeux de données" and "1 réutilisations".

On the right side, there is a sidebar titled "SOMMAIRE" with a numbered list: 1. Description, 2. Fichiers (84), 3. Ressources communautaires (5), 4. Réutilisations (52), and 5. Discussions (0).

# Exemple de données ouvertes : Les accidents de la route



# Exemple de données ouvertes : les accidents de la route



Carte par Joël Gombin dans Marsactu

# Exemple de données ouvertes : les inspections des restaurants

 [data.gouv.fr](https://data.gouv.fr) Se connecter | S'enregistrer

Recherche Q

[Données](#) [Réutilisations](#) [Organisations](#) [Actualités](#) [À propos](#) [Nous contacter](#)

Accueil > Jeux de données > Résultats des contrôles officiels sanitaires

## Résultats des contrôles officiels sanitaires : dispositif d'information « Alim'confiance »

Mis à jour le 27 août 2018 — Licence Ouverte / Open Licence

[Producteur](#) [Métadonnées](#) [Actions](#)



**Ministère de l'Agriculture et de l'Alimentation**  
Le ministre de l'agriculture et de l'alimentation prépare et met en œuvre la politique du Gouvernement dans le domaine de l'agriculture, de la forêt et du bois. Il prépare et met en œuvre la politique de l'alimentation en liaison avec le ministre de l'économie, des finances et du commerce...

