

## grep und sed

### Aufgabe 1: grep

Laden sie die Datei grepdata.txt aus Stud.IP herunter, und bearbeiten Sie die folgenden Aufgaben auf dieser Datei. Benutzen Sie dabei den Befehl egrep

- a) Geben Sie allen Zeilen aus, die eine Telefonnummer mit extension beinhalten (x oder X, gefolgt von 4 Zahlen)
- b) Geben Sie alle Zeilen aus, die mit drei Zahlen, gefolgt von einem Leerzeichen beginnen.
- c) Geben Sie alle Zeilen aus, die ein Datum enthalten
- d) Geben sie alle Zeilen aus, die einen Vokal enthalten, gefolgt von einem einzelnen weiteren Zeichen, gefolgt von dem gleichen Vokal (findet "eve" und "adam", aber nicht "vera"). Hinweis: "(" und ")"
- e) Geben Sie alle Zeilen aus, die nicht mit S beginnen
- f) Geben Sie alle Zeilen aus, die CA in Groß- oder Kleinbuchstaben enthalten
- g) Geben Sie alle Zeilen aus, die eine E-Mail-Adresse beinhalten, aus, gefolgt von der Zeilennummer
- h) Geben Sie alle Zeilen aus, die das Wort de als ganzes Wort enthalten

### Aufgabe 2: sed

Laden Sie die Datei docbook.txt aus Stud.IP herunter, und bearbeiten Sie die folgenden Aufgaben auf dieser Datei:

- a) Zeilen mit <article> und </article> sollen gelöscht werden
- b) Ersetzen Sie <title> mit Title:, und </title> mit nichts
- c) Ersetzen Sie alle <para> und </para> tags nichts. Wenn die resultierende Zeile leer ist, löschen Sie die Zeile.
- d) Ersetzen Sie alle <emphasis> und </emphasis> tags mit einem Stern \*
- e) Ersetzen Sie überall web mit Web
- f) Ersetzen Sie Zeilen, die mit <listing> beginnen, mit --begin listing. Ersetzen Sie Zeilen, die mit </listing> beginnen, mit --end listing
- g) Zwischen <listing> und </listing>, machen Sie folgendes
  - Ersetzen Sie Vorkommen von &lt; mit <
  - Ersetzen Sie Vorkommen von &gt; mit >
  - Ersetzen Sie Vorkommen von &amp; mit &

### Aufgabe 3

Benutzen Sie sed und grep, um sich alle Links aus einer HTML-Datei ausgeben zu lassen. Holen Sie sich dafür zunächst eine HTML-Datei, z.B. so:

wget <https://wiki.ubuntuusers.de/grep/>