

# Machen Sie Ihren Doktor-Title! Neue Serie: SEO für Einsteiger – Teil 1



## Machen Sie Ihren Doktor-Title! Neue Serie: SEO für Einsteiger – Teil 1 – [websiteboosting.com](http://websiteboosting.com)

Seit Anbeginn der Suchmaschinenoptimierung und sogar noch lange, bevor es diese Bezeichnung überhaupt gab, war der Title ein sehr wichtiges Element für Suchmaschinen. Es darf durchaus erstaunen, dass heute, mehr als 20 Jahre später, noch immer viele Websitebetreiber den Text des Titles eines...

Seit Anbeginn der Suchmaschinenoptimierung und sogar noch lange, bevor es diese Bezeichnung überhaupt gab, war der Title ein sehr wichtiges Element für Suchmaschinen. Es darf durchaus erstaunen, dass heute, mehr als 20 Jahre später, noch immer viele Websitebetreiber den Text des Titles eines Dokuments dem Zufall, mehr oder weniger untalentierten und/oder ungeschulten Editoren überlassen oder des Jammers nicht müde werden, dass man doch nicht für jede Seite einen individuellen Title verfassen könne. Dabei sind die wenigen Wörter des Titles auch heute noch extrem relevant für das Ranking eines Dokuments! Und wer glaubt, bereits genügend über den Title zu wissen, den laden wir herzlich ein, dieses Wissen kritisch zu überprüfen und ggf. zumindest nachdenklich aufzustoßen. Richtig gute SEOs kokettieren nicht selten damit, dass das Texten eines Titles „Chefsache“ sei. Nach dem Lesen dieses Beitrags werden Sie hoffentlich mit ihnen übereinstimmen, dass dies sooo weit hergeholt nicht ist. Aufgrund vieler Anfragen und Anregungen unserer Leser starten wir erneut eine kleine Serie mit SEO-Basics für Einsteiger und motiviert Fortgeschrittene. Im ersten Teil in dieser Ausgabe finden Sie alles Wichtige und

**viele Tipps zu einem der wichtigsten Bestandteile der Suchmaschinenoptimierung, dem Title.**

Eines gleich vorweg: Auch der beste Title bringt Ihnen keine Toprankings, sofern die Gesamtkomposition an Signalen in Richtung Google nicht stimmig ist. Umgekehrt ist es extrem schwer, ein Seite-1-Ranking zu bekommen, wenn man im Title unpassenden oder nicht stimmigen Text verwendet.

## **Wow – Sie treffen heute den Menschen Ihres Lebens!**

Stellen Sie sich doch bitte einmal vor, Sie säßen in einer Bar und entdeckten den Menschen (Frau, Mann, Divers) Ihres Lebens alleine an einem Tisch sitzend. Am liebsten würden Sie hingehen und ihn ansprechen. Das geht aber nicht, denn die Bar hat strenge Regeln. Gäste an Esstischen darf man nicht einfach so ansprechen bzw. belästigen. Und jetzt steht dieser Mensch auf und geht zur Toilette. Dabei muss er an Ihnen vorbei, direkt an der Bar. Sie sitzen zusammen mit neun weiteren Einzelpersonen an dieser Barzeile. Wenn der Mensch Ihres Lebens mit Ihnen Augenkontakt aufnimmt beim Vorbeigehen, haben Sie etwa zwei bis drei Sekunden Zeit, etwas zu sagen. Das ist Raum für sechs bis acht Worte. Natürlich überlegen Sie sich vorher ganz genau, was Sie sagen werden – oder etwa nicht? Sie würden kein Wort dem Zufall überlassen – oder etwa nicht? Und wenn Sie *\*wirklich\** clever sind, testen Sie sogar, welche Worte die nötige Aufmerksamkeit erzeugen würden, damit dieser Mensch bei Ihnen stehen bleibt, lächelt und anfängt, mit Ihnen zu sprechen, Ihnen eine erste Chance auf „mehr“ gibt. Würden Sie sagen: „Na, auch hier?“ Oder: „Sektchen“? Würden Sie sagen: „Toilette – Toilette günstig besuchen, Toilette finden Sie hier?“ Gerade der letzte Satz kommt Ihnen völlig verblödet vor? Aber Moment – dieser Mensch wollte doch die Toilette aufsuchen. Was spräche dagegen, ihm genau diese Worte mehrmals um die Ohren zu werfen? O. k., eine ernste Antwort auf diese

Frage können wir uns alle zusammen sparen.

Überträgt man die fiktiv beschriebene Situation auf die Kontakthanbahnung von Unternehmen mit potenziellen Interessenten bzw. Kunden, wird schnell klar, warum dieses Beispiel gewählt wurde. Jedem ist klar, dass Standardsprüche in Bars nur bei Menschen funktionieren würden, bei denen man das gar nicht wirklich will. Man muss sich Mühe geben, damit es bei denen klappt, mit denen man in Kontakt kommen möchte.



Abbildung 1: Bei vielen Unternehmen werden Titleinhalte offenbar extrem uninspiriert getextet. Schema F. Ich bin etwas Besonderes – daher texte ich ganz genauso wie alle anderen? Aber: Klick gefälligst mich?

Aber warum nur dringt diese einfache Erkenntnis nicht in die verantwortlichen Online-Abteilungen durch? Dieselben Menschen, die sich in einer Bar einen wirklich interessant wirkenden Spruch einfallen lassen würden, texten auf ihren Webseiten im Titel so Uninspirierendes wie: „Sandalen online kaufen/Domainname“? Ist es falsch verstandenes SEO? Fehlende Zeit? Oder die klassische Abwehrformel von Shops: „Wir haben [beliebige Zahl über 100 einsetzen] Artikel und können doch nicht für jeden einen eigenen Title texten!“?

*„Sieht ein Suchender Ihren Title in einem Suchergebnis, ist das der ‚Moment of Truth‘ für Sie!“*

Was auch immer der Grund für falsch und/oder langweilige getextete Title ist: Sie sollten sich klarmachen, dass diese eine Zeile noch immer nicht nur wichtig für das Ranking ist, sondern auch für die Klickentscheidung von Suchenden. Mit einem guten Text bekommen Sie bei gleichem Ranking mehr Klicks und damit am Ende mehr Umsatz. Die Anzeige von Title und Description (siehe nächste Ausgabe) ist der „Moment of Truth“, also der Moment der Wahrheit. Zeigt Google sie zusammen mit neun anderen unbezahlten Treffern im sog. organischen Bereich in einem Suchergebnis an, bekommen sie ein bis zwei Sekunden

Augenüberflug des Suchenden und er oder sie entscheidet sich spontan, ob es seine/ihre Zeit wert ist, Ihre Seite zu erkunden. Auf diese Einschätzung zählt die Formulierung ganz entscheidend ein. Standard-Textgedöns hält sicherlich viele Suchende ab, wie sie in der Search-Console von Google als Diskrepanz zwischen guter Position, Impression (also wie oft in der Suche angezeigt) und Klicks (wie viele haben tatsächlich auf Ihren Title geklickt) leicht ersehen können.

## Von vorne: Was ist der „Title“ überhaupt?

Laut Konvention ist `<title>` ein Element eines in HTML erzeugten Dokuments, also einer Webseite. Er zählt damit nicht zu den Meta-Tags wie z. B. der Description und dem völlig nutzlosen und trotzdem ständig befüllten „Keyword“-Meta-Tag.

Man findet den Title oben im `<head>` des Quelltexts (rechte Maustaste, „Quelltext anzeigen“ – oder ähnlich, je nach Browser). Ist er dort nicht sofort ersichtlich, drücken Sie STRG und die Taste f (Eselsbrücke: f wie finden). Jetzt ploppt ein kleines Suchfenster auf, in das man den Suchtext eintippt, also hier dann „title“ (ohne Anführungszeichen). So lässt sich bei umfangreichen Quelltexten gezielt finden, was man sucht.



Abbildung 2: Die Aussage in den ersten Dokumenten von Google zu SEO sind auch heute noch korrekt

Ein Beispiel: So sieht der Title im Quelltext einer HTML-Seite technisch korrekt aus:

```
<title>websiteboosting.com</title>
```

Im US-Patent Nr. 20070022110A1 von 2004 kann man nachlesen, dass die Wörter im Title gut als Rankingfaktor verwendet werden (können). Das Patent stammt allerdings nicht von Google, auch wenn es immer wieder im Zusammenhang mit den Rankingfaktoren von Google genannt wird.



Abbildung 3: Streng genommen falsch: Der Title ist ein fester Bestandteil einer HTML-Seite und eben kein Meta-Tag

Man kann sich den Title gut als den Namen eines Dokuments vorstellen oder auch als den kurzen, aber sprechenden Titel eines Buches, der auf dem Buchrücken steht. Sieht man ihn in der Buchhandlung im Regal zusammen mit anderen Büchern, entscheidet er in der Regel, ob wir das Buch in die Hand nehmen und einen Blick hineinwerfen. Dieses Herausnehmen entspricht dem Klick auf Ihr Suchergebnis. Sind die Inhalte nicht gut, wird das Buch wieder zurückgestellt oder der Zurück-Button des Browsers geklickt. Der Title lockt also nur an, überzeugen muss dann aber eben auch der Content. Aber wenn schon der Title nicht lockt – wie soll dann der beste je gesehene Content verkauft werden können? Richtig, nie bzw. viel zu selten.

## **Für wen ist der Title und wo sieht man ihn überhaupt?**

Auch wenn ein Seitenbesucher den Title also solches meist gar nicht bemerkt, weil er nur im Tab des Browsers erscheint oder als Textvorschlag für einen Favoriteneintrag bzw. ein Bookmark, schreiben Sie ihn trotzdem ausschließlich für Menschen. Oft ist man der irrigen Meinung, Titles müssten für Suchmaschinen und vor allem für Google geschrieben werden. Dies lässt völlig außer Acht, dass Google ausschließlich die Nutzer im Fokus hat und Dinge so gewichtet, wie diese Nutzer das präferieren. Es würde keinen Sinn machen, dies nach anderen Regeln zu werten. Ein Suchergebnis ganz oben mit einem schlecht lesbaren oder irrelevanten Title würde ja gar nicht geklickt werden und dort nur Platz beanspruchen. Verwerfen Sie daher am besten den Gedanken, für eine Maschine texten zu müssen, gleich wieder. Dieser entspringt einem falschen Verständnis, wie Google seine Kriterien fixiert. Nämlich (fast) nie nach eigenen Überlegungen, wie man es gerne hätte,

sondern eben streng nach dem, was Google groß gemacht hat und groß hält: Was will der User? Betrachten Sie Google immer nur als eine Brücke zum Besucher, die hilft, den Kontakt herzustellen.

Den Title-Eintrag einer Webseite kann man wie erwähnt im Quelltext einsehen (Abbildung 4) und im Tab-Feld des Browsers (Abbildung 6). Die wichtigste Stelle ist allerdings sicherlich die Darstellung in Suchergebnissen. In der Regel zeigt Google hier bis auf sehr seltene Ausnahmen den Title-Eintrag als Überschrift an (Abbildung 5). Ab und zu werden auch die Ankertexte eingehender (Back-)Links zu einer Seite als Title verwendet. Das passiert immer dann, wenn Google aus irgendeinem Grund dem vom Websitebetreiber hinterlegten Title misstraut oder z. B. jede Seite den gleichen Title trägt und dieser somit nicht aussagekräftig sein kann. Wird eine Seite als Link in Social-Media-Netzwerken geteilt, wird der Title als Überschrift verwendet, sofern für ein Netzwerk nichts speziell anderes hinterlegt wurde. Dies variiert jedoch von Plattform zu Plattform.



Abbildung 4: Den Title-Eintrag findet man im Quelltext oben im <head> einer Webseite



Abbildung 5: Im Suchergebnis wird der Title als blaue Überschrift angezeigt – in der Regel...



Abbildung 6. Der Title erscheint für einen Nutzer nur im Browser in der Fensterleiste oben bzw. in einem Tab als Bezeichnung

## **Sie wissen nicht, wo Sie den Title ändern/erzeugen können?**

Leider ist bei Content-Management-Systemen nicht immer klar, welcher Eintrag im Backend dann tatsächlich in das Title-Element geschrieben wird bzw. welches Formularfeld das

wirklich korrekte dafür ist. Dazu kommt erschwerend, dass bei Unternehmen oft nicht genügend Know-how vorhanden ist, um mit den Editor-Werkzeugen richtig umzugehen, die ihnen eine Webagentur im übertragenen Sinn auf den Tisch geworfen hat. Hier kann der eindeutige Rat nur lauten: Erkundigen Sie sich oder finden Sie selbst heraus, welcher Feldeintrag den Title „erzeugt“. Dazu sehen Sie einfach wie oben beschrieben im Quelltext nach und suchen dann dieses Textmuster in Ihrem Editor. Gibt es mehrere Felder mit gleichem Inhalt, variieren Sie diese einfach und publizieren diese eine Seite neu. Rufen Sie diese dann im Browser auf und sehen sich im Quelltext an, wo genau die eben formulierten Texte „hingewandert“ sind. Identifizieren Sie den Text im <title>-Eintrag und ab jetzt wissen Sie genau, welches Editorfeld zum Title gehört.

Ein Beispiel. Der Quelltext zeigt z. B. zweimal den Eintrag „websiteboosting.com“ (Ausschnitt):

```
<meta name="generator" content="TYP03 CMS">
<meta property="og:title" content="websiteboosting.com">
<title>websiteboosting.com</title>
<meta name="robots" content="index,follow">
```

Dann ändern Sie für diese Seite die generell infrage kommenden Felder um z. B. in „Hier ist Feld X“ und in das andere Feld „Hier ist Feld Y“. Nach der Publikation/Aktualisierung der Seite finden Sie im Quelltext dann:

```
<meta name="generator" content="TYP03 CMS">
<meta property="og:title" content="Hier ist Feld X">
<title>Hier ist Feld Y</title>
<meta name="robots" content="index,follow">
```

Sie wissen jetzt, dass dort, wo Sie in Ihrem Editor „Hier ist Feld Y“ eingetragen haben, der eigentliche Title-Eintrag im Quelltext generiert wird. Voilà.

