Préambule

Les expressions ici utilisent la syntaxe reconnue par python qui suit le standard POSIX (à quelques exceptions près [1)]).

Capture

Les parenthèses ( ) de groupement permettent de délimiter un ensemble d'alternatives et d'opérer des captures. Le plus souvent si nous réalisons une capture nous voudrons pouvoir y faire référence via un nom que nous aurons choisi.

En python la syntaxe qui permet cela est la suivante

```
(?P<nom_du_groupe_de_capture>...)
```

La sous-expression régulière (ici symbolisée par ...) qui aura été trouvée sera accessible par le nom ***nom_du_groupe_de_capture***

Dès lors nous pouvons y faire référence dans la suite de l'expression régulière elle-même :

```
(?P=nom_du_groupe_de_capture)\\
```

Mais le plus souvent nous voudront surtout y faire référence lors d'un post formatage.
Tout de suite un petit exemple.

Numéro de facture

Soit le texte extrait suivant

```
RueDuCommerce SA
Facture 44-50, avenue du Capitaine Glarner
93585 Saint-Ouen Cedex
EzDEV
                                                                        2
Rue Adolphe Pégoud
90130 PETIT-CROIX
        Numéro de facture : 56497        Référence client : 100132
        Date de la facture : 04/11/2015
```

Cet exemple illustre bien l'un des cas que vous retrouvez le plus souvent. Il s'agit de récupérer un numéro (ici de facture).

Nous pourrions être tentés de faire une recherche avec le filtre de champ

```
Numero de facture : .*
```

puis un post formatage

```
@keepdigits(@self)
```

Ce qui pourrait fonctionner si nous spécifions correctement la longueur du champ.

Soit le filtre de champ suivant

```
Numero de facture : (?P<numero>\d+)
```

Pour ne garder que le numero on fera ceci dans le champ post de formatage :

```
@caststr(@<numero>)
```

1)
Article wikipedia sur les expression rationnelles
https://fr.wikipedia.org/wiki/Expression_rationnelle#Python

From:
http://wiki.ezdev.fr/ - **EzGED Wiki**

Permanent link:
http://wiki.ezdev.fr/doku.php?id=cold:tutoriaux:regex&rev=1446640992

Last update: **2023/03/17 09:56**