

Suchen / Extrahieren

I/O Umleitung

- Die I/O-Umleitung ermöglicht es dem Benutzer, Informationen von oder zu einem Befehl mithilfe einer Textdatei umzuleiten.

```
$ echo "Hello World!" > text
$ cat text
Hello World!
$ echo "Hello to you too!" >> text
$ cat text
Hello!
Hello to you too!
```

Wichtig

- `>` Beim zweiten Aufruf wird die gleiche Datei mit dem neuen Text überschrieben
- `>>` Beim zweiten Aufruf wird Text am Ende hinzugefügt

Std Error

- Die Error Ausgabe kann mit `2>` umgeleitet werden:

```
$ find /usr games 2> text-error
```

Std Input

- Diese Art der Umleitung dient dazu, einem Befehl Daten aus einer bestimmten Datei statt über die Tastatur zu übergeben:

```
$ cat < text
Hello!
Hello to you too!
```

Kombination von StdErr und StdIn

- `&>` Kann als Kombination von stdin und stderr verwendet werden:

```
$ find /usr admin &> newfile
$ cat newfile
/usr
/usr/shar
```

Pipes

- Um die Ausgabe von einem Befehl mit den Eingang des nächsten zu verbinden, werden Pipes verwendet:

```
$ cat /etc/passwd | less
root:x:0:0:root:/root:/bin/bash
```

```
daemon:x:1:1:daemon:/usr/sbin:/bin/sh
bin:x:2:2:bin:/bin:/bin/sh
```

Grep

- Zum schnellen suchen von Matches in Text input (StdIn, File, Folder)

```
$ grep bash /etc/passwd
root:x:0:0:root:/root:/bin/bash
user:x:1001:1001:User,,,:/home/user:/bin/bash
```

- **-i** Die Suche ist case-insensitiv, ignoriert also Groß-/Kleinschreibung.
- **-r** Die Suche ist rekursiv (sie sucht in allen Dateien innerhalb des angegebenen Verzeichnisses und seiner Unterverzeichnisse).
- **-c** Die Suche zählt die Anzahl der Treffer.
- **-v** Kehrt die Übereinstimmung um und gibt Zeilen aus, die nicht dem Suchbegriff entsprechen.
- **-E** Schaltet erweiterte reguläre Ausdrücke ein (benötigt von einigen fortgeschritteneren Metazeichen wie |, + und ?). grep hat viele andere nützliche Optionen. Schauen Sie in die Man Page, um mehr darüber zu erfahren.

Regular Expressions

- Reguläre Ausdrücke sind äußerst nützlich bei der Extraktion von Daten aus Textdateien durch die Angabe von Mustern.
- Bei der Arbeit mit regulären Ausdrücken ist es sehr wichtig zu beachten, dass jedes Zeichen zählt und dass das Muster mit dem Ziel geschrieben wird, eine bestimmte Zeichenfolge abzubilden, die als Zeichenkette oder String bezeichnet wird.
- Die folgende Liste beschreibt die Metazeichen der regulären Ausdrücke, die zur Bildung der Muster verwendet werden:
- **.** Übereinstimmung eines einzelnen Zeichens (außer Zeilenumbruch).
- **[abcABC]** Übereinstimmung eines beliebigen Zeichens innerhalb der Klammern.
- **[^abcABC]** Übereinstimmung eines beliebigen Zeichens außer denen innerhalb der Klammern.
- **[a-z]** Übereinstimmung eines beliebigen Zeichens im angegebenen Bereich.
- **[^a-z]** Übereinstimmung eines beliebigen Zeichens außer denen im angegebenen Bereich.
- **sun|moon** Übereinstimmung mit einer der angegebenen Zeichenketten.
- **^** Zeilenbeginn
- **\$** Zeilenende

```
$ echo "aaabbb1" > text.txt
$ echo "abab2" >> text.txt
$ echo "noone2" >> text.txt
$ echo "class1" >> text.txt
$ echo "alien2" >> text.txt
$ cat text.txt
aaabbb1
abab2
noone2
class1
alien2

$ grep "ab" text.txt
aaabbb1
abab2
$ grep "[ab]" text.txt
```

aaabbb1
abab2
class1
alien2