

Documentation : Analyseur de Fréquence des Mots + Nuage des mots

Introduction

L'application « Analyseur de Fréquence des Mots » permet d'analyser la fréquence des mots dans un texte ou un fichier, d'exclure certains mots, d'appliquer des règles de lemmatisation et de présenter les résultats sous forme de tableau et de nuage de mots interactif.

Structure du projet

Fichiers principaux

- **Code.php** : Contient les fonctions de traitement des données et l'analyse de fréquence des mots.
 - **Interface.php** : Fournit une interface utilisateur pour interagir avec les fonctionnalités de l'application.
 - **Interface.css** : Définit le style de la page web.
-

Détails des fichiers

Code.php

1. Fonction : `readFileContent($filePath)`

- **Description** : Lit et renvoie le contenu d'un fichier donné.
- **Entrée** :
 - `$filePath` : Chemin du fichier.
- **Sortie** : Contenu du fichier sous forme de texte.

2. Fonction : `readJsonFile($filePath)`

- **Description** : Lit un fichier JSON et le decode en tableau associatif.
- **Entrée** :
 - `$filePath` : Chemin du fichier JSON.
- **Sortie** : Tableau associatif contenant les données JSON ou `null` en cas d'erreur.

3. Fonction : `lemmatizeWord($word, $lemmatizationRules)`

- **Description** : Applique des règles de lemmatisation à un mot donné.
- **Entrée** :
 - `$word` : Mot à lemmatiser.
 - `$lemmatizationRules` : Tableau associatif contenant les règles de lemmatisation.
- **Sortie** : Lemma du mot ou le mot d'origine si aucune règle n'est trouvée.

4. Fonction : `analyzeFrequency($text, $exclusionPath, $lemmatizationPath)`

- **Description** : Analyse la fréquence des mots dans un texte en excluant certains mots et en appliquant des règles de lemmatisation.
- **Entrée** :
 - `$text` : Texte à analyser.
 - `$exclusionPath` : Chemin du fichier JSON contenant les mots d'exclusion.
 - `$lemmatizationPath` : Chemin du fichier JSON contenant les règles de lemmatisation.
- **Sortie** : Tableau des mots fréquents triés par ordre de fréquence ou un message d'erreur si un fichier est invalide.

5. Déroulement :

1. Charge les mots à exclure et les règles de lemmatisation depuis des fichiers JSON.
 2. Nettoie le texte en supprimant les caractères non pertinents.
 3. Applique la lemmatisation et exclut les mots non désirés.
 4. Calcule la fréquence des mots restants et retourne les résultats triés.
-

Interface.php

1. Interface Utilisateur

- Un champ de texte pour entrer ou modifier le contenu.
- Une option pour télécharger un fichier texte.
- Une option pour sélectionner le nombre de mots à afficher.

2. Traitement des Données

- Charge le texte à partir d'un champ ou d'un fichier.
- Utilise la fonction `analyzeFrequency` pour analyser le texte.
- Affiche les résultats sous forme de tableau et de nuage de mots.

3. Composants Affichés

- **Tableau des mots fréquents** : Liste des mots triés par fréquence.
- **Nuage de mots** : Représentation graphique des mots, leur taille variant en fonction de leur fréquence.

4. Scripts JavaScript

- Génération dynamique du nuage de mots avec `D3.js`.
 - Options de forme pour le nuage : cercle ou rectangle.
 - Toggle pour afficher ou masquer le tableau.
-

Fonctionnalités Clés

1. **Prise en charge des fichiers JSON :**
 - Exclusion des mots définis.
 - Application de règles de lemmatisation.
 2. **Nettoyage du texte :**
 - Suppression des caractères non pertinents.
 - Conversion en minuscules.
 3. **Analyse de Fréquence :**
 - Calcul rapide et précis des fréquences.
 - Résultats triés par ordre décroissant.
 4. **Visualisation :**
 - Tableau dynamique des mots fréquents.
 - Nuage de mots interactif.
-