

260528_Essai sur \\NASMAISON\foxy\Géographie\

Workflow geo_loud — Optimisation des fichiers lourds
Contexte

Le dossier \\NASMAISON\foxy\Géographie\ contient 1 312 fichiers (PDF et EPUB) pour un total de 32,9 Go. Parmi eux, 127 fichiers dépassent 75 Mo et représentent 13,6 Go — soit 41 % du volume total concentré sur moins de 10 % des fichiers.

L'objectif est de sortir ces fichiers lourds pour les optimiser (compression PDF), puis de les remettre exactement à leur emplacement d'origine sans perdre l'arborescence.

Architecture du système

Fichier	Rôle
---------	------

nas_indexer.py	Scanne le NAS et crée la base SQLite
----------------	--------------------------------------

nas_geographie.db	Référentiel de tous les fichiers avec métadonnées et statut
-------------------	---

nas_geo_loud.py	Script checkout/checkin — déplace et restaure les fichiers lourds
-----------------	---

La base SQLite est la source de vérité : elle connaît à tout moment l'emplacement original de chaque fichier et son statut.

Colonnes de traçabilité (table fichiers)

Colonne	Description
---------	-------------

chemin_absolu	Chemin d'origine complet sur le NAS
---------------	-------------------------------------

chemin_geo_loud	Chemin dans le dossier de travail après déplacement
-----------------	---

statut	original / deplace / restaure
--------	-------------------------------

date_deplacement	Horodatage du déplacement
------------------	---------------------------

date_restoration	Horodatage de la restauration
------------------	-------------------------------

Workflow complet

Étape 1 — Indexation initiale

1

```
python nas_indexer.py "\\NASMAISON\foxy\Géographie" --output nas_geographie.db
```

Résultat : 1 312 fichiers indexés en ~22 secondes, 0 erreur.

Étape 2 — Export (checkout)

Déplace les 127 fichiers > 75 Mo vers \\NASMAISON\foxy\geo_loud\ :

1

2
3
4

```
python nas_geo_loud.py export \  
--db nas_geographie.db \  
--dest \\NASMAISON\foxy\geo_loud \  
--confirmer
```

Ce que fait le script :

- Interroge la base pour tous les fichiers statut = 'original' et taille > 75 Mo
- Renomme chaque fichier {id:06d}_{nom_original} pour éviter les collisions de noms
- Déplace physiquement le fichier vers geo_loud/
- Met à jour la base : statut = 'deplace', chemin_geo_loud, date_deplacement

Le préfixe numérique garantit l'unicité même si deux sous-dossiers contiennent un fichier de même nom.

Étape 3 — Optimisation PDF

Traiter les fichiers dans \\NASMAISON\foxy\geo_loud\ avec l'outil d'optimisation. Les fichiers peuvent être traités dans n'importe quel ordre et en plusieurs sessions.

Étape 4 — Restauration (checkin)

1
2
3

```
python nas_geo_loud.py restaurer \  
--db nas_geographie.db \  
--dest \\NASMAISON\foxy\geo_loud
```

Ce que fait le script :

- Lit tous les enregistrements statut = 'deplace' dans la base
- Pour chaque fichier présent dans geo_loud/ : déplace vers son chemin_absolu d'origine
- Recrée les sous-dossiers si nécessaire
- Met à jour : statut = 'restaure', date_restoration
- Ignore les fichiers absents de geo_loud/ (pas encore optimisés)

Étape 5 — Vérification du statut

1

```
python nas_geo_loud.py statut --db nas_geographie.db
```

Affiche :

1
2
3
4

Statut	Fichiers	Total Go
--------	----------	----------

deplace	127	13.577
original	1 185	19.275

Vues SQL utiles (Metabase / DB Browser)

1
2
3
4
5
6
7
8

-- Fichiers encore dans geo_loud

```
SELECT nom, ROUND(taille_octets/1048576.0,1) AS mo, date_deplacement, chemin_absolu
FROM fichiers WHERE statut = 'deplace'
ORDER BY taille_octets DESC;
```

-- Bilan par statut

```
SELECT statut, COUNT(*) AS nb, ROUND(SUM(taille_octets)/1073741824.0,3) AS go
FROM fichiers GROUP BY statut;
```

Paramètres disponibles

Paramètre	Défaut	Description
-----------	--------	-------------

--db	nas_geographie.db	Chemin vers la base SQLite
--dest	\\NASMAISON\foxy\geo_loud	Dossier de travail
--seuil	75	Seuil en Mo pour la sélection
--confirmer	(off)	Bypass la confirmation interactive

Ré-indexation après optimisation

1

```
python nas_indexer.py "\\NASMAISON\foxy\Géographie" --output nas_geographie.db --vider
```

Fichiers du projet

Fichier	Emplacement
---------	-------------

nas_indexer.py	C:\Users\eliob\Documents\nas_indexer.py
nas_geo_loud.py	C:\Users\eliob\Documents\nas_geo_loud.py
nas_rapport.py	C:\Users\eliob\Documents\nas_rapport.py
nas_geographie.db	C:\Users\eliob\Documents\nas_geographie.db

Revision #1

Created 2026-05-30 15:53:57 UTC by Julien

Updated 2026-05-30 15:55:05 UTC by Julien