

Inhalt

Vorwort	XI
1 Einführung in reguläre Ausdrücke	1
Definition regulärer Ausdrücke	1
Suchen und Ersetzen mit regulären Ausdrücken	6
Tools für das Arbeiten mit regulären Ausdrücken	8
2 Grundlagen regulärer Ausdrücke	27
2.1 Literalen Text finden	28
2.2 Nicht druckbare Zeichen finden	30
2.3 Ein oder mehrere Zeichen finden	33
2.4 Ein beliebiges Zeichen finden	37
2.5 Etwas am Anfang und/oder Ende einer Zeile finden	39
2.6 Ganze Wörter finden	44
2.7 Codepoints, Eigenschaften, Blöcke und Schriftsysteme bei Unicode	47
2.8 Eine von mehreren Alternativen finden	59
2.9 Gruppieren und Einfangen von Teilen des gefundenen Texts	61
2.10 Vorher gefundenen Text erneut finden	64
2.11 Teile des gefundenen Texts einfangen und benennen	66
2.12 Teile der Regex mehrfach wiederholen	69
2.13 Minimale oder maximale Wiederholung auswählen	72
2.14 Unnötiges Backtracking vermeiden	75
2.15 Aus dem Ruder laufende Wiederholungen verhindern	78
2.16 Etwas auf Übereinstimmung prüfen, ohne es dem Gesamtergebnis hinzuzufügen	81
2.17 Abhängig von einer Bedingung eine von zwei Alternativen finden	87
2.18 Kommentare für einen regulären Ausdruck	90

2.19	Literalen Text im Ersetzungstext nutzen.	92
2.20	Einfügen des Suchergebnisses in den Ersetzungstext	95
2.21	Teile des gefundenen Texts in den Ersetzungstext einfügen.	96
2.22	Suchergebniskontext in den Ersetzungstext einfügen.	100
3	Mit regulären Ausdrücken programmieren.	103
3.1	Literale reguläre Ausdrücke im Quellcode	109
3.2	Importieren der Regex-Bibliothek.	115
3.3	Erstellen eines Regex-Objekts.	117
3.4	Optionen für reguläre Ausdrücke setzen.	123
3.5	Auf eine Übereinstimmung in einem Text prüfen.	131
3.6	Auf eine vollständige Übereinstimmung einer Regex mit einem Text prüfen	137
3.7	Auslesen des übereinstimmenden Texts	142
3.8	Position und Länge der Übereinstimmung ermitteln	149
3.9	Teile des übereinstimmenden Texts auslesen	154
3.10	Eine Liste aller Übereinstimmungen erhalten	162
3.11	Durch alle Übereinstimmungen iterieren	167
3.12	Übereinstimmungen in prozeduralem Code überprüfen	173
3.13	Eine Übereinstimmung in einer anderen Übereinstimmung finden	177
3.14	Alle Übereinstimmungen ersetzen.	181
3.15	Übereinstimmungen durch Teile des gefundenen Texts ersetzen.	189
3.16	Übereinstimmungen durch Text ersetzen, der im Code erzeugt wurde ..	194
3.17	Alle Übereinstimmungen innerhalb der Übereinstimmungen einer anderen Regex ersetzen.	200
3.18	Alle Übereinstimmungen zwischen den Übereinstimmungen einer anderen Regex ersetzen.	203
3.19	Einen String aufteilen	208
3.20	Einen String aufteilen und die Regex-Übereinstimmungen behalten	217
3.21	Zeile für Zeile suchen	222
4	Validierung und Formatierung	227
4.1	E-Mail-Adressen überprüfen	227
4.2	Nordamerikanische Telefonnummern validieren	233
4.3	Internationale Telefonnummern überprüfen	239
4.4	Klassische Datumsformate validieren	241
4.5	Klassische Datumsformate exakt validieren	245
4.6	Klassische Zeitformate validieren	250
4.7	Datums- und Uhrzeitwerte im Format ISO 8601 validieren	252

