Procédure après intégration de données

Listes des étapes à suivre après une intégration de données (Gn2Pg ou via scripts).

Vérifier la cohérence des données

- Vérifier le bon fonctionnement de l'import : nombre de données dans la synthese, modification du nombre de données en ligne, ...

Ajouter les zones géographiques manquantes

- Pour PACA :
  - Vérifier que les nouvelles données sont bien dans le territoire et toutes associées aux mailles utilisées pour floutage, sinon ajouter les zones géo (COM, DEP, M1, M5, M10) manquantes.
  - Placez vous dans le dossier des scripts : '' cd ~/data/area-outside/data/sql/update/ ''
  - Lister les observations sans zone géo liées : '' psql -h localhost -U geonatadmin -d geonature2db -f ./01_* ''
  - Charge les zones géo concernées par le SINP :
    - les communes et départements : `./02_load_admin_areas.sh geonature2db '' * Si nécessaire, installer le paquet git-lgs : ''sudo apt install git-lfs`
      - Installer LFS sur le dépôt : `git lfs install`
      - Télécharger les fichiers gérés avec LFS du dépôt : `git lfs pull`
    - les mailles M1, M5 et M10 : `./02_load_meshes.sh geonature2db '' * Insérer les zones géo manquantes : * les zones administratives : '' psql -h localhost -U geonatadmin -d geonature2db -f ./03_insert_admin_areas* '' * les mailles : '' psql -h localhost -U geonatadmin -d geonature2db -f ./03_insert_meshes* '' * Établir les liens manquant entre synthese et ref_geo pour * les zones administratives : ''psql -h localhost -U geonatadmin -d geonature2db -f ./04_after_insert_admin_areas*`
    - les mailles : `psql -h localhost -U geonatadmin -d geonature2db -f ./04_after_insert_meshes*`
  - Lister les observations sans zone géo liées après rétablissement des zones géo manquantes : `psql -h localhost -U geonatadmin -d geonature2db -f ./05_*`
    - Comparer les tables `tmp_outside_*` et `tmp_outside_after_*`, il devrait y avoir moins d'observations dans les tables `*_after_*`.
  - Mettre à NULL les géométries des observations non liées aux mailles M5 pour éviter leur apparition dans la synthese : '' psql -h localhost -U geonatadmin -d geonature2db -f ./06_*''
- Pour AURA :
  - Lancer les scripts SQL du dossier `~/data/area-outside/data/sql` :
    - '' 01_create_outside_all.sql '' : pour créer la table contenant toutes les observations hors territoire
    - '' 03_fix_outside_geom.sql '' : pour mettre à NULL toutes les géométries des

observations hors territoire.

Mise à jour images INPN dans Taxhub

- Mettre à jour des images INPN si de nouveaux taxons sont apparus...
- Voir la procédure de mise à jour des images de l'INPN.
- Sur le serveur `db-srv` :
  - Mettre à jour la table `taxonomie.bib_noms` avec les éventuels nouveaux noms présents dans la synthèse : '' psql -h localhost -U geonatadmin -d geonature2db -c "INSERT INTO taxonomie.bib_noms (cd_nom, cd_ref) SELECT DISTINCT s.cd_nom, t.cd_ref FROM gn_synthese.synthese AS s JOIN taxonomie.taxref AS t ON s.cd_nom = t.cd_nom WHERE NOT s.cd_nom IN (SELECT DISTINCT cd_nom FROM taxonomie.bib_noms); " ''
- Sur le serveur `web-srv` :
  - Se placer dans le dossier du script de Taxhub : `cd ~/www/taxhub/data/scripts/import_inpn_media`
  - Activer l'environnement virtuel : `source venv/bin/activate`
  - Lancer le script : `python import_inpn_media.py`
  - Vérifier la présence des nouveaux médias dans la table `taxonomie.t_medias`.
  - Si tout c'est bien passé, désactiver l'environnement virtuel : `deactivate`
  - Le script ajoute les photos en tant que "secondaire" pour désigner celle avec le plus petit id_media (=~ au hasard) comme "principale" pour les nouveaux taxons, utiliser la requête suivante :
    - Pour PACA :
      - Soit via le fichier `set_first_img.sql` (recommandé) sur serveur `db-srv` dans `~/data/maintenance/data/sql/` avec : `psql -h localhost -U geonatadmin -d geonature2db -f ./set_first_img.sql`
      - Soit directement : `psql -h localhost -U geonatadmin -d geonature2db -c "WITH first_media AS (SELECT MIN(id_media) AS first_id_media_founded, cd_ref FROM taxonomie.t_medias GROUP BY cd_ref) UPDATE taxonomie.t_medias AS tm SET id_type = 1 FROM first_media AS fm WHERE tm.id_media = fm.first_id_media_founded AND tm.cd_ref = fm.cd_ref ;"`
    - Pour AURA :
      - Utiliser la requête qui tient compte des images du PIFH présente sur la page suivante.
- **NOTES** :
  - Les vues matérialisées de l'Atlas nécessitent d'être mise à jour pour prendre en compte les nouveaux médias.

