

# Projet Python

*Dans le cadre du MS, il vous est demandé de choisir un sujet parmi les 5 proposés ci-dessous. Le projet sera codé en Python. Le livrable devra comporter une interface graphique afin d'effectuer des choix le cas échéant. Le projet pourra se faire en groupes, mais avec un maximum de 5 étudiants par groupe. Le même sujet pourra être choisi par plusieurs groupes.*

## Sujet 1 : Extraction de données d'un site web

On souhaite écrire un programme auquel on communique une URL web, ainsi qu'une expression régulière, et le programme va chercher toutes les informations sur les pages du site qui matchent cette expression. Si aucune expression régulière n'est saisie, le programme proposera les plus courantes (@IP, numéro de téléphone, horaire, date ...). Il proposera aussi un export des résultats en CSV.

## Sujet 2 : Analyse de logs multicritères

On souhaite écrire un programme qui va ouvrir un fichier de logs (avec entête), et trier les données : vous proposerez un menu déroulant avec les différents éléments présents. Le programme permettra les recherches multicritères, ainsi que la sauvegarde des résultats, et l'export en CSV.

## Sujet 3 : Analyse statistique d'un fichier

On souhaite écrire un programme qui va ouvrir un fichier texte (rockyou.txt) contenant des mots de passe à raison d'un mot de passe par ligne, et produire des statistiques sur l'usage des caractères (majuscules, minuscules, numériques, symboles), la longueur minimale, maximale et moyenne, la fréquence d'usage des lettres dans ce fichier.

## Sujet 4 : Analyse des métadonnées

Dans ce projet, on souhaite définir un utilitaire auquel on communique un répertoire du disque contenant des fichiers, le programme devra extraire les métadonnées des fichiers présents dans le répertoire. Les fichiers pourront être des fichiers audio, vidéo, image, pdf, office. Les résultats seront sauvegardés.

## Sujet 5 : Génération de logs

Dans ce projet, on souhaite créer des fichiers de logs d'application courantes ou des logs systèmes (maximum 10 types de logs). Il faudra pour ce faire, proposer dans un menu déroulant, de choisir le type de logs, entrer les dates de début, dates de fin de log. Générer les données entre ces deux dates, et préciser des éléments de customisation. Par exemple, si on considère des logs APACHE, vous pourrez donner le choix à l'utilisateur de cocher les méthodes http qu'il souhaite voir apparaître, avec le pourcentage de présence de chacun. Vous proposerez des chemins aléatoires basés sur de vrais mots, vous pourrez par exemple consulter (<http://www.listedemots.com/>). Bien entendu, figureront des adresses IP aléatoires, et l'utilisateur pourra choisir l'OS et le navigateur figurant dans les logs ...

