

# 1 La récupération des données

Parler de la récupération des données depuis les deux APIs discutés ainsi que le fichier YAML de configuration.

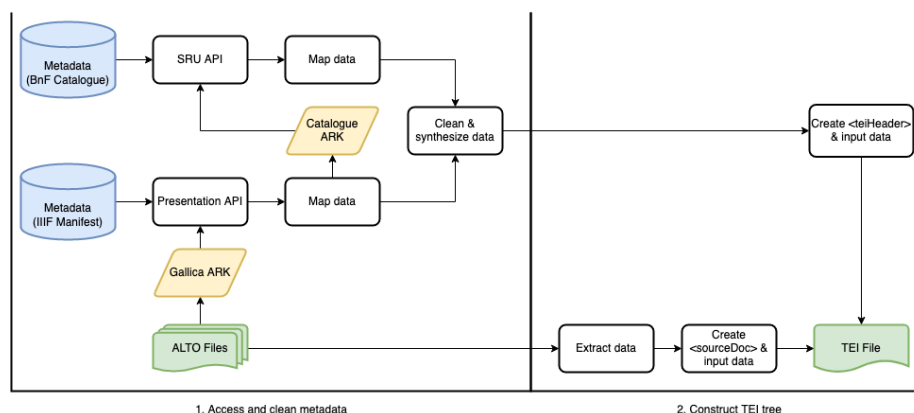


Figure 1: Workflow

## 1.1 Du manifest IIF à un dictionnaire Python

Montrer le mapping des données du manifest au dictionnaire Python.

## 1.2 De l'Unimarc à un dictionnaire Python

Montrer le mapping des données du catalogue au dictionnaire Python.

# 2 L'analyse des données

Parler de la stratégie d'atténuation des risques en sélectionnant les données fiables. Les métadonnées du catalogue général de la BnF sont utilisées uniquement si le même exemplaire physique du document numérisé sur Gallica a bien été trouvé. Sinon, on risque de mettre les données d'un autre exemplaire de l'oeuvre que celui qui a été transcrit et donc introduire des fausses données, tel que le cote ou même l'éditeur et la date de publication de l'exemplaire.

# 3 Le modèle du <teiHeader>

Montrer le mapping des données au <teiHeader>.