# Wie gestaltet man einen guten Title?

Was empfiehlt eigentlich Google für die Gestaltung des Titles? Das ist nachzulesen unter <http://einfach.st/gtip>. Dort heißt es, man solle aussagekräftige Title verwenden, die den Inhalt der jeweiligen Seite kurz und knapp möglichst gut beschreiben. Irrelevante Wörter und Inhalte sollte man auf jedem Fall vermeiden. Alles, was CMS oder Shopsysteme oft automatisch einfüllen wie „Homepage“, „Seite 1“ oder „unbenannt“, sollte man auf jeden Fall mit sinnhaftem Text überschreiben. Auch das vor 20 Jahren übliche „Herzlich willkommen auf unserer Homepage“ wirkt bei vielen Suchenden wohl eher altmodisch und von vorgestern.



Abbildung 7: Welchen Eindruck erzeugen wohl solche Title beim Suchenden?



Abbildung 8: Leider zeigt Google ab und an auch völlig unpassende Ergebnisse im Ranking – hier sind die Title für die – dann unterbleibenden – Klicks extrem hilfreich



Abbildung 9: Die Texte auf Buchrücken wurden meist wohlüberlegt und Wort für Wort ausgewählt – so sollten auch Title gestaltet werden

*„Es gibt Tests, die uns gezeigt haben, dass die Menschen geschriebene Titles bevorzugen.“ Gary Illyes, Google*

Ebenso wichtig ist es laut Google, dass jede Seite einen individuellen Titel bekommt. Die Verwendung gleicher Inhalte für mehrere oder gar alle Seiten einer Domain ist also unbedingt zu vermeiden.

Ein Beispiel von unserer eigenen Website:

<title>websiteboosting.com</title>

Finden Sie diesen Titel gut gestaltet für eine Startseite einer Domain? Nein, sicher nicht. Der wiederholt nur den Domainnamen und enthält daher keinerlei relevante Information, wenn er in den Suchergebnissen auftaucht. Noch nicht einmal „Startseite“ wurde als zusätzliche Information für Klickende hinterlegt. Wir leiden unter den gleichen Problemen wie so viele Unternehmen. Das wird halt einfach „erledigt“ und am besten ist es, wenn das CMS da gleich irgendwas automatisch reinschreibt, dann muss man sich erst gar nicht damit beschäftigen.

Wenn Sie an das Bar-Beispiel zu Anfang dieses Beitrags denken, wird noch mal klar, dass man sich tatsächlich überlegen muss: „Warum ich?“ Warum soll jemand unter vielen auf dieses Suchergebnis klicken. Natürlich kommen auch mit diesem Title viele Klicks – und zwar genau von denen, die das Magazin oder das frühere Buch „Website Boosting“ schon kennen. Aber jemand, der Know-how im Online-Bereich sucht? Dem „Website Boosting“ noch rein gar nicht sagt?

- Know-how satt für Online – Website-Boosting-Magazin?
- SEO, SEA, Web-Usability – Wissen für Ihren Erfolg?
- Die Zeitschrift für Ihren Weberfolg – Website Boosting?
- Seit über 10 Jahren feinstes Online-Know-how – guckst du?
- ...

Im vorliegenden Fall müsste man noch nicht einmal überlegen oder gar testen. Alle Vorschläge wären sicher besser, als nur „websiteboosting.com“ auszuspielen. Natürlich muss bzw. sollte der Title SEO-relevante Wörter enthalten, das darf man nicht aus den Augen verlieren. Bei der Startseite ist allerdings in der Regel die Marke bzw. der Domainname selbst so stark, dass man mehr Freiheiten hat.

DEN richtigen, wahren Title gibt es in der Regel leider nicht. Bedenken Sie, dass eine URL natürlich auch für

unterschiedliche Suchbegriffe und vor allem ggf. auch für verschiedene Suchabsichten ranken kann. Möchte jemand kaufen (Sandalen kaufen), wirkt ein Text für eine eher informationsgesteuerte Suche (welche Sandalen trägt man diesen Sommer) nicht optimal – und umgekehrt. Insofern muss man beim Texten oft Kompromisse eingehen, sofern man bisher nicht wirklich eine gute und vorausschauende Planung/Strukturierung beim Erstellen von Webseiten hatte. Insofern gibt es auch nicht DEN einen, alles einschließenden Ratschlag oder Tipp, der für alle gleich gilt.

## Ihre Checkliste für die Gestaltung eines Title.

- Texten Sie immer **aussagekräftig und prägnant**.
- Vermeiden Sie **überflüssige Keywords und leere Füllwörter**.
- Vermeiden Sie **wiederholte Title, Title aus Textbausteinen oder eine automatisierte Erstellung**.
- Jede einzelne Seite sollte einen **einzigartigen Title** haben (Uniqueness)
- Länge: Bis max. 60 Zeichen (das sind 512 Pixel). Google schneidet zu lange Titles in den Suchergebnissen je nach Gerät ab, **Wichtiges** sollte daher immer ganz **vorne** stehen.
- Tipp: Das „**Abschneiden**“ kann natürlich an der richtigen Stelle auch **Neugierde** erzeugen.
- Großschreibung für ganze Wörter oder gar den gesamten Title ist schlechte Etikette, schwerer zu lesen bzw. zu erkennen und wirkt meist nach marktschreierischem „Brüllen“.
- Verwenden Sie wohlüberlegt Attribute je nach Produkt oder angebotener Dienstleistung wie z. B. „vollständig“, „umfassend“ oder „verständlich“. Auch die üblichen Aufzählungen a là „10 Tipps“ oder „Mit 2 Klicks zum ...“

bringen mehr Klicks.

- Bei aktuellen Beiträgen ggf. den Monat und das Jahr im Title hinterlegen.
- Wenn die Marke *\*wirklich\** bekannt ist: Unbedingt am Anfang nennen, um das Markenvertrauen in Klicks umzuwandeln.
- Vergleichen Sie die anderen Suchergebnisse, die bereits ranken. Seien Sie anders. Verwenden alle anderen bestimmte Begriffe wie z. B. „ultimativ“, verwenden Sie diese keinesfalls und suchen Sie nach besseren Alternativen.
- In der Regel funktionieren Emotionen gut, v. a. Ärger (vermeiden), Angst oder Neugierde.
- Bauen Sie sich ein eigenes Schema auf wie z. B. Keyword/Vorteil/Marke/Kategorie.

## Wo kann ich das Aussehen meiner Title simulieren bzw. testen?

Im Web gibt es eine Menge „Simulatoren“, die für eine URL oder einen Text zeigen, wie er in den Suchergebnissen dargestellt werden würde, wie z. B.:

- [app.sistrix.com/de/serp-snippet-generator](https://app.sistrix.com/de/serp-snippet-generator)
- [searchwilderness.com/tools/pixel-length](https://searchwilderness.com/tools/pixel-length)
- [serpsim.com](https://serpsim.com)



Abbildung 10: Der SERP Snippet Generator von SISTRIX kann auch einfach „nur“ prüfen und gibt Hinweise (Quelle: Sistrix.com)



Abbildung 11: SERPsim.com zur Prüfung, wie der Title mobil oder auf dem Desktop (1) wirkt

Aber auch in vielen SEO-Tools wie z. B. in Ryte.com, ahrefs, sitebulb, Screaming Frog und vielen anderen kann man sich

ansehen, wie eine einzelne URL im Suchergebnis mit Title und Description dargestellt würde. Als Hinweis bleibt zu erwähnen, dass fast alle dieser Tools ein klein wenig anderes rechnen, was die Länge bzw. den angezeigten Text angeht. Was die Länge angeht, sollten Sie also nur Anpassungen vornehmen, wenn Ihre Textlänge deutlich aus dem Rahmen fällt. Ist Ihr Title zu lang, prüfen Sie, ob wichtige Wörter möglichst weit vorne stehen und dann auch auf möglichst vielen Geräten gesehen werden. Ein paar Wörter „hinten“ schaden nicht. Und ein zu kleinliches Einkürzen kann Sie spätestens dann wieder einholen, wenn mehr und mehr über Sprachsuchen Ergebnisse von Maschinen vorgelesen werden. Dann ist ein klein wenig mehr Info sicherlich durchaus hilfreich und Sie müssen dann nicht ergänzen, was sie jetzt gelöscht haben. Ist der Title zu kurz, stellt das oft eine verpasste Chance dar, noch mehr zu überzeugen und auch einige wichtige Wörter (für die Rankingalgorithmen) unterzubringen.

## **Was tue ich, wenn ich zu viele Seiten habe?**

Immer dann, wenn etwas in zu großer Zahl auftritt, um sich mit allem einzeln im Detail zu beschäftigen, hilft das sog. Pareto-Prinzip. Es besagt vereinfacht ausgedrückt, dass man mit 20 % Aufwand 80 % des Erfolges steuert. Diese Faustformel trifft erstaunlich oft tatsächlich die Realität. Nicht selten sind bei Domains sogar nur 2 % der Seiten für mehr als 80 % der Besucher über Rankings verantwortlich.

### **Tipp 1: Filtern Sie!**

Filtern Sie sich einfach über Ihr Tracking-Tool wie z. B. Google Analytics diese Seiten mit viel organischem Traffic von Google heraus und kontrollieren Sie die Titleeinträge auf Klick-Affinität. Stellen Sie sich die ehrliche Frage: Würde ich bei Suchen auf diesen Text klicken, wenn es nicht mein

eigenes Unternehmen wäre? Macht der Text neugierig? Enthält er einen echten Klickanreiz? Sehen Sie sich die gesamten Suchergebnisse incl. Ihres Textes für eine entsprechende Suche bei Google an. Stechen Sie heraus oder gehen Sie im Textbrei der anderen mit unter? Das Wichtigste: Seien Sie ehrlich zu sich selbst!

## **Tipp 2: Holen Sie harte Daten!**

Wenn Sie Titles wegen Tipp 1 ändern: Sehen Sie sich vor der Änderung die Klickmetriken in der Google Search Console an. Dort können Sie für jede URL und jedes Keyword, für das Sie ranken, für einen Zeitraum von 16 Monaten sehen, ob und was sich verändert. Notieren Sie sich, für welche (wichtigen) Keywords die URLs ranken, bei denen Sie die Title optimiert haben. Nach den Änderungen sehen Sie nach einiger Zeit in den Suchergebnissen nach, ob Ihr neuer Title dort schon angezeigt wird. Wenn ja, warten Sie noch einige Wochen, damit aussagekräftige Zahlen zustande kommen, und kontrollieren Sie in der Search Console, ob sich die Klickrate verbessert hat. Behalten Sie aber im Auge, dass eine Veränderung der durchschnittlichen Position auch immer eine Änderung in der Klickrate zur Folge hat, weil Ergebnisse weiter unten oder weiter oben natürlich im Verhältnis fast automatisch mehr Klicks bekommen. Ist das Ranking aber im Schnitt in etwa gleich geblieben, können Sie ganz gut erkennen, was Ihre Änderung bewirkt hat.

## **Tipp 3: Lernen Sie daraus!**

Stellen Sie bei der Befolgung von Tipp 2 fest, dass die Klickrate und damit die Anzahl der Besucher gestiegen ist, lernen Sie aus den eigenen Optimierungen! Was hat sehr gut funktioniert? Was gut, aber nicht so stark? Können Sie geänderte Textelemente identifizieren, die zu deutlich mehr Klicks geführt haben? Dann weiten Sie diese auch auf andere URLs aus. Damit gehen Sie in jedem Fall datengesteuert vor und

nicht aus dem Bauch heraus – mit Letzterem liegt man in der Online-Branche leider oft komplett daneben.

## **Tipp 4: Kalkulieren und überzeugen Sie!**

Versuchen Sie, Ihre Daten betriebswirtschaftlich zu verwenden. Ermitteln Sie das Delta der Mehrbesuche auf Ihrer Website. Wenn Sie Ihre Conversion-Rate kennen, also das Verhältnis von Besuchern zu Käufern (oder was Ihren Zielkriterien entspricht), können Sie diesen zusätzlichen Umsatzwert zumindest rechnerisch grob beziffern. Ein einfaches Beispiel: Sie haben eine Conversion-Rate von 5 %. Das heißt, von 100 Besuchern kaufen fünf etwas ein (oder schicken ein Kontaktformular ab, rufen Sie an, abonnieren den Newsletter etc. – je nach Ziel). Der durchschnittliche Warenkorb liegt bei 200 €. Also haben Sie monatlich bei 100 Besuchern 1.000 € Umsatz. Die Klickrate auf Ihr Suchergebnis lag vor der Optimierung laut Search Console bei 10 %, nach der Optimierung bei 20 %. Bei gleichen Rankings haben Sie also Ihre Besucherzahl von Google verdoppelt. Lassen Sie uns annehmen, dass es vorher 1.000 Besucher pro Monat und damit 10.000 € Umsatz waren. Jetzt sind es 2.000 Besucher und 20.000 € Umsatz.

Geht man von einer Gewinnmarge von 20 % aus, haben Sie mit den Änderungen 2.000 € mehr Gewinn gemacht (20 % von 10.000 € = 2.000 € und von 20.000 € wären das 4.000 €, also 2.000 € mehr).

Jetzt haben Sie etwas zum Kalkulieren in der Hand. Pro Monat 2.000 € sind 24.000 € mehr Gewinn pro Jahr. Stellen Sie diesem Zusatzgewinn (oder dem Zusatzumsatz) die Kosten der Änderungen gegenüber. Liegen diese deutlich darunter – und das tun sie in den allermeisten Fällen –, stellt sich die Frage nach der Sinnhaftigkeit weiterer Optimierungen nicht mehr. Sie ist nachgewiesen. Überlegen Sie, wer Optimierungen an weiteren Seiten vornehmen kann, den Gewinn-/Umsatzhebel kennen Sie jetzt und können auch kritischen Vorgesetzten (bringt das

wirklich was?) gegenüber richtig und mit Geld argumentieren. Wichtig: Sie kommen jetzt nicht mehr als Bittsteller nach Budget daher, die immer nur Geld wollen für undefinierte Tätigkeiten. Sie können vorrechnen, was ein investierter Euro pro Jahr an Return-on-Investment bringt. Der Chef, der hier noch Nein sagt, ist aller Voraussicht nach falsch auf seinem Posten...

