



Projet de valorisation des données relatives aux voies de Lyon

*Coupler Wikidata, Wikimedia Commons et
OpenStreetMap*



Rapide état des lieux, janvier 2019

- Fichier **FANTOIR**, relativement complet mais pas d'information géographique
- Données du portail **Open Data Lyon** : beaucoup de lacunes mais quelques données intéressantes comme les voies à venir
- **OpenStreetMap**, très complet, quelques manques. Avantage lyonnais : la plupart des voies sont mappées en tant que relation
- **Wikidata/Wikipédia** : quelques dizaines d'entités Wikidata ; quelques articles fr.wikipedia



Objectifs et démarche

- Bénéficier des informations de chaque source (OSM, FANTOIR, Open Data Lyon) pour construire un **jeu de données**
- Créer un jeu de données qui permet d'accéder à toutes les informations disponibles (Wikipédia, images de la voie, OSM, éventuellement la géométrie de la voie) consolidées dans **Wikidata**



Les données d'une voie dans Wikidata

- Exemple de la [rue Garibaldi](#) à Lyon
- Nature de l'élément : voie (rue) ou une des sous-classes de “voie de communication” ([SPARQL](#))
- Propriétés retenues : nom, description, coordonnées approximatives du milieu de rue, arrondissement, pays, nommé en référence de, code FANTOIR, relation OSM, image, image du panneau de nom de lieu (+ éventuellement wp, commons)
- On pourrait par la suite aller plus loin, par exemple avec la propriété “Limitrophe de” : [rue René-Goscinny](#) à Paris



Wikidata dans OpenStreetMap

- Précision préalable : mapping des voies par relation à Lyon
- Le tag Wikidata plutôt mis sur la relation et non sur les ways, exemple de la rue Garibaldi à Lyon :

<https://www.openstreetmap.org/relation/3847929>



Les étapes nécessaires

1. Consolider un jeu de données, un fichier CSV, à partir de FANTOIR, d'OpenStreetMap, des données issues du Grand Lyon
2. Vérifier les doublons (les entités déjà présentes dans Wikidata)
3. Faire tourner un script d'insertion des données dans Wikidata avec les propriétés suivantes : nom, coordonnées, localisation administrative, Pays, code FANTOIR, code de la relation OSM



Un test d'insertion : les stations Vélo'v

- Aucun doublon possible ; création de l'entité [station de vélopartage](#)
- Obtention des données ouvertes disponibles en plusieurs endroits dont [ici](#) sous format CSV
- Script Python d'insertion, [ici](#)
- Requête SPARQL des stations Vélo'v à illustrer, [ici](#)
- La requête équivalente Over-Pass (OSM), [ici](#)
- Tag Wikidata sur le node OSM :
<https://www.openstreetmap.org/node/653889184>



Étape 1 : obtenir la liste des voies depuis OSM

- OpenStreetMap : requête par arrondissement ; [exemple](#) :
<http://overpass-turbo.eu/s/GDz>
- Obtention de la liste des noms de rues depuis le json obtenu : `cat voies.json | egrep -o '"name": "(.*?)"' | sed 's/"name": //' | sort | uniq > voies.txt`



Étape 2 : obtenir la liste des voies depuis FANTOIR

- On obtient le fichier des voies FANTOIR, par exemple depuis le site OpenDataSoft, exemple pour le [1er arrondissement](#) ; clé : code INSEE
- On consolide un fichier CSV avec ce type de [script](#) qui ressemble à cela :

FANTOIR	VOIE	NOM	Code Annulation
6900190004X	rue	Rue des chênes	



Étape 3 : consolidation du fichier csv entre OSM, FANTOIR et DATALYON

- On élimine les doublons
- On élimine les voies disparues (code annulation FANTOIR, données [DATALYON](#))
- On élimine les voies déjà présente sur Wikidata
- On ajoute manuellement la coordonnée géographique (milieu approximatif) et l'ID de la relation OSM



Étape 4 : ajouter les données dans Wikidata

Script d'insertion Wikidata similaire à celui des Vélo'v

- “Coordination” en ligne : [page projet](#) sur Wikipédia francophone :
- [Documentation](#) en cours de rédaction sur WikiLivres (Wikibooks) : inclut les scripts OSM, FANTOIR et insertions Wikidata



Le suivi en direct de l'illustration des voies

- **Bilan à mi-juin 2019**
 - arrondissements 1er, 2e, 3e, 4e, 6e, 7e, 8e : exhaustif sur Wikidata
 - arrondissements aux voies illustrées proche de 100 % : 2e, 6e et 7e
 - arrondissements aux voies illustrées proche de 75 % : 1er, 3e, et 8e
- Une rue est considérée comme illustrée si une vue générale existe sur Commons et qu'elle est utilisée sur Wikidata
- Requêtes : [places uniquement](#) et [voies hors places](#)



Un projet d'illustration qui avance tout seul ?

- D'une certaine manière, oui.
- Effet secondaire positif : une fois la rue elle même illustrée et les bâtiments historiques également (MH etc.), la photographie s'oriente vers le "quotidien" ce qui bénéficie à la cartographie :
 - les devantures de magasins, de restaurants
 - le mobilier urbain (bouches incendie, armoires de rue, distributeurs, bancs, arbres etc.)
 - les plaques commémoratives, les dates d'architectes
 - etc.

Exemples de “bénéfices” illustrations Commons => OSM (a)

- Le projet OSMhydrant utilise ce type d'image.
- [Category:Red fire hydrants in Lyon](#)
- <https://www.openstreetmap.org/node/6237630996>



Exemples de “bénéfiques” illustrations Commons => OSM (b)

- Armoire de rue
- Caméra de surveillance
- Plaque commémorative
- Date et mention de l'architecte





A faire encore...

- L'insertion exhaustive des voies manquantes sur Wikidata
- L'illustration exhaustive sur Wikimedia Commons
- La génération d'un fichier GeoJSON reprenant les données issues de Wikidata, l'ID de la relation OSM et surtout la géométrie de la relation
- Discussions pour mettre un tel jeu de données à disposition sur le portail open data du Grand Lyon



Remerciements

Merci à Romain Behar et Sébastien Gathier avec qui j'ai présenté ce projet aux JDLL 2019 et qui ne pouvaient pas être présents à SOTMFR 2019.

Merci à Romain, Justine et Sophie pour les relectures des slides.

[Benoît Prieur](#) sur OpenStreetMap ; [@benprieur](#) sur Twitter