

Automatisation des flux de données au sein du SINP : moissonnage des données GeoNature

Dans le cadre du dispositif SINP, les structures régionales ou nationales doivent intégrer un nombre important de données de manière très régulière. Dans la majorité des cas, ces imports sont effectués manuellement. La plateforme régionale ou la structure productrice de donnée propose un export de ces données à un format défini au standard SINP et la plateforme nationale ou les plateformes régionales intègrent ces données directement dans leur outil.

Aujourd'hui plus de 70 structures agissant pour la biodiversité sont équipées de l'outil open source GeoNature (dont certains niveaux régionaux et le niveau national du SINP). Cet outil permet la collecte, la gestion, la diffusion et la valorisation des données de biodiversité. Il semble alors intéressant d'automatiser l'insertion de données venues de cet outil.

Dans le cadre de l'AAP SINP, le Parc National des Écrins propose de soumettre un projet en trois étapes :

1/ Phase 1 : Amélioration des performances et des fonctionnalités du module d'export de GeoNature

Le module d'export de GeoNature (https://github.com/PnX-SI/gn_module_export), est conçu pour satisfaire les problématiques métiers d'export et d'échange de données. Les SINP, gérant souvent plusieurs millions de données, doivent être équipés d'outils performants qui répondent aux besoins d'échanges de données et de métadonnées.

2/ Phase 2 : Amélioration de l'outil GN2PG

L'outil GN2PG : <https://github.com/lpoaura/GN2PG> récupère au sein d'une base de données PostgreSQL cible, les données mises à disposition via un export d'un GeoNature source. Fonctionnel, le téléchargement des données s'effectue aisément. Plusieurs problématiques demandent encore à être traitées:

- **Les performances de l'outil** sont encore une fois un enjeu de taille pour les SINP qui s'échangent plusieurs milliers/millions de données. L'outil doit être amélioré pour absorber des transferts importants.

- **Automatiser les actions de GN2PG.**

- **Les métadonnées** doivent être le cœur des échanges de données entre les acteurs. Aujourd'hui, l'outil récupère les données et les métadonnées telles que formatées dans le GeoNature source. Une mise en cohérence avec des projets portés par des acteurs régionaux doit être envisagée.

3/ Phase 3 : Intégration de GN2PG en tant que module de GeoNature et développement d'une interface associée

Actuellement l'outil GN2PG est une application Python téléchargeable via l'installer PIP de Python. L'objectif est de le rendre générique et accessible aux gestionnaires des bases de données des acteurs de biodiversité. Ainsi l'objectif est de proposer :

- **Une installation packagée simple**

- **Une interface de l'outil permettant la gestion et le suivi des flux de données**


- Système de configuration des connexions entre outils

- Système de suivi des téléchargements/imports :


- Système d'alerte

Cette interface pourra fonctionnellement ressembler à celle-ci :


Tableaux de bords de vos échanges de données




1397985
OBSERVATIONS
RÉCUPÉRÉES



9564
TAXONS



9
ORGANISMES
CONNECTÉS



147
JEUX DE DONNÉES

Effectuer un import manuel**

Connecter une nouvelle source de données

Nom de la source/Projet ↓	Organisme* ↓	Fréquence des imports	Dernier import ↑	Nb de données	Alerte
Asso 1 – Limite PACA	Asso 1	Mensuelle – Jour 01	01/06/2021	120 455	
Asso 1 – Limite AuRA	Asso 1	Mensuelle – Jour 01	01/06/2021	20 978	
Asso 2	Asso 2	Manuelle	01/03/2021	830	
Asso 3	Asso 3	Hebdomadaire – Lundi	07/06/2021	6976	

[Accéder à la liste complète des imports ?***](#)

* L'organisme sera associé à son identifiant du référentiel national des organismes. A terme l'interface pourra être liée à l'API mise à disposition par le SINP.

** À tout moment, il pourra être possible de lancer un moissonnage manuel des données d'une source déjà connectée. Par exemple, le flux de l'association 3, planifié de manière hebdomadaire pourra être lancé à n'importe quel moment par un utilisateur.

*** Une historisation des imports permettra leur traçabilité.

Voici un schéma simplifié des grandes phases de notre proposition :

