

FICHE : LANCEMENT "KICKOFF" DU PROJET

IDENTIFIER VOTRE CHOIX DE SUJET

Titre du projet : Multi correcteur automatique

URL si applicable : <https://www.cordial.fr/enligne.php>

DÉCRIVEZ VOTRE PROJET (2 PARAGRAPHES MINIMUM)

Correcteur en python ayant pour but d'aider des personnes à corriger soit du texte, soit des calculs. Le correcteur se baserait sur un système de parsing en utilisant une librairie pour la correction de langage et la syntaxe des calculs (à déterminer à l'issue de la recherche successive). Une autre librairie pour donner la réponse au calcul entré (à déterminer à l'issue de la recherche successive). Une interface graphique sera intégrée (à déterminer à l'issue de la recherche successive).

Le projet aura pour but d'intégrer principalement une correction de la langue française. La deuxième itération verra une mise en place du correcteur et calculateur mathématique. La saisie de l'expression mathématique ressemblera à celle du site internet dCode (voir plus bas). L'interface graphique permettrait de choisir entre une correction de texte ou une correction d'expression (à voir si une détection automatique du type d'expression pourra être mise en place).

NOMMEZ 3 COMPÉTITEURS À VOTRE PROJET

qui vous inspireront pour les fonctionnalités (avec les liens)

Cordial	https://www.cordial.fr/enligne.php
dCode	https://www.dcode.fr/
MiniMath	http://www.minimath.net/index_fr.htm

LISTER LE MATÉRIEL DE CE PROJET

- Un ordinateur avec connexion et python
-
-

Est-ce que tous les éléments matériels nécessaires au projet sont en votre possession ? Sinon, lesquels ?

Aimeriez-vous que le département vous prête un composant, une pièce ? Si oui laquelle ?

IDENTIFIER DEUX ASPECTS TECHNOLOGIQUES

Aspects : Serveurs, Services, Algorithmes, Web, IoT, etc

Aspect 1 : Algorithmes

Aspect 2 : Services

PRÉCISEZ L'APPRENTISSAGE TECHNIQUE QUI SE FERA PENDANT LE PROJET

soyez aussi précis que possible (nom de la technologie, de la librairie, url)

Technologie : Lark, PyQt5, Pandas et OpenLexicon, Numpy

URL : <https://github.com/lark-parser/lark>,
<https://www.riverbankcomputing.com/static/Docs/PyQt5/index.html#>,
<https://github.com/chrplr/openlexicon>, <https://pandas.pydata.org/>,
<https://numpy.org/>

Quelle utilisation en ferez-vous à l'intérieur du projet ?

Lark me permettra de créer la grammaire pour le parseur en français. L'application permettra de trouver les fautes de grammaire, conjugaison et orthographe. J'utiliserai la base de mots OpenLexicon (accessible grâce à la librairie Pandas) pour obtenir une correction du mot possible, et Lark pour la correction de grammaire et conjugaison.

OpenLexicon sera utilisée pour accéder aux données de la base Lexique, via internet. Avec Pandas, Lexique sera lue sous son format .tsv et stockée dans une variable locale et accédée via internet (voir le README de la page d'accueil de OpenLexicon pour la méthode). Dans le cadre d'une possible future amélioration, j'aimerais créer une détection du wifi, pour permettre d'utiliser le fichier de base de donnée localement et faire de la correction même sans connexion. Cependant dans le cas présent, j'opte pour un accès en ligne pour limiter le poids de l'application.

Pour la correction d'expressions mathématiques, j'utiliserai la librairie numpy pour faire les calculs. Le principe de correction sera simplement une comparaison avec le résultat entré par l'utilisateur et celui trouvé avec python.

L'interface graphique sera faite avec PyQt5 car elle peut se faire rapidement. L'application comprendra deux onglets, un pour la correction du langage et l'autre pour la vérification mathématique.

ILLUSTRATION

Si une illustration peut nous aider à mieux comprendre le projet, merci de l'inclure ici.

Remplacer l'image

Vous pouvez inclure plus de matériel directement dans Github.

=> Après avoir rempli cette fiche : veuillez accepter l'assignment GitHub, convertir ce document en pdf et le remettre dans Github directement.