

1

C'est quoi ?

Apparu aux États-Unis et au Royaume-Uni dans les années 2000, en France en 2010...

Dans la logique des logiciels libres, les données libres !

Publiées par les collectivités voire parfois des entreprises ou des collectifs citoyens, des données numériques libres d'utilisation.

IR 4 - Open Data 2 Big Data ESAIP 20

2

Objectifs

Open Data Goal

- Bénéfices d'une démarche Open Data (principe et utilité)
- Quelles données diffuser : freins, bénéfices et méthodologie
- Aspects juridiques et financiers d'un projet Open Data
- Facteurs de succès d'un projet Open Data (stratégie, organisation, moyens, ...)

Open Data Organisation

- ✓ Créneaux d'environ 15 h
- ✓ Théorie appuyé par le retour d'expérience
- ✓ Travaux Pratique individuel et groupe systématiquement rapporté à i.bougarram@gmail.com
- ✓ Evaluation sur table

IR 4 - Open Data 2 Big Data ESAIP 20

3




D'abord, un peu d'histoire.....une longue histoire !
De la mutation industrielle au customizing

- ✓ Appliquez votre histoire
- ✓ Dominez votre histoire
- ✓ Faites vivre votre histoire

IR 4 - Open Data 2 Big Data ESAIP 20

4

Données ouvertes :



Le principe

l'avez-vous compris ?

5




Mise à disposition des données publiques
des administrations, organismes publics,
entreprises et citoyens



Pour les
citoyens



Pour les
entreprises




Pour les
collectivités


IR 4 - Open Data 2 Big Data ESAIP 20

6

► Réutilisation libre et gratuite



Licence
adaptée



Formats ouverts
et normalisés

► IR 4 – Open Data 2 Big Data ESAIP 20 ►

7

► Pour servir à la fois :

► À une gouvernance des états : transparente, participative, collaborative


► De « matière première » à des applications citoyennes, des produits et services de l'économie numérique mobile

► IR 4 – Open Data 2 Big Data ESAIP 20 ►

8


01 Présentation

- ▣ Les formats
- ▣ Les licences
- ▣ Etalab (data.gouv.fr)
- ▣ Open Data sur Angers / Nantes



1

9



Présentation - Les formats

Contenus statiques


- Tableurs : CSV / XLS
- JSON / XML
- Documents : PDF
- Images : JPG
- Zones : SHP et KML

Contenus dynamiques, sous forme d'API

- JSON/XML

IR 4 - Open Data 2 Big Data ESAIP 20

10



Présentation - Les licences

Le principe : définition par l'Open Knowledge Foundation
 « donnée qui peut être librement utilisée, réutilisée et redistribuée par quiconque - sujette seulement, au plus, à une exigence d'attribution et de partage à l'identique »



- Disponibilité et accès
- Réutilisation et Redistribution

-> licence Open Database License (ODbL)

Parfois des licences restrictives, loin de l'esprit Open Data
 Exemple la licence de la Base d'Adresses Nationale (BAN) : licence à « Usage Propre uniquement », avec interdiction de toute autre utilisation
 Revente des données est donc interdite (donc met en danger juridique toute application payante)

IR 4 - Open Data 2 Big Data ESAIP 20

11


Présentation – Etalab (data.gouv.fr)

« Service du Premier ministre en charge de l'ouverture des données publiques et du gouvernement ouvert. »

- Organisme d'état (dépend directement du premier ministre),
- Créé en 2011,
- Missions : centraliser, harmoniser les données, promouvoir
- Responsable du portail data.gouv.fr

Portail data.gouv.fr

- 19.983 informations gratuites et réutilisables
- centralise les données des portails locaux ou spécialisés



IR 4 - Open Data 2 Big Data ESAIP 20

12



Présentation – Open Data sur Angers et Nantes

Héberge les données de Métropole, de département, de la région

Portail Open Data Angers <https://data.angers.fr>
 Portail Open Data Nantes <https://data.nantesmetropole.fr/>
 Portail Open Data « votre ville de naissance » ?

