

HOMBERT Sarah

BOURON Priscilla

M1 MIAGE Alternance

# Compte-rendu Distanciel Open Data

X1IP020 : Données sur le web



UNIVERSITÉ DE NANTES

# SOMMAIRE

<b>SOMMAIRE</b>	<b>1</b>
<b>Introduction</b>	<b>2</b>
<b>Les données</b>	<b>3</b>
Modèle de données	3
Circuit Loire à vélo	3
Etablissement “Accueil vélo”	3
Analyse des données	4
Typologie et accessibilité	4
Réutilisabilité	4
<b>Les applications</b>	<b>4</b>
<b>Retour sur le projet</b>	<b>6</b>
<b>Conclusion</b>	<b>6</b>
<b>Sitographie</b>	<b>7</b>

## Introduction

L'open data, ou base de données ouverte est un ensemble d'informations dont l'utilisation est à disposition des usagers. Les organisations privées ou publiques sont incitées à mettre à disposition leurs données sur le web. Ce mouvement a pris de l'ampleur à tel point que les organisations publiques doivent désormais mettre en oeuvre la démarche pour ouvrir leurs données.

La mise en place de l'open data permet une nouvelle utilisation de ces données pour en tirer une valeur supplémentaire, soit à but analytique, soit pour répondre à un besoin particulier.

Notre projet a pour but de donner de la valeur à des jeux de données bruts mis à disposition de tous, notamment en les croisant pour faire ressortir de nouvelles informations. La thématique du projet concerne le circuit de "Loire à vélo", itinéraire de tourisme à vélo qui longe la Loire. Il s'agit de mettre en lien les données de cet itinéraire avec les données des établissements de la marque nationale "Accueil vélo". Cette marque permet de recenser les hébergements, réparateurs, restaurateurs ou encore des sites de visite le long des itinéraires cyclables.

## 1. Les données

### a. Modèle de données

#### i. Circuit Loire à vélo

Les données concernant le circuit de la Loire à vélo sont stockées en tracé : il s'agit d'une suite d'enregistrement de données de localisation, avec pour chaque enregistrement des métadonnées associées. Elles sont distribuées par le site de données ouvertes de la région Pays de la Loire.

Les données comprennent chacune les champs suivants :

- Localisation : les coordonnées 2D
- Geo Shape : un type de données pour l'utilisation de l'API GeoShape qui permet de faciliter l'utilisation de données de localisation sur des cartes dynamiques.
- Code et Libellé du département
- Nom et longueur du tronçon
- Etat : renseigne la catégorie de voie proposée (par exemple, route partagée ou piste dédiée aux cyclistes)
- Longitude

Un champ supplémentaire est utilisé : "st\_length", mais aucune description n'est précisée, il n'est donc pas possible de préciser correctement ce qu'il représente.

#### ii. Etablissement "Accueil vélo"

Le distributeur des données est le département Loire Atlantique. Ces données sont plus classiques : on y trouve des informations usuelles concernant les établissements listés.

Les champs répertoriés sont les suivants (liste non-exhaustive) :

- l'identifiant, le nom, les adresses principales et secondaires de l'établissement
- Le code postal, la commune, le cedex
- Les coordonnées téléphoniques, postales et autres
- Les tarifs
- La classification principale(hôtel, chambres d'hôtes...) plus potentiellement une secondaire
- L'appartenance à un certains nombres de marques et labels concernant les itinéraires et dispositifs de vélos ainsi que la restauration ou l'hôtellerie
- La validation de la présence ou non d'un restaurant, parking, piscine ou autres équipements
- Les langues parlées par le personnel de l'établissement
- Les capacités en personnes et en chambres
- La latitude et la longitude.

## b. Analyse des données

### i. Typologie et accessibilité

Les deux jeux de données ont des **degrés de complexité très différents**. Les données concernant les établissements "Accueil vélo" sont **compréhensibles au premier abord**, et exploitables facilement à partir d'un simple tableau. Ce n'est pas le cas de l'autre jeu de données : bien que les métadonnées associées apportent quelques précisions. Du fait du **caractère géographique** des données, ces dernières prennent **peu de sens lorsqu'on les observe telles qu'elles**. Il est donc nécessaire d'avoir un prétraitement pour en comprendre le sens.

Du point de vue de l'accessibilité, les données sont disponibles dans différents formats : CSV, JSON et Excel pour les deux jeux de données, plus des formats de fichiers géographiques spécifiques pour le tracé de la Loire à Vélo. Cela représente déjà une certaine accessibilité, mais de nouveau, on a besoin d'avoir une API supplémentaire pour exploiter les données géographiques.

