

TP 08

TP portant sur l'utilisation basique des expressions régulières

Objectifs : comprendre les notions fondamentales relatives à l'utilisation des expressions régulières.

Exercice 1 :

Cet exercice n'est pas à rendre mais il est indispensable pour préparer le livrable.

Proposez un motif pour sélectionner toutes les lignes d'un fichier :

1. qui contiennent :
 - a. des chiffres ou des majuscules ;
 - b. des points ;
 - c. trois points (...)
 - d. des nombres hexadécimaux "isolés", c'est à dire qu'ils ne sont pas collés à d'autres caractères dans la ligne ;
 - e. un mot d'au moins 12 caractères alphanumériques ;
 - f. exactement 5 lettres a (pas nécessairement successives) ;
 - g. des crochets (] ou [) ;
 - h. que des lettres a ou des espaces. Il peut y avoir des a, des espaces ou un mélange des deux mais pas autre chose.
 - i. quelque chose qui ressemble à une adresse IP (valide ou pas)
2. qui sont :
 - a. vides ; (uniquement le caractère de saut de ligne)
 - b. blanches ; (au moins un séparateur en plus du caractère de saut de ligne)
 - c. non vides.
3. qui ne contiennent pas :
 - a. de lettre a ;
 - b. des espaces ;
 - c. des chiffres décimaux.
4. qui débutent par un numéro de téléphone au format 01 64 25 89 78
5. idem mais à la place des espaces, on peut avoir des , ou des tirets (-)
6. idem 5 mais le numéro de téléphone débute nécessairement par 0. Ce 0 peut être entouré de parenthèses.
7. qui terminent par un numéro de téléphone au format 0 164 258 978, les espaces pouvant être remplacés par des points ou des tirets et le premier chiffre pouvant être entre parenthèses.

TP 08

TP portant sur l'utilisation basique des expressions régulières

Exercice 2 :

On teste les réponses de l'exercice 1 en scannant un fichier d'entrée et en produisant un fichier de sortie.

Fichier d'entrée : RegExpTP8_Ex2.txt.

Fichier de sortie : SortieRegExpTP8_Ex2.txt. Ce fichier de sortie est à votre disposition.

Lorsque vous rendrez le script, nous le lancerons et devrons trouver exactement le même fichier de sortie.

Vous rendrez un dossier compressé contenant votre script nommé TP8_Exo2.py et le fichier d'entrée. Nous lancerons le script qui produira le fichier de sortie dans le même dossier, il portera le nom indiqué plus haut. Il ne nous restera plus qu'à vérifier que votre fichier de sortie coïncide rigoureusement avec celui fourni.

Factorisation du code.

On attend deux fonctions afficheTitre et afficheLignes

afficheTitre attend le titre à afficher pour chaque item :

```
afficheTitre("1.g : lignes contenant des crochets ( ] ou [ )");  
et produit l'affichage correspondant :
```

```
*****  
1.g : lignes contenant des crochets ( ] ou [ )  
*****
```

(lignes 145 à 147 du fichier de sortie).

Il y a autant d'étoiles que de caractères au-dessus et en-dessous de la ligne à afficher. Copiez les 3 lignes dans un éditeur de textes basique et vous verrez que c'est vrai ;-).

afficheLignes attend l'expression régulière et la liste des lignes à tester :

```
afficheLignes('[\n][\n]',lignes);  
# RegExp entre quotes, c'est plus lisible qu'avec guillemets  
et produit l'affichage correspondant :
```

```
cette ligne contient des [  
] cette ligne aussi  
capitaine [] est tres mechant  
***** 3 lignes trouvees *****
```

(lignes 148 à 151 du fichier de sortie).