

# GREP-Referenz

Auszug aus dem Buch **InDesign automatisieren – Keine Angst vor Skripting, GREP & Co.** von Gregor Fellenz  
Erschienen beim dpunkt-verlag. Die Homepage zum Buch: <http://www.indd-skript.de>

In dieser Referenz sind URLs mit einem Linkverkürzer abgekürzt: ↗ 76 wird zu <http://indd-skript.de/76>

## Metazeichen für die Suche

### Reservierte Zeichen

Zeichen, die im Suchbegriff mit einem \ maskiert werden müssen: \ . \* + ? ( ) { } [ ^ \$ | ~

### Standardzeichenklassen

Wenn ein Großbuchstabe verwendet wird, wird genau das Gegenteil gefunden. \D findet alles außer Ziffern.

GREP	Beschreibung
.	Alle Zeichen außer Umbruchzeichen
\u	Alle Großbuchstaben (uppercase). Findet keine elektronisch versal gestellten Texte.
\l	Alle Kleinbuchstaben (lowercase)
[\u\l]	Beliebiger Buchstabe
\w	Alle Wortzeichen (inklusive Zahlen und Unterstrich _ ohne Bindestrich -)
\d	Alle Ziffern (digits) 0 bis 9. Keine Brüche und nur elektronisch hoch- oder tiefgestellte Zahlen.
\s	Leerräume und Umbruchzeichen. InDesign CS3/Mac findet nur Tabulatoren, Umbrüche und das Leerzeichen, keine Festabstände.
\r	Alle Umbruchzeichen (Absatzende) außer \n (harter Zeilenumbruch). Die Umkehrung mit \R funktioniert nicht, keine echte Zeichenklasse.

### Eigene Zeichenklassen

Innerhalb eigener Zeichenklassen müssen die folgenden Zeichen maskiert werden: ] ^ - \ ~

GREP	Beschreibung
[]	Alle Zeichen innerhalb der eckigen Klammern. Es können Bereiche angegeben werden [a-zA-Z].
[^]	Eigene Zeichenklasse negieren

### Zeichen per Codepoint oder Unicode-Name

GREP	Beschreibung
\x{####}	Vierstelliger Unicode-Codepoint, z. B. \x{0020} (Leerzeichen)
\N{Unicode-Name}	Zeichen über englischen Unicode-Namen, z. B. \N{Space} (Leerzeichen)
[ [.Unicode-Name. ] ]	Wie \N{Unicode-Name}

## Wichtige Unicode-Properties

Unicode-Properties können ab CS4 eingesetzt werden. Mit \P erfolgt die Negation der Properties. In eigenen Zeichenklassen können keine Unicode-Properties eingesetzt werden.

Kurzform	Lange Form	Beschreibung
<b>\p{L*}</b>	<b>\p{letter}</b>	<b>Alle Buchstaben der vier folgenden Klassen</b>
\p{Ll}	\p{lowercase_letter}	Alle Kleinbuchstaben
\p{Lu}	\p{uppercase_letter}	Alle Großbuchstaben
\p{Lt}	\p{titlecase_letter}	Anfangsbuchstaben, die sich von Großbuchstaben unterscheiden (spezielle Schriftsysteme, siehe ↗ 76 oder ↗ 77)
\p{Lm}	\p{modifier_letter}	Diakritische Zeichen, die das vorangehende Zeichen modifizieren (Unicode 02B0–02FF)
\p{Lo}	\p{letter_other}	Buchstaben, die nicht von den oben genannten Klassen erfasst werden
<b>\p{Z*}</b>	<b>\p{separator}</b>	<b>Alle Leerräume und Umbruchzeichen ohne Tabulator</b>
\p{Zs}	\p{space_separator}	Alle Leerräume ohne die Umbruchzeichen und ohne Tabulator
\p{Zl}	\p{line_separator}	Line-separator (Unicode 2028) (wird beim XML-Export verwendet)
\p{Zp}	\p{paragraph_separator}	Paragraph-separator (Unicode 2029) (wird beim XML-Export verwendet)
<b>\p{S*}</b>	<b>\p{symbol}</b>	<b>Mathematische, Währungs- und sonstige Symbole</b>
\p{Sm}	\p{math_symbol}	Mathematische Symbole
\p{Sc}	\p{currency_symbol}	Währungssymbole
\p{Sk}	\p{modifier_symbol}	Bedeutungsverändernde Symbole wie Akzente
\p{So}	\p{other_symbol}	Wingdings, Dingbats aus verschiedenen Unicode-Bereichen
<b>\p{N*}</b>	<b>\p{number}</b>	<b>Alle Ziffern</b>
\p{Nd}	\p{decimal_digit_number}	Ziffern von 0 – 9
\p{Nl}	\p{letter_number}	Römische Ziffern mit den Codepoints (2150–218F)
\p{No}	\p{other_number}	Hoch- und tiefgestellte Ziffern und Brüche
<b>\p{P*}</b>	<b>\p{punctuation}</b>	<b>Satzzeichen, Striche, Anführungen usw.</b>
\p{Pd}	\p{dash_punctuation}	Divis, Gedankenstrich usw., keine bedingten Trennstriche oder automatischen Trennungen
\p{Ps}	\p{open_punctuation}	Öffnende Klammern wie {{{
\p{Pe}	\p{close_punctuation}	Schließende Klammern
\p{Pi}	\p{initial_punctuation}	Öffnende Anführungen, aber nicht "
\p{Pf}	\p{final_punctuation}	Schließende Anführungen, aber nicht "
\p{Po}	\p{other_punctuation}	Sonstige Satzzeichen: „:?!/% & ...