## **Tipp 5: Rankingplus – für Experten!**

Die Erklärung der genauen Umsetzung würde hier den Rahmen leider sprengen, aber wer einigermaßen gut mit Tools umgehen kann, für den mag die Idee hinter dem folgenden Tipp sicherlich recht hilfreich sein. Besorgen Sie sich eine Liste mit Ihren Rankings nach Keyword und URL aus dem SEO-Tool Ihrer Wahl. Bei SISTRIX finden Sie diese z. B. in der Unterrubrik „Keywords“. Rechts oben ist dann der „Export“-Button für den Download einer Excel-Liste. Filtern/sortieren Sie die Liste so nach Rankingpositionen, dass die guten Rankings (z. B. Position 1–5) nicht angezeigt werden. Diese brauchen Sie ja nicht optimieren. Jetzt müssten Sie für jede URL den Title holen. Das geht sehr geschmeidig z. B. mit SEO-Tools für Excel. Die Formel `=@HtmlTitle(B1)` holt den Title einer URL, die in der Zelle B1 steht. Wenn Sie diese Formel nach unten kopieren, steht für jede URL der Titledtext in einer Spalte. Anschließend prüfen Sie, z. B. mit der Funktion `=SUCHEN(A1;B1)`, ob das rankende Keyword in Zelle A1 im Text des Titles in Zelle B1 enthalten ist. Die Funktion `=SUCHEN` hat gegenüber `=FINDEN` den Vorteil, dass Groß-/Kleinschreibung ignoriert wird. Trotzdem muss man noch auf Leerzeichen und Bindestriche achten, aber mit ein paar Excel-Kenntnissen bekommt man das auch recht einfach in den Griff. Wer SEO-Tools für Excel nicht zur Verfügung hat, kann sich den Title auch mit anderen Tools wie z. B. dem Screaming Frog holen und via `=SVERWEIS` die erzeugte Datei und die aus dem Rankingtool matchen. Das Ergebnis ist eine Liste, die zeigt, welche Rankings schon gut, aber noch nicht top sind und bei denen das

Keyword eben (noch) nicht im Title vorhanden ist. Ergänzt man das Keyword dort im Title, kann man in der Regel einen kleinen Sprung in den Rankings nach oben erzeugen. Gerade bei Ergebnissen auf Position 11 bis 15 kann ein Sprung auf Seite eins durchaus „plötzlich“ Traffic bringen, der bekanntlich von Seite zwei fast nie zu verzeichnen ist. Eine einfache Möglichkeit also, „low hanging fruits“ zu ernten.

## Tipp

Die Inhalte dieses Beitrags und weitergehende Infos finden sie auch bei Website Boosting TV unter <http://einfach.st/wsbtv5> erklärt. Abonnieren Sie uns doch einfach kostenlos auf YouTube – dann verpassen Sie ab jetzt keinen Videobeitrag mehr!

## Fazit

Sie haben eine genaue Anweisung erwartet, welches Wort Sie wie und wo ändern sollten? Das ist aus verschiedenen Gründen nicht sinnvoll, allein schon deswegen, weil Sie ja „anders“ als die anderen sein sollten, hervorstechen sollen. Je mehr Sie nach Schema F vorgehen und je weniger Sie ihre eigene Kreativität für die Spezifika Ihrer Branche und Angebote nutzen, desto mehr entfernen Sie sich von einem guten Ziel. Denken Sie für Ihre wichtigsten Seiten an die Bar. Der Suchende geht vorbei, fixiert alle zehn nacheinander und Sie haben nur diese eine, kurze Chance, die nötige Aufmerksamkeit zu bekommen. Wenn Sie das verbummeln, bekommen Sie definitiv keine weitere Möglichkeit zu zeigen, dass Sie gut sind. Behalten Sie daher diesen wichtigen Moment of Truth immer im Kopf!

---

# Wie Sie verkaufen, auch wenn Sie teurer sind



## Wie Sie verkaufen, auch wenn Sie teurer sind – [websiteboosting.com](https://www.websiteboosting.com)

Zugegeben, aktuell geht es der Digitalwirtschaft sehr gut. Zum einen wissen viele Kunden um die Wichtigkeit ihres digitalen Auftritts Bescheid, zum anderen unterstützt der Staat zusätzlich mit zahlreichen Förderprogrammen Investitionen – und damit auch so manche Agentur. Doch egal, wie die...

**Zugegeben, aktuell geht es der Digitalwirtschaft sehr gut. Zum einen wissen viele Kunden um die Wichtigkeit ihres digitalen Auftritts Bescheid, zum anderen unterstützt der Staat zusätzlich mit zahlreichen Förderprogrammen Investitionen – und damit auch so manche Agentur. Doch egal, wie die aktuellen Konstellationen sind: Es wird immer Agenturen geben, die teurer als andere sind. Vielleicht gehören auch Sie zu den Höherpreisigen. Dann ist die entscheidende Frage: Wie erfolgreich schließen Sie Aufträge ab, wenn Sie nicht der billigste Anbieter sind? Oder knicken Sie gar am Ende der Verhandlung ein und geben zahlreiche Leistungen gratis, weil Sie nicht in der Lage sind, Ihren Preis auf Augenhöhe zu erklären?**

So manche Gründer orientieren sich bei der Preisfindung an ihren Mitbewerbern: Was nehmen diese pro Stunde? Wie sehen deren Angebote aus? Im Ergebnis ist dann deren Angebot für Außenstehende oft austauschbar. Denn wenn Webseiten, Imagebroschüren und auch die Mitarbeiter alle die gleichen Phrasen dreschen – woran soll sich dann ein Kunde orientieren? Erst recht dann, wenn ein potenzieller Abnehmer wenig Ahnung von der Materie hat und zwangsläufig einen objektiven Vergleich zwischen beispielsweise drei Agenturen nicht machen

kann, wird sich dieser oft für den günstigsten entscheiden müssen. Denn wie soll ein Kunde seinen eigenen Kollegen erklären, dass er nicht das billigste Angebot gekauft hat, sondern das teurere für 10.000 Euro mehr, wenn er das Angebot und das Konzept dahinter nicht wirklich versteht? Sie können davon ausgehen, dass manche Kunden bei Ihrem günstigeren Mitbewerber kaufen, obwohl Sie eigentlich den Auftrag wirklich verdient hätten, weil Sie ihm nicht ausreichend verdeutlichen konnten, dass Sie Ihren Preis wirklich(!) wert sind.

## **Am Anfang steht die Bedarfsanalyse**

Das Wichtigste im Verkaufsgespräch, die Bedarfsanalyse, wird selbst von Profi-Verkäufern oft unterschätzt. Eine schlechte Bedarfsanalyse rächt sich spätestens bei der Preisverhandlung. Denn wie wollen Sie argumentieren, wenn Sie nicht wissen, was Ihrem Kunden wichtig ist? Bei einer guten Bedarfsanalyse laufen Sie bildlich gesprochen im Kopf des Anfragenden herum, um diesen vollumfänglich zu verstehen. Sie wissen nach dieser Analyse, was Ihr Kunde will und nicht will. Sie wissen danach, woran Ihr Kunde festmacht, dass Sie seinen Auftrag gut erledigt haben – und dass er Ihren Preis respektiert und akzeptiert. Aufgrund der Antworten können Sie dann auch für sich entscheiden, ob Sie diesen Kunden überhaupt haben wollen – oder nicht. Oder etwas spitzer formuliert: Sie müssen nach der Bedarfsanalyse ganz klar für sich entscheiden, ob Sie es verantworten können, den Auftrag zu übernehmen. Denn wenn zwei Personen einen Vertrag unterschreiben, dann müssen beide Seiten auch liefern.

Teilen Sie die Bedarfsanalyse auf wenigstens(!) zwei Termine auf. Den ersten Termin führen Sie online oder telefonisch durch. Finden Sie bei diesem Gespräch Antworten auf mindestens folgende fünf Fragen:

- Was haben Sie genau vor?
- Weshalb?

- Ab wann soll das Ganze fertig sein?
- Welche Erfahrungen haben Sie bisher mit Agenturen gemacht?
- Wie sind Ihre finanziellen Vorstellungen?

Hauptziel dieses Erstgespräches ist es, dass Sie danach ganz klar sagen können: „Ja, es macht Sinn, weitere Zeit in das Projekt zu investieren“, oder: „Dafür sind wir nicht die Richtigen. Hier können wir es nicht verantworten, den Auftrag zu übernehmen.“ Im letzteren Fall sollten Sie den Kunden nicht im Regen stehen lassen, sondern im Idealfall an eine andere Agentur weiterempfehlen.

## Kunden wollen Sicherheit

Aber nicht nur Sie müssen den Anfragenden „abklopfen“, sondern Sie müssen sich auch richtig inszenieren. Und zwar so, dass im zeitlichen Verlauf der Verkaufsabschluss zu Ihren kalkulierten Preisen das Ergebnis ist. Es bringt wenig, wenn Sie mit Interessenten langweilige Logosammlungen Ihrer zufriedenen Kunden durchgehen. Denn nur weil Sie Logos von Kunden veröffentlichen, bedeutet das ja noch lange nicht, dass Sie diese Kunden auch wirklich begeistert haben. Darüber hinaus sagt ein Logo auch recht wenig über die Größe des Projekts aus. Fragen Sie sich besser, welche Sorgen und Ängste potenzielle Kunden haben:

- Haben die mich wirklich verstanden?
- Können die das wirklich?
- Wie verhalten die sich, wenn Probleme auftreten?
- Gibt es unangenehme Überraschungen (finanzieller oder zeitlicher Art)?
- Was passiert, wenn ich während der Zusammenarbeit merke, dass ich mich für den falschen Anbieter entschieden habe?
- Sollte ich jetzt kaufen – oder bekomme ich diese

Leistung auch irgendwo anders billiger oder besser?

- Wie viel „Luft“ ist noch in den Preisen?
- Was sagen die anderen Agenturen, die ich auch angefragt habe, wenn ich mich jetzt gegen diese entscheide?
- Wie werden meine Kollegen reagieren, wenn ich die falsche Agentur beauftrage?
- Wäre es besser, gar keine Agentur zu beauftragen – und alles so zu lassen, wie es jetzt ist?

Auf diese Fragen, die sich nahezu jeder potenzielle Kunde mehr oder weniger bewusst stellt, geben die meisten Anbieter von sich aus kaum gute Antworten. Zwangsläufig verlaufen so manche Anfragen im Sand. Denn wer hat noch nicht beim Nachfassen Sätze gehört wie: „Wir lassen erst mal alles so, wie es ist“, oder: „Wir haben jetzt andere Themen auf der Agenda. Melden Sie sich gerne in einem halben Jahr wieder, dann haben wir dafür wieder den Kopf frei“?

*Der Preis ist egal, wenn die Gegenleistung stimmt.*

## **Vorbereitung Preisverhandlung**

Egal, ob Sie mit neuen Kunden verhandeln oder mit bestehenden: Bereiten Sie sich auf Preisgespräche immer vor. Stellen Sie sich folgende Fragen:

1. Wie verdient der Kunde eigentlich sein Geld? Was sind auf Sicht der nächsten Monate und Jahre Probleme und Aufgaben, auf die er zusteuert?
2. Wie hoch ist eigentlich sein tatsächliches Einkaufsvolumen? Arbeitet er beispielsweise noch mit einer weiteren Agentur zusammen?
3. Was ist der Grund, dass der Kunde den Preis verhandeln möchte – und nicht etwas anderes?
4. Was sind unsere Ziele bei der Preisverhandlung?
5. Wie lange soll der verhandelte Preis gültig bleiben?

6. Wer nimmt alles an der Preisverhandlung teil?
7. Was lief bisher gut bei der Zusammenarbeit? Was nicht?
8. Mit welchen Möglichkeiten kann die Beziehung zu den Verhandlungspartnern verbessert werden?
9. Was sind mögliche Argumente aus Kundensicht, um einen niedrigeren Preis zu bekommen?
10. Welche Sonderleistungen wurden in der Vergangenheit nicht berechnet?
11. Mit welchen Argumenten verteidigen bzw. erklären Sie Ihren Preis?
12. Welche sinnvollen Alternativen zu einem Preisnachlass können dem Kunden angeboten werden?
13. Was sind mögliche Nachteile für den Kunden, wenn er nicht mit uns zusammenarbeitet?
14. Wie kann dem Kunden geholfen werden, Geld zu sparen oder mehr Geld zu verdienen?
15. Was ist der Plan B, wenn es bei der Verhandlungsrunde zu keinem Ergebnis kommt?

## **Verkäufer dürfen keine Schönredner sein**

Sie sind für das verantwortlich, was Sie verkaufen, aber auch für das, was Sie nicht verkaufen. Das muss Ihnen ganz klar sein. Sie dürfen sich niemals aus der Verantwortung mit Sätzen wie: „Hätte der Kunde ja sagen können, dass er das so meint“, oder: „Hat er mir ja nicht gesagt, dass ihm x wichtig ist“, aus der Verantwortung stehlen. Denn im Zweifelsfalle bekommen Sie immer die Schuld, auch wenn Sie diese objektiv vielleicht gar nicht haben. Oder glauben Sie, dass ein Kunde, der sich von Ihnen schlecht beraten fühlt, in seinem Umfeld sagt: „Ja, und dann habe ich mich für die Agentur x entschieden. Denen habe ich leider vergessen zu sagen, dass mir a und b sehr wichtig sind. Das hat den gesamten Zeitplan zerfetzt und zu Mehrkosten von 30.000 Euro geführt“? Nein, Ihr Kunde wird in

seinem Umfeld sagen: „Geh bloß nicht zu der Agentur x. Die können das nicht. Labern rum, kommen aber nicht in die Pötte. Letztlich musste ich denen noch 30.000 Euro nachschießen, weil die keine Ahnung haben.“ Klingt hart, aber so ist die Realität.

Darum müssen Sie nicht nur im Rahmen der Auftragsklärung vieles zur Verständnisförderung hinterfragen, sondern gegebenenfalls auch klare Grenzen ziehen. Beispielsweise mit Sätzen wie: „Nein, so machen wir das nicht, weil ...“, oder: „Wir könnten das in so einer einfachen Version machen, wie Sie das wünschen. Das wollen wir aber nicht, weil Sie dann nicht Ihre Ziele erreichen. Und letztlich können wir es uns nicht erlauben, Sie besseren Wissens ins offene Messer laufen zu lassen. Wir machen das also entweder richtig oder sonst lieber gar nicht. Denn wir haben auch eine Verantwortung dafür, dass Sie Ihr Geld gut investieren.“ Für manche Anbieter sind das „scharfe Aussagen“, die sie sich leider nicht zu sagen trauen. Doch wenn Sie dies wertschätzend und wohlwollend und nicht arrogant rüberbringen, dann wird so manch ein Kunde denken: „Die wissen, was sie wollen. Die scheinen ihr Geschäft zu verstehen. Also sind das wohl die richtigen Partner für mich, die ihr Geld wert sind.“

## **Verkaufen ist Erwartungsmanagement**

Viele Kunden sind sich gar nicht darüber im Klaren, wie viel Arbeit und auch Kosten auf sie zusätzlich zukommen. Als schlechter Verkäufer werden Sie das auch von sich aus nicht kommunizieren. Als souveräner Verkäufer durchaus. Es geht nicht darum, den Kunden zu belehren, sondern ihn vielmehr zu informieren. Denn je offener die Karten sind, mit denen Sie spielen, umso harmonischer verläuft auch die Geschäftsbeziehung. Das bedeutet, dass Sie nicht nur einen klaren Zeitplan mit Ihrem Kunden gemeinsam(!) erstellen, bis wann er was geliefert haben muss, sondern dass Sie auch offensiv und rechtzeitig kommunizieren, wenn etwas von Ihrer Seite aus nicht eingehalten werden kann.

