



GenderedNews

L'Équipe



Mathilde Aguiar
Cheffe de projet



Oumaima Hajji
SCRUM Master



Rokiatou dite Rose
Sidibe
Développeuse

Contexte et sujet

Contexte

Inégalité du temps de parole,
représentation des femmes dans les
médias.

Sujet

GenderedNews : un site Web qui recense le
taux de masculinité présent dans les médias.

Objectifs

Améliorer la solution déjà existante en
remplaçant la solution Métabase et en
travaillant sur l'aspect traitement du langage
naturel

Travaux à réaliser

Remplacer la technologie Metabase

Alternative envisagée : Apache superset

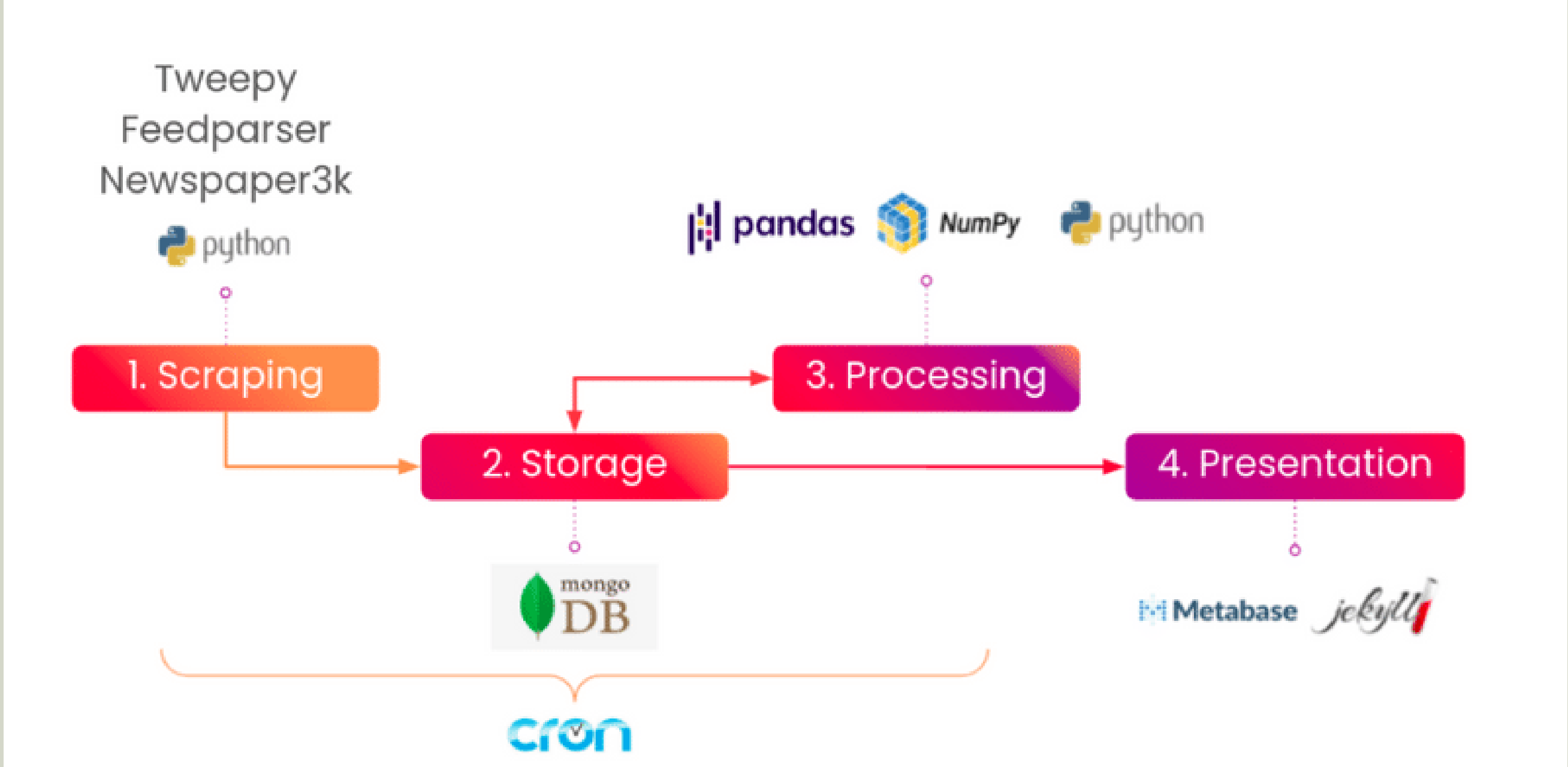
Améliorer la reconnaissance des noms NRE