Die Liste ist eine Auswahl aus der vollständigen Übersicht von Peter Kahrel in *GREP in InDesign CS3/CS4*. Unter ↗ 78 sind die genauen Inhalte der Blöcke, von denen sich viele Properties ableiten, beschrieben.

## Posix-Ausdrücke

Die Tabelle enthält Ausdrücke, die die Standardzeichenklassen erweitern. Die Negation erfolgt mit [^[:alnum:]]. Posix-Ausdrücke können nicht in eigenen Zeichenklassen verwendet werden.

GREP	Beschreibung
[[:alnum:]]	Buchstaben und Zahlen (alphanumeric characters) [\u\1\d]
[[:alpha:]]	Groß- und Kleinbuchstaben (alphabetic characters) [\u\1]
[[:blank:]]	Leerzeichen und Tabulator [ \t]
[[:punct:]]	Satzzeichen, Striche, Anführungen usw. (punctuation) [ , ; ! ? \ - ... ]
[[:graph:]]	Sichtbare Zeichen [ ^ \t\n\r ]
[[:xdigit:]]	Hexadezimale Ziffer [ 0-9a-fA-F ]
[[:unicode:]]	Alle Zeichen, deren Codepoint größer 255 (Hex 00FF) ist (alle Zeichen außer den Unicode-Bereichen Basis-Lateinisch und Lateinisch-1, Ergänzung)
[[:a=]]	Buchstabenäquivalent, alle Formen des Buchstabens [ ä å â ... ]

## Positionen

GREP	Beschreibung
^	Anfang des Absatzes
\$	Ende des Absatzes
\b	Wortgrenze
\B	Keine Wortgrenze
\<	Wortanfang
\>	Wortende
\A	Anfang des Textabschnitts
\z oder \Z	Ende des Textabschnitts. Unter CS3 ignoriert \z die Textabschnitte in Tabellenzellen.

## Wiederholungen

Wiederholungszeichen	Beschreibung
?	Null oder ein Mal
*	Null oder mehrere Male
+	Ein oder mehrere Male
??	Null oder ein Mal (kürzeste Entsprechung/non greedy)
*?	Null oder mehrere Male (kürzeste Entsprechung/non greedy)
+?	Ein oder mehrere Male (kürzeste Entsprechung/non greedy)
{n}	Der Suchbegriff muss exakt n-mal vorkommen.
{n,m}	Der Suchbegriff-Ausdruck muss mindestens n-mal und darf maximal m-mal vorkommen.
{n,}	Der Suchbegriff-Ausdruck muss mindestens n-mal und darf öfter vorkommen.
{n,}?	Der Suchbegriff-Ausdruck muss mindestens n-mal und darf öfter vorkommen (non greedy).

## Entsprechungen – Fundstellen in der Suche markieren

GREP	Beschreibung
()	Fundstelle, markierter Unterausdruck
(?:)	Fundstelle nicht markieren, nicht markierter Unterausdruck
	Variante, oder (entweder oder)
\1...\9	Fundstelle/markierten Unterausdruck innerhalb der Suche verwenden

## Look Around Assertions/Umgebung

Assertion	Beschreibung
(?<=Präfix)	Positives Lookbehind
(?!Präfix)	Negatives Lookbehind
(?=Suffix)	Positives Lookahead
(?!Suffix)	Negatives Lookahead

## Modifizierer

GREP	Beschreibung
(?i)	Nicht zwischen Groß- und Kleinschreibung unterscheiden (case insensitive)
(?-i)	Zwischen Groß- und Kleinschreibung unterscheiden (Standardeinstellung)
(?s)	Gesamten Textabschnitt durchsuchen (der . trifft auch Umbruchzeichen; single line mode)
(?-s)	Suchbegriff muss innerhalb eines Absatzes liegen (Standardeinstellung)
(?x)	Leerzeichen im Ausdruck ignorieren
(?-x)	Leerzeichen im Ausdruck nicht ignorieren (Standardeinstellung)
(?:)	Klammer wird nicht als Fundstelle gezählt
(?#comment?)	Kommentar
(?m)	^ und \$ treffen vor bzw. nach einem harten Zeilenumbruch \n und einem Umbruchzeichen \r (multi line mode; »Mehrzeilig« ein; Standardeinstellung)
(?-m)	^ und \$ treffen nur am Anfang und Ende des Textabschnitts (multi line mode; »Mehrzeilig« aus)