4.8	Eingabe auf alphanumerische Zeichen beschränken	257
4.9	Die Länge des Texts begrenzen	260
4.10	Die Zeilenanzahl eines Texts beschränken	265
4.11	Antworten auswerten	269
4.12	US-Sozialversicherungsnummern validieren	271
4.13	ISBN validieren	274
4.14	ZIP-Codes validieren	281
4.15	Kanadische Postleitzahlen validieren	282
4.16	Britische Postleitzahlen validieren	282
4.17	Deutsche Postleitzahlen validieren	283
4.18	Namen von „Vorname Nachname“ nach „Nachname, Vorname“ umwandeln	285
4.19	Kreditkartennummern validieren	288
4.20	Europäische Umsatzsteuer-Identifikationsnummern	294
5	Wörter, Zeilen und Sonderzeichen	301
5.1	Ein bestimmtes Wort finden	301
5.2	Eines von mehreren Wörtern finden	304
5.3	Ähnliche Wörter finden	306
5.4	Alle Wörter außer einem bestimmten finden	310
5.5	Ein beliebiges Wort finden, auf das ein bestimmtes Wort nicht folgt	312
5.6	Ein beliebiges Wort finden, das nicht hinter einem bestimmten Wort steht	313
5.7	Wörter finden, die nahe beieinanderstehen	317
5.8	Wortwiederholungen finden	323
5.9	Doppelte Zeilen entfernen	325
5.10	Vollständige Zeilen finden, die ein bestimmtes Wort enthalten	330
5.11	Vollständige Zeilen finden, die ein bestimmtes Wort nicht enthalten	332
5.12	Führenden und abschließenden Whitespace entfernen	333
5.13	Wiederholten Whitespace durch ein einzelnes Leerzeichen ersetzen	336
5.14	Regex-Metazeichen maskieren	337
6	Zahlen	341
6.1	Integer-Zahlen	341
6.2	Hexadezimale Zahlen	345
6.3	Binärzahlen	348
6.4	Führende Nullen entfernen	349
6.5	Zahlen innerhalb eines bestimmten Bereichs	350
6.6	Hexadezimale Zahlen in einem bestimmten Bereich finden	357

6.7	Gleitkommazahlen	359
6.8	Zahlen mit Tausendertrennzeichen	363
6.9	Römische Zahlen	364
7	URLs, Pfade und Internetadressen	367
7.1	URLs validieren	367
7.2	URLs in einem längeren Text finden	371
7.3	URLs in Anführungszeichen in längerem Text finden	373
7.4	URLs mit Klammern in längerem Text finden	374
7.5	URLs in Links umwandeln	376
7.6	URNs validieren	377
7.7	Generische URLs validieren	379
7.8	Das Schema aus einer URL extrahieren	385
7.9	Den Benutzer aus einer URL extrahieren	386
7.10	Den Host aus einer URL extrahieren	388
7.11	Den Port aus einer URL extrahieren	390
7.12	Den Pfad aus einer URL extrahieren	392
7.13	Die Query aus einer URL extrahieren	396
7.14	Das Fragment aus einer URL extrahieren	397
7.15	Domainnamen validieren	398
7.16	IPv4-Adressen finden	400
7.17	IPv6-Adressen finden	403
7.18	Einen Pfad unter Windows validieren	417
7.19	Pfade unter Windows in ihre Bestandteile aufteilen	420
7.20	Den Laufwerkbuchstaben aus einem Pfad unter Windows extrahieren ..	425
7.21	Den Server und die Freigabe aus einem UNC-Pfad extrahieren	426
7.22	Die Ordernamen aus einem Pfad unter Windows extrahieren	427
7.23	Den Dateinamen aus einem Pfad unter Windows extrahieren	430
7.24	Die Dateierweiterung aus einem Pfad unter Windows extrahieren	431
7.25	Ungültige Zeichen aus Dateinamen entfernen	432
8	Markup und Datenaustausch	435
8.1	Tags im XML-Stil finden	442
8.2	-Tags durch ersetzen	459
8.3	Alle Tags im XML-Stil außer und entfernen	463
8.4	XML-Namen finden	466
8.5	Einfachen Text durch Ergänzen von <p>- und - Tags nach HTML konvertieren	473
8.6	Ein bestimmtes Attribut in Tags im XML-Stil finden	476

8.7 Tags vom Typ <table> ein Attribut „cellspacing“ hinzufügen, die es noch nicht haben.	481
8.8 Kommentare im XML-Stil entfernen.	484
8.9 Wörter in Kommentaren im XML-Stil finden.	488
8.10 Ändern der Feldbegrenzer in CSV-Dateien	493
8.11 CSV-Felder aus einer bestimmten Spalte extrahieren	496
8.12 Sektionsüberschriften in INI-Dateien finden	500
8.13 Sektionsblöcke in INI-Dateien finden.	502
8.14 Name/Wert-Paare in INI-Dateien finden	503
Index	505