

SAÉ 1.05 — Traitement des données (Gros fichiers)

IUT de Colmar – Département Réseaux & Télécommunications

Encadrant : O. ECKLE

Date : Janvier 2025

Objectif de la SAÉ

L'objectif de cette SAÉ est de développer un **outil de reporting** qui permet d'**analyser un disque** afin de localiser les **100 fichiers les plus volumineux** dans une arborescence donnée. Ces fichiers peuvent ensuite être **visualisés graphiquement**, avec la possibilité de générer un **script de suppression PowerShell**.

Ce travail répond à un **besoin réel des administrateurs systèmes**, notamment lors de situations critiques où le disque est saturé.

Fonctionnalités attendues

1. Sélection d'un répertoire de base

- Un script Python permet de sélectionner un répertoire via une **interface PyQt5** (`QFileDialog.getExistingDirectory()`).
- Le chemin est renvoyé au script PowerShell.

2. Analyse récursive des fichiers

- Un second script Python analyse l'arborescence complète et retourne une **liste des fichiers** avec leur taille.

Exemple de résultat :

json

CopierModifier

```
[  
    ["C:\\INSTALL\\...\\JUKEBOX0.DDB", 43134976],  
    ...  
]
```

○

3. Filtrage et tri

- Les fichiers sont triés par taille décroissante.
- On ne conserve que les **100 plus gros fichiers** dépassant une **taille seuil** configurable (ex. 10 Mo).

4. Export JSON

- Les données sont stockées dans un fichier JSON.
- Les chemins Windows sont correctement échappés (\\).

5. Affichage graphique

- Un troisième script Python affiche les résultats sous forme de :
 - **Camembert** (QChart / QPieSeries)
 - **Légendes** avec cases à cocher (25 par page)
 - **Onglets** PyQt5 (QTabWidget)
 - **Bouton** générant un script PowerShell de suppression

6. Génération de script PowerShell

- Si des cases à cocher sont sélectionnées, un bouton permet de **générer un script PowerShell** qui supprime les fichiers sélectionnés.

Technologies utilisées

- **Python 3**
 - `pathlib` pour gérer les chemins multiplateformes
 - `json` pour exporter les résultats
 - `PyQt5`, `PyQtChart` pour l'IHM
 - **PowerShell** pour l'exécution et la suppression
-

Compétences mobilisées

Compétence	Description
Recherche de fichiers	Utilisation de parcours récursifs (<code>rglob</code>)
Programmation Python	Scripts modulaires et orientés objets
Traitement de données	JSON, tri, filtrage, slicing
Interface utilisateur	Conception graphique avec PyQt5
Sécurité système	Mise en garde sur les suppressions de fichiers sensibles
Scripting PowerShell	Automatisation et interaction avec Python