# Feilschen macht Spaß

Was sagen Sie eigentlich spontan, wenn Ihr (potenzieller) Kunde zu Ihnen sagt: „Am Preis müssen wir noch was machen?“ Und Ihre Kollegen? Viele Verkäufer sind nicht in der Lage, auf diese elementare Frage souverän(!) zu antworten. So machen dann viele Anbieter ihr eigenes Unvermögen zum Problem des Kunden. Denn wenn dieser eine Antwort hört, die ihm nicht das Gefühl gibt, dass der Preis so in Ordnung ist, dann wittert er schnell die Chance auf Preisnachlässe. Es ist elementar, dass Sie nicht nur gute Antworten haben, sondern diese auch glaubwürdig (Stichworte sind hier Stimme, Modulation, Blickkontakt, ...) kommunizieren.

Es gibt viele Gründe, warum Kunden um Preisnachlässe feilschen:

- Test, ob noch „Luft“ in den Preisen ist.
- Sicherheit, nicht zu viel zu bezahlen.
- Mitbewerber bietet günstiger an.
- Einfach andere Vorstellungen vom Wert des Angebots.
- Begrenztes Budget.
- Gebot der Wirtschaftlichkeit, den besten Preis zu bekommen.
- Einkäufer haben die Aufgabe, Preise zu drücken.
- Oft kommen Nachlässe nur, weil man danach fragt.
- Spaß und Lust am Feilschen.
- Macht ausleben.

Insbesondere die letzten beiden Faktoren spielen die Hauptrolle bei Preisverhandlungen. Darum sollten Sie sich vor voreiligen Zugeständnissen hüten. Denn andernfalls denkt so manch ein Kunde: „Das war ja einfach. Ich habe einmal gefragt und bekomme jetzt schon einen Preisnachlass. Wie viel geht der denn nun noch runter, wenn ich ihn noch zwei Wochen zappeln lasse?“

*Wenn der Kunde nicht feilscht, dann hast du zu billig angeboten.*

## **Antworten auf „zu teuer!“**

Ein Angebot muss nicht zwangsläufig zu teuer sein, nur weil der Kunde dies sagt. Möglicherweise sind die vorgestellte Lösung und der erwartete Nutzen für den Kunden zu gering und deswegen im Verhältnis zur Leistung zu teuer. Hier ein paar Formulierungsbeispiele zur Inspiration.

1. Natürlich kann ich Ihnen beim Preis entgegenkommen. Soll ich Ihnen A oder lieber B aus dem Angebot rausrechnen?
2. Wie kommen Sie darauf?
3. Wenn Sie jetzt ein wenig mehr investieren, rechnet es sich mittelfristig besser für Sie. Schauen Sie hier ...
4. Ja, billig sind wir nicht. Billig wäre es, wenn ...
5. Ja, zuerst scheint der Preis ein wenig hoch zu sein. Lassen Sie uns bitte noch einmal zusammenfassen, welches Leistungspaket für Sie dahintersteckt.

## **Antworten auf „Mitbewerber ist billiger!“**

Einkäufer müssen sich eventuell auch intern rechtfertigen, weshalb sie nicht bei dem billigsten, sondern bei einem anderen gekauft haben. Darum braucht der Einkäufer häufig nur gute Argumente vom Verkäufer, um sich intern ggf. erklären zu können. Formulieren Sie doch mal folgende Ideen in Ihren Worten um und probieren Sie aus:

1. Unser Mitbewerber ist sein Geld durchaus wert. Mehr allerdings auch nicht. Darf ich Ihnen kurz aufzeigen, was uns entscheidend von diesem abhebt?

2. Jetzt mal Hand aufs Herz. Wir würden uns doch nicht so lange unterhalten, wenn es Ihnen nur um den Preis ginge, oder?
3. Oh, das ist interessant. Haben Sie schon herausgefunden, woran die sparen, um solche Preise machen zu können? Denn auch die können ja nicht vom Verschenken leben, oder?

## Über den Preis kann jeder reden

Auch wenn es respektlos klingt, aber „zu teuer!“ kann jeder sagen. Auch jemand, der vom Thema gar keine Ahnung hat. Im schlimmsten Falle versteht also Ihr Gegenüber gar nicht, was Sie mit seinem Geld für ihn alles Gutes tun wollen. Er sieht nur Ihr Angebot, Ihren Preis – und denkt: „Das Einzige, was ich hier verstehe, ist, dass die viel Geld haben wollen. Also rede ich nur über den Preis.“

Dazu wenden manche Kunden auch interessante Tricks an. Das ist legitim. Denn auch Verkäufer nutzen ja mehr oder weniger bewusst psychologische Tricks, um den Abschluss zu machen.

- Appell über die Beziehung: Wenn man sich kennt, dann kann man doch auch nett zueinander sein. Also warum nicht einen Preisnachlass geben?
- Guter Cop, böser Cop: Einer auf der Kundenseite spielt den wohlwollenden Kunden, der andere den kritischen. Vorher haben sich beide aufgrund Ihres schriftlich eingereichten Angebots über 20.000 Euro auf 18.500 Euro verständigt. Nun fordert der böse Cop von Ihnen, vielleicht auch mit etwas Drama und spitzer Argumentation, einen Preis von 15.000 Euro. Der gute Cop vermittelt zwischen Ihnen wohlwollend und schlägt 18.500 Euro vor. Erfreut über diesen Kompromissvorschlag stimmen Sie diesem Preis zu. Ein Preiszugeständnis, welches Sie ohne den bösen Cop vermutlich niemals gegeben hätten.

- Salamitaktik: Hier fordert der Kunde immer weitere Zugeständnisse. Stimmen Sie ihm beispielsweise zu, dass die Einweisung seiner Mitarbeiter nicht berechnet wird, wird er gleich den nächsten Punkt nachverhandeln, beispielsweise die monatliche Wartungspauschale. Denken Sie, nachdem Sie hier auch entgegengekommen sind, dass es nun endlich den Auftrag gibt, wird der nächste Punkt aufs Korn genommen.

Viele Anbieter erkaufen sich Aufträge über Rabatte bzw. Leistungen, die dann „ausnahmsweise“ nicht berechnet werden. Das rächt sich in der unternehmerischen Bilanz schnell. Denn das, was nicht eingenommen wird, kann auch nicht zum Gewinn beitragen oder für unternehmerisch wichtige Investitionen beispielsweise in die Mitarbeiterentwicklung genutzt werden.

## **Verkaufen Sie souverän**

Alle Mitarbeiter mit Kundenkontakt müssen(!) kommunikativ geschult sein. Andernfalls besteht das Risiko, dass diese nicht nur eine Zumutung für die Kunden werden, sondern auch eine Gefahr für Ihre gesamte Unternehmung. Denn Menschen mit Kundenkontakt beeinflussen entscheidend Ihre unternehmerischen Bilanzen. Von diesen ist es abhängig, ob der Kunde den Preis leicht akzeptiert, ob er aggressiv reklamiert oder Sie mit Überzeugung weiterempfiehlt. Dass Sie Ihren Job beherrschen, setzt ein Kunde voraus. Er möchte aber die Sicherheit haben, dass Sie ein Geschäftspartner auf Augenhöhe sind und Ihren Preis wert sind.

Wer sich seiner eigenen Stärken bewusst ist, diese auch so erklären kann, dass (potenzielle) Kunden sie leicht verstehen, darüber hinaus auch wertschätzende Antworten parat hat, wenn es mal einen kritischen Einwand gibt, wird eine gute Beziehung aufbauen – und den Auftrag machen.

---

# Datenanalysen mit R und der Google Analytics API



## Datenanalysen mit R und der Google Analytics API – websiteboosting.com

Das Programm bzw. die Programmiersprache R erfährt in den letzten Jahren eine steigende Nachfrage. Neben dem erfreulichen Aspekt, dass es Open Source und damit komplett kostenfrei ist, kann R durch Pakete flexibel erweitert werden. Neben klassischen Datamining-Paketen und fortgeschrittenen...

Das Programm bzw. die Programmiersprache R erfährt in den letzten Jahren eine steigende Nachfrage. Neben dem erfreulichen Aspekt, dass es Open Source und damit komplett kostenfrei ist, kann R durch Pakete flexibel erweitert werden. Neben klassischen Datamining-Paketen und fortgeschrittenen Analysemethoden wie neuronalen Netzen existieren diverse Möglichkeiten, den Google-Kosmos mit R anzusprechen. Die Anbindung der API von Google Analytics, Search Console bzw. der komplette Zugriff auf Sheets/Docs sind exemplarische Beispiele hierfür. Daten können in beliebiger Kombination abgefragt, aufbereitet, analysiert und visualisiert/exportiert werden.

## Analysen mit R – „flexibel wie ein Schweizer Taschenmesser“

R wird seit vielen Jahren durch eine große Community kontinuierlich weiterentwickelt. Vergleichbar mit Add-ins bei Excel können in R weitere Funktionen und Schnittstellen zum

Standardsystem hinzugefügt werden. So schön diese Flexibilität und Leistungsstärke klingt (und tatsächlich ist), so unspektakulär ist die Nutzerfreundlichkeit des Systems: Eingabe von Programmcode anstelle klickbarer Icons und Mausklicks à la Excel. Was jetzt als Nachteil klingt, ist gleichzeitig ein Riesenvorteil: Sofern der Programmcode erstellt ist, kann dieser jederzeit „abgespielt“ werden, die entsprechenden Analysen/Visualisierungen dauern wenige Sekunden/Minuten und können unternehmensweit geteilt werden. Jeder Schritt ist transparent und durch den Einsatz von Variablen voll flexibel – einmalig die ID des Google Analytics-Kontos angeschaut und alle Analysen können automatisch von Neuem starten. R ist quasi ein „großes Excel-Makro“, das die Daten auf Befehl Schritt für Schritt verarbeitet.

Die Software R kann unter [www.r-project.org/](http://www.r-project.org/) kostenfrei heruntergeladen werden und ist nach der Installation sofort einsatzbereit. Es empfiehlt sich, zusätzlich RStudio zu installieren ([www.rstudio.com/products/rstudio/download/](http://www.rstudio.com/products/rstudio/download/)), welches einen deutlichen Mehrwert in der Bedienung des Systems bietet (Abb. 1).



Abb. 1: RStudio „bettet“ die Software R benutzerfreundlich ein und macht die Bedienung komfortabler

Prinzipiell funktioniert R mittels Eingabe und sequenzieller Verarbeitung von Programmcode. Eingaben können sowohl in Form klassischer Rechenoperationen erfolgen (bspw.  $3+3$ ) als auch als Definition von Formeln, Funktionen, Variablen, Vektoren und Dataframes. Der Vorteil der Nutzung von Variablen ist, dass diese nach Definition zur gesamten Laufzeit zur Verfügung stehen und im weiteren Verlauf darauf zurückgegriffen werden kann. Füllt man bspw. die Variable „ga\_id“ (Variablennamen sind frei wählbar) mit der Datenansichts-ID von Google Analytics und verweist im weiteren Verlauf der Berechnungen auf die Variable anstelle einer fix definierten Nummer, kann die identische Analyse durch Neudefinition der Variable

flexibilisiert werden. Variablen werden in R mittels der Syntax „Variablenname <- Variableninhalt“ definiert.

## googleAnalyticsR – das flexible R-Paket

Damit R auf die API von Google Analytics zugreifen kann, muss ein Paket installiert werden (genereller Befehl: `install.packages(„Paketname“)`). Die Installation des Pakets erfolgt nach Befehlseingabe binnen Sekunden, da die Pakete automatisch aus dem Internet geladen werden. Die Erweiterung „dependencies = TRUE“ bewirkt, dass passende Pakete zusätzlich installiert werden, insb. `googleAuthR` zur OAuth-Authentifizierung des Google-Nutzerkontos mit R (siehe Abb. 2).



Abb. 2: Einfache Installation von `googleAnalyticsR` nebst weiteren Paketen direkt in R

Nachdem die gewünschten Pakete installiert wurden, stehen diese prinzipiell zur Verfügung. Damit sie tatsächlich auch genutzt werden können, müssen sie mit dem Befehl `library(„Paketname“)` geladen werden. Das heißt, nach Absetzen des Befehls „`library(„googleAnalyticsR“)`“ besteht eine Verbindung von R zur Google Analytics API. Der Befehl „`ga_auth()`“ bzw. „`gar_auth()`“ aus dem Paket `googleAuthR` ermöglicht die Authentifizierung von R mit dem jeweiligen Google-Konto über das OAuth-Verfahren. Ein erstmaliger Aufruf leitet zum Google-Konto-Log-in nebst Angabe der übergebenen Berechtigungen (siehe Abb. 3).



Abb. 3: Zugriff des R-Pakets auf Google-Konto

Sofern das Google-Konto bei Google Analytics mindestens das Recht „Bearbeiten“ auf eine Property hat, kann im weiteren Verlauf direkt die Google API genutzt werden. Wie angesprochen ist es sinnvoll, mit Variablen zu arbeiten und die gewünschte

Auswahl an Metriken und Dimensionen in Vektoren zu speichern. Der Befehl „`ga_account_list()`“ erzeugt eine Übersicht aller Analytics Property's nebst Datenansichten (Views) im jeweiligen Google-Konto. Als Beispiel kann eine Variable „`account_list`“ mit genau dieser Übersicht befüllt (Befehl `account_list <- ga_account_list()`) und anschließend aufgerufen werden (siehe Abb. 4). Auf Basis dieser Liste kann bspw. eine Variable „`ga_id`“ dynamisch mit der zweiten Datenansicht befüllt werden. Hierzu wird das Feld `viewId` mit dem Dollarzeichen (\$) aus `account_list` angesprochen und die zweite Position adressiert. Alternativ könnte die Variable `ga_id` selbstverständlich direkt mit der jeweiligen View-Nummer befüllt werden.



