

# A

## Algorithme

C'est une méthode de traitement des données. Ce traitement est automatisé : la machine l'effectue à partir de données et en produit d'autres (composante technique). Ce traitement est produit par une personne ou un service en fonction d'un objectif (composante sociale). Par exemple, le calcul du jour de la semaine pour une date donnée. La loi pour une République numérique impose la mention de la finalité d'un traitement algorithmique.

## Anonymisation

L'anonymisation des données consiste à en modifier la structure afin de rendre très difficile ou impossible la « ré-identification » des personnes (physiques ou morales) ou des entités concernées.

## API : (Application programming interface)

Les interfaces de programmation permettent à un logiciel de fournir des services ou des données à un autre logiciel de manière simple.

# B

## Base de données

C'est un ensemble de données organisé dans le but de faciliter leur usage. Une base de données contient un ensemble d'informations structurées permettant de les stocker pour opérer des traitements et fournir des services.

## Big Data (ou grandes données)

Cette expression désigne un ensemble de données très volumineux qui doit être traité par des outils spécifiques.

# C

## Cada

La Commission d'accès aux documents administratifs est une autorité administrative indépendante chargée de veiller à la liberté d'accès aux documents administratifs et aux archives publiques ainsi qu'à la réutilisation des informations publiques.

## Catalogue

Le catalogue opendata d'une collectivité rassemble les métadonnées de description des jeux de données qu'elle publie.

## Code source

Le code source est la version lisible d'un programme informatique, écrit par un programmeur ou une programmeuse. Il peut s'agir de programmes complexes ou de quelques lignes.

## Coffre-fort électronique

Système informatique permettant de garantir l'intégrité dans le temps de données numériques

## Contrôle Scientifique et Technique (CST)

Contrôle exercé par un représentant de l'Etat au moyen d'inspections et de visites de conseil portant sur les conditions de gestion, de collecte, de sélection et d'élimination ainsi que sur le traitement, le classement, la conservation et la communication des archives.

## Copie fiable

Reproduction numérisée d'un document dont la conformité juridique est reconnue selon les conditions du décret n° 2016-1673 du 5 décembre 2016.

## Copie fidèle

Reproduction fidèle à l'identique d'un document original.

## **Crowdsourcing (ou production participative)**

C'est l'utilisation de la créativité, de l'intelligence et du savoir-faire d'un grand nombre de personnes, en sous-traitance, pour réaliser certaines tâches traditionnellement effectuées par un employé ou un entrepreneur. Par exemple, les contenus de Wikipédia et d'Open Street Map sont réalisés en crowdsourcing.

## **CSV (Comma Separated Values)**

C'est un nom d'un format de fichier contenant des données textuelles. Les données sont lisibles par un très grand nombre d'outils : les tableurs, les éditeurs de texte... L'organisation du fichier étant répandue et simple, elle peut être plus facilement traitée par un programme.

# **D**

## **Datavisualisation (ou "Dataviz")**

Il s'agit de représentation graphique de données. Quelques formes simples et connues de visualisation de données sont le "camembert", l'histogramme, le nuage de points. La visualisation de données peut s'appuyer sur différentes sources de données. Elle a pour objectif de rendre les données plus lisibles et compréhensibles.

## **Délégué à la protection des données (DPD)**

Il est chargé de mettre en œuvre la conformité au règlement européen sur la protection des données au sein de l'organisme qui l'a désigné s'agissant de l'ensemble des traitements mis en œuvre par cet organisme.

## **Document administratif**

Sont considérés comme administratifs tous les documents produits ou reçus par une administration publique (administrations d'Etat, collectivités territoriales, établissements publics). Il en va de même pour les documents détenus par les organismes privés chargés d'une mission de service public pour autant qu'ils sont liés, par leur nature, leur objet, ou leur utilisation à la gestion de cette mission.

## **Donnée**

En informatique, une donnée est la représentation d'une information dans un programme : soit dans le texte du programme (code source), soit en mémoire durant l'exécution. Les données, souvent codées, décrivent les éléments du logiciel tels qu'une entité (chose), une interaction, une transaction, un évènement, un sous-système, etc.

### **Donnée chaude**

Donnée mise à jour régulièrement.

### **Donnée froide**

Donnée qui ne change pas dans le temps.

