

Übung 3: Werkzeuge (AWK und sed)

1. Kopieren Sie die Datei `/home/public/vorkurs_ss18/de_patents.tsv` in Ihr Vorkurs-Verzeichnis. Die Datei enthält Metadaten für deutsche Patente in fünf TAB-separierten Spalten (Patentnummer, Datum, Titel, Firma, Erfinder (einzelne Erfinder sind durch „ ,“ getrennt)).

Hinweis: Die meisten dieser Aufgaben lassen sich mit AWK oder mit Unix Tools lösen. Probieren Sie beides aus! Achten Sie darauf, dass AWK standardmäßig Leerzeichen als Spaltentrenner interpretiert, die Datei aber TAB-separiert ist.

2. Lassen Sie sich nur die 1. und 3. Spalte anzeigen.
3. Lassen Sie sich die 5 Spalten in einer anderen Reihenfolge ausgeben.
4. Geben Sie jetzt nur jede 3. Zeile aus.
5. Schreiben Sie für jedes Patent der Firma Siemens nur den Titel des Patents heraus (achten Sie auf Großschreibung!).
6. Ersetzen Sie die Spaltentrenner durch " ||| ".
7. Schreiben Sie eine Pipe, die die einzelnen Erfinder absteigend sortiert nach der Anzahl ihrer Erfindungen ausgibt. Wie heißt die Frau mit den meisten Patenten?