

WikiTrans

Sujet 4

L'objectif principal du projet est de développer un site web de scrapping de contenu sur Wikipedia avec sa conversion de l'anglais au Français et Vice-versa.

Wikipedia permet à tous les internautes de proposer des suggestions et des modifications du contenu des articles convertis dans d'autres langues, dans une logique participative. WikiTrans doit donc fournir à tous les internautes des fonctionnalités d'interaction directe avec la plateforme Wikipédia à travers des interfaces pour écrire et modifier des articles par la consommation de services Web fournis par Wikipedia.

Additionnellement, Wikipédia existe aujourd'hui en plus de 300 langues, dans une apparence unie, mais avec de grandes variations de contenus. Chose qui nécessite d'assurer une consistance linguistiques inter-langues en utilisant des technologies liées au Natural Language Processing.

Le système peut également proposer des suggestions du contenu déjà entré et validé par d'autres internautes par le biais du Machine Learning., mais aussi la comparaison sémantique du contenu pour la suggestion du contenu manquant dans d'autres langues.

Comptant plus de six millions et demi d'articles fin 2022, WikiTrans devrait assurer la gestion d'une grande quantité de contenu sur le cloud par le biais de la Big Data et le Business Intelligence.

