

SAE 1.03

Guide d'utilisation de l'extracteur de données OCR

Prédispositions :

- Posséder une version fonctionnelle de Docker sur son ordinateur
- Déposer les images (au format .png) dont on veut extraire les données dans le répertoire ./scriptvolume/
- Ouvrir un terminal dans le dossier contenant le script et ses fichiers complémentaires

Configuration :

Avant d'exécuter le script il y a plusieurs tâches à effectuer :

- Modifier le fichier config.txt pour sélectionner les informations extraites voulues.

Par exemple avec la configuration :

ID=1
PRENOM=1
NOM=1
MAIL=1
TEL=1
ADR=1

On souhaite garder toutes les informations extraites du fichier .png sélectionné. Là où on aurait remplacé les 1 par des 0 dans les champs qu'on ne souhaite pas conserver.

- Attribuer les droits UNIX d'exécution au fichier script avec la commande :

```
chmod +x ./script
```

Exécution du script :

- Exécuter la commande suivante dans le terminal :

```
./script <nom_image> <motif_recherché>
```

<nom_image> : nom de l'image que vous avez préalablement déposé dans le répertoire ./scriptvolume/ dont vous souhaitez extraire des informations (« sample.png » par exemple).

<motif_recherché> : argument facultatif qui permet sa recherche dans les données extraites sans discerner les majuscules des minuscules. Attention cependant, le fichier obtenu à la fin de l'exécution du script contiendra UNIQUEMENT les lignes d'information qui présentent ce motif.

Ce motif étant utilisée dans un filtre UNIX nommé egrep, je vous renvoie vers [cette documentation](#) pour connaître la majorité des cas particuliers de motifs. (par exemple « marco | johan » représente ou marco ou johan)

- Attendre SANS INTÉRAGIR avec le terminal jusqu'à l'affichage du résultat dans un format paginé (appuyer sur « q » pour quitter l'affichage).
- Attendre SANS INTÉRAGIR avec le terminal jusqu'à la fin de l'exécution du script symbolisée par la ligne de texte :
« fin du script :^) »

Après l'exécution :

- Récupérer le fichier « final.txt » qui contient les données extraites en fonction des options sélectionnées

En cas d'erreur ou de blocage :

Dans le cas où vous rencontreriez une erreur ou un plantage du script lors de son exécution, appuyez simultanément sur « Ctrl » et « C » pour arrêter le script puis entrez respectivement les deux commandes suivantes dans le terminal :

```
docker container stop doctess
```

```
docker container rm doctess
```

Cela aura pour effet de supprimer le conteneur créé pendant l'exécution du script et vous pourrez relancer ce dernier (ces commandes peuvent prendre plusieurs secondes à s'exécuter)