

"TECHNO DU BIG-DATA »

Open Data

Stéphane Derrode, Dpt MI -
stephane.derrode@ec-lyon.fr



OPEN DATA

Open Data, c'est quoi?

[TEDx : L'open data à la loupe](#)

- L'**open data** ou **donnée ouverte** est une donnée numérique dont l'accès et l'usage sont laissés libres aux usagers. Elle peut être d'origine publique ou privée, produite notamment par une collectivité, un service public ou une entreprise. Elle est diffusée de manière structurée selon une méthode et une licence ouverte garantissant son libre accès et sa réutilisation par tous, sans restriction technique, juridique ou financière.
- L'ouverture des données (*open data*) est à la fois un mouvement, une philosophie d'accès à l'information et une pratique de publication de données librement accessibles et exploitables.
- Elle s'inscrit dans une tendance qui considère l'information publique comme un bien commun dont la diffusion est d'intérêt public et général.

La loi (en France...)

- **1978** Loi n. 78-753 / Crédation de la Commission d'accès aux documents administratifs (CADA). Objectif : faciliter et contrôler l'accès des particuliers et des organisations privées aux documents administratifs.
- **2007** Directive européenne INSPIRE, visant à établir une infrastructure d'information géographique dans la Communauté européenne pour favoriser la protection de l'environnement. Elle s'applique aux données géographiques détenues par les autorités publiques, dès lors qu'elles sont sous forme électronique et qu'elles concernent l'un des 34 thèmes figurant dans les annexes de la directive.
La directive INSPIRE a été transposée dans le droit français par l'ordonnance 2010-1232 du 21 octobre 2010, apportant des modifications au code de l'environnement.
- **2015** Loi n. 2015-1779 ("Valter") / Instauration d'un principe de gratuité par défaut de l'information publique.
- **2016** Loi n. 2016-1321 ("pour une République Numérique", "Lemaire") / Ouverture par défaut des données publiques et d'intérêt général (disposition entrée en vigueur le 7 octobre 2018), service public de la donnée, . . .

L'open data en France

- <https://www.data.gouv.fr>
- <https://www.data.gouv.fr/fr/organizations/>
- <https://geo.data.gouv.fr>
- <http://www.opendatafrance.net>
- <https://airtable.com/shrmZOh5If14Q1PVN/tblwkJPsyayeH5IX>
- <https://airtable.com/shr3IYfHZW1cVTS3s/tbIVtTOudK5Ygpsmj>
- <https://dataactivist.coop/fr/>



Quelles données ?

Toutes les données structurées et produites par des acteurs publics ou privés dans le cadre d'une mission de service public. Cette définition s'étend aux données décrivant l'espace, les services publics et l'exercice politique, produites par des acteurs non institutionnels : associations, citoyens engagés, acteurs économiques ou académiques.

Sont exclus de l'opendata

- Les données à caractère personnel

« Constitue une donnée à caractère personnel toute information relative à une personne physique identifiée directement ou indirectement par référence à un ou plusieurs éléments qui lui sont propres » (la loi 78-17 du 6 janvier 1978 art 2)

Ex : nom, adresse, photographie, num. tél., num. d'identifiant, adresse IP, adresse électronique...

- les documents administratifs dont la consultation ou la communication porterait atteinte :

A-6a) Au secret des délibérations du Gouvernement et des autorités responsables relevant du pouvoir exécutif ;

A-6d) A la sûreté de l'Etat, à la sécurité publique ou à la sécurité des personnes;

A-10c) Ou sur lesquels des tiers détiennent des droits de propriété intellectuelle.

Données publiques versus Big Data

Données publiques : services publics, intérêt général, ouvertes

- Effectifs: d'agents municipaux, d'élèves
- Horaires : bus et tram en temps réel,
- Patrimoine : arbres, équipements sportifs, points lumineux, bornes fontaines ...
- Éléments liés à l'exploitation : livres empruntés dans les bibliothèques, budget et compte administratif, marchés publics
- Description du territoire : photo aérienne, altimétrie, description des rues,
- Reflet de la vie locale : mariage, naissance, prénoms, élections, agendas...

Big data: données personnelles, comportement individuel, données privées

- Traces (téléphone mobile, connexion)
- Navigation internet, réseaux sociaux, messagerie,
- Achats (CB), déplacements individuels ou collectifs (badge, carte)
- Consommations individuelles (eau, électricité, ..)
- Objets connectés

 PIONNIERS
MOS

CAPITALE

The image shows a screenshot of the Brocas OpenData website. The header features the 'Brocas OpenData' logo with a green globe icon. Below the header, there's a search bar with the placeholder 'Rechercher une commune'. A large, bold, black diagonal banner across the center of the page reads 'COMMUNES RURALES'. In the bottom left corner, there are icons for 'Département', 'Wikipedia', and 'Carte'. On the right side, there's a sidebar with links like 'Accès à la carte', 'Qui sommes-nous', 'Méthodologie', 'Politique OPENDATA', 'Géographie', 'Communes', 'Contactez-nous', 'Mentions légales', and 'Plan de site'. The background of the main content area shows a scenic view of a village with houses and a church tower under a blue sky.

