

PROJET BAN  
COMPTE-RENDU

---

**Objet de la réunion :**

Audio BAN Team  
IGN / La Poste / Etalab

---

**Date** 20/09/2016

---

**Participants :**

ETALAB : Christian

IGN : Elodie, Olivier

LA POSTE : Laurent, Frédéric, Christophe

validation du dernier CR

Rédaction : Elodie

**Amélioration du processus d'initialisation :**

Actuellement, ce processus est lent ; à améliorer. Le code actuel ne tire pas parti du multi-thread autant qu'il pourrait. Ce n'est pas bloquant et n'a aucun impact pour le fonctionnement de l'API au final, mais ceci ralentit la phase d'init et de tests.

Nouvelle base : Etalab puis envoie un mini-dump (départements 06 et 33) d'ici demain après-midi.

Améliorations :

- rapprochement de voies et donc diminution du nombre de doublons : (100 000 voies rapprochées par « chemin – CH » et 200 000 restantes : ). Test de rapprochement en cours sur les noms en utilisant l'extension pg\_trgm (trigram) de postgresql similarity(text1, text2) : similarité entre 2 chaînes, retour entre 0 et 1, considéré comme OK à partir de 0.8
- ajout d'étapes d'initialisation pour intégrer les groupes et les housenumbers La Poste et IGN non appariés avec les données DGFIP e

**Amélioration des anomalies identifiées sur les données producteur :**

Question : améliorations à faire en amont ou aval de la BAN ?

A priori, en amont.

Christophe a identifié des mauvais appariements La Poste/IGN dans l'export IGN (<https://github.com/BaseAdresseNationale/ban/issues/154>). L'analyse est à poursuivre (action Christophe et Olivier). L'idée est d'améliorer les données IGN avant de charger la BAN avec celles-ci.

### **Amélioration des données producteur envoyées pour l'initialisation :**

Prise en compte des adresses/voies détruites ou annulées dans les données Ceci permettrait par exemple de récupérer environ 200 000 voies (parfois avec code fantoir et identifiant Poste). Pas de chaînage facilement exportable des données IGN détruites.

Olivier fait un échantillon pour Christian et Christophe sur 33, 06 pour la fin de semaine.

Remarque : les données « détruites » du fantoir sont actuellement chargées dans la ban : on a « Fantoir => annul » : référence objets détruits.

### **Cas des extension de numéros avec confusion B/BIS T/TER.**

L'IGN confirme bien que des B peuvent correspondre en fait à des BIS et que des doublons B+BIS peuvent être présents alors que seul B ou BIS existe sur le terrain.

Il est convenu de conserver en priorité l'extension fournie par La Poste, et à défaut celle provenant des données DGFIP extraites par les scripts BANO.

Cette logique est à intégrer dans la prochaine init.

### **Question de la traçabilité des sources sur HN et sur groupe:**

Actuellement, c'est la dernière source qui est enregistrée dans attributs sur housenumber et group, malgré le fait que ces entités peuvent avoir été fourni par la DGFIP, La Poste, l'IGN, une commune. L'utilisateur peut savoir qu'un groupe ou un housenumber provient de La Poste ou l'IGN en regardant les identifiants IGN et La Poste de ces objets. Pour les groupes, il y a le code fantoir. Une simple interrogation du champ attributs ne suffit pas pour récupérer les sources multiples. . On a aussi le détail dans les versions successives des objets, source n'indiquant finalement que la dernière source utilisée lors de l'init qui a modifié l'objet. Olivier ouvre un ticket github sur ce sujet pour avoir un avis plus approfondi des uns et des autres.

### **Question du stockage des appariements N-M dans la BAN :**

La recette de l'initialisation a mis en évidence des cas d'appariement N-M. C'est à dire un lieu-dit représenté par N noms différents à la DGFIP (Ex : pechot sud, pechot nord, pechot est, pechot ouest) et par un seul nom dans les données de La Poste et l'IGN (pechot). Comment stocker ce lieu-dit dans la BAN. Christian fait une proposition en séance. Olivier va ouvrir un ticket github pour expliquer le problème et échanger sur la bonne solution à retenir.

### **Evolution technique de Falcon à Flask :**

Pas d'info, ni sur l'avancement de l'évolution technique ( a priori ne modifie pas l'installation de l'API), ni sur le code pour donner la route des ressources étendues.

#### **Demande de token :**

Maël Reboux a fait une demande de token pour accéder à l'instance initialisée et commencer à développer un prototype. → OK pour les membres de la BAN Team.

#### **Guichet adresse :**

Redirection permanente du guichet adresse faite.

#### **Livraison BAN v0 hebdo :**

Etalab signale que la dernière livraison n'avait pas mis à jour les DOM. Ceci est dû à une refonte interne côté IGN avec la base DOM qui est en cours de fusion avec la base métropole.

Etalab signale aussi que le département 975 (St Pierre et Miquelon) est vide depuis plusieurs mois.

#### **Infos complémentaires :**

Sprint prévu avec des américains (Département des transports) cette semaine pour tester ADDOK avec leurs données (200 millions d'adresses).

Remontée de demandes pour un outil de recherches d'adresse avec offre packagée, style ADDOK (10 à 15 millions de requêtes par semaine).

#### **Agenda :**

- Dump 33 et 06 prévu le 21/09 puis recette IGN et LP.
- France entière : début octobre puis recette utilisateurs et IGN/LP avec fusions de communes
- France entière définitive : prévu fin de l'année si recette OK

<b>Prochain Point BAN Team le Mardi 27/09/2016 à 14h00</b>
--

#### **Relevé de décisions :**

Actions	Responsables	échéance
Écriture de ticket github	Olivier	
Correction du processus d'initialisation de la BAN suite à la recette (chargement du COG 2016, chargement des données IGN/LP manquantes...)	Etalab	Pour la prochaine initialisation
Correction de l'export IGN (BDUNI/SGA) : ajout	IGN	Fin septembre

des sources, des détruits		
Amélioration des appariements LP/IGN dans les données IGN (analyse + correction si possible)	IGN avec appui LP	Mi-octobre
Correction de la chaine de diffusion Ban v0 (guichet adresse) → DOMs	IGN	Fin septembre
Envoi d'une documentation sur la méthode d'import	Etalab	Mi septembre (avant la prochaine initialisation pour en discuter)
Envoi du mini-dump	Etalab	Mercredi 21 septembre
Recette des mini-dumps	IGN, La Poste	fin septembre
Envoi d'une nouvelle initialisation de la base france entière	Etalab	Fin septembre
Recette de la base france entière	IGN, La Poste, Utilisateurs	Fin octobre ?
Création d'un monitoring	Etalab, IGN, La Poste	Avant l'ouverture de la BAN
CGU	Etalab, La Poste, IGN	Pour le prochain COPIL
Partenariats	Etalab	Pour le prochain COPIL
Tests de rapprochement	Etalab, La Poste, IGN	Fin septembre
Envoi d'objets détruits	IGN	Fin septembre
Extension javascript pour bbox et geojson	IGN	Fin septembre
Evolution vers flask	Etalab	19/09