Association nantaise Libertic
 « Promouvoir la démocratie en ligne à travers l'ouverture des données publiques. »
 Hébergée à la Cantine du Numérique (proche Hôtel Dieu)

OpenStreetMap Nantes
 Très actif sur Nantes
 Organise des Cartographies : « animation consistant à créer des données OSM - donc libres - et à initier de nouveaux contributeurs au projet OSM ».

IR 4 – Open Data 2 Big Data ESAIP 20

13




02 Jeux de données

2

- ▣ Principaux fournisseurs
- ▣ La Poste
- ▣ SNCF
- ▣ WikiData
- ▣ Portail Open Data Nantes
- ▣ Autres données

IR 4 – Open Data 2 Big Data ESAIP 20

14



Jeux de données – Principaux fournisseurs

Les principaux fournisseurs français de données :

- INSEE : statistiques, base SIRENE
- IGN : délimitation des zones
- ADEME : statistiques sur l'environnement
- CAF
- OpenStreetMap : carte
- RATP
- La Poste
- SNCF

IR 4 – Open Data 2 Big Data ESAIP 20

15


OpenData
 Les données libérées

Jeux de données – La Poste


 LE GROUPE LA POSTE


Quelques exemples de données :

- Liste des boîtes aux lettres de rue France métropolitaine et DOM
- Base officielle des codes postaux
<https://www.data.gouv.fr/fr/datasets/base-officielle-des-codes-postaux>
- Liste des bureaux de poste, agences postales et relais poste
- Liste des services disponibles en bureaux de poste, agences postales et relais poste
- Calendrier d'ouverture des bureaux de poste, agences postales et relais poste


Portail : <http://datanova.legroupe.laposte.fr>

IR 4 – Open Data 2 Big Data ESAIP 20

16


OpenData
 Les données libérées

Jeux de données – SNCF


SNCF OPEN DATA

Exemples d'API :

- calcul d'itinéraire en train de gare à gare
- recherche d'horaires planifiés de gare en gare
- recherche automatique de noms de gare
- grille horaire de ligne
- retards et suppressions de train

Licence SNCF Open Data
 Un volume d'extraction de 20 requêtes par minute et par terminal connecté

Portails : <https://data.sncf.com/> et <http://doc.navitia.io/>

IR 4 – Open Data 2 Big Data ESAIP 20

17


OpenData
 Les données libérées

Jeux de données – WikiData

Données brutes Wikidata



Infobox Wikipédia



Base de données libre et collaborative


- Données stockées au format RDF (développé par le W3C)
RDF: Chaque information est un triplet élément - propriété - valeur
- Langage de requêtes SPARQL (développé par le W3C)
- Licence Creative Commons CC0 → réutilisables par tous

Exemple de données : les résultats de courses cyclistes
 Les éditeurs renseignent sur Wikidata. Les résultats apparaissent dans Wikipédia.

En cours : intégration du Wiktionnaire (dictionnaire collaboratif) et de Wikimedia Commons (fichiers libres)

IR 4 – Open Data 2 Big Data ESAIP 20

18


 OpenData
Les données libérées

Jeux de données – exemple de l'Open Data Nantes

Portail unique de la ville, la métropole, le département, la région
 Nantes (et sa métropole) : 200 jeux de données
 Département : 190 jeux de données
 Pays de Loire : 340 jeux de données

Exemples de données :

- Menus des cantines scolaires de la ville de Nantes
- Aires de covoiturage de Nantes Métropole
- API : getDisponibiliteParkingsPublics
- API : getFluiditeAxesRoutiers
- API : getInfoTraficTANPrevisionnel
- API : getInfoTraficTANReel
- API : getTempsParcours



IR 4 – Open Data 2 Big Data ESAIP 20

19

 OpenData
Les données libérées

Jeux de données – Autres données

 Prix du carburant
 • <http://prix-carburants.gouv.fr/>

 Carte des Bornes de Recharge pour Véhicules Électriques

 Base de données sur les produits alimentaires
 • Open Food Facts (76893 produits)

IR 4 – Open Data 2 Big Data ESAIP 20

20

03 Exemples d'application

- OpenStreetMap
- Mieux trier à Nantes

IR 4 – Open Data 2 Big Data ESAIP 20

21

03 Exemple d'application

3

- OpenStreetMap
- Mieux trier à Nantes

22

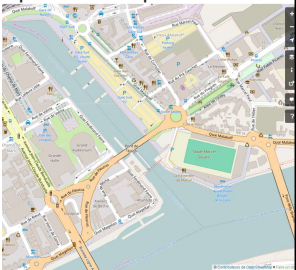
Exemples d'application - OpenStreetMap

Le Wikipédia de la carte
Alternative à Google Maps

Des applications de visualisation :
OsmAnd, MapsMe...