Pour autant, la disponibilité pour tous et chacun préconisée par l'open data est respectée dans une certaine mesure dans le sens où le distributeur permet à l'utilisateur de visualiser les données sur une carte sur le site.

Les données traitées sont des données dites "statiques" : elles sont peu fluctuantes, et exploitables sans besoin de rafraîchissement courant.

Les données du circuit de la Loire à Vélo sont mises à jour régulièrement, la dernière modification remontant au 1 octobre 2019. La fréquence de mise à jour n'est pas précisée, mais nous pouvons attester que les données sur lesquelles nous travaillons ont été traitées il y a peu.

On ne peut pas en dire autant pour ce qui est des établissements Accueil Vélo : il est indiqué qu'un dernier traitement a été effectué en 2018, alors que dans les informations de modification, la date remonte à 2012. De plus, une recherche sur le site officiel France vélo tourisme a permis de constater que certains établissements n'étaient pas présents dans la liste, de même que sur le site Tourisme Loire Atlantique, qui est indiqué comme référence des données dans les informations concernant le jeu de données. Il y a donc un clair **manque de maintenance** sur ces données.

### ii. Réutilisabilité

Les jeux de données respectent la licence Open Data Commons : **Open Database License** (ODbL). Cette dernière permet de copier, distribuer, réutiliser, modifier ou encore adapter les données concernées. Cette licence est une des licences à privilégier pour des données ouvertes. Il n'y a donc aucune restriction particulière tant qu'on applique aux réutilisations faites les mêmes préceptes que ceux préconisés par la licence, et qu'on spécifie quelle licence de données est utilisée.

## 2. Les applications

### a. Applications existantes

Les distributeurs proposent une section similaire sur leurs jeux de données : les réutilisations. Pour autant, aucune n'est spécifiée, sur l'un comme sur l'autre.

Nous avons cependant trouvé des services sur les sites Loire à Vélo et France Vélo Tourisme qui proposent de visualiser le circuit de la Loire à vélo tout en pouvant sélectionner un certains nombre de services à afficher sur la carte, dont les établissement Accueil vélo. Nous pouvons constater que le distributeur de données ouvertes de cartographie OpenStreetMap est cité, ce qui nous permet de savoir d'où viennent les données du tracé de la Loire à vélo, en revanche, rien n'indique la provenance des informations concernant les établissements cités sur le site. De plus, ces données concernent l'ensemble du circuit, ce qui n'est pas le cas de notre étude. En effet, notre jeu de données se cantonne uniquement au département Loire Atlantique.

Enfin, il n'apparaît pas exister d'application mobile à part entière qui croisent ces types de données. Les applications existantes concernent principalement des itinéraires et balades cyclistes de manière générale.

### b. Notre application : les établissements accueil vélos sur le circuit de la Loire à vélo

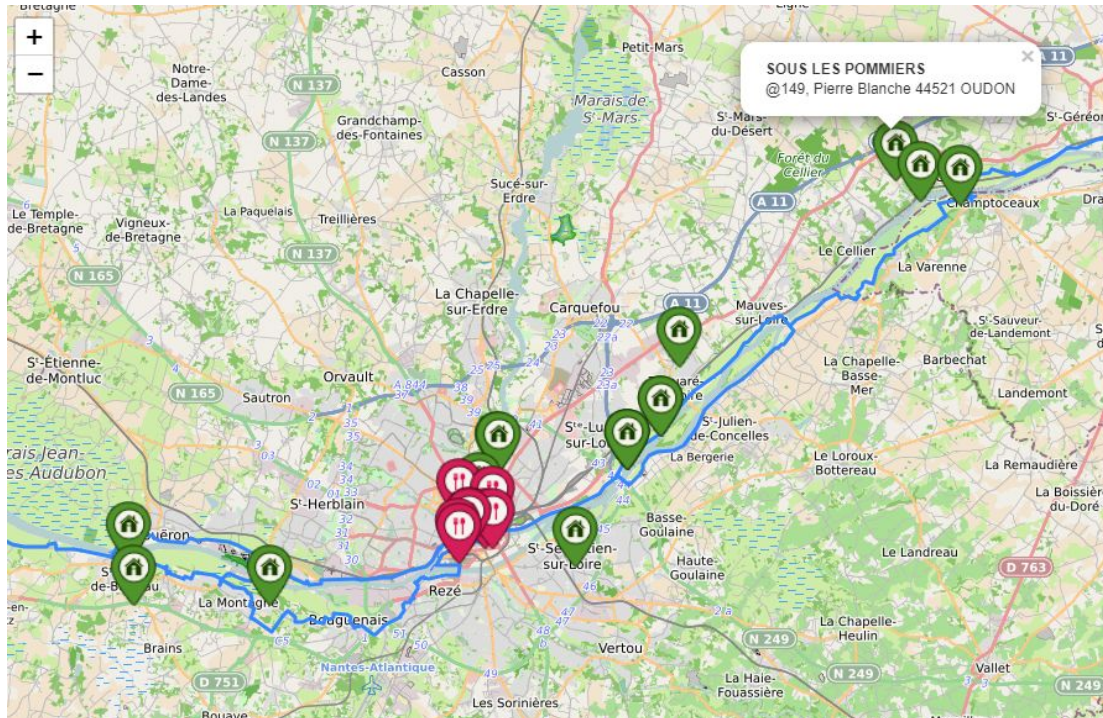
Notre choix s'est porté sur un service similaire à ceux existants, mais centré sur le département Loire Atlantique.

