

Installation de l'environnement

1. Aller dans le répertoire "deliver_dgs_pdf"
2. Lancer la commande `conda env create -f env_starclay_plumber.yml`
3. Lancer la commande `conda activate starclay_plumber`

Exécution des tests

Cette exécution permet de vérifier que l'installation fonctionne bien sur le jeu de test.

1. Aller dans le répertoire "deliver_dgs_pdf"
2. Lancer la commande `python -m unittest discover -v`

Si le test renvoie "OK", c'est que tout fonctionne.

Exécution complète

1. Aller dans le répertoire "deliver_dgs_pdf"
2. Placer les fichiers PDF à traiter dans le dossier "working_dir"
3. Lancer la commande `python main_plumber.py`

Les fichiers en sortie seront dans deux répertoires:

- `processed_dir` : contient une extraction de chaque page de tous les fichiers PDF.
- `concat_dir` : contient la table complète extraite pour chaque fichier PDF.

Remarque : Si le traitement est interrompu en cours, il suffit de relancer la commande 3.