Abb. 4: R-Befehle zum Abruf von R

Mit dem Befehl „`google_analytics`“ wird die API angesprochen und die entsprechenden Werte werden in R ausgegeben. Sofern bspw. der Datumbereich sowie die gewünschten Metriken und Dimensionen nebst der View-ID im Befehl `google_analytics` übergeben werden, werden die Daten unmittelbar abgerufen. Aufgrund des Abrufs der Daten über die API kann ggf. das Sampling von Daten umgangen werden. Hierzu bedarf es beim Aufruf des Befehls `google_analytics` der Ergänzung um den Parameter `anti_sample = TRUE`. Die Funktion teilt dabei den API-Aufruf in mehrere Teilaufrufe auf, damit möglichst ungefilterte Daten über die API abgerufen werden. Weiterhin kann mit dem Parameter `max = „Zahl“` die maximale Anzahl von abzurufenden Daten limitiert werden (siehe Abb. 5).

Tipp: Es besteht für Google Analytics ein sog. „Dimensionen- und Metriken-Explorer“, der anzeigt, welche Dimensionen und Metriken generell kombiniert werden dürfen und wie die entsprechenden Felder in der API heißen (bspw. `ga:bounceRate` für die Absprungrate in %): [einfach.st/gareferenzen](http://einfach.st/gareferenzen).

Das Paket `googleAnalyticsR` nutzt ein allgemeines Kontingent an API-Aufrufen von Google Analytics. Bei intensiver Nutzung und zur Sicherstellung der Verfügbarkeit sollte ein eigenes API-

Kontingent über Google Developers beantragt werden (<https://console.developers.google.com>). Nach Registrierung bei Developers stehen diverse Google APIs zur Verfügung. Als Anpassung in R müssen vor der Authentifizierung die Client-ID sowie der geheime Clientschlüssel als Option angegeben werden (`options(googleAuthR.client_id = „xx.apps.googleusercontent.com“)` sowie `options(googleAuthR.client_secret = „xx“)`). Beide Informationen sind nach der Registrierung und Aktivierung der Analytics API abrufbar. Bei Google Analytics bietet die API 50.000 Abfragen pro Tag/User bzw. alle 100 Sekunden 100 Abfragen.

## Flexibler Export und Datenvisualisierungen mit ggplot2

Sofern das Ergebnis des Abrufs mittels „google\_analytics“ in eine Variable/Dataframe gespeichert wird, können neben der Durchführung statistischer Auswertungen (bspw. `summary(„Variable“)`) die Daten in ein spezifisches Dateiformat gespeichert und dann mit anderen Programmen weiterverarbeitet werden. Eine CSV-Datei wird über den Befehl „write.csv“ erzeugt (siehe Abb. 5).

Tipp: Perfekt harmoniert R auch mit Google Sheets, was damit mittelbar eine nachgelagerte Datenvisualisierung mit Google Data Studio ermöglicht. Das Szenario wäre folgendes: Daten werden aus Google Analytics und anderen Quellen in R analysiert und verarbeitet, automatisch über das R-Paket „googlesheets“ in ein Google Spreadsheet übertragen und anschließend über Data Studio visualisiert und ggf. mit weiteren Datenquellen ergänzt.



Abb. 5: Daten aus der API in R analysieren und als .csv speichern

R hat herausragende Möglichkeiten der Datenvisualisierung. Es existiert eine Vielzahl an Paketen, welche die Daten mit

weiteren Analysen grafisch anreichern. Pakete wie „RColorBrewer“ erzeugen u. a. Heatmaps als Datenvisualisierung. Das Paket „ggplot2“ (`install.packages("ggplot2")`) sowie Aktivierung über `library("ggplot2")` bietet Schaubilder, die in Excel nicht existent sind. Mittels des Befehls `geom_boxplot` wird bspw. ein Boxplot erstellt, `geom_line` / `geom_smooth` erzeugt ein Liniendiagramm nebst nicht linearer Trendlinie (siehe Abb. 6).



Abb. 6: Datenvisualisierungen mit dem Paket ggplot2

Tipp: Die Visualisierung bzw. Berichterstellung ist in R sehr umfassend durch sogenanntes R Markdown nutzbar. Es wird einmalig der Programmcode erstellt bzw. ein Bericht konzipiert, welcher anschließend als HTML-Seite, PDF oder Word-Dokument ausgegeben wird. Damit sind aufwendiges „Exceln“ und formatieren obsolet, die Corporate Identity ist in hohem Maße in den Berichten standardisierbar.

Richtig interessant werden Analysen und Visualisierungen, wenn diese auf Rohdaten beruhen. Mittels der Analytics API können viele Informationen verarbeitet werden, die es später zu clustern gilt. Eine sinnvolle Ergänzung ist deshalb, die Google-Analytics-Daten durch benutzerdefinierte Dimensionen zu erweitern. Wird jedem Datensatz ein exakter Zeitstempel mitgegeben, jede Session mit einer Session-ID in den Daten ergänzt sowie die Cookie-ID als Klammer für den Besucher verstanden, so könnte eine Customer-Journey-Analyse mit Rohdaten erfolgen. Simo Ahava hat in seinem Blog eine detaillierte Anleitung für diese benutzerdefinierten Dimensionen erstellt (<http://einfach.st/simohava>). Es ist sehr ratsam, vor der Integration den jeweiligen Datenschutzverantwortlichen zu konsultieren!

In diesem Sinne: Geben Sie R eine Chance! Im ersten Schritt hat es eine „Old-School“-Anmutung, da Befehle über eine Console eingegeben werden müssen. Mit Neugier und Training werden die Abläufe Tag für Tag vereinfacht und Webanalyse bzw.

Datenvisualisierung kann wirklich auf „Knopfdruck“ erfolgen – so besteht viel mehr Freiraum und Zeit, über die Ergebnisse nachzudenken bzw. diese zu interpretieren, anstatt Zeit für die Datenaufbereitung ineffizient zu vergeuden. Und zuallerletzt: Das System wird permanent (kostenlos) weiterentwickelt– für jede Anforderung existiert bestimmt ein Paket.

Noch ein Literaturhinweis: Das kostenlose englischsprachige E-Book „Using Google Analytics with R“ von Michal Brys ist unter <http://michalbrys.com/book/> in diversen Formaten abrufbar und bietet einen hervorragenden Einstieg in die Nutzung der Analytics API!

---

## Das Powertool R – hands-on!



### Das Powertool R – hands-on! – websiteboosting.com

Da der Beitrag von Tobias Aubele in der letzten Ausgabe (Datenabzug von Google Analytics über das Tool R) auf große Resonanz gestoßen ist, möchten wir all diejenigen hier, die noch keinen Kontakt zu dem kostenlosen Tool R hatten bzw. ihn bisher gemieden haben, eine kurze und leicht verständliche...

**Da der Beitrag von Tobias Aubele in der letzten Ausgabe (Datenabzug von Google Analytics über das Tool R) auf große Resonanz gestoßen ist, möchten wir all diejenigen hier, die noch keinen Kontakt zu dem kostenlosen Tool R hatten bzw. ihn bisher gemieden haben, eine kurze und leicht verständliche Einführung geben.**

**Keine Sorge, es wird in diesem Beitrag nicht zu technisch. Und die Beispiele sind so gehalten, dass Sie einen schnellen und**

effizienten Einblick in das Tool bekommen, sofern Sie nicht nur lesen, sondern sich einfach selbst mal an die Tasten setzen. Wie bei vielen mächtigen Tools ist es oft die Anfangshürde der Bedienung, an der man scheitert. Gelingt es, diese zu überwinden, eröffnet sich wie so oft ein wahres Eldorado an Vereinfachungen für die eigene Arbeit und/oder zusätzliche Möglichkeiten für nötige Analysen oder die Gewinnung wichtiger Erkenntnisse.

Eigentlich ist R als Statistiktool konzipiert und bekannt. Aber im Lauf der Jahre ist durch eine schier unendliche Anzahl an Funktionsbibliotheken die Anwendungsbreite förmlich explodiert. So kann man z. B. Ergebnisse automatisch in HTML-Webseiten übertragen oder Textmining betreiben. Zwei zentrale Stärken von R sind wohl zum einen die Anzahl der Daten, die man damit sehr schnell verarbeiten kann. Zum anderen kann man aber alles, was man für ein Projekt oder eine Analyse experimentell durchgeführt hat, auf Knopfdruck ganz oder teilweise mit anderen Daten wiederholen. Richtig eingesetzt, kann das gerade im Online-Marketing sehr viel Zeit und damit Kosten sparen. Natürlich sind Sie nach der Lektüre des Beitrags kein R-Spezialist – aber dann können Sie für sich entscheiden, ob Sie dem Tool eine Chance geben und sich etwas tiefer darin eingraben. Im Web gibt es viele Tutorials, bei Udemy einen deutschen Kurs dazu und am Ende dieses Beitrags finden Sie Literatur für den Einstieg und für Fortgeschrittene zum Thema Data Science.

Erwarten Sie persönlich, dass die Menge der Daten künftig an Ihrem Arbeitsplatz eher zunehmen wird? Wird die Zahl der Datenquellen eher größer und haben diese wahrscheinlich unterschiedliche Strukturen? Schreitet die Digitalisierung bei Ihnen gut voran? Dann, ja dann kann es sich wirklich lohnen, sich frühzeitig mit so einem Tool auseinanderzusetzen, das Sie bei der Bewältigung gut unterstützt. Und wie immer gilt: Wer früher dran ist, kann den anderen die lange Nase zeigen, während die sich wundern, wie man das Kaninchen mal wieder aus dem (neuen) Hut gezaubert hat ...

# Was ist R überhaupt?

R selbst ist eigentlich eher eine Programmiersprache als ein Tool und ähnelt Python. Über das ebenfalls kostenlose RStudio erhält R eine komfortable und übersichtliche Benutzeroberfläche, die Einsteigern etwas den Schrecken nimmt. Ursprünglich für statistische Anwendungen gedacht, ist R mittlerweile sehr viel mehr und hat große Stärken, wenn man Daten auswerten, umwandeln oder visualisieren muss. Sie möchten wissen, wie oft das Wort „Westernstiefel“ auf einer bestimmten Webseite vorkommt? Kein Problem – mit den entsprechenden kostenlosen Erweiterungspaketen (Librarys) ist das mit einer Befehlszeile erledigt. Sie brauchen einen Überblick über die Struktur der Shop-Umsatzdaten des letzten Jahres? Dazu laden Sie die Umsatzdatei ein und der Befehl „sum“ gibt Ihnen im Bruchteil einer Sekunde statistisch relevante Informationen. Vorausgesetzt, Sie wissen, was ein Median oder eine Standardabweichung ist. Mit anderen Worten sehen Sie mit einer Zeile Code, ob z. B. der häufig falsch verwendete Mittelwert als Kennzahl bei Ihren Daten überhaupt aussagekräftig ist.

*„Windows, Mac oder Linux? Geht alles!“*

Wer tiefer einsteigt, kann mit R auch Anwendungen für Machine Learning fahren oder Data Mining betreiben, indem man im einfachsten Fall über die Daten Regressionsanalysen laufen lässt. Bereits mit zwei bis drei Befehlszeilen kann man sich durch Umwandlung von Daten einen visuellen Überblick über mögliche Abhängigkeiten verschaffen. Ist der Umsatz tatsächlich vom Wetter abhängig und wenn ja, wie stark? Wie wirken sich Preisveränderungen von Mitbewerbern auf den eigenen Absatz aus und wie eine Erhöhung des Ads-Budgets? Gibt es Abhängigkeiten, die man bisher noch gar nicht auf dem Schirm hatte und wo es sich lohnt, tiefer zu graben?

R und RStudio sind für Windows, Mac OS und Linux verfügbar.

# Was ist das Besondere an R?

Wie schon erwähnt, machen die vielen Funktionsbibliotheken (Librarys) R erst richtig nützlich und breit anwendbar. Diese lassen sich sehr simpel bei Bedarf einbinden und stehen fortan im Tool zur Verfügung. Damit kann man dann z. B. bestimmte Dateiformate lesen und erzeugen, Daten direkt in HTML-Vorlagen ausgeben, Anbindungen an verschiedene Datenbanken managen, linguistische Analysen durchführen oder auch Marktforschung betreiben. Und das Ganze völlig kostenlos, sehr stabil und performant.

Eine Besonderheit gegenüber anderen Programmiersprachen liegt darin, dass man in R mit sog. Vektoren arbeitet. Das erleichtert den Umgang mit Daten und es sind nicht wie sonst gesondert programmierte Schleifen nötig. Einen Vektor kann man sich als komplette Zeile (oder Spalte) in Excel vorstellen, der sich direkt verarbeiten lässt. Im Prinzip lässt sich so z. B. der komplette Quellcode einer Webseite in einen Vektor einlesen und über den Namen des Vektors ansprechen.

*„R ist kostenlos, stabil und performant.“*

Wie Sie später noch sehen werden, ist es recht einfach, alle Befehle, die man in die Console eingegeben hat, später wieder aufzurufen und ggf. bei Bedarf in Summe in einen individuellen Programmcode zu überführen. Das erleichtert das Verständnis, die Erstellung eines kleinen Programms und auch die Fehlersuche. Hier greift das gleiche Grundprinzip wie beim Arbeiten mit Excel, wenn man kein ausgebildeter bzw. fachkundiger Programmierer ist. Man bearbeitet Daten Schritt für Schritt und sieht jeweils sofort das Ergebnis und ob alles wie gewünscht passt. Nach Eingabe einer fehlerhaften Formel erscheint sofort eine Fehlermeldung – nicht erst, wenn später ein Programm läuft. Erst wenn alles passt, speichert man sich die einzelnen Befehlszeilen ab und lässt später direkt alle auf einmal ablaufen.

R und RStudio enthalten eine umfassende Hilfe: Durch Eingabe eines Fragezeichens, gefolgt von einem Funktionsbefehl wird die Hilfe aufgerufen, die die Verwendung genau erklärt. Auch im Reiter „Help“ (Abbildung 2; Ziffer 8) ist ein kleines Suchfeld integriert, das die Suche erleichtert. Empfehlenswert für einen strukturierten Einstieg ist es allerdings, in ein entsprechendes Fachbuch zu investieren. Eine kleine Auswahl finden Sie am Ende dieses Beitrags.