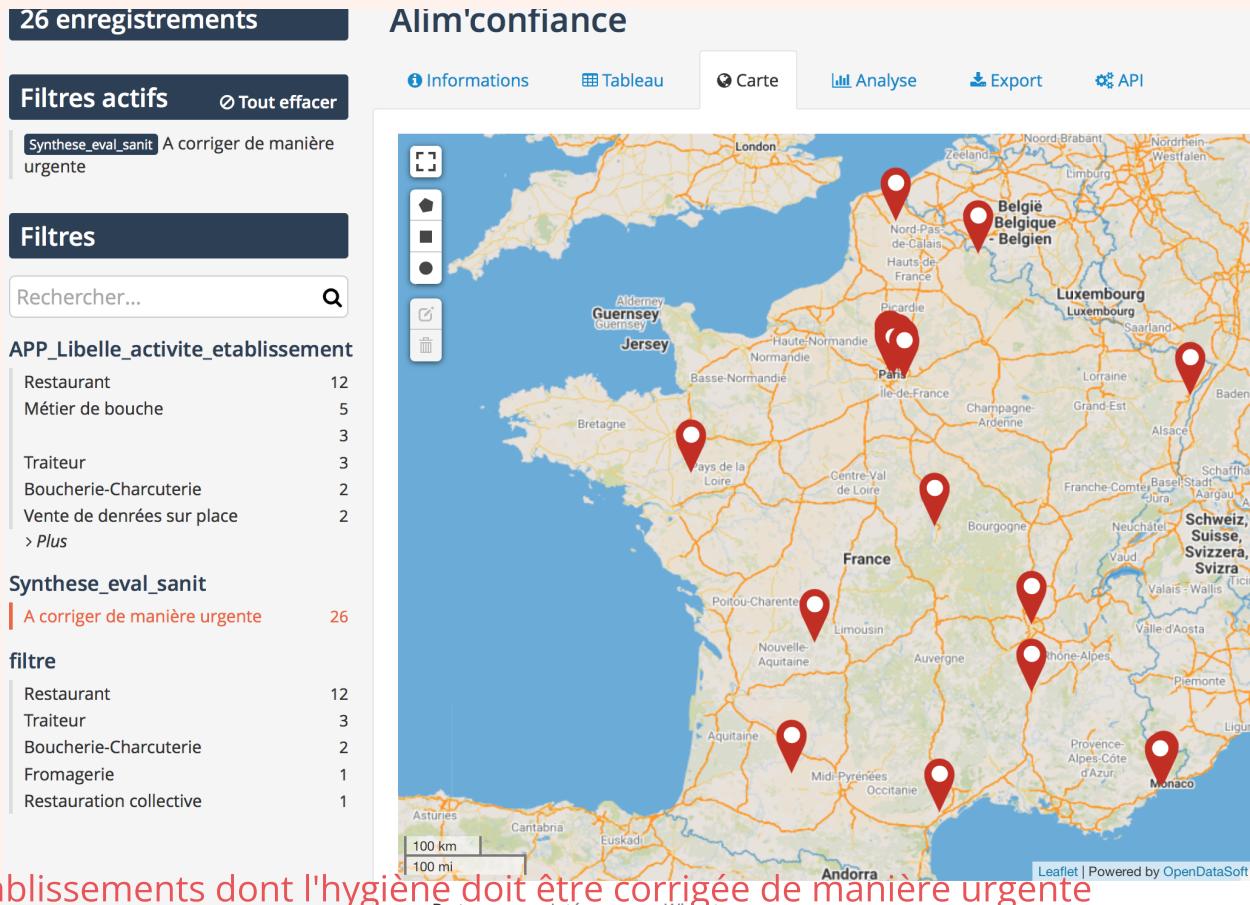
201 jeux de données 2 réutilisations

**SOMMAIRE**

1. Description
2. Fichiers (7)
3. Ressources communautaires (4)
4. Réutilisations (7)
5. Discussions (10)

### Résultats des contrôles officiels sanitaires : dispositif d'information « Alim'confiance »

# Exemple de données ouvertes : les inspections des restaurants



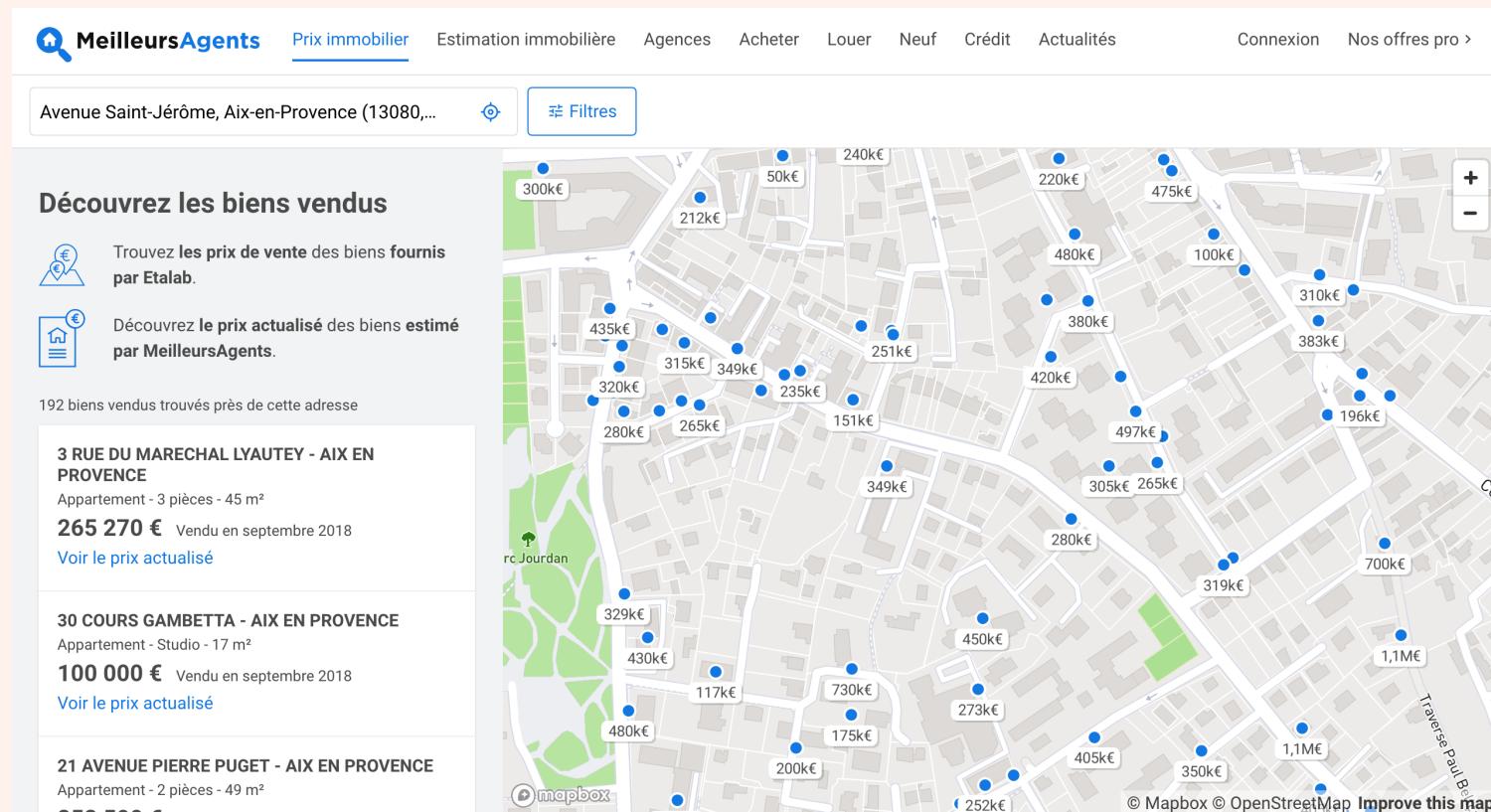
# Exemple de données ouvertes : le répertoire national des élus