Saisies manuelles (préférées aux imports de masse)

Licences :
ODbL pour les données, CC-by-sa pour les cartes créées avec les données



IR 4 - Open Data 2 Big Data ESAIP 20

23

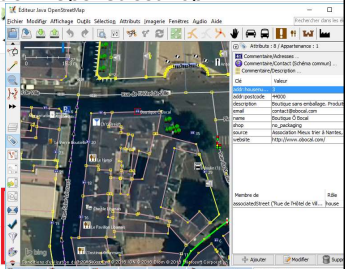
Exemples d'application - OpenStreetMap

Contribuer

Consulter le dictionnaire des balises
<http://wiki.openstreetmap.org/wiki/FR:Balises>

Éditer avec sur openstreetmap.org
(éditeur intégré au navigateur)

Éditer avec le logiciel libre JOSM



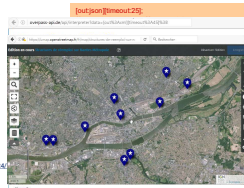
IR 4 - Open Data 2 Big Data ESAIP 20

24

Exemples d'application - OpenStreetMap


Comment utiliser ces données ? Un exemple

1. Faire des requêtes GeoJSON sur la base O.S.M.
GeoJSON : format ouvert pour coder une donnée géographique
Exemple : <http://overpass-turbo.eu/s/bDS>
2. Exporter les données brutes depuis l'API Overpass
Exemple : [https://overpass-api.de/api/interpreter?data=\[out:json\]\[out:geojson\]\[out:json\]\[3:45E,38:52\]](https://overpass-api.de/api/interpreter?data=[out:json][out:geojson][out:json][3:45E,38:52])
3. UMap permet de créer des cartes.
Fonds OpenStreetMap + données distances de OverPass.
Exemple : https://umap.openstreetmap.fr/fr/map/par-roschets-et-liebourg_-74843f13c461562a



IR 4 – Open Data 2 Big Data ESAIP 20

25


OpenData
Les données libérées

Exemples d'application – Mieux trier à Nantes



Objectif : trier plus régulièrement, plus soigneusement
regrouper tout ce qui concerne le tri sur Nantes Métropole

Créé par l'association « Mieux trier à Nantes »

Utilise les données Open Data

Logiciel libre « Mieux trier » : code, sources, photos

Application mobile

www.mieuxtrieranantes.fr


IR 4 – Open Data 2 Big Data ESAIP 20

26

Exemples d'application – Mieux trier à Nantes

IR 4 – Open Data 2 Big Data ESAI2P

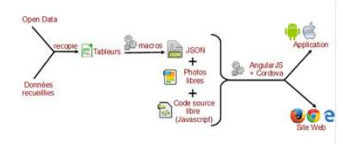
27


OpenData
 Les données libérées

Exemples d'application – Mieux trier à Nantes

Données utilisées (portail Open Data Nantes)

- Jours et modes de collecte à domicile → 3100 rues
- Colonnes enterrées et aériennes → 1297 conteneurs
- Horaires des déchèteries/écopoints (et déchets admis)
- Les lieux/horaires des distribution Trisars → 175 lieux



IR 4 – Open Data 2 Big Data ESAIP 20

28


OpenData
 Les données libérées

Exemples d'application – Mieux trier à Nantes

Données fournies au portail Open Data Nantes

Mieux trier à Nantes fournit la liste des structures de ré-emploi
<http://data.nantes.fr/donnees/detail/structures-de-re-emploi-de-loire-atlantique/>



IR 4 – Open Data 2 Big Data ESAIP 20

29

Remarques

Collectivités/entreprises publiques :

- nombreuses données intéressantes.
- de plus en plus de collectivités.

Les citoyens : très actifs.

Entreprises privées : peu de données.