Maintenance de la base de données

- **ATTENTION** : dans les versions récentes de Postgresql (v14+) cette manipulation n'est plus vraiment nécessaire. Le mécanisme automatique de nettoyage de la base semble efficace. Une "blocage" de la base de données a été constaté lors de son utilisation automatisé dans le cadre du SINP AURA.
- Sur le serveur `db-srv` :
  - Vérifier la place restante :

```
df -h |grep /dev/sda1 ; du -hs /var/lib/postgresql/
```

- Lancer le script de maintenance de la synthese :

```
psql -h localhost -U geonatadmin -d geonature2db -f
~/data/shared/data/sql/synthese_maintenance.sql
```

Mettre à jour les profils

- Sur l'instance `web-srv` :

```
ssh geonat@web-<region>-sinp
```

  - Lancer la mise à jour des vues matérialisées des profils :
    - activer une session Screen :

```
screen -S gn-update-profiles
```

    - se placer dans GeoNature et activer le venv :

```
cd ~/geonature ; source backend/venv/bin/activate
```

    - lancer la mise à jour des vues matérialisées :

```
geonature profiles update
```

    - désactiver le venv :

```
deactivate
```

    - Se détacher de la session Screen : `Ctrl-a + Ctrl-d`

Mettre à jour l'Atlas

- Sur le serveur `web-srv` :
  - Passer l'Atlas en maintenance : '' mv ~/www/maintenance/atlas/maintenance.disable ~/www/maintenance/atlas/maintenance.enable ''
- Sur le serveur `db-srv` :
  - Lancer la mise à jour de l'Atlas :
    - activer une session Screen :

```
screen -S atlas-update
```

    - se placer dans le dossier de maintenance :

```
cd ~/data/maintenance/data/sql/
```

    - lancer la mise à jour des vues matérialisées de l'Atlas :

```
 psql -h localhost -U geonatadmin -d gnatlas -f
```

```
./atlas_refresh.sql
```

    - Se détacher de la session Screen : `Ctrl-a + Ctrl-d`
- Sur le serveur `web-srv` :
    - Supprimer la maintenance de l'Atlas : '' mv ~/www/maintenance/atlas/maintenance.enable ~/www/maintenance/atlas/maintenance.disable ''

Extraire les indicateurs liés à l'import

- Pour PACA :
    - En local, se placer dans le dossier `indicators/` du dépôt `sinp-<region>-data` avec :

```
cd /<path>/sinp-paca-data/extracts/indicators
```

    - Lancer la commande :

```
cat ./observations_count_by_imports.sql | ssh geonat@db-paca-sinp
'export PGPASSWORD="<db-user-pwd>" ; psql -q -h localhost -p 5432
-U gnreader -d geonature2db' > ./$(date +'%F')_obs_by_imports.csv
```

From:
https://wiki-sinp.cbn-alpin.fr/ - **CBNA SINP**

Permanent link:
https://wiki-sinp.cbn-alpin.fr/database/procedure-apres-integration-data?rev=1734375894

Last update: **2024/12/16 19:04**