Nous souhaitons donc faire une application cartographique centrée sur la Loire-Atlantique, où apparaîtront : le circuit de la Loire à vélo du département, ainsi que les lieux où dormir, se restaurer, visiter, ou encore réparer son vélo à proximité du circuit. Il est ensuite possible de visualiser les informations détaillées concernant un établissement particulier, à l'aide d'une interface simplifiée et intuitive.

Notre choix s'est porté sur un site web, bien qu'il aurait été intéressant de réaliser une application mobile. Nous avons en effet plus de compétences en ce qui concerne les langages html et Javascript que sur la plate forme de développement mobile android.

Pour ce qui est de l'exploitation des données, nous avons utilisé le jeu de données au format json pour les établissements Accueil vélo, et le jeu de données au format GeoJson pour le circuit de la Loire à vélo. Nous n'avons pas réalisé de traitement supplémentaire sur ces derniers. Le format GeoJson rend l'exploitation des données beaucoup plus simple avec des APIs existantes.

Notre maquette finale présente ainsi une page principale qui affiche la carte centrée sur Nantes avec le tracé de la Loire à vélo et les établissements indiqués avec des marqueurs autour de ce trajet (plus ou moins 5km à la ronde).



Carte de notre application web

### 3. Retour sur le projet

Le choix du projet a été compliqué, ne sachant pas quel sujet choisir, ni quelles données pourraient être pertinentes. Notre choix s'est porté sur des données à cartographier, le besoin fonctionnel lié à cette représentation étant intéressant.

Après études des données, nous avons pu former notre réflexion autour de l'exposé à présenter, et du "Proof of concept" à développer.

Il est important de noter que le rendu de notre application est intéressant, mais malheureusement les données manquent de rafraîchissement et de ce fait il serait compliqué de l'utiliser telle qu'elle.

### Conclusion

Ce projet nous a permis de mettre en lumière les différents concepts de l'open data avec un exemple d'étude de données ouvertes. On a pu constater les avantages et inconvénients de travailler sur ce type de données, et également qu'il reste encore des améliorations à faire par les distributeurs du côté du maintien des données.

## Sitographie

- data.paysdelaloire.fr, le 25/09/2019, "Parcours de la Loire à Vélo" - [https://data.paysdelaloire.fr/explore/dataset/234400034\\_parcours-de-la-loire-a-velo/custom/](https://data.paysdelaloire.fr/explore/dataset/234400034_parcours-de-la-loire-a-velo/custom/)
- data.loire-atlantique.fr, le 25/09/2019, "Accueil vélo en Loire Atlantique" - [https://data.loire-atlantique.fr/explore/dataset/793866443\\_accueil-velo-en-loire-atlantique/information/](https://data.loire-atlantique.fr/explore/dataset/793866443_accueil-velo-en-loire-atlantique/information/)
- France vélo Tourisme, le 01/10/2019, "Les services Accueil Vélo" - <https://www.francevelotourisme.com/accueil-velo>
- Géo-shape - <https://www.elastic.co/guide/en/elasticsearch/reference/current/geo-shape.html>  
01/10/2019
- Tutoriel 1 - <https://nouvelle-techno.fr/actualites/2018/05/11/pas-a-pas-inserer-une-carte-openstreetmap-sur-votre-site> 21/10/2019
- Tutoriel 2 - <https://gist.github.com/alanpilloud/b0ecfaf6717ed18d8e4796fe800b43cd>
- Tutoriel 3 - <https://www.bwap.ch/geolocalisation-avec-javascript>
- Documentation API Leaflet pour cartes - <https://leafletjs.com>