*„R ist sehr flexibel bei der Datenausgabe.“*

R hat sehr flexible Ausgabeformate. So können Ergebnisse, Tabellen, Grafiken und anderes über Erweiterungspakete z. B. als CSV, TXT, XLS, Word, PowerPoint oder auch direkt in HTML ausgegeben werden (Abbildung 1). Ebenso ist es natürlich möglich, die Daten in angebundene Datenbanken zu schreiben. Damit entfällt der sonst übliche Anwendungsbruch, weil man z. B. in Excel arbeitet, Diagramme erzeugt und diese dann für eine Präsentation in PowerPoint kopiert. Wer seine Berichte modern lieber gleich per HTML ins Intranet oder ins Web schießen möchte, kann das ebenfalls tun. Analyse, Aufbereitung und Dokumentation können also in einem einzigen Tool erledigt und alle nötigen Schritte dazu reversibel zentral mit abgespeichert werden.



Abbildung 1: Auszug für gängige Ausgabe- bzw. Exportformate in R

*„Nichts geht mehr verloren – kein einziger Arbeitsschritt.“*

Vielleicht ist einer der größten Vorteile von R, dass Sie alles zusammen in einer Umgebung halten. Wie laufen üblicherweise Analysen im weitesten Sinne in Unternehmen ab? Man hat ein Set von Daten in verschiedenen Formaten und nicht selten strukturell unterschiedlich. Man kopiert bzw. liest diese dann oft in Excel ein, wandelt bestimmte Daten um, verdichtet, sortiert, filtert und modifiziert sie, führt

weitere Konsolidierungen und Berechnungen durch und erzeugt diverse Diagramme oder Übersichtstabellen. Werden diese in einen Bericht gegossen oder präsentiert, kopiert man sie in Word oder PowerPoint. So weit, so gut. Im nächsten Monat/Quartal/Jahr oder beim nächsten Projekt steht dann eine Wiederholung des Ganzen an. Dabei entsteht oft das Problem, dass man nicht nur all diese Arbeiten erneut durchführen muss. Noch öfter steht man vor den fertig modifizierten Tabellen und weiß nicht mehr, wie man von den Ursprungsdaten dort hingelangt ist, da nicht mehr alle Formeln vorhanden sind bzw. neu erdacht und ausprobiert werden müssen. Alles in allem eine gigantische und auch ärgerliche Verschwendung von Zeit und Ressourcen.

Wenn Sie sich darauf einlassen, R etwas näherzukommen, gehört dies künftig der Vergangenheit an. Denn in R speichern Sie bei Bedarf jeden einzelnen Schritt – das Öffnen der Datei(en), jede Bearbeitung und schließlich auch die Erzeugung von Ergebnissen, Kennzahlen, Diagrammen und visuellen Auswertungen. Alles wird in einer Umgebung abgelegt und kann jederzeit nachvollzogen und reproduziert werden. Wenn Sie tiefer in R eintauchen, werden Sie sogar in der Lage sein, jedwede Auswertung automatisch per HTML oder PowerPoint neu zu erzeugen. Im einfachsten Fall kopieren Sie eine immer gleich benannte Datei mit den aktuellen Monatszahlen in das Arbeitsverzeichnis von R und starten Ihre aufgezeichnete Programmierung. Das war es dann auch schon. Ein kompletter und durchaus auch optisch anspruchsvoller Bericht steht Sekunden später vollautomatisch im Intranet.

## **Wo finde ich R und RStudio?**

Die Installation von R ist recht einfach. Es lässt sich kostenlos auf der Website von

[cran.r-project.org](http://cran.r-project.org)

herunterladen, und zwar als Windows-, Mac-OS- und

Linuxversion. R selbst ist wie erwähnt nicht so komfortabel zu bedienen und daher empfiehlt es sich für die meisten Nutzer, gleich noch RStudio dazu zu installieren. RStudio muss *nach* R installiert werden, sucht dann aber automatisch nach der R-Installation und integriert das Tool in eine übersichtlichere Benutzeroberfläche mit vielen nützlichen zusätzlichen Funktionen. Die Basisversion ist ebenfalls frei und unter [rstudio.com](http://rstudio.com) erhältlich:

[einfach.st/rstudio](http://einfach.st/rstudio).

Wer den entsprechenden Rechner hat, sollte sich gleich die 64-Bit-Version installieren bzw. verwenden, da gerade rechenintensive Analysen deutlich schneller sind und auch die üblichen Restriktionen beim Datenvolumen unter 32 Bit entfallen.

Für normale Anwendungen ist RStudio kostenlos. Es gibt allerdings auch kommerzielle Lizenzen mit Support und weiteren Features sowie eine Serverversion.

Beim ersten Start prüft RStudio bzw. R noch, ob auf den Rechner eine aktuelle Java-Version installiert ist, was aus Sicherheitsgründen sowieso auf jedem Rechner Pflicht sein sollte. Anschließend präsentiert sich R integriert in RStudio aufnahmebereit für die ersten Befehle bzw. Gehversuche.



Abbildung 2: Der Startbildschirm von RStudio

## Legen Sie jetzt einfach los

Nach der erfolgreichen Installation kann man sofort mit R bzw. RStudio arbeiten (Abbildung 2). Im linken Teil befindet sich die Console (Ziffer 1), in der die Befehle eingetippt werden. Geben Sie nach dem Prompt (der blinkende Cursor nach dem „>“ Zeichen) einfach

2+5

ein und drücken Sie Return (Ziffer 2). Als Antwort erhalten Sie das Ergebnis dieser einfachen Rechenoperation. Gibt man einen Befehl falsch ein, z. B. „Print“ mit einem Großbuchstaben, erscheint eine Fehlermeldung, die hier zur Demonstration erzwungen wurde. Normalerweise erscheint nämlich bereits beim Tippen ein Vorschlagfenster mit den gültigen Befehlen für Funktionen (Ziffer 4). Das bedeutet, man kann sich den Rest einer Anweisung wie z. B. „print.data.frame“ sparen und den gewünschten Befehl einfach übernehmen.

## Tipp

Wenn der Cursor im Consolenfenster blinkt, können Sie mit den Cursortasten hoch und runter die letzten eingetippten Befehle holen bzw. in diesen blättern. Das hilft insbesondere bei Tippfehlern, damit nicht alles erneut eingegeben werden muss. Einfach mit der Cursortaste-oben den letzten (falschen) Befehl holen, an der entsprechenden Stelle ausbessern und mit Return erneut auslösen. Zudem zeichnet R jeden Befehl in einer Historie auf, sodass er jederzeit auch später einfach noch mal verwendet werden kann.

Variablen weist man mit „<-“ ganz einfach einen Wert zu. Das machen wir hier nur einmal beispielhaft, um die Vorzüge des RStudios zu zeigen. Ziffer 5 zeigt eine solche Zuweisung:

```
wsb <- 100
```

weist der Variable „wsb“ damit den Wert 100 zu. Also erst der gewünschte Name, dann die Zuweisungszeichen und anschließend der Wert. Im rechten oberen Teil von RStudio findet man dann unter dem Reiter „Environment“ (Ziffer 6 und das angezeigte Fenster mit Ziffer 6) genau diese und natürlich alle anderen Variablen wieder, die man im Lauf einer Sitzung definiert hat. Das dient der Kontrolle, welchen Wert eine Variable gerade hat. Wenn Sie jetzt als Befehlszeile

gefolgt von Return eingeben, sehen Sie unter Environment rechts oben den neuen Wert 105 für diese Variable. Prinzipiell kann man ganzen Datenpools (Vektoren, Matrizen, Data Frames) Variablennamen zuweisen, was das Handling und die Lesbarkeit durch einfache Namensvergabe enorm vereinfacht. Dies aber nur als Hinweis.

Unter dem Reiter „History“ (Ziffer 3) werden alle Befehle aufgezeichnet, die man in die Console eingegeben hat. Mit einem Mausklick lassen sie sich einfach wiederholen bzw. in die Console übernehmen. Später lassen sich diese Befehle dann ganz einfach in einem kleinen Programmcode überführen, der nacheinander abläuft.

Das charmante Prinzip ist also für den Einsteiger, dass er zunächst Schritt für Schritt nach jeder Eingabe explorativ prüfen kann, ob alles so richtig ist, und erst danach fasst man quasi die einzelnen Befehle zu einem ablauffähigen Programm zusammen und kann dieses dann mit einem einzigen Befehl starten.

## **Einfach mal machen!**

Alle hier gezeigten Beispiele sind so gehalten, dass Sie möglichst einfach selbst erste Schritte zum Ausprobieren unternehmen können. Wir haben daher oft auf branchenübliche Zahlenbeispiele verzichtet, weil R einige Datenquellen zum Experimentieren bereits mitliefert. Sie müssen sich also nicht erst mit dem Erstellen und Einlesen von Datensätzen herumschlagen, nur um überhaupt eine Grundlage zum Testen zu haben. Ob Sie nun die Kelchlänge von Lilienarten oder den Benzinverbrauch von Autos visualisieren oder Marketingausgaben oder Backlinks, ist von der Methode her ja austauschbar. Hier liegt der Fokus tatsächlich darin, Sie sofort in die Lage zu versetzen, einfach einmal mit dem Tool herumzuspielen.

Ein Anwendungsbeispiel: Sie nutzen regelmäßig einen oder mehrere Datensätze, z. B. aus einem SEO-Tool, mit denen Sie arbeiten müssen. Diese enthalten aber unerwünschte (Sonder-)Zeichen, wie Umlaute als &szlig“ statt einem „ß“ oder schlicht den typischen Punkt statt des deutschen Kommas als Kommastelle. Das kann man natürlich auch in Excel erledigen, aber man muss alle Schritte einzeln und jedes Mal durchführen. Über R lädt man sich so einen Datensatz, verändert ihn mit den entsprechenden Anweisungen einmalig und prüft, ob alles korrekt ist. Anschließend speichert man sich die Befehle einfach ab und kann sie später direkt aufrufen. Der Datensatz muss dann nur noch geladen werden, das entsprechend passende erstellte Programm wird gestartet und der Datensatz entweder manuell oder sogar schon vom eigenen Programm abgespeichert – fertig. Prinzipiell lassen sich damit alle Daten auf einfache Weise verändern, bereinigen, auf Konsistenz prüfen oder einfach nur einzelne Zeichen tauschen.

Im rechten unteren Bereich von RStudio (Ziffer 8) hat man Zugriff auf Datenfiles, Grafiken (Plots), Packages (Erweiterungen) sowie auf eine Hilfe und einen Viewer.



Abbildung 3: In RStudio findet man unter dem Menüpunkt „Help“ nützliche Cheatsheets



Abbildung 4: In R selbst (nicht in RStudio) sind umfangreiche PDF-Handbücher hinterlegt (englisch)

## Mitgelieferte Testdateien

Praktischerweise beinhaltet die Installation von R bereits einige Dateien mit Daten, sodass man die ersten Gehversuche unternehmen kann, ohne erst einmal selbst Daten einlesen zu müssen. Eine dieser Beispieldatensätze ist „state.x77“. Er lässt sich mit dem Befehl

```
View(state.x77)
```

anzeigen. Links oben im Fenster werden jetzt alle 50 US-Staaten mit einigen Spalten wie „Einwohnerzahl“, „Durchschnittseinkommen“ etc. angezeigt (Abbildung 5).



Abbildung 5: R liefert einige Beispieldaten automatisch mit Ein einfacher Befehl zeigt Ihnen direkt statistisch relevante Daten dieses Datensatzes an (Abbildung 6):

```
summary(state.x77)
```

R zeigt für jede Spalte gesondert die statistisch wichtigsten Strukturinformationen an. Für eine spätere Datenanalyse oder die Bildung von Kennzahlen ist dies sehr nützlich, weil man sofort u. a. die Streuung ablesen kann. Ein extremes Beispiel zur Verdeutlichung: Hatten 200 Verkäufe den Wert 1 € und weitere 200 den Wert 100 €, dann macht die Verwendung eines Mittelwerts (50 €) nicht wirklich Sinn.



Abbildung 6: Statistisch relevante Strukturdaten der Datei „state.x77“

Der Befehl

```
state.x77[11,2]
```

greift z. B. direkt im Datensatz „state.x77“ auf die elfte Zeile und die zweite Spalte zu und liefert 4963 als das Durchschnittseinkommen in Hawaii zurück (in Abbildung 5 zu sehen). Genauso können über Spalten oder Zeilen natürlich entsprechende weitere einfache Berechnungen oder Analysen durchgeführt werden.

Die ebenfalls in R mitgelieferte Datei „state.name“ beinhaltet nur die Namen der 50 US-Staaten. Welche dieser Namen besteht aus mehreren Wörtern wie z. B. New Jersey? Hier hilft die Funktion „grep“, mit der man in diesen Daten ganz einfach nach einem Leerzeichen sucht. Beachten Sie bitte, dass sich der Name der verwendeten Datendatei jetzt ändert bzw. eine andere

Datei (state.name) zugrunde liegt (Testen Sie doch mal View(state.name) um zu sehen, was der Datensatz beinhaltet.)

```
state.name[grep(" ", state.name)]
```

Als Ergebnis erhält man in der Ausgabe die zehn Staaten mit einem „Doppelnamen“. Tauschen Sie das Leerzeichen z. B. gegen den Wortbestandteil „New“ aus, erhalten Sie die vier Bundesstaaten, die eben diese Zeichenfolge enthalten, der Befehl lautet dann:

```
state.name[grep(„New“, state.name)]
```

Diese einfachen Beispiele sollen nur verdeutlichen, wie leicht die Extraktion und Filterung in Datenbeständen ist und welches Prinzip dahintersteckt. Natürlich kann man solche Aufgaben bei so kleinen Datensätzen auch gut und einfach z. B. in Excel erledigen. Werden die Daten jedoch umfangreicher, macht das schon deutlich mehr Arbeit, die zudem prinzipiell jedes Mal neu durchgeführt werden muss. Die Befehle bzw. Befehlssequenzen hier in R können jedoch abgespeichert und zu jedem späteren Zeitpunkt einzeln oder in Summe aufeinanderfolgend gestartet werden.

Tipp: Geben Sie den Befehl

```
datasets::
```

ein und Sie erhalten in einem Pop-up-Fenster die in R integrierten Versuchsdatensätze angezeigt, die Sie per Mausklick einfach übernehmen können (Abbildung 7). Geben Sie den Befehl

```
data()
```

ein, erscheinen im linken oberen Fenster alle Namen und eine jeweils kurze Beschreibung aller Datensets, die bereits installiert sind.