## Metazeichen für die Ersetzung

Zeichen, die bei der Ersetzung maskiert werden müssen: \ \$ ~

### Fundstellen verwenden

GREP	Beschreibung
\$0	Gesamten Suchtreffer in der Ersetzung verwenden
\$1 ... \$9	Fundstelle/markierten Unterausdruck in der Ersetzung verwenden
~c	Zwischenablage einfügen
~C	Zwischenablage unformatiert einfügen

## Sonderzeichen für Suche und Ersetzung

In den folgenden Tabellen sind zusätzlich die Unicode-Codepoints und ggf. die Eigenschaft in der Aufzählung SpecialCharacters enthalten. Beachten Sie, dass einige Spezialzeichen auf Unicode-Codepoints liegen, die eine andere Bedeutung haben. Die Tabelle enthält die Codepoints, wie sie von InDesign verwendet werden.

### Leerräume

GREP	Zeichen	Beschreibung	Unicode	SpecialCharacters.
[ ]	·	Leerzeichen, definiert über die Schriftart	0020	
\t	»	Tabulator	2009	
~S	^	Geschütztes normales Leerzeichen, variabel bei Blocksatz	00A0	NONBREAKING_SPACE
~s	^	Geschütztes normales Leerzeichen fester Breite	202F	FIXED_WIDTH_NONBREAKING_SPACE
~m	—	Geviert, Größe der Schrift (12 pt = 12 pt Breite)	2003	EM_SPACE
~>	—	Halbgeviert	2002	EN_SPACE
~3	•	Drittelgeviert	2004	THIRD_SPACE
~4	•	Viertelgeviert	2005	QUARTER_SPACE
~%	:	Sechstelgeviert	2006	SIXTH_SPACE
~<	˘	Achtelgeviert	2009	THIN_SPACE
~	˘˘	24tel-Geviert	200A	HAIR_SPACE
~.	!	Interpunktionsleerzeichen, Breite eines Satzzeichens in der aktuellen Schrift	2008	PUNCTUATION_SPACE
~/	#	Ziffernleerzeichen, Breite einer Ziffer in der aktuellen Schrift	2007	FIGURE_SPACE
~f	˘	Ausgleichsleerzeichen, treibt Blocksatz aus	2001	FLUSH_SPACE

## Umbruchzeichen

GREP	Zeichen	Beschreibung	Unicode	SpecialCharacters.
\r		Alle Umbruchzeichen außer harter Zeilenumbruch, in der Ersetzung Standardumbruchzeichen		
~b	¶	Standardumbruchzeichen, Zeilenschaltung	000D	
\n	↵	Harter Zeilenumbruch	000A	FORCED_LINE_BREAK
~M	↵	Spaltenumbruch	000D	COLUMN_BREAK
~R	⏏	Rahmenumbruch	000D	FRAME_BREAK
~P	⏏	Seitenumbruch	000D	PAGE_BREAK
~L	↵	Umbruch für ungerade Seiten	000D	ODD_PAGE_BREAK
~E	↵	Umbruch für gerade Seiten	000D	EVEN_PAGE_BREAK
~k		Bedingter Zeilenumbruch	200B	DISCRETIONARY_LINE_BREAK

## Spezialzeichen

GREP	Zeichen	Beschreibung	Unicode	SpecialCharacters.
~h	\	Verschachteltes Format hier beenden	0003	END_NESTED_STYLE
~F		Marker für Fußnote	0004	
~i	†	Einzug bis hierhin	0007	INDENT_HERE_TAB
~y	⏏	Nach rechts austreibender Tabulator	0008	RIGHT_INDENT_TAB
		Marker für Tabelle	0016	
~v		Textvariable	0018	TEXT_VARIABLE
~#		Alle Seitenzahlen		
~N		Automatische Seitenzahl	0018	AUTO_PAGE_NUMBER
~X		Nächste Seitenzahl	0018	NEXT_PAGE_NUMBER
~V		Vorherige Seitenzahl	0018	PREVIOUS_PAGE_NUMBER
~x		Abschnittsmarke	0019	SECTION_MARKER
~-	-	Bedingter Trennstrich	00AD	DISCRETIONARY_HYPHEN
~k		Bedingter Zeilenumbruch	200B	DISCRETIONARY_LINE_BREAK
~j	~	Verbindung unterdrücken	200C	ZERO_WIDTH_NONJOINER
	‡	Verbindung erzwingen	200D	ZERO_WIDTH_JOINER (ab CS5)
~a	⌘	Marker für verankertes Objekt	FFFC	
~l	⏏	Marker für Indexeintrag	FEFF	

