

# Reguläre Ausdrücke

*Hinweis:*

- Die Testdaten sind jeweils in eine positive und eine negative Menge aufgeteilt. Die Elemente in der positiven Menge sollen abgedeckt werden, die in der negativen nicht.
- Sie können Ihre Lösungen auf der Seite `regexr` testen.

## Aufgaben

### 1. Einfache Datumsangaben (einfach)

Schreiben Sie einen regulären Ausdruck, der eine einfache Version von Datumsangaben findet. Dabei sollen der Tag und der Monat jeweils aus zwei und die Jahresangabe aus vier Zahlen bestehen. Eine Validierung auf kalendarische Korrektheit (zum Beispiel, dass der 30.02.2019 oder der 99.17.3005 nicht existieren) sind nicht erforderlich.

Sie können beispielsweise folgende Testdaten benutzen: \* Positiv: 12.01.1999, 99.17.3006, 33.11.1931 \* Negativ: 1.1.1999, 01.01.700, 3.12.2011, 29.3.2018, 1.1.1, 12.10, 11.2.

### 2. Komplexe Datumsangaben (mittel)

Die Datumsangaben aus Aufgabe 1 werden erweitert, so dass die Tages- und Monatsangabe auch einstellig und die Jahreszahl zwei- bis vierstellig sein kann.

Testdaten: \* Positiv: 12.01.1999, 99.17.3006, 33.11.1931, 1.1.1999, 01.01.700, 3.12.2011, 29.3.2018 \* Negativ: 1.1.1, 12.10, 11.2.

### 3. Semantische Versionierung (mittel)

(Semantische) Versionsnummern bestehen aus einer Haupt-, einer Neben-, einer Revisions- und einer optionalen Buildnummer, die jeweils mit einem Punkt beziehungsweise im Fall der Revisions- und Buildnummer mit einem Minus getrennt werden. Jede der Nummern kann mehrstellig sein und führende Nullen besitzen.

Testdaten: \* Positiv: 2.3.5-0041, 0.0.1, 1.3.3-7, 1.12.2-21 \* Negativ: 1.0, 1..10, 1.0-3

#### 4. Postleitzahlen (einfach)

Deutsche Postleitzahlen bestehen aus fünf Ziffern ohne führende Nullen.

Testdaten: \* Positiv: 97070, 90210, 10115 \* Negativ: 01234, 8911, 654321

#### 5. Mailadressen (schwer)

Eine Mailadresse besteht aus einem lokalen und einem Domänenteil, die durch ein @-Zeichen getrennt sind. Der Domänenteil wiederum besteht aus einem Hostnamen und einer Top-Level-Domain, die durch einen Punkt getrennt sind.

Der Einfachheit halber beschränken wir die Zeichen, die in einer Mailadresse vorkommen dürfen, auf alle Klein- und Großbuchstaben, alle Ziffern, Minus, Punkt und Unterstrich (\_). Jeder Teil besteht aus mindestens einem dieser Zeichen.

Testdaten: \* Positiv: james.bond@mi5.uk, jason\_bourne@cia.gov, jack.bauer@24.com \* Negativ: xyz@localhost, @web.de, test.net

## Musterlösungen

### 1. Einfache Datumsangaben:

- `\d\d\.\d\d\.\d\d\d\d`
- `\d{2}\.\d{2}\.\d{4}`
- Anstatt `\d` kann auch `[0-9]` verwendet werden.

### 2. Komplexe Datumsangaben:

- `\d\d?\.\d\d?\.\d\d\d?\d?`
- `[0-9]{1,2}\.[0-9]{1,2}\.[0-9]{2,4}`
- ...

### 3. Semantische Versionierung:

- `\d+\.\d+\.\d+(-\d+)?`

### 4. Postleitzahlen:

- `[1-9][0-9]{4}`

### 5. Mailadressen

- `([a-z] | [A-Z] | [0-9] | _ | - | \. )+ @ ( [a-z] | [A-Z] | [0-9] | _ | - )+ \. ( [a-z] | [A-Z] | [0-9] | _ | - | \. )+`