IR 4 – Open Data 2 Big Data ESAIP 20

30

4

Données ouvertes :

Enjeux Et motivations

31

► Pour les individus et collectifs citoyens, les associations

- Accès direct à l'information
- Transparence
- Démocratie plus "humaine"
- Développement personnel et économique

► IR 4 - Open Data 2 Big Data ESAIP 20 ►

32

► Pour les Entreprises

- Création de nouveaux services
- Accès à une information agrégée et efficace
- Valorisation en termes d'image
- Amélioration des données via le *crowdsourcing*
- Gains financiers (par économies ou croissance)

► IR 4 - Open Data 2 Big Data ESAIP 20 ►

33

► Pour les administrations

- Accès à une information agrégée et efficace
- Valorisation de leur travail
- Amélioration des données via le *crowdsourcing*

...Et réponse à une obligation légale

► IR 4 – Open Data 2 Big Data ESAIP 20 ►

34

5

Les applications Existantes Et potentielles

35

3 exemples d'applications :

- Handimap (Rennes, France)
- Crimespotting (san francisco, USA)
- Home Sales (New York City, USA)

Dérives possibles :

- Impact sur un quartier (image, prix du loyer)
- Spam ciblé

► IR 4 – Open Data 2 Big Data ESAIP 20 ►

36

Et potentiellement :

- ▶ Optimisation du dernier kilomètre (transporteurs, livreurs de pizzas, etc)
- ▶ assistant de déplacements (mashup route/bus/pollution/neige par exemple)
- ▶ Suivi des dépenses collectives, avec possibilité d'invalidation par les utilisateurs

...Et plein d'applications encore inconnues !

IR 4 - Open Data 2 Big Data ESAIP 20

37

L'absence d'ouverture Les freins à l'open data

6

38

L'open data effraie ?



Les freins peuvent être :

- ▶ Culturels,
- ▶ Financiers,
- ▶ Méthodologiques,
- ▶ Un blocage Humain,
- ▶ La peur de l'échec.



En l'absence d'ouverture :

- ▶ Crowdsourcing,
- ▶ Parsing manuel,
- ▶ Collecte manuelle,
- ▶ Applications non officielles (zone grise)

IR 4 - Open Data 2 Big Data ESAIP 20

39

7

La stratégie De mise en place De l'open data

40

Un processus itératif

Partir d'un
existantObtenir des
soutiensDévelopper
un MVP

5 bonnes pratiques :

- ▮ S'inspirer des initiatives voisines
- ▮ Animer la communauté
- ▮ Soutenir les initiatives
- ▮ S'appuyer sur des experts
- ▮ Accepter l'échec pour progresser

IR 4 - Open Data 2 Big Data ESALP 20

41

8

Notre Collaboration Avec le Grand Lyon

42

► Première phase

- Prise de contact à l'occasion du barcamp #offwww2012
- Découverte du projet opendata du Grand Lyon
- Premiers rendez-vous

Planning initial : printemps 2012 pour un premier lot, 2013 pour un lot plus conséquent

► IR 4 - Open Data 2 Big Data ESAIP 20 ►

43

► Deuxième phase

- Consultation du collectif sur des propositions de données à ouvrir en priorité
- Travail collaboratif pour évaluer les fournisseurs pour de nombreux jeux de données
- Participation aux tests de réutilisation

Bilan : ouverture très liée à des décisions politiques, mais ouverture tout de même !

► IR 4 - Open Data 2 Big Data ESAIP 20 ►

44

Exemple :

**Les données
Ouvertes dans
Les transports**

10

45

Trois acteurs, trois avancements :

- ▮ SNCF, ouverture progressive
- ▮ Grand Lyon, en cours de négociation
- ▮ TCL, situation bloquée

Face au blocage :

- ▮ Solutions **curatives** :
Lobbying, sensibilisation
- ▮ Solutions **palliatives** :
Crowdsourcing...



IR 4 - Open Data 2 Big Data ESAIP 20


46

11

Exemple :
L'accompagnement
citoyen

47

Un objectif :
Enrichir l'expérience de l'élection pour
chaque citoyen.

 **Sources de données
administratives**

Data-visualization pour offrir une
solution de référence et accessible aux
citoyens.

IR 4 - Open Data 2 Big Data ESAIP 20

48
