



De l'intérêt des Humanités Numériques pour la recherche

Utiliser le numérique pour le traitement des sources



But du cours

- Apporter une **culture générale** numérique
- Présenter des **outils** pour aider à chacune des étapes du traitement des sources

**Apporter une culture générale
numérique**

Pour la recherche



Culture numérique dans la recherche

Les Humanités Numériques
sont un *buzzword* qui
permettent de débloquent des
fonds



Culture numérique dans la recherche

Les Humanités Numériques
sont un *buzzword* qui
permettent de débloquent des
fonds

Pouvoir communiquer avec des
ingénieurs et exprimer ses
attentes



Culture numérique dans la recherche

Les Humanités Numériques sont un *buzzword* qui permettent de débloquer des fonds

Pouvoir communiquer avec des ingénieurs et exprimer ses attentes

Intérêt pour la diffusion de la recherche dans une perspective d'*Open data*

Pour l'enseignement



Culture numérique pour l'enseignement

Pouvoir illustrer des concepts et
des phénomènes historiques



Culture numérique pour l'enseignement

Pouvoir illustrer des concepts et
des phénomènes historiques

Donner une perspective
analytique sur les sources

Et pour les autres ?



Culture numérique pour le reste...

Comprendre

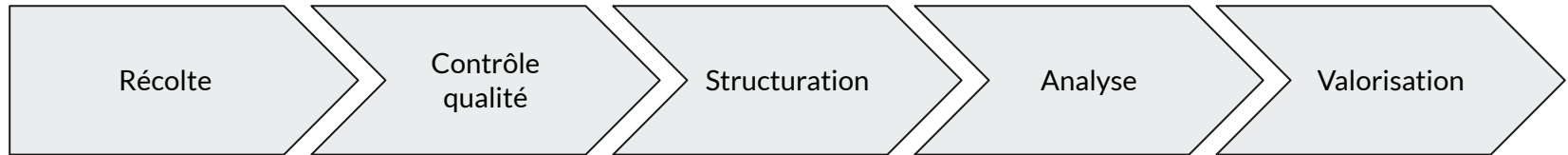
Utiliser

Impressionner

Présenter des outils



Intervenir à chaque étape du traitement des sources



Récolte



Récolte

Transcription automatique

Web scraping

Dépôt de données



Récolte

Transcription automatique

Transkribus

From the page

E-scriptorium

Kraken

Web scraping

Import.io

Scrapy

Webscraper

ParseHub

Dépôt de données

Gallica

Archive.org

Europeana

Salle des inventaires virtuelles

Contrôle qualité



Contrôle qualité

Nettoyage des données

Repérage des irrégularités

Vérification des sources



Contrôle qualité

Nettoyage des données

Dataïku

OpenRefine

Repérage des irrégularités (doublons, manques, etc.)

Heurist

Kibana

Vérification des sources

Zotero

Altmetric

Contrôle qualité



Structuration

Modélisation

Base de données

Uniformiser les données



Structuration

Modélisation

MySQL Workbench

LucidChart

Base de données

Heurist

SQLite

Uniformiser les données

TEI/EAD

Dublin Core

Analyse



Analyse

Statistiques

Visualisation

Textométrie



Analyse

Statistiques

R

Omniscope

Stata

Visualisation

Tableau

RawGraph

Palladio

Textométrie

TXM

Voyant Tools

Médite

Valorisation



Valorisation

Publication en ligne

Édition numérique

Ouverture de la recherche



Valorisation

Publication en ligne

Omeka

Heurist

Wordpress

Édition numérique

XML TEI

IIIF

Overleaf

Ouverture de la recherche

Isidore

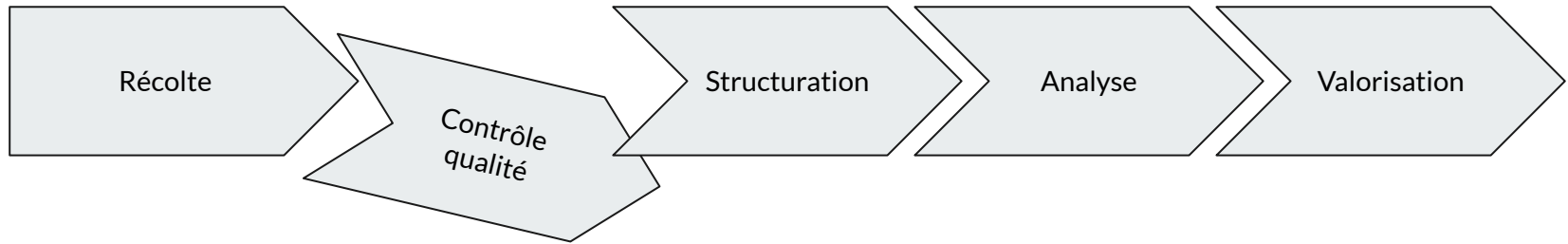
HAL

FAIR

Comment incorporer ça à son travail ?



Interdépendance de chacune des étapes





Organiser les sources pour sa recherche

Documents ayant des traits
communs

Corpus quantifiable et
qualifiables

Prendre du recul par rapport à
ses sources

Des exemples concrets ?



Élément de notation

Présentation d'un livrable

- Utilisation d'un outil numérique pour le traitement des sources
- Montrer qu'on a exploité les fonctionnalités de l'outil

Explication de la démarche

- Comment on a procédé : restitution des étapes pour la réalisation du livrable
- Difficultés rencontrées

Analyse de l'intérêt pour la recherche

- Explication de l'intérêt de la méthode
- Critique et limite de l'outil

Sources textuelles



Traiter des sources textuelles

Aider à la transcription

Analyse textuelle

Éditer un texte



Exemples de livrables

Aider à la transcription

Document pdf transcrit avec
Transkribus

Analyse textuelle

Ensemble de visualisations
produites avec Voyant Tools

Éditer un texte

Document textuel annoté avec
XML TEI

Sources audiovisuelles



Traiter des ressources audiovisuelles

Structuration du corpus

Visualisations des ressources

Analyse sur les formats



Exemples de livrables

Structuration du corpus

Création d'un modèle
conceptuel pour ordonner les
sources

Visualisations des ressources

Création de graphiques
représentant les proportions
des interventions

Analyse sur les formats

Compte rendu critique sur les
formats de conservation utilisés

Sources hétérogènes



Traiter des sources hétérogènes

Nettoyer les données sources

Conceptualiser l'ensemble des
ressources

Publier le corpus sur le web



Exemples de livrables

Nettoyer les données sources

Propriéter un tableur de données avec Dataïku

Conceptualiser l'ensemble des ressources

Former une base sur Heurist avec différentes tables pour décrire la diversité des sources

Publier le corpus sur le web

Créer une collection de ressources sur Omeka

La modélisation des données : Un point central

Ensemble de photographies de stèles funéraires

La mode féminine au XVIIIe

Corpus d'instruments d'observation astronomique du XVe
