

Automatisation de l'extraction
de données de librairies en
ligne.

- Problématiques du suivi des prix des livres d'occasion proposé par les librairies en ligne :

- Nombreuse référence chez différents revendeurs.
- Les produits de différents revendeurs doivent pouvoir être comparés.
- Les données recueillies doivent être sauvegardées dans un format accessible.
- Le code doit pouvoir s'intégrer dans une pipeline ETL.
- Chaque site de revendeur à sa propre architecture.

- Problématiques du suivi des prix des livres d'occasion proposé par les librairies en ligne :
Solution proposée:
-

- Nombreuse référence chez différents revendeurs.

Ciblage des données extraites

- Les produits de différents revendeurs doivent pouvoir être comparés.

Transformation des données extraites en données standard

- Les données recueillies doivent être sauvegardées dans un format accessible.

Création d'objets standard

- Le code doit pouvoir s'intégrer dans une pipeline ETL.

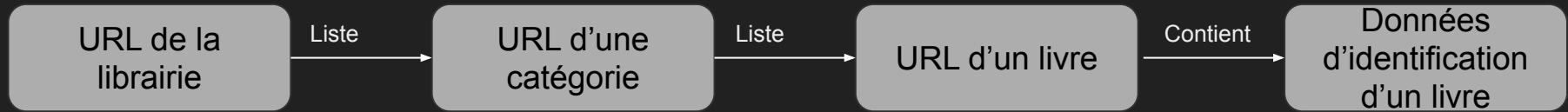
Extract, Transform, Load (extraire, transformer, charger)

- Chaque site de revendeur à sa propre architecture.

Extraire et traiter les données au cas par cas en fonctions du site cible

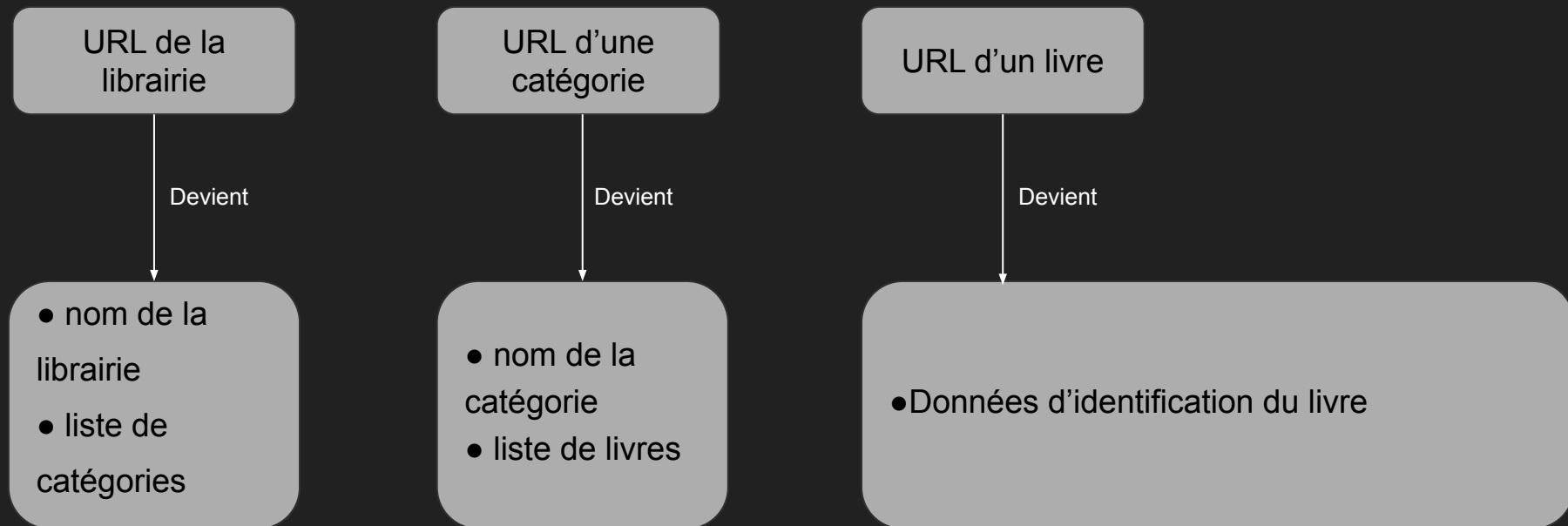
- Nombreuse référence chez différents revendeurs.

Ciblage des données extraites



- Les produits de différents revendeurs doivent pouvoir être comparés.

