

# **Projet Python3: Application de webscrapping d'articles et d'image**

**Licence 3 Developpement d'application - UCAO-UUT**

**Date de présentation : (à décider en classe)**



## **Enoncé :**

Une entreprise spécialisée dans la data science cherche à mettre en place une application capable de collecter des actualités sur internet. Cependant, elle remarque que chaque site internet a sa propre façon de présenter l'information. Elle souhaite donc rendre son application adaptable à tous les sites web d'actualités en recherchant une solution robuste et facile à adapter, développée en Python 3, qui est le langage phare de l'entreprise.

Votre mission consiste à créer une application en Python 3 qui recherche des informations sur différents sites web de news à l'aide de la bibliothèque **Scrapy** et les enregistre dans une base de données en tant qu'articles. L'application devra gérer dynamiquement les sites web et permettre l'ajout de nouveaux sites facilement, même avec des architectures différentes. Cela signifie que si l'entreprise souhaite collecter des informations sur un site web récemment publié, vous devriez être en mesure d'ajouter ce site sans modifier le comportement interne de l'application initiale.

Vous devrez également mettre en place une **interface** (API REST ou programme CLI) pour accéder aux informations sur les articles et les sites web utilisés. Cette interface devra permettre la création, la lecture, la mise à jour et la suppression des sites web de l'application, ainsi que la sélection et la suppression des articles collectés.

## **Remarques:**

Vous devrez rédiger un cahier des charges simplifié du projet pour détailler correctement le projet.

Vous êtes libre d'ajouter des fonctionnalités supplémentaires à votre choix, comme l'ajout de catégories sur les articles pour les filtrer.