07/01/2025

Timothy MARCIA, Elias AIT HASSOU

Master 1 Informatique - Université Lumière Lyon II

Rapport de projet

Développement d’un moteur de recherche en Python

Table des matières

[Présentation du projet 2](#_Toc187152999)

[Partie conceptuelle 2](#_Toc187153000)

[Mise en place du projet 2](#_Toc187153001)

[Validation du logiciel 2](#_Toc187153002)

[Maintenance du logiciel 2](#_Toc187153003)

# Présentation du projet

Ce projet a pour objectif de développer un moteur de recherche textuel sur des données touristiques en utilisant le langage de programmation Python. Le moteur de recherche permet aux utilisateurs de rechercher des activités qu'ils souhaitent réaliser et retourne des descriptions de villes où ces activités sont disponibles. Ce logiciel repose sur la collecte, le traitement, le stockage, l’analyse et la restitution des données.

Pour constituer le corpus de données, nous avons exploité l'API WikiVoyage. Cette API fournit des descriptions des activités touristiques proposées dans différentes villes à travers le monde. Ces données textuelles représentent la base de notre moteur de recherche.

L’objectif principal est de proposer une interface permettant :

* À l’utilisateur de saisir une activité spécifique qu’il aimerait réaliser.
* D'analyser les descriptions des villes contenues dans le corpus
* De restituer les résultats sous forme d’une liste des villes pertinentes accompagnées de leurs descriptions.

# Partie conceptuelle

## L’environnement de travail

Pour le développement de notre moteur de recherche, nous avons choisi un environnement de travail adapté aux exigences de collecte, traitement et d’analyse des données textuelles.

Le projet a été implémenté en Python, étant le langage de programmation étudié lors des séances de cours. Sa simplicité, et ses bibliothèques dédiées au traitement des données textuelles ont été essentielles pour mener à bien ce projet.

Les différentes bibliothèques utilisées ont des rôles spécifiques décrit ci-après :

* **Requests** pour interagir avec l’API WikiVoyage
* **Pandas**
* **JSON** pour lire et écrire des fichiers JSON tels que les réponses d’API ou nos fichiers de stockage des données.
* **NLTK** pour récupérer une liste des termes à faible valeur ajoutée dans la langue anglaise nous permettant de nettoyer le corpus.
* **NumPy**
* **Tkinter** pour réaliser l’interface graphique finale permettant d’effectuer une rechercher et afficher les résultats.

## Les données

## Diagramme des classes

# Mise en place du projet

# Validation du logiciel

# Maintenance du logiciel

# Conclusion et perspectives