Transformation des données extraites en données standard



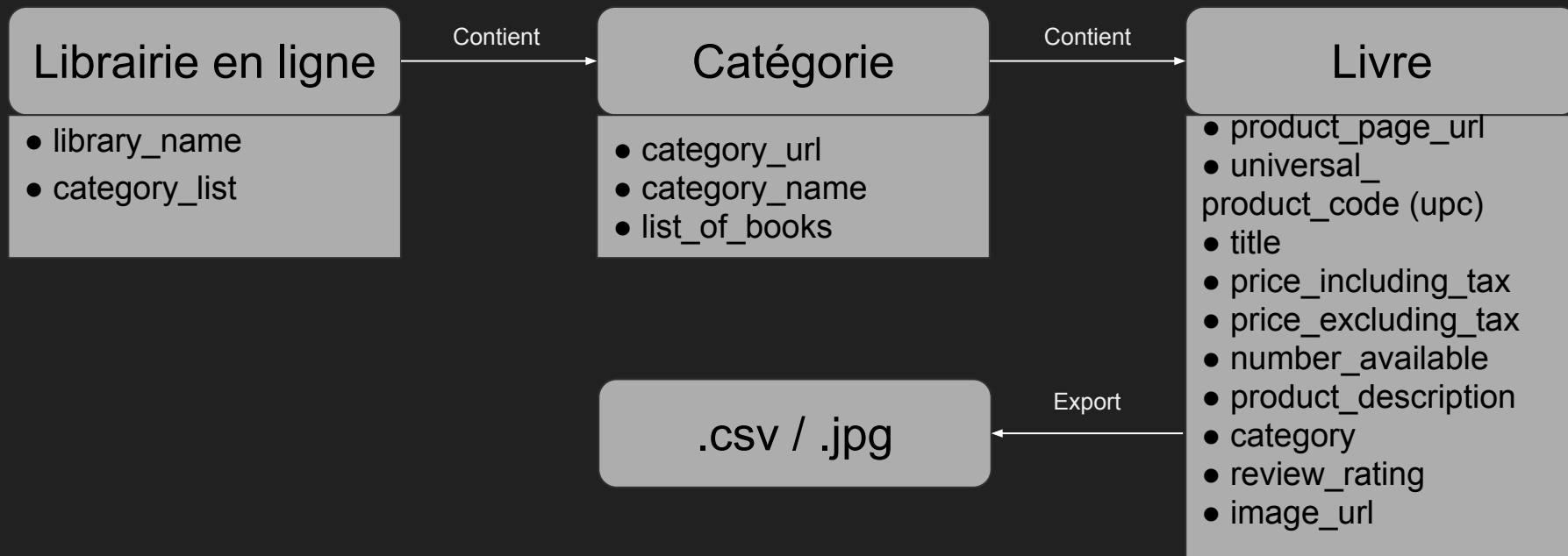
- Les données recueillies doivent être sauvegardées dans un format accessible.

Création d'objets standard



- Les données recueillies doivent être sauvegardées dans un format accessible.

Création d'objets standard



- Le code doit pouvoir s'intégrer dans une pipeline ETL.

Extract, Transform, Load (extraire, transformer, charger)

Extract

Ciblage des données
extraites

Transform

Transformation des données
extraite en données standard

Load

Création d'objets standard

Extract

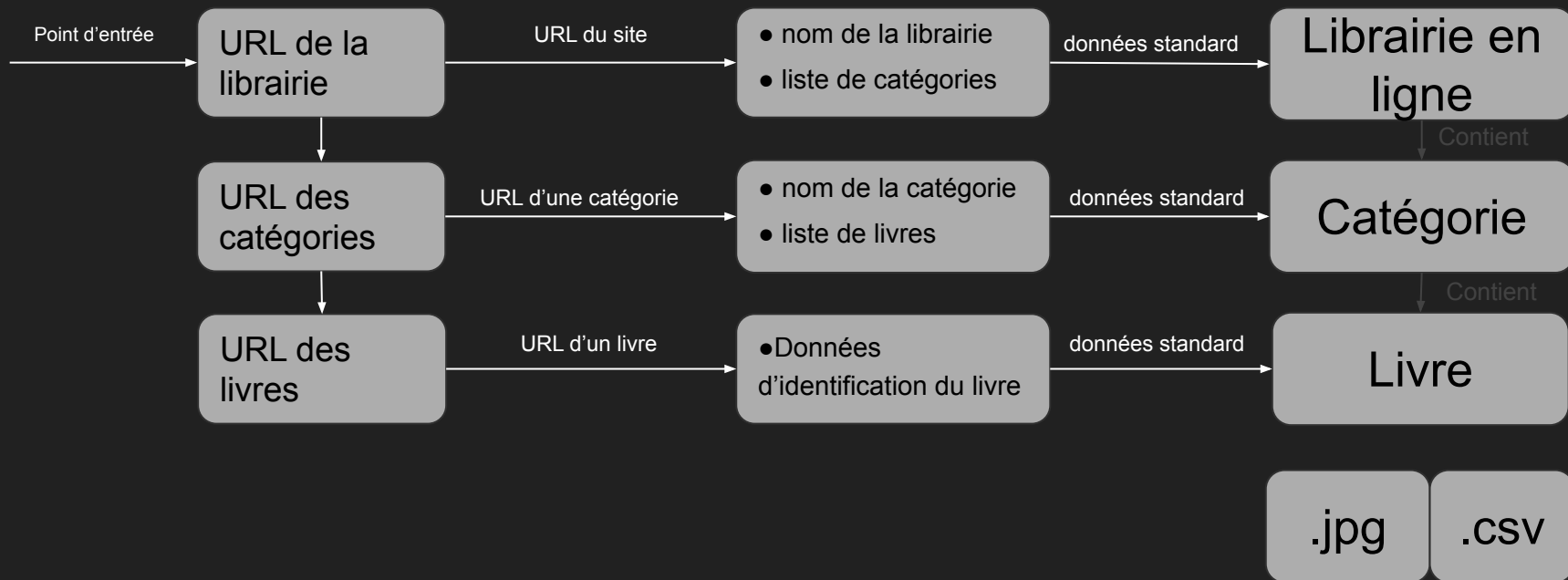
Ciblage des données
extraites

Transform

Transformation des données
extraite en données standard

Load

Création d'objets standard



- Chaque site de revendeur à sa propre architecture.

Extraire et traiter les données au cas par cas en fonctions du site cible