Reconnaissance et différenciation suivant le contexte

Ajouter de nouvelles sources

Enrichir le nombre de sources étudiées

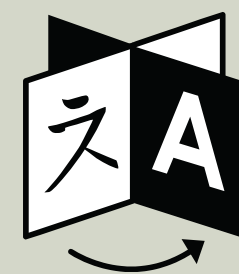
Architecture





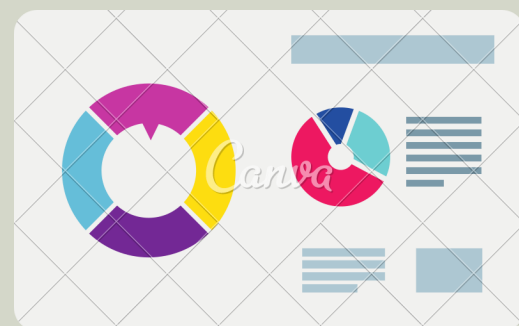
Base de données

Mongodb, SQLite



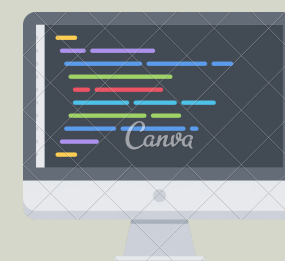
TALN

SpaCy (pipelines NER), docanno



Visualisation de données

Apache Superset, Metabase



Langage de programmation

Python, SQL

Gestion de projet

Kanban

Définition des tâches par la SCRUM Master et utilisation d'un trello kanban pour le suivi des tâches.

Réunions hebdomadaires

Réunions hebdomadaires avec les porteurs pour faire les point sur les tâches réalisées et futures.

Oumalma

- faire la migration vers SQLite/Postgre
- tuto superset
- tester dataset sur les données genderednews
- Comprendre la lib Spacy et le NER (cf mail)
- lire le papier nlp
- regarder comment on fait l'algo nlp

+ Ajouter une carte

Rose

- Résoudre le problème du front local
- tester les autres sources
- Tester l'algo des captures de prénoms déjà existant
- Calculs de perfs des scripts Doc Anno
- Tester le site "doc Anno"

+ Ajouter une carte

Mathilde

- faire la migration vers SQLite/Postgre
- Deploy la version actuelle en local du site
- Faire l'UML de la BD SQLite
- Algo avec Spacy

+ Ajouter une carte

Priorité

- Comprendre la lib Spacy et le NER (cf mail)
- Faire marcher la page web test
- Faire un algo pour déterminer si un nom est propre ou commun/ éliminer les confusions
- Calculs de perfs entre les différents algo
- Faire un modèle complet
- Faire des tests utilisateurs

+ Ajouter une carte

Done

- Trouver une alternative à Metabase - production d'une doc
- Comprendre les scripts Doc Anno
- Regarder le link entre MongoDB
- Ajouter une ressource de la liste des journaux
- Tester les scripts exemples (test ttes ensemble lundi)
- Créer un exemple avec Apache Superset
- Lire la doc sur git
- Lire la doc sur git
- Ajouter une ressource de la liste des journaux

+ Ajouter une carte

Tâches restantes

- 1 Implémentation de la base de données relationnelle
- 2 Implémentation de scripts NER
- 3 Faire des tests utilisateurs de notre future interface

Merci de votre
attention