## **Données de référence**

Ce sont des données fréquemment utilisées par de multiples acteurs publics et privés, et dont la qualité et la disponibilité sont critiques pour ces utilisations, comme, par exemple, les données des référentiels géographiques de l'État.

## **Données ouvertes**

Ce sont des données numériques 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.

## **Donnée pivot (ou donnée-clé)**

C'est une donnée qui permet de relier plusieurs jeux de données, comme, par exemple, le *numéro SIRET d'une entreprise*.

## **Donnée publique**

Donnée en lien avec une mission de service public (produite par l'Etat, les collectivités, les personnes de droit public ou privé en charge d'une mission de service public) et concernant des éléments non sensibles,

non personnels, non soumis au droit d'auteur, sans lien avec la sécurité ou la sûreté publique.

## **Droit de communication**

Les administrations sont tenues de publier en ligne ou de communiquer les documents administratifs qu'elles détiennent aux personnes qui en font la demande. Cela ne s'applique qu'à des documents achevés et ne concerne pas les documents préparatoires à une décision administrative tant qu'elle est en cours d'élaboration. Dans le cas où la demande permet de bénéficier d'une décision individuelle créatrice de droits, les documents sont communicables à l'auteur de cette demande dès leur envoi à l'autorité compétente pour statuer sur la demande. Le droit de communication perdure, même si le document est déposé aux archives. Le droit de communication cesse lorsque les documents font l'objet d'une diffusion publique.

# **E**

## **Empreinte numérique**

Chaîne de caractères unique obtenue en appliquant un algorithme de hachage à un fichier numérique.

# **F**

## **Fichier numérique / copie numérique**

Fichier informatique généré suite à la numérisation.

# **G**

## **GAFAM**

C'est l'abréviation de Google, Amazon, Facebook, Apple, Microsoft : il s'agit des entreprises les plus puissantes de l'internet et accessoirement celles qui détiennent et/ou manipulent le plus de données.

## **Géolocalisation**

Technologie permettant de déterminer la localisation d'un objet ou d'une personne avec une certaine précision. La technologie s'appuie généralement sur le système GPS ou sur les interfaces de communication d'un téléphone mobile.

## **Gouvernance de la donnée**

C'est un ensemble de principes et de pratiques qui visent à assurer la meilleure exploitation du potentiel des données.

# **H**

## **Hackathon**

Désigne un événement durant lequel des personnes volontaires, de spécialités variées, se réunissent pendant une période de temps donnée afin d'imaginer, en mode collaboratif, des services ou produits à partir de données en Open data par exemple.

## **Horodatage**

Mécanisme consistant à associer une date et une heure précise à un événement ou une information.

# **I**

## **Indexation**

Opération consistant à associer à un contenu binaire ou à un ensemble de contenus binaires une ou plusieurs métadonnées.

## **Information**

Tout ensemble de sons, d'images, de données ou de documents intelligible par l'homme, qu'il soit fixé sur un support ou non.

## **Interopérabilité**

Capacité d'un logiciel à fonctionner avec d'autres logiciels existants ou futurs et ce sans restriction d'accès et de mise en œuvre.

# **J**

## **Jeu de données (dataset)**

C'est un ensemble de données qui forme un tout. Par exemple, la liste des installations sportives.

# **L**

## **Licence**

Fixe les conditions de la réutilisation des informations publiques. Ces conditions ne peuvent apporter de restrictions à la réutilisation que pour des motifs d'intérêt général et de façon proportionnée. Elles ne peuvent avoir pour objet ou pour effet de restreindre la concurrence. Pour ses données, la Ville a choisi la Licence Ouverte.

# **M**

## **Métadonnée**

C'est une information descriptive liée à une donnée. Par exemple, la date de production de la donnée, son producteur, son format, sa licence constituent des métadonnées. Pour qu'un jeu de données soit facilement accessible et réutilisable, la qualité des métadonnées joue un rôle déterminant.

# **O**

## **Open data**

Désigne un mouvement d'ouverture et de mise à disposition des données produites et collectées par les services publics. La finalité consiste en la possibilité, pour toute personne, de réutiliser librement et gratuitement les informations publiées, à d'autres fins que celles de la mission de service public pour les besoins de laquelle les documents ont été produits ou reçus.

## **Open data France**

C'est une association loi de 1901, créée le 9 octobre 2013. Elle a pour objectifs de regrouper et soutenir les collectivités territoriales françaises engagées activement dans une démarche d'ouverture des données.