The screenshot shows the RNE (Répertoire National des Élus) page on the data.gouv.fr platform. At the top, there's a navigation bar with links for 'Données', 'Réutilisations', 'Organisations', 'Actualités', 'À propos', and 'Nous contacter'. On the right, there are 'Se connecter' and 'S'enregistrer' buttons, and a search bar with a magnifying glass icon. Below the navigation, a breadcrumb trail shows 'Accueil > Jeux de données > RNE'. The main title 'Répertoire national des élus' is displayed with the acronym 'RNE' in red. To the right of the title is a button for '50 favoris'. A sub-header indicates the data was last updated on '1 février 2022 — Licence Ouverte / Open Licence version 2.0'. Below this, there are three tabs: 'Producteur' (selected), 'Métadonnées', and 'Actions'. The 'Producteur' section features a logo for 'Ministère de l'Intérieur' and a brief description: 'Placé au cœur de l'État, le ministère de l'Intérieur assure la permanence et la continuité de l'État. Cette fonction régaliennes se concrétise par le rôle majeur et les services rendus par le réseau des préfectorales et des sous-préfectorales aux citoyens sur tout le territoire national. Au quotidien,...'. It also shows statistics: '630 jeux de données' and '1 réutilisations'. To the right, a 'SOMMAIRE' section lists five items: '1. Description', '2. Fichiers (9)', '3. Ressources communautaires (2)', '4. Réutilisations (15)', and '5. Discussions (77)'.

# Exemple de données ouvertes : les valeurs foncières

The screenshot shows the data.gouv.fr website interface. At the top, there is a navigation bar with the French Republic logo, the URL "data.gouv.fr", and links for "Se connecter" and "S'enregistrer". A search bar is also present. Below the navigation, there are menu links: "Données", "Réutilisations", "Organisations", "Actualités", "À propos", and "Nous contacter". The main content area shows the breadcrumb "Accueil > Jeux de données > DVF". The title "Demandes de valeurs foncières" is displayed with a small "DVF" icon. To the right, there is a button with "94 favoris" and a star icon. Below the title, it says "Mis à jour le 8 avril 2022 — Licence Ouverte / Open Licence version 2.0". There are three tabs: "Producteur", "Métadonnées", and "Actions", with "Actions" being the active tab. The "Producteur" section shows the logo of the "Ministère de l'économie, des finances et de la relance" and a brief description of its role in economic, financial, budgetary, and fiscal policy. It also shows "451 jeux de données" and "2 réutilisations". The "Actions" section has a sidebar titled "SOMMAIRE" with a list of five items: 1. Description, 2. Fichiers (9), 3. Ressources communautaires (4), 4. Réutilisations (49), and 5. Discussions (79).

# Exemple de données ouvertes : les valeurs foncières



# **Exemple de données ouvertes : Santé Publique France sur la Covid-19**

# Exemple de données ouvertes : Campagnes photographiques des Monuments Historiques

**data.culture.gouv.fr**  
la plate-forme de données ouvertes du ministère de la Culture

316 enregistrements

Filtres actifs Ø Tout effacer

REGION BRETAGNE  
NOM\_DEPT FINISTERE

Filtres

Rechercher...