Abbildung 7: Automatisch mitgelieferte Datensätze zum Ausprobieren mit Inhaltsbeschreibung

Noch ein weiteres kleines Beispiel zeigt Abbildung 8. Für die Variable „Zeichenkette“ wird zunächst hilfsweise ein Text „Die Website Boosting ...“ eingelesen. Der Befehl „strsplit“ zerlegt dann den Text in einzelne Wörter bzw. trennt nach einem Leerzeichen. Stellen Sie sich vor, Sie hätten eine Exceltabelle mit mehreren Tausend Zellen, die jeweils den Text einer Webseite enthalten, und für semantische Analysen müsste jedes Wort extrahiert werden. In R liest man dazu vereinfacht erklärt über ein Erweiterungspaket ein Set von URLs mit einer einzigen Funktion in Vektoren ein und ein weiterer Befehl extrahiert aus jedem Vektor jedes einzelne Wort. Bereits mit einigen Grundkenntnissen in R lässt sich so durchaus enorm Zeit sparen.

Unter [einfach.st/rdatasets](http://einfach.st/rdatasets) finden Sie übrigens eine recht gut kommentierte Übersicht (englisch) mit Beispielen und Beispielcode für die mitgelieferten Data Sets.



Abbildung 8: Mit einem einzigen Befehl Texte in Wörter zerlegen

### **Erweiterungspakete installieren**

Wie erwähnt werden für R sehr viele Erweiterungspakete im Web angeboten, die sich recht einfach integrieren lassen. Diese muss man nur einmal zu R hinzuladen und kann sie später dann nach dem jeweiligen Programmstart bei Bedarf aktivieren und die dort hinterlegten Funktionen nutzen. Eine erste gute Quelle finden Sie auf [cran.r-project.org](http://cran.r-project.org). Die dort ladbaren Pakete kann man direkt über R laden. Dafür muss bzw. sollte man natürlich den Namen der Erweiterung kennen. Diese werden meist in Anleitungen, Beispielen oder in der Literatur erwähnt. Die Installation kann dann per Befehlszeile über die Funktion „install.packages“ erfolgen (hier die Erweiterung „htmltidy“ zum Umgang mit HTML-Dokumenten):

```
install.packages(„htmltidy“)
```

oder über das Menü oben in RStudio unter Tools/Install Packages, wie in Abbildung 9 zu sehen ist. Auch hier erscheint beim Tippen bereits ein kleines Pop-up mit einer entsprechenden Auswahl. Lässt man den Haken „Install dependencies“ (in der Abb. verdeckt unter dem Pop-up) drin, werden sog. abhängige Packages gleich mit installiert. Die Erweiterung „htmltidy“ greift z. B. unter anderem auf die Erweiterung „xml“ und „htmlwidgets“ zu und löst damit auch gleich deren automatische Integration aus. In der Console erscheint dann eine entsprechende Meldung.

## Tipp

Manchmal zickt R bei der Eingabe von Erweiterungspaketen über die Kommandozeile mit einer Fehlermeldung zurück, weil die Groß-/Kleinschreibung nicht passt. Daher empfiehlt es sich tatsächlich, das (einmalige) Laden von Packages – aus Sicht eines Programmierers sicher unehrenhaft, aber eben einfacher – über das Menü von RStudio wie in Abbildung 9 gezeigt zu erledigen.

Die meisten Packages sind recht gut dokumentiert. So findet man für das eben erwähnte Package weitere Infos unter [cran.r-project.org/web/packages/htmltidy/index.html](https://cran.r-project.org/web/packages/htmltidy/index.html) und dort unter „Downloads“ auch ein PDF mit einer genauen Funktionsbeschreibung.

Möchte man die Funktionen eines einmal installierten Packages in R verwenden, aktiviert man es ganz einfach mit der Anweisung:

```
library(packagename) – im Beispiel also: library(htmltidy)
```



Abbildung 9: Erweiterungen lassen sich sehr einfach installieren

Einen Überblick über installierte Erweiterungen findet man in RStudio im rechten unteren Fenster unter dem Reiter „Packages“ (siehe Abbildung 2, Ziffer 8).

## **Eine CSV-Datei einlesen, modifizieren und wieder ausgeben**

Zum Einlesen von Dateien ist es nützlich, gleich zwei Erweiterungen zu installieren: „readr“ und „readxl“ (für das Excelformat). Dazu geben Sie einfach die beiden Befehlszeilen ein

```
install.packages(„readr“)
```

```
install.packages(„readxl“)
```

und aktivieren anschließend zumindest die Erweiterung „readr“:

```
library(readr)
```

Erstellen Sie zum Ausprobieren eine einfache CSV-Datei z. B. über Excel oder einen Texteditor mit mehreren Zeilen und Spalten. In die erste Zeile schreiben Sie durch Kommata getrennt die Spaltenüberschriften und darunter ebenfalls durch Kommata getrennt die Datensätze, wie z. B.

```
URL,Statuscode,Inlinks  
www.meine-domain.de,200,1.250  
www.deine-domain.com,200,3.243  
www.ihre-domain.at,200,13.283
```

Der Punkt bei den Inlinks-Zahlenwerten ist absichtlich gewählt, er soll durch ein Komma getauscht werden, um das Prinzip der Datenmodifikation zu zeigen. Erzeugen Sie in RStudio über „File“ – „New Projekt“ ein neues Projekt und speichern Sie dieses auf dem Rechner unter einem beliebigen Ort ab. Der angegebene Projektname ist dann automatisch der Verzeichnisname und Ihr „Working Directory“. In dieses

Verzeichnis speichern Sie dann die CSV-Datei ab, damit Sie diese durch einen einfachen Befehl öffnen und später speichern können. Alternativ kann man mit R natürlich auch auf alle anderen Verzeichnisse/Dateiorte zugreifen, aber die Dateien zu Anfang in einem eigenen Working Directory zu speichern, erleichtert die Arbeit etwas. In dem Beispiel hier wurde als Working Directory „WSB Einführungsartikel“ gesetzt und rechts unten unter „Files“ taucht die Datei beispiel.csv nach der Erzeugung bzw. dem Abspeichern dann auch entsprechend auf (Abbildung 10).

Der Befehl

```
getwd()
```

gibt übrigens das aktuelle Verzeichnis aus.



Abbildung 10: Legen Sie über ein neues Projekt einen Speicherort zum Testen an  
Über den Befehl

```
beispiel <- read.csv(„beispiel.csv“)
```

weisen Sie jetzt ganz einfach der Variable „beispiel“ (oder auch einem beliebigen anderen Namen) über die Funktion „read.csv“ die Datei beispiel.csv zu. Da die Datei im Working Directory liegt, brauchen Sie keine gesonderten Pfadangaben. Den Inhalt dieser Datei kann man anschließend einfach über den Variablennamen ansprechen. Tippen Sie z. B.

```
print(beispiel)
```

ein, wird der Inhalt der eingelesenen Datei ausgegeben (Abbildung 11). Probieren Sie auch den Befehl `view(beispiel)` und statt der Ausgabe in der Console geht links oben ein Datenfenster auf, das die Daten sauber in Tabellenform anzeigt.



## Abbildung 11: Dateien einlesen und anzeigen lassen

Geht es nicht nur um Datenanalysen, sondern werden Daten auch modifiziert oder neue Daten erzeugt, die man speichern möchte, geht auch das recht einfach mit dem Befehl `write.table`:

```
write.table(datensatz, „dateispeichername.csv“)
```

Möchten Sie z. B. die Daten des integrierten Beispiels „state.x77“ in eine CSV-Datei namens „test.csv“ in das aktive Arbeits-/Projektverzeichnis speichern, lautet der Befehl also:

```
write.table(state.x77, „test.csv“)
```

## Abbildungen erzeugen

Natürlich lassen sich Grafiken in Excel recht einfach erzeugen. R hält jedoch deutlich mehr Möglichkeiten und nach Meinung vieler auch bessere Grafiken vor. Wer z. B. schon einmal versucht hat, aus einem Datensatz ein vernünftiges Histogramm zu erzeugen, wird sich über die fehlende Möglichkeit geärgert haben, die automatische Clusterung von Excel zu umgehen. Hat man stark abweichende Werte, sind Histogramme dort praktisch wegen der Spreizung nicht zu verwenden. In R stellt das Definieren von Clustern für die Histogrammsäulen kein Problem dar. Ebenso wenig wie die (befehlsorientierte) Formatierung einer Abbildung. Einmal definiert, lässt sie sich für Folgeabbildungen mit einer Zeile Code in immer demselben eigenen Design erzeugen.



## Abbildung 12: Anriss des mitgelieferten Datensatzes „mtcars“

Ein Histogramm erzeugt man in R mit dem einfachen Befehl „hist“ gefolgt von dem gewünschten Datensatznamen in Klammer dahinter. Eine der mitgelieferten Datenbeispiele ist die Datei `mtcars`. Mit dem nun schon bekannten Befehl

```
print(mtcars)
```

können Sie einen schnellen Blick darauf werfen. Möchte man nun

den Benzinverbrauch aller dort gelisteten Modelle (erste Spalte, mpg, also miles per gallon) in einem Histogramm darstellen, benötigt man nur eine Befehlszeile:

```
hist(mtcars$mpg, col = „blue“)
```

Das Prinzip ist recht leicht ersichtlich. Aus den Daten „mtcars“ soll aus der Datenspalte „mpg“ ein Histogramm erzeugt werden (Befehl „hist()“), dessen Balken blau sind („blue“). Das Ergebnis zeigt Abbildung 13 auf der rechten Seite – und im Vergleich links das Histogramm, das Excel mit den gleichen Daten erzeugt. R teilt die Datencluster statistisch vernünftiger auf und erstellt ein durchaus differenziertes Bild. Wenn Sie die Clusterbildung selbst beeinflussen möchten, weil Sie z. B. feinere Abstufungen brauchen oder einfach nur auch zeigen wollen, dass unter zehn Gallonen pro Meile eben kein Wert und damit keine Säule vorhanden ist, geht das über eine einfache Erweiterung des Befehls mit Angabe der gewünschten Intervalle „breaks“ (der Übersichtlichkeit halber wurde auf die Farbangabe hier verzichtet). Probieren Sie es einfach aus:

```
hist(mtcars$mpg, breaks = c(5,10,15,20,25,30,35))
```

Selbstverständlich könnte man die Intervalle auch mit einer eigenen Formel statt mit manuell eingegebenen Intervallen wie hier im Beispiel berechnen lassen.



Abbildung 13: Vergleich eines unbearbeiteten Histogramms mit identischen Daten in Excel und R

Aber auch ein schneller optischer Blick auf strukturelle Eigenheiten von Datenreihen ist möglich. Dazu lässt man sich eine oder mehrere Datenreihen als Boxplot ausgeben. Zur Demonstration kann man ganz einfach erneut auf das bereits genutzte Datenset „mtcars“ zugreifen. Möchte man mehr über den Einfluss der Anzahl der Zylinder eines Motors auf den Verbrauch bzw. auf die Streuung der Daten wissen, setzt man beides einfach in Beziehung. Der Befehl dazu ist

```
boxplot(mpg ~ cyl, data=mtcars)
```

In Abbildung 12 ist ersichtlich, dass es zwei Spalten im Datenset gibt, die mit „mpg“ (Miles per Gallon) und „cyl“ (Zylinder) überschrieben sind. Der Befehl „boxplot“ bezieht sich auf die angegeben Datenquelle, setzt die beiden Datenreihen dann ganz einfach in Beziehung und gibt Abbildung 14 aus.



Abbildung 14: Statistische Kennwerte visualisieren

Dort erkennt man auf den ersten Blick, wie stark die Verbrauchsdaten des nach Zylinder geclusterten Verbrauchs streuen. Würde man statt mtcars auf einen eigenen eingelesenen Datensatz verweisen und zwei Spaltenüberschriften auswählen, müsste man den boxplot-Befehl einfach nur entsprechend anpassen und mehr nicht. Dazu wären in R nur zwei Befehlszeilen nötig – der zum Einlesen und der Befehl zur entsprechenden Ausgabe. Eine Sache von wenigen Sekunden.

## Kleines Cheatsheet für statistische Funktionen

Durchschnitt/Mittelwert: `mean()`

Median (zentraler Wert): `median()`

Varianz: `var()`

Standardabweichung: `sd()`

Quantile: `quantile()`

Bandbreite von–bis: `range()`

*„Vor der Verwendung von Daten ist es immer hilfreich, einen statistisch motivierten Blick auf die Struktur dieser Daten zu werfen.“*

Wer es eine Spur härter, aber keinesfalls komplexer haben möchte, kann auch problemlos mehrere Datenreihen miteinander korrelieren lassen bzw. optisch über Streudiagramme prüfen, ob

es einzelne Abhängigkeiten bzw. Korrelationen gibt. Dies lässt sich gut bzw. schnell und einfach am Beispiel des Datensets „iris“ zeigen. Dort sind für drei Schwertlilienarten in Summe 150 Datensätze mit jeweils Länge und Breite der Kelchblätter (Sepal) und der Kronenblätter (Petal) hinterlegt. Mit

```
View(iris)
```

wird das Dataset angezeigt (Abbildung 15).

Geben Sie jetzt doch einfach den Befehl

```
plot(iris[-5])
```

ein. Sie erhalten zwölf Streudiagramme aller möglichen Kombinationen zwischen Kelchblättern und Kronenblättern nach Länge und Breite (Abbildung 16). In den beiden rot markierten Diagrammen lässt sich unmittelbar ein fast linearer Zusammenhang bezüglich Länge und Breite bei den Kelchblättern erkennen, während bei den Kronenblättern und zwischen Kronen- und Kelchblättern kein solcher offensichtlich ist.



Abbildung 16: Korrelieren einzelne Daten miteinander?

## Fazit

Tauschen Sie jetzt einfach geistig oder auch real in der Praxis die letzten Beispiele wie z. B. das florale Dataset mit echten Daten aus Ihrem Arbeitsumfeld aus – dann erkennen Sie zumindest in Ansätzen den Wert und die Transparenz, die R in wenigen Minuten bringen kann. Und das ganz ohne wirklich tiefergehende Programmierkenntnisse und nur mit dem Wissen über einige einzelne Befehlszeilen. Natürlich bleibt der Nutzen tatsächlich überschaubar, wenn Sie bei der Lektüre nur dieses Beitrags stehen bleiben. Aber stellen Sie sich doch mal vor, Sie steigen jetzt motiviert und in Erahnung der Kenntnis des Potenzials von R Stück für Stück tiefer ein. Sie haben noch nicht einmal ein halbes Prozent von dem ausprobiert, was

R leisten kann, wenn Sie es richtig bespielen. Was nun am Ende mehr nützt – die Rationalisierungseffekte für einfacheres und schnelleres Arbeiten oder neue Erkenntnisse und bessere, wirklich datengestützte Entscheidungen –, bleibt Ihren Überlegungen und Ihrer Motivation überlassen.