# **P**

## **Producteur de la donnée**

Entité responsable de la création de la donnée jusqu'à sa validation. Le producteur de données a en charge le contrôle des données

# **Q**

## **Quantified self (mesure de soi)**

C'est un mouvement qui regroupe les outils, les principes et les méthodes permettant à chacun de mesurer ses données personnelles, de les analyser et éventuellement de les partager. Les outils du quantified self peuvent être des objets connectés, des applications mobiles ou des applications Web (Wikipédia juillet 2017 :

[https://fr.wikipedia.org/wiki/Quantified\\_self](https://fr.wikipedia.org/wiki/Quantified_self)). Le fait de mesurer le nombre de pas effectués par jour constitue une pratique de quantified self.

## R

### Redevabilité algorithmique

Devoir, pour l'administration publique qui fait usage d'un système algorithmique, de rendre des comptes concernant ses choix technologiques, les usages en résultant, et leurs éventuelles conséquences.

### Répertoire des informations publiques

Les administrations qui produisent ou détiennent des informations publiques tiennent à la disposition des usagers un répertoire des principaux documents dans lesquels ces informations figurent.

### Réutilisation des informations publiques

C'est une utilisation par un tiers à d'autres fins que celles de la mission de service public pour laquelle les documents ont été produits ou reçus.

## S

### Self data

Le self data désigne la production, l'exploitation et le partage de données personnelles par les individus, sous leur contrôle et à leurs propres fins : pour mieux se connaître, prendre de meilleures décisions, se faciliter la vie, etc.

### Service public d'archives

Service ayant pour missions de collecter, de conserver, d'évaluer, d'organiser, de décrire, de communiquer, de mettre en valeur et de diffuser des archives publiques (article R212-4-1 du code du patrimoine).

### Service public de la donnée

Le service public de la donnée créé par l'Article 14 de la loi pour une République numérique vise à mettre à disposition, en vue de faciliter leur réutilisation, les jeux de données de référence qui présentent le plus fort impact économique et social. Il s'adresse principalement aux entreprises et aux administrations pour qui la disponibilité d'une donnée de qualité est critique. Les producteurs et les diffuseurs prennent des engagements auprès de ces utilisateurs. La mission Etalab est chargée de la mise en oeuvre et de la gouvernance de ce nouveau service public. Elle référence l'ensemble des données concernées (Extrait : <https://www.data.gouv.fr/fr/reference>).

## SIG

**Le système d'information géographique est un système d'information conçu pour recueillir, stocker, traiter, analyser, gérer et présenter tous les types de données spatiales et géographiques.**

### Smart city (ou ville intelligente)

**La ville intelligente est un nouveau concept de développement urbain. Il s'agit d'améliorer la qualité de vie des citoyens en rendant la ville plus adaptative et efficace, à l'aide de nouvelles technologies qui s'appuient sur un écosystème d'objets et de services.**

### Socle Commun de Données Locales (SDCL)

Définit un ensemble de jeux de données prioritaires, communs aux collectivités territoriales et homogénéisés pour une publication en open data : Base adresse locale, catalogue, délibérations, équipements, IRVE, marchés publics, subventions.

## T

## **Tiers-archiviste**

Personne morale ou physique se chargeant, pour le compte d'un tiers, de gérer et d'assurer la conservation de ses archives.

## **Transparence algorithmique**

Devoir, pour l'administration publique qui fait usage d'un système algorithmique, d'informer les citoyens de l'utilisation d'un tel système à leur endroit, et droit de ces citoyens de demander des explications sur son fonctionnement.

# **U**

## **URL (Acronyme anglais de Uniform Resource Locator)**

Désigne l'adresse web d'un site Internet. Elle se présente sous la forme d'une série de caractères permettant l'identification d'un site Internet et/ou d'une page web.

# **X**

## **Xml (Extensible Markup Language)**

Désigne un langage informatique utilisé, entre autres, dans la conception des sites Web et pour faciliter les échanges d'informations sur Internet. Il est un des langages possible pour la publication des données en Open data.

### **Sources :**

- **Opendata France**
- **Administrateur général de la Donnée**
- **la Cada**
- **CNIL**
- **Wikipedia**
- **InfoWebMaster**
- **SIAF**