

Reguläre Ausdrücke einfach erklärt

Zusammenfassung von Matthias Dieter Wallnöfer

Quelle: <https://www.slideshare.net/arnolambert/regex-presentation-1677711>

Einleitung

- Zwei Standards:
 - **PCRE** (Perl-Compatible)
 - **Posix** (Simple Regular Expression) gegliedert in *Basic* (Tool grep) bzw. *Extended* (egrep)
- Darum je nach Interpreter (Editor, Kommandozeilentool, Programmiersprache) unterschiedlich unterstützt, es hilft nur ausprobieren und Dokus durchlesen
- Die angegebenen Beispiele laufen unter Eclipse

Literale

a matcht „a“

Zeichenklassen

- Gruppen wie [abcdefgh0123]
- Intervalle wie [a-h0-3]
- Inverse Intervalle bzw. Zeichen [^i-z4-9]
- Auch [gG]uide[lL]ine für „guideline“ oder „GuideLine“

Abkürzungen für die Zeichenklassen

- . → beliebiges Zeichen
- \d → [0-9]
- \w → [a-zA-Z_0-9]
- \s → Leerzeichen + [\t\r\n] (Tabulator, Wagenrücklauf und Zeilenvorschub)
- \D → [^0-9]
- \W → [^a-zA-Z_0-9]
- \S → [^\t\r\n]

Nichtdarstellbare Zeichen

- \t, \r, \n
- \xFF (hexadezimaler ASCII-Code), Beispiel \x0A für Zeilenvorschub
- \uFFFF (Unicode-), Beispiel \u00F6 für „ö“
- ^ Beginn (der Zeile)
- \$ Ende (der Zeile), Beispiel: ^hallo\$

Quantifikatoren

- {min, max} / {min, } / { , max} / {exact}
- ? → {0, 1}
- + → {1, }
- * → {0, }
- | → oder

Gruppen

Beispiel: ((red|white) (king|queen)) matcht „red king“, „white king“, „red queen“, „white queen“

Gruppensubstitution

Beispiel CSV-Datei Spalten vertauschen

Suche und Zuweisung der Spalten in zwei Gruppen: `^(\\w*);(\\w*)$`
Ersetzen durch: `$2;$1`

Beispiel Java-Code

`variablenName.someMethod()` →
`((TypeName)variablenName.someMethod())`

Suche und Zuweisung einer Gruppe: `(\\w+\\.someMethod\\(\\))`
Ersetzen durch: `((TypeName)$1)`

Beispiel Umdrehen der Könige und Königinnen

Suche und Zuweisung in drei Gruppen: `((red|white) (king|queen))`

Ersetzen durch (`$1` wäre die gesamte äußere Gruppe): `$3 $2`