**# 11/10/2024**

Mot clef et résumé d'un texte

Plusieurs textes .txt en anglais, qui traitent de l'éthique de l'IA.

|  |  |
| --- | --- |
| Collecter les pdf/html | OK |
| Parser | Problème avec HTML, supprime certains espaces |
| CLeaning | Reste quelques docs dégueu (CSS). Regex ne marche pas. |
| TF-IDF sur chaque phrase pour trouver les phrases les plus importantes en fonction des mots. | Introuvable ? Paul l’a fait normalement. |
| Trouver les mots clefs dans le corpus (mots les plus fréquents ? Mots donnés par le prof ? TF-IDF sur chaque document puis mots les plus fréquents ?) | Alex + Julian |
|  |  |
|  |  |
|  |  |