

Études d'approfondissement Open Data

Simon Planès (RICM5)

Polytech Grenoble

11 octobre 2013

Sommaire

1 Les données ouvertes

- Historique
- Définition
- Applications

2 Aspects techniques

- Licences
- Web des données
- Description des données
- Utilisation des données

3 Démonstration

Historique de l'ouverture des données

- DUDH :

La Société a le droit de demander compte à tout Agent public de son administration.

- 2007 : Sunlight Foundation

- USA : Obama + data.gov (2009)

- France :

- data.gouv.fr (2011)

- pionniers : Paris, Rennes

- 2013 : obligation pour les collectivités locales de plus de 3500 habitants

Qu'est-ce qu'une « donnée ouverte » ?

Données ouvertes

- libre disponibilité pour tous et chacun
- sans restriction de copyright, brevets ou d'autres mécanismes de contrôle

⇒ **informations publiques brutes utilisables par tous**

123opendata.com

Objectif premier : servir la **démocratie**

Exemples d'applications et conséquences

Applications

- cartographie, transports, ...
- à différentes échelles

Conséquences

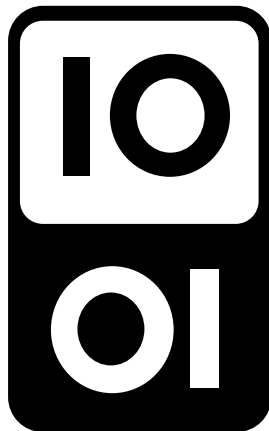
- publiques et sociales
ex : transparence des échanges économiques, adresses des délinquants sexuels répertoriés
- économiques
ex : concurrence équitable, baisse de la consommation énergétique



OpenStreetMap

Les licences

- Open Database License
- Open Knowledge Foundation : 3 types de licences
- Creative Commons
- Licence Ouverte (France)



Licence Ouverte

Le Web des données (2006)

Définition du W3C

Données **structurées** et **liées** entre elles pour pouvoir être exploitées par les machines (équivalent de l'hypertexte pour les données)

Règles (Tim Berners-Lee)

- 1 utiliser des URIs pour nommer les choses
- 2 utiliser des URIs HTTP pour les consulter
- 3 donner des informations utiles sur les URIs (standards RDF, SPARQL)
- 4 inclure des liens vers d'autres URIs

Aboutir à un **Web intelligent**

Description des données

Resource Description Framework

Définition (wikiversity.org)

Un document structuré en RDF est un **modèle de graphe** destiné à décrire de façon formelle les ressources Web et leurs métadonnées, de façon à permettre le traitement automatique (par des machines) de telles descriptions.

Document RDF = ensemble de **triplets** (*sujet, prédicat, objet*)



Interrogation des graphes RDF

SPARQL Protocol and RDF Query Language

- recommandation du W3C
- généralement utilisé avec un langage de programmation (ex : PHP)

```
PREFIX rdfs: <http://www.
w3.org/2000/01/rdf-
schema#>
SELECT ?subject ?label
WHERE { ?subject rdfs:
label ?label } LIMIT
100
```

exemple : DBpedia

```
{{Infobox Town AT |
name = Innsbruck |
image_coa = InnsbruckWappen.png |
image_map = Karte-tirol-I.png |
state = [[Tyrol]] |
regbek = [[Statutory city]] |
population = 117,342 |
population_as_of = 2006 |
pop_dens = 1,119 |
area = 104.91 |
elevation = 574 |
lat_deg = 47 |
lat_min = 16 |
lat_hem = N |
lon_deg = 11 |
lon_min = 23 |
lon_hem = E |
postal_code = 6010-6080 |
area_code = 0512 |
licence = I |
mayor = Hilde Zach |
website = {http://innsbruck.at} |
}}
```

Innsbruck




Country	Austria
State	Tyrol
Administrative region	Statutory city
Population	117,342 (2006)
Area	104.91 km²
Population density	1,119 /km²
Elevation	574 m
Coordinates	47°16' N 11°23' E
Postal code	6010-6080
Area code	0512
Licence plate code	I
Mayor	Hilde Zach
Website	www.innsbruck.at

dbpedia.org

Démonstration : Yahoo ! Pipes

