

Kurs: Sed/Awk-Programmierung (inkl. Reguläre Ausdrücke)

© T. Birnthaler, OSTC GmbH (www.ostc.de)

(Version 1.1 vom 22.12.2004)

Dauer: 2-3 Tage

Motivation: Sed und Awk sind 2 "programmierbare" Linux/UNIX-Standard-Werkzeuge zur Text- und Datenverarbeitung, die mit relativ geringem Aufwand zu erlernen und einzusetzen sind (im Gegensatz z.B. zu Perl!). Zusammen mit der Shell sind beide Tools für praktisch alle Aufgaben in diesem Bereich geeignet.

In beiden Sprachen werden Reguläre Ausdrücke häufig zur Textsuche und -modifikation eingesetzt. Kenntnisse in Regulären Ausdrücken sind generell sinnvoll, da sie inzwischen in vielen anderen Bereichen (Linux/UNIX-Tools, Editoren, Programmiersprachen) sogar unter Windows einsetzbar sind.

Kenntnisse des Awk sind aus folgenden Gründen sinnvoll: vollständige und leistungsfähige Programmiersprache, leicht zu erlernen, standardmäßig auf jedem Linux/UNIX-System vorhanden, Reguläre Ausdrücke in erweiterter Form werden unterstützt, Vorläufer vieler anderer Sprachen (z.B. Perl!).

Ziel: Sie sind nach dem Kurs in der Lage, als Anwender, Datenverarbeiter, Systemadministrator oder Software-Entwickler gut und effizient mit Regulären Ausdrücken, Sed und Awk umzugehen und sie vor allem gerne einzusetzen.

Hinweis: Der Kurs ist für jedes Linux/UNIX-System geeignet.

Inhalt:

- Einführung in Linux/UNIX und Skript-Sprachen
- Reguläre Ausdrücke
 - Unterschiede zu Shell-Metazeichen

- Escape-Sequenzen
- Standard-Metazeichen
- Erweiterte Metazeichen
- Metazeichen im Ersetzungsteil
- Anwendung
 - * Suchen
 - * Suchen + Ersetzen
- Sed
 - Aufruf (Parameter und Optionen)
 - Funktionsweise (automatische Leseschleife)
 - Programmstruktur
 - Adressen
 - Blöcke
 - Kommentare
 - Operationen
 - * Grundbefehle
 - * Test und Verzweigung
 - * Zwischenpuffer
 - * Mehrzeilen-Verarbeitung
- Awk
 - Eigenschaften und Einsatzgebiete
 - * 10 typische Einzeiler
 - Funktionsweise (automatische Leseschleife und automatische Feldtrennung)

- Aufruf (Parameter und Optionen)
- Konstanten
- Programmstruktur (Muster + Aktion)
- Datentypen
 - * Automatische Datentyp-Konvertierung
 - * Defaultwerte
- Automatische Variablen
- Operatoren
- Kontrollstrukturen
- Assoziative Arrays (Hashes)
 - * Standard-Arrays
 - * Mehrdimensionale Arrays
- Eigene Funktionen
- Weitere Standard-Variablen
- Standard-Funktionen
 - * Datei-Funktionen
 - * Zeichenketten-Funktionen
 - * Mathematische Funktionen
 - * Zeit-Funktionen
- Spezialitäten des Gawk

Schulungsunterlagen:

- Skript
- Lösungsblätter zu allen Übungen
- Zusammenfassung der wichtigsten Linux/UNIX-Befehle

- Zusammenfassung der wichtigsten Vi-Befehle
- Ausführliche Linux/UNIX-Befehlsübersicht
- Reguläre Ausdrücke — Beschreibung und Anwendung
- Der Editor Sed — Einführung, Tipps und Tricks
- Die Programmiersprache Awk — Einführung, Tipps und Tricks

Voraussetzungen:

- Muss: Allgemeine Computer-Kenntnisse (Tastatur, Maus, Grafische Oberfläche, Editor)
- Muss: Erfahrung in irgendeiner Programmiersprache
- Optional: Linux/UNIX-Grundlagen
- Optional: Arbeiten auf der Kommandozeile (analog MS-DOS)

Abgrenzung:

- Keine grafische Oberfläche (KDE oder GNOME)
- Keine root-Rechte notwendig