COMMUNES RURALES

ORIGINAUX

The screenshot shows the official French government data portal, data.gouv.fr. At the top left is the French flag and the website address. The top right features the text "Plataforme ouverte des données publiques françaises". Below the header is a search bar with the placeholder "Recherche". A sidebar on the left lists various government departments like Agriculture, Economy, and Education. The main content area has a blue background with the slogan "Partagez, améliorez et réutilisez les données publiques". A large black diagonal banner across the middle contains the word "L'ETAT" in white. Below the slogan are several data visualization cards, including one showing a map of France with colored regions and another showing a grid of colored squares representing data. The bottom right corner contains a small image of a person and some text about data reuse.

L'ETAT

ASSOCIATION



DEPARTEMENTS



PRIVES

Gisement de données : exemples

- **Collectivités territoriales :**

<https://fr.wikipedia.org/wiki/OpenDataFrance>

- **EtaLab (Ministères) :** <http://data.gouv.fr>

- **Citoyens :**

- Data4citizens (<https://open.data4citizen.com>),
- Regards Citoyens (<https://www.regardscitoyens.org>),
- DataPublica (<https://www.data-publica.eu>)

- **Instituts :**

- INSEE : <https://www.insee.fr/fr/statistiques>
- INPI : <https://www.inpi.fr/fr/services-et-prestations-domaine/open-data>

- **Agences de l'air, de l'eau, etc. :**

- Pollution (échelle mondiale) : <https://openaq.org>
- Pollution (région Auvergne-Rhône-Alpes) :

<https://www.atmo-auvergnerhonealpes.fr/donnees/telecharger>

[TEDx :How we found the worst place to park in New York City.](#)

Formats

- **Pour être Opendata, les données doivent être :**
 - au format électronique,
 - accessible sur internet,
 - structurées,
 - à un format ouvert (et de préférence non propriétaire).
- **Types de format des fichiers disponibles**
 - Données alpha-num. : csv, ods, xls, doc, xml, JSON...
 - Un document au format pdf, même full text –et pas image- est lisible par une machine mais n'est ouvrable que par Acrobat/Adobe, donc non Opendata
 - Données géographiques : Shape, GéoJSON, KML, KMZ,
 - Format propriétaires non ouverts : DWG et DXF (Autocad)
 - Format spécifiques :
 - Domaine Transport : GTFS, Chouette, NetEX, Neptune...
 - GPX (ouvert), ...

Accès par API

- **Données temps-réel**
 - Horaire, charges, disponibilité, ...
- **Données statiques** mais à jour à tout instant
 - Accès direct à une cellule, à une ligne d'un tableau par ex.
- Il existe des **portails** qui ont systématisé la publication des données sous forme de fichiers ET d'API

The screenshot shows a web browser window for data.toulouse-métropole.fr. The page title is "DATA.toulouse-métropole" and the subtitle is "Le portail des données ouvertes du territoire". The navigation menu includes Démarche, Licence, Données, Cartes, API (highlighted), Communes, Réutilisations, and Demande. Below the menu, there are two tabs: "V1" and "V2 (beta)". Under "V1", there are links for "Console" and "Documentation". The main content area is titled "Lister les jeux de données" and shows the endpoint "GET /api/datasets/1.0/search". It includes a search bar with fields for "Requête en texte intégral" and "Code langue de 2 lettres", and a dropdown for "rows".

The screenshot shows a web browser window for opendata.paris.fr. The page title is "PARISDATA" and the subtitle is "MAIRIE DE PARIS". The navigation menu includes Les données, L'API (highlighted), La licence, and La démarche. Below the menu, there are tabs for "Cartographie" (selected) and "V2 (beta)". The main content area is titled "Console API" and says "La référence SWAGGER complète est disponible en [JSON](#) et [YAML](#)". It shows a tree view of API operations under "root". The "root" node has a "GET /" operation. The "metadata" node has three "GET" operations: "{source}/metadata_templates", "{source}/metadata_templates/{metadata_template_type}", and "{source}/metadata_templates/{metadata_template_type}/{metadata_template_name}".

Outils d'exploitation

Données Alphanumériques

- Suites bureautiques habituelles : OpenOffice, Libre Office,
 - Et les outils commerciaux habituels (Google, MS, ...)
- De nombreux outils en ligne : ConversionTools Services (xml<>csv...)
- Les langages de programmation habituels : C, Python...
- <https://openrefine.org> pour l'exploration et le retraitement des données
- Frise Chronologique

Données géographiques

- QGIS : vraie suite logicielle de géomatique (SIG), convivial
- UMAP, Carto, MapServer, GéoServer : manipulation des cartes

Données cartographiques

- **Visualisation**

- OSM: <https://www.openstreetmap.org>
Démonstration : <http://tools.geofabrik.de>

- **Edition / Modification**

- <http://openstreetmap.fr/>
 - <http://open.mapquest.fr>
 - <https://umap.openstreetmap.fr/fr/>

- **Exploitation**

- Routing : <http://www.yournavigation.org>,
<http://map.project-osrm.org>,
 - OverPassTurbo : <http://overpass-turbo.eu>