## Variablen

GREP	Zeichen	Beschreibung	Unicode	SpecialCharacters.
~v		Alle Variablen	0018	TEXT_VARIABLE
~Y		Lebender Kolumnentitel (Absatzformat)	0018	
~Z		Lebender Kolumnentitel (Zeichenformat)	0018	
~u		Benutzerdefinierter Text	0018	
~T		Letzte Seitenzahl	0018	
~H		Kapitelnummer	0018	
~S		Erstellungsdatum	0018	
~o		Änderungsdatum	0018	
~D		Ausgabedatum	0018	
~l		Dateiname	0018	
~J		Metadatenbeschriftung	0018	

## Anführungszeichen

GREP	Zeichen	Beschreibung	Unicode	SpecialCharacters.
"		Alle doppelten Anführungszeichen		
'		Alle einfachen Anführungszeichen		
~"	"	Gerade Anführungen, Zollzeichen	0022	DOUBLE_STRAIGHT_QUOTE
~{		Öffnende doppelte Anführung (sprachabhängig)		DOUBLE_LEFT_QUOTE
~}		Schließende doppelte Anführung (sprachabhängig)		DOUBLE_RIGHT_QUOTE
~'	'	Gerades einfaches Anführungszeichen	0027	SINGLE_STRAIGHT_QUOTE
~[		Öffnende einfache Anführung (sprachabhängig)		SINGLE_LEFT_QUOTE
~]		Schließende einfache Anführung (sprachabhängig)		SINGLE_RIGHT_QUOTE

## Symbole

GREP	Zeichen	Beschreibung	Unicode	SpecialCharacters.
~_	—	Geviertstrich	2014	EM_DASH
~=	-	Gedankenstrich, Halbgeviertstrich	2013	EN_DASH
\-	-	Bindestrich	002D	
~-	-	Bedingter Trennstrich	00AD	DISCRETIONARY_HYPHEN
~~	-	Geschützter Trennstrich	2011	NONBREAKING_HYPHEN
~8	•	Aufzählungszeichen	2022	BULLET_CHARACTER
~2	©	Copyright-Symbol	00A9	COPYRIGHT_SYMBOL
~e	...	Ellipse, Auslassungszeichen	2026	ELLIPSIS_CHARACTER
~7	¶	Absatzmarke	00B6	PARAGRAPH_SYMBOL
~r	®	Eingetragene Marke	00AE	REGISTERED_TRADEMARK
~d	™	Trademark-Symbole	2122	TRADEMARK_SYMBOL
~6	§	Paragrafenzeichen	00A7	SECTION_SYMBOL
\\	\	Backslash, umgekehrter Schrägstrich	005C	
\^	^	Caret-Zeichen, Zirkumflex	005E	
°	°	Gradzeichen	00B0	DEGREE_SYMBOL

## A 1 Ressourcen und Literatur

Die folgende Übersicht enthält Internetquellen, Ressourcen und Bücher, die ich für die Automatisierung von InDesign hilfreich finde. Naturgemäß ist diese Liste nicht vollständig. Unter [79](#) finden Sie eine stets aktuelle Version dieser Liste.

### Reguläre Ausdrücke in InDesign

Neben den in Unterkapitel vorgestellten Ressourcen kann ich noch die folgenden Quellen und Bücher empfehlen.

Gregor Fellenz  
InDesign automatisieren –  
Keine Angst vor Skripting GREP & Co  
<http://www.indd-skript.de>  
dpunkt-verlag, 2011

Peter Kahrel  
GREP in InDesign CS3/CS4  
O'Reilly, 2008

Jochen Uebel  
Unter [81](#) finden Sie regelmäßige Blog-Beiträge zum Thema GREP (und InDesign).

Laurent Tournier  
Der Blog [82](#) ist französischsprachig. Hier finden Sie interessante GREP-Techniken.

InDesigns GREP-Funktion basiert auf der Boost-Library. Die Referenz finden Sie unter: [83](#)

### Reguläre Ausdrücke

Jeffrey E. F. Friedl  
Reguläre Ausdrücke  
O'Reilly, 2007, 3. Auflage [84](#)

Jan Goyvaerts & Steven Levithan  
Regular Expressions Cookbook  
O'Reilly, 2009 [85](#)

Auf Jan Goyvaerts Seite [86](#) finden Sie neben einer Einführung auch sehr anspruchsvolle Informationen und Diskussionen zum Thema Reguläre Ausdrücke.