

# Übersicht wichtige Regex-Engines

- GNU Basic Regular Expressions (u.a. grep in Defaultmodus)
- Python-Bibliothek re
- Java-Bibliothek java.util.regex
- Perl Compatible Regular Expressions (u.a. Perl, R, grep -P)
- Boost.Regex (u.a. Notepad++)

# Vergleich von Regex-Engines

Feature	GBRE	Python	Java	PCRE	Boost
Metazeichen ? + {   ( ) müssen nach \ stehen	Ja	Nein	Nein	Nein	Nein
Rekursives Matching	Nein	Nein	Nein	Ja	Ja
Look-ahead und Look-behind	Nein	Ja	Ja	Ja	Ja
Regex wie (a ab abc) matchen erstes od. längstes	längst.	erstes	erstes	erstes	längst.
Unicode-Kategorien und Properties	Nein	Nein	Ja	Ja	Nein
POSIX-Klassen	Ja	Nein	Nein	Ja	Ja
Abkürzungen in [ ]*	Nein	Ja	Ja	Ja	Ja
\X für ein sichtbares Zeichen**	Nein	Nein	Ja	Ja	Ja
Kommentare im Regex	Nein	Ja	Nein	Ja	Ja

\*z.B. [%&\w] als ungefähr äquivalent zu [%&A-Za-z0-9]

\*\*Manche Symbole (z.B. Emoji mit Hautfarbenmodifizierer oder bestimmte Formen arabischer Schriftzeichen) werden, obwohl sie nur ein sichtbares Zeichen darstellen, intern als eine bestimmte Kombination mehrerer Unicode-Zeichen gespeichert. \X matched ein *sichtbares* Zeichen, egal aus wie vielen Unicode-Zeichen es besteht.

GBRE-Regex sind auf *sehr* langen Dokumenten i.A. deutlich schneller

Vergleich von noch vielen anderen, weniger benutzten Features:

<https://www.regular-expressions.info/reference.html>