PAYS

France 316

REGION

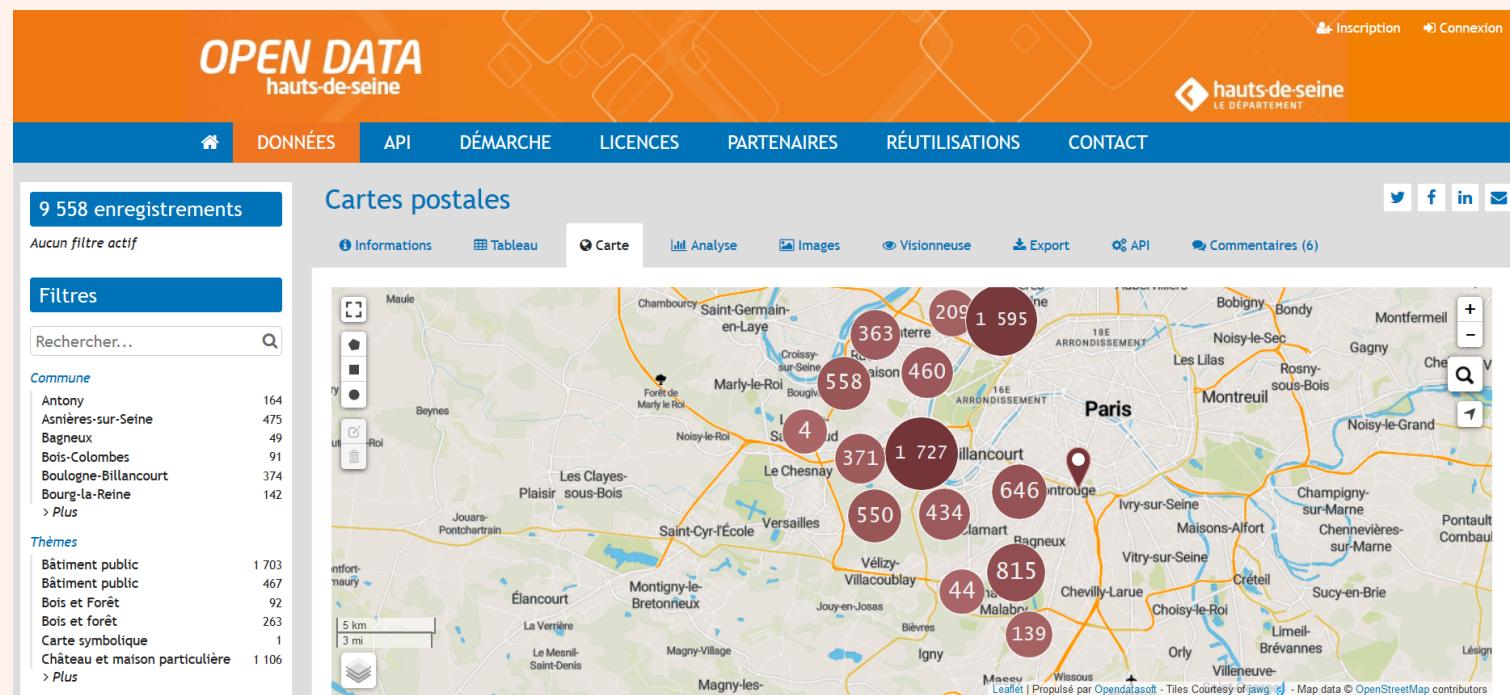
BRETAGNE 316

Photographies série « Monuments historiques » de 1851 à 1914

Informations Tableau Carte Analyse Images Export API

# Exemple de données ouvertes : Les AD des Hauts-de-Seine

Un projet mélant open data et indexation collaborative



# **A la semaine prochaine !**

## **Pour commencer**

### **Le témoignage d'une archiviste impliquée dans un projet d'open data**

[Lien vers la vidéo](#)

# Ressources

# Ressources

- **Guide pratique dédié à l'ouverture des données publiques culturelles**  
Publié par le ministère de la Culture, juillet 2019
- **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

# Ressources

- **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**

# Merci !

Contact : [samuel@dataactivist.coop](mailto:samuel@dataactivist.coop) [annelaure@dataactivist.coop](mailto:annelaure@dataactivist.coop)