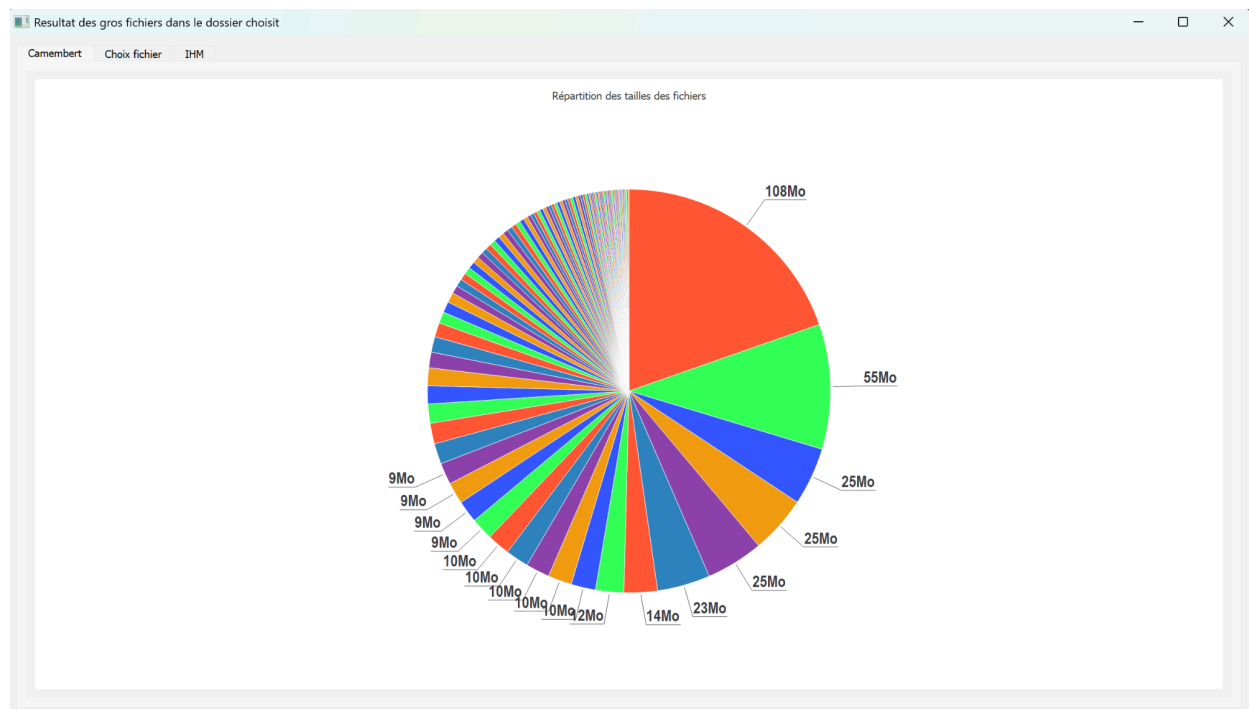
Travail réalisé

- Développement de 3 scripts Python modulaires.
- Création d'un script PowerShell d'intégration.
- Réalisation d'une interface graphique complète.
- Test de la compatibilité multiplateforme (Windows, Linux, MacOS).
- Mise en œuvre de contrôles pour éviter la suppression de fichiers critiques.

Fichiers développés

- `Analyse_Gros_Fichiers.py`
- `Affichage_Camembert.py`
- `Selection_Repertoire.py`
- `Creation_Onglets.py`
- `Creation_Camembert.py`
- `Creation_Legendes.py`
- `Creation_Boutons.py`
- `Analyse_Gros_Fichiers.ps1`

Preuve / Travail :



Resultat des gros fichiers dans le dossier choisi

Camembert Choix fichier IHM

<input checked="" type="checkbox"/>	V2 saÃ©.mp4 (113774829 octets)
<input type="checkbox"/>	videofond.mp4 (57896024 octets)
<input checked="" type="checkbox"/>	PrÃ©sentation Oral.pptx (26749301 octets)
<input checked="" type="checkbox"/>	dd.zip (26690135 octets)
<input type="checkbox"/>	eee.zip (26629924 octets)
<input type="checkbox"/>	R1.03-RLEA.AvecAnimations.pdf (24561087 octets)
<input type="checkbox"/>	20241209_095649521_iOS (1).png (15473233 octets)
<input type="checkbox"/>	20241209_100226853_iOS.png (13063793 octets)
<input type="checkbox"/>	20241209_095850723_iOS (1).png (11370436 octets)
<input type="checkbox"/>	20241209_102940163_iOS (1).png (11049977 octets)
<input type="checkbox"/>	20241209_102940163_iOS.png (11038156 octets)
<input type="checkbox"/>	20241209_075324411_iOS.png (10733718 octets)
<input type="checkbox"/>	20241209_095939152_iOS.png (10591643 octets)
<input type="checkbox"/>	20241209_075323588_iOS.png (10366414 octets)
<input type="checkbox"/>	20241209_095949525_iOS.png (10206255 octets)
<input type="checkbox"/>	20241209_075538124_iOS.png (10013028 octets)
<input type="checkbox"/>	20241209_095901220_iOS.png (9985453 octets)
<input type="checkbox"/>	20241209_102946038_iOS (1).png (9560131 octets)
<input type="checkbox"/>	20241209_102946038_iOS.png (9552191 octets)
<input type="checkbox"/>	20241209_095352696_iOS.png (8904910 octets)
<input type="checkbox"/>	20241209_075517677_iOS.png (8313807 octets)
<input type="checkbox"/>	20241209_075640218_iOS.png (8259565 octets)
<input type="checkbox"/>	Rapport.docx (7224578 octets)
<input type="checkbox"/>	Rapport SAE 13.docx (7223189 octets)
<input type="checkbox"/>	CM 109 Etudiant.pdf (6586340 octets)
<input type="checkbox"/>	PrÃ©sentation (1).pptx (5283410 octets)
<input type="checkbox"/>	1.pcap (5056267 octets)
<input type="checkbox"/>	R104_TP3_2024.docx (4644086 octets)
<input type="checkbox"/>	toon-showing-blank-placard.jpg (3765924 octets)
<input type="checkbox"/>	Sujet SAE 105.pdf (3546307 octets)
<input type="checkbox"/>	2.pcap (3275240 octets)