

## Grep, le chercheur de caractères.

Vous avez sans doute déjà été amené à rechercher une suite de caractères. Il existe une commande qui offre cette fonction essentielle : `grep` pour Global Regular Expression Print.

### Les réponses positives.

```
# grep oui participation.list
```

Pour l'occasion imaginez que vous organisez un anniversaire et que vous avez enregistré les réponses de vos invités dans un fichier texte nommé `participation.list` confectionné avec `vi` et enregistré dans votre dossier maison. Dans ce texte, chaque ligne comporte le nom de l'invité et sa réponse séparés par une tabulation. La commande ci-dessus répond par la liste des lignes contenant le mot 'oui'.

### Ignorer la casse

```
# grep -i oui participation.list
```

Vous n'avez pas systématiquement écrit 'oui', mais parfois 'Oui' ou encore 'OUI' pour ce qui est de la réponse de vos invités. Résultat la liste obtenue par la première commande est incomplète. Pour demander à `grep` d'ignorer la casse, de ne plus faire de différence entre majuscules et minuscules, utilisez l'option `-i`.

### Compter les bonnes réponses.

```
# grep -i -c oui participation.list
```

Imaginez un instant que ce n'est plus votre anniversaire, mais la prochaine conférence de Steve Jobs que vous organisez. Obtenir une liste est une bonne chose, savoir combien de spectateurs vont venir est mieux. Pour cela ajoutez à votre ligne de commande `-c` pour compter. La réponse sera alors un nombre et plus une liste de lignes.

### Ne chercher que les mots

```
# grep -i -w -c oui participation.list
```

Voici qu'un des participants se prénomme Louis. Si vous utilisez la commande précédente il sera décompté parmi les réponses positives quoi qu'il arrive car sa ligne contient bien la suite de caractères 'oui' même si cette suite fait partie du mot Louis. Pour que `grep` ne prenne en compte que les 'oui' qui constituent un mot vous pourriez entourer votre oui par des espaces de part et d'autres. Une solution plus élégante consiste à utiliser l'option `-w` (word) qui indique à `grep` que vous recherchez un mot.

### Limiter la recherche aux oui en fin de ligne

```
# grep -ic 'oui$' participation.list
```

Pour éviter l'erreur due à Louis, vous pourriez utiliser le `$`. Lorsqu'il se trouve en fin de critère de recherche, il signifie 'fin de ligne' (mais pas 'retour à la ligne'). Ailleurs il ne désigne rien d'autre que le caractère dollar. Pour l'interpréteur de commande, le dollar indique un nom de variable. Pour éviter que le shell n'interprète ce dollar vous entourez votre critère de recherche entre apostrophes. Ainsi `grep` décomptera les oui uniquement lorsqu'ils sont situés en fin de ligne.

## Recherche multiple

```
# grep -Eiwc 'ok|oui|yes' participation.list
```

Voilà, qu'un tiers est intervenu sur votre liste, mais n'a pas respecté votre convention d'écriture non/oui, d'ailleurs il a aussi bien utilisé un yes, Yes que un ok ou encore OK. Pour contourner cette difficulté vous allez faire appel au mode étendu -E qui vous donne l'opportunité de spécifier plusieurs critères de recherche séparé par des barres verticales. Pour inverser la recherche, connaître ceux qui n'ont dit ni oui, ni yes, ni ok, vous utiliserez l'option -v pour écrire `grep -Eiwcv 'ok|oui|yes' participation.list`

## Trouver l'information

```
# dmesg | grep -A 10 sda
```

Le principal intérêt de `grep` est de filtrer un flot d'information issu d'une autre commande. Par exemple celles retournées par la commande `dmesg`. Pour ne voir que les dix lignes qui suivent la présence de "sda" vous utilisez l'option -A (after) suivie du nombre de lignes voulues. La barre verticale, appelée pipe, fait en sorte que `grep` traite ce qui provient de `dmesg` et non d'un fichier comme dans les exemples précédents.

## Visualiser les résultats dans vi

```
# vi $(grep -l MOTIF *.txt)
```

Vous désirez ouvrir une liste de fichier avec `vi`. Cette liste est issue d'une recherche effectuée par `grep`, l'option -l signifie ne donner que la liste des noms de fichier, pour la suite vous aurez vous même compris qu'il s'agit de la liste des fichier dont le suffixe est .txt et dont une des lignes au moins contient l'expression régulière "MOTIF".