---

# **Vitamin D: Datenanbindung über R an Google Schnittstellen**



## **Vitamin D: Datenanbindung über R an Google Schnittstellen – [websiteboosting.com](https://www.websiteboosting.com)**

R ist nicht nur Open Source (damit kostenlos) und anpassungsfähig durch vielfältige Pakete, sondern auch offen für andere (technische) Verbindungen. Diese Verbindungen, vulgo API, ermöglichen, flexibel Daten an der Ursprungsquelle abzufragen, entsprechend weiterzuverarbeiten und ggf. modifiziert...

R ist nicht nur Open Source (damit kostenlos) und anpassungsfähig durch vielfältige Pakete, sondern auch offen für andere (technische) Verbindungen. Diese Verbindungen, vulgo API, ermöglichen, flexibel Daten an der Ursprungsquelle abzufragen, entsprechend weiterzuverarbeiten und ggf. modifiziert wieder zurückzuliefern. Google bietet eine Vielzahl von API, die in R eingebunden eine enorme Individualisierung und Automatisierung von analytischen Prozessen und Visualisierungen ermöglichen. Die Excel-Alternative Google Sheets in Kombination mit Google Data Studio ist bereits eine unkomplizierte Lösung für vielfältige

**Visualisierungen. Gepaart mit der statistischen Leistung von R eröffnet sich durch die weitere Kombination mit API nochmals ein deutlicher Zugewinn an Flexibilität! Tobias Aubele zeigt, wie Sie diese Schnittstellen nutzen können.**

R wird kontinuierlich von einer großen Community weltweit weiterentwickelt, Gleiches gilt für die Dienste von Google wie Maps & Co. Unglaubliche Mengen an Daten liegen in diversen Datensilos und „sehnen“ sich danach, verarbeitet bzw. interpretiert zu werden und einen Nutzen zu stiften. Durch die Vielzahl an Silos und steigenden Datenmengen bedarf es einer effektiven und effizienten Verarbeitung, damit letztlich genügend Zeit für die Interpretation der Daten und das Ableiten von Maßnahmen übrig bleibt. Obwohl ggf. sehr zeitaufwendig, wird in Fachbereichen gerne Microsoft Excel eingesetzt, da dort Daten mittels Formeln und Funktion aus diversen Quellen flexibel zusammengeführt und verarbeitet werden können. Leider sind Daten teilweise veraltet, ehe sie letztlich über manuelle Zusammenfassungen und Verdichtungen für den eigentlichen Nutzungszweck zur Verfügung stehen. Sofern direkt API (Programmierschnittstellen) genutzt werden, können aktuelle Daten unmittelbar am Ursprung abgeholt und zurückgeliefert werden und bieten langfristig ggf. ein hohes Potenzial für automatisierte Abläufe.

## **Programmcode online verfügbar!**

Damit Sie Codezeilen nicht einzeln abtippen müssen, sind zwei der Skripte aus diesem Beitrag direkt online für Sie zum Download hinterlegt:

Das Skript zum Knowledge Graph unter  
[www.websiteboosting.com/code/r-knowledge](http://www.websiteboosting.com/code/r-knowledge)

Das Skript für Google Sheets unter  
[www.websiteboosting.com/code/r-sheets](http://www.websiteboosting.com/code/r-sheets)

# Die API – aktuelle Daten direkt vom Ursprung

Über die Google API können diverse Google-Dienste direkt über R angesprochen werden. Nach einer Registrierung unter [console.cloud.google.com](https://console.cloud.google.com) kann die gewünschte API, wie bspw. die des Google Knowledge Graph, einzeln aktiviert und unmittelbar getestet bzw. genutzt werden (siehe Abb. 1).



Abb. 1: Aktivierung einer Google API

Weiterhin können anschließend die notwendigen Anmeldedaten wie ein API-Schlüssel oder eine OAuth-Client-ID erstellt werden (siehe Abb. 2). Damit wird sichergestellt, dass der Datenzugriff nur von autorisierten Nutzern ermöglicht wird.



Abb. 2: Erstellung von gesicherten Anmeldemechanismen

Mittels der API werden die Daten direkt aus der Datenbank von Google abgerufen. Es stehen damit nicht nur aktuelle, sondern auch ggf. erweiterte Daten zur Verfügung. Das heißt, Informationen über Albert Einstein, welche im Graphen am rechten Bildschirmrand angezeigt werden (siehe Abb. 3), stehen direkt zur weiteren Verarbeitung zur Verfügung. Die API ermöglicht die Extraktion weiterer Daten, sollten mit dem Suchbegriff mehrere Entitäten (Personen, Objekte) in Verbindung gebracht werden wie bspw. die Person Hans Albert Einstein.



Abb. 3: Google Knowledge Graph von Albert Einstein

Das zur Datenextraktion notwendige R-Paket lautet *GoogleKnowledgeGraphR*, entwickelt von Daniel Schmeh (zur Dokumentation siehe das PDF unter: [einfach.st/r4kgraph](https://einfach.st/r4kgraph)). Mittels des individuellen API-Keys (Generierung siehe Abb. 2) werden mit dem Befehl

```
gkg_raw
```

die gewünschten Informationen in diversen Sprachen abgerufen. Das heißt, es stehen neben der Beschreibung unmittelbar die URL des Bildes, die URL des Eintrages sowie Informationen zu den Nutzungsrechten (dennoch prüfen!) zur Verfügung (siehe Abb. 4). Ohne die jeweilige Suchabfrage bei Google starten zu müssen, können die Ergebnisse direkt in R abgerufen und weiterverarbeitet werden. Dies kann zu einer weiteren Vereinfachung und damit zur Zeitersparnis führen.



Abb. 4: Aufruf der Knowledge Graph API über R

Ähnlich diesem Prozess erfolgt die generelle Nutzung von Google API (mit R). Es bedarf eines API-Schlüssels bzw. einer Autorisierung sowie eines korrespondierenden Programms (hier R-Paket). In der Praxis können nahezu alle Google API, wie bspw. die Google Translator API oder die Google Maps/Places API, abgefragt werden. Vereinzelt sind die Services kostenpflichtig mit einem flexiblen Abrechnungsmodell wie bspw. auf Basis der Anzahl an API-Abfragen. Damit könnten Textübersetzungen bzw. Abfragen zu Distanzen, Öffnungszeiten, Adressen etc. von Restaurants und anderen Google-Places-Einträgen vollzogen werden. Die entsprechenden R Packages heißen *googleway* und *googleLanguageR*.

## Kollaboration mit Google Sheets

Das Besondere an Google Sheets ist der weltweite Zugriff, die große Rechenleistung sowie die gleichzeitige Zusammenarbeit mit mehreren Personen. Weiterhin können Daten aus diversen Quellen zusammengeführt (vergleichbare Excel-Funktionen SVERWEIS bzw. INDEX/VERGLEICH) und letztlich ein Bericht sehr flexibel gestaltet, individualisiert und geteilt werden. Daher erscheint es sinnvoll, die Möglichkeiten von Google Sheets sowie ggf. der nachgelagerten Visualisierung mit Google Data Studio zu nutzen.

Daten in R können an Google Sheets bzw. von Google Sheets

übermittelt werden. Hierzu sind die Pakete *googlesheets* (PDF zur Dokumentation siehe: [einfach.st/r4sheets](http://einfach.st/r4sheets)) bzw. *googleAuthR* notwendig bzw. hilfreich. Nach der Installation des Paketes über den Befehl

```
install.packages("googlesheets")(siehe Abb. 6)
```

erfolgt die Freigabe bzw. Authentifizierung des R-Pakets mit Google Drive über

```
gs_auth() (siehe Abb. 5).
```



Abb. 5. Autorisierung des R-Pakets

Nach erfolgreicher Authentifizierung kann mit dem Befehl

```
gd_user()
```

der aktuelle Nutzer sowie der Berechtigungsstatus abgefragt werden. Der Befehl

```
gs_new
```

ermöglicht, ein neues Sheet anzulegen sowie den Namen bzw. Arbeitsblattnamen zu vergeben. Ab diesem Zeitpunkt können Daten eingegeben bzw. das Sheet genutzt werden. Der Befehl

```
gs_delete
```

löscht das Sheet entsprechend. Das jeweilige Sheet muss vor einer Aktion immer exakt adressiert werden, wozu diverse Möglichkeiten zur Verfügung stehen (siehe Dokumentation). Die exakte Art ist über die URL via

```
gs_url (siehe Abb. 6),
```

sofern Namen einzigartig sind, auch hilfsweise über den Titel (Befehl: *gs\_title*). Soll das Sheet im Browser aufgerufen werden, muss der Befehl

```
gs_browse
```

mit dem Namen des Sheets (im Beispiel in der Variable *gs*

gespeichert) und dem Blattnamen (relative Nummer bzw. absoluter Name) ausgeführt werden.



Abb. 6: Erstellung eines Google Sheets innerhalb von R

## Google Sheets mit Daten füllen

Prinzipiell können auch Daten aus Google Analytics in R abgerufen und analysiert werden (siehe den ersten Teil dieses Beitrags in der letzten Ausgabe #53). Dem Google Sheet „Analytics“ kann über den Befehl

```
gs_ws_new
```

ein neues Arbeitsblatt mit dem Namen Oktober hinzugefügt werden (siehe Abb. 7) und anschließend mit den Google-Analytics-Daten befüllt werden (Abruf der Daten über den Befehl *google\_analytics* aus dem R-Paket *googleAnalyticsR*). Die neuen Informationen sollen ab der Position (anchor) A1 zeilenweise (*byrow* = FALSE) eingefügt werden. Sollten dann im späteren Verlauf weitere Zeilen hinzugefügt werden (und dabei die alten Daten nicht überschrieben werden), kann ein Arbeitsblatt ausgelesen (Befehl *gs\_read*), die aktuelle Anzahl an benutzten Zeilen festgestellt (*nrow*) und anschließend über die R-Funktion *paste0* ein neuer Start ermittelt (im Beispiel Zelle A38) und an die Variable *gs\_position* übermittelt werden. Ein weiterer Aufruf der Funktion *gs\_edit\_cells* fügt die Daten demnach nicht wieder an der Position A1 ein, sondern nutzt als initiale Startposition den Inhalt der Variable *gs\_position*.



Abb. 7: Daten an Google Sheets aus R übertragen

## Highlight: spezifische Google-

# Sheets-Funktionen

Mittels Austausches von Information an Google Sheets (lesen und schreiben) kann ein flexibler Datenaustausch mit R erfolgen. Ein weiterer Vorteil von Sheets ist die Mannigfaltigkeit spezieller Funktionen. Über die Funktion

```
=IMPORTXML(VERKETTEN("http://www.meineSeite.de";A2);"//h1")
```

kann eine XPath-Abfrage generiert werden, welche Elemente aus beliebigen Webseiten direkt ausliest (hier der H1-Text; siehe Abb. 8). Mittels einer weiteren Funktion

```
=GOOGLETRANSLATE(D2;"de";"en")
```

wird der Inhalt einer Zelle in die englische Sprache übersetzt. Die Funktion Verketteten ermöglicht in diesem Zusammenhang, URL-Fragmente aus Google Analytics zu einer kompletten URL zusammenzuführen. Die Daten in Google Sheets können im weiteren Verlauf (manuell) beliebig verändert und anschließend wieder in R eingelesen/ weiterverarbeitet oder mit einer Visualisierung in Google Data Studio verbunden werden.



Abb. 8: Nutzung von Google-Sheets-spezifischen Funktionen

Google Data Studio hat diverse Verbindungen zu Datenquellen, was für die Analyse und Aufbereitung eine hohe Flexibilität bietet. Weiterhin können beliebige Google-Tabellen als Datenquelle initiiert und anschließend in die Berichte integriert werden – bspw. jeweils als eigene Seiten. Die Anbindung geschieht über zwei Schritte. Erstens muss über den Google Tabellen-Connector das jeweilige Tabellenblatt gewählt und somit als Datenquelle definiert werden (siehe Abb. 9). Zweitens müssten anschließend die damit geschaffenen Datenverbindungen mit einem spezifischen Bericht verbunden und die jeweiligen Datenfelder ausgewählt werden. Zur Visualisierung stehen diverse Möglichkeiten zur Verfügung (siehe Abb. 10).



Abb. 9: Auswahl von Google-Tabellen als Data-Studio-Connector  
Innerhalb des Berichtes besteht die Möglichkeit, mit mehreren Seiten zu arbeiten und damit bspw. die Sheets-Integration als Detailbericht einer großen Analyse zu nutzen. Weiterhin können einzelne Informationen/Spalten aus Sheets ausgewählt und in diversen Visualisierungsmöglichkeiten auf dem Bericht verortet werden.



Abb. 10: Berichtsaufbau mit Sheets-Integration

## Fazit

Zusammenfassend lässt sich sagen, dass R, Google Sheets, Google Data Studio etc. für sich hervorragende Möglichkeiten bieten. Die wahre Stärke zeigt sich, wie fast immer, in der Kombination verschiedener Faktoren/Quellen. Daten aus diversen Quellen (schnell) zusammenzufassen, aus unterschiedlichen Blickwinkeln zu betrachten und unkompliziert einem definierten Empfängerkreis zur Verfügung zu stellen, sind nur punktuelle Vorteile.

In diesem Sinne: Verbinden Sie Ihre Informationen, kombinieren Sie Daten – ggf. trägt ja der erste Anschein.

---

# R4SE0: Automatisierte Reports mit R generieren (Teil 1)

# R-Leuchtungen! Teil 9



## R-Leuchtungen! Teil 9: Title-Geddon – war da was bei Ihnen? – websiteboosting.com

In den Ausgaben bis 54 bis 58 der Website Boosting konnten Sie in der Serie „R4SEO“ von Patrick Lürwer nachvollziehen, wie man die kostenlose Software R verwendet, was sie leistet und wie man sie nutzbringend für die eigene Arbeit für SEO bzw. die Aufklärung im Online-Marketing einsetzen kann. R...

**In der letzten Ausgabe der Website Boosting gab Tobias Aubele einen sehr guten Einblick in die Potenziale von R, um auf einfache und reproduzierbare Weise Daten über die Google-Analytics-API abzufragen. Diese neue Artikelserie von Patrick Lürwer baut darauf auf und zeigt, wie mittels R Reports erstellt werden können, die sich zu einem bestimmten Zeitpunkt automatisch aktualisieren und per Mail verschicken. Im ersten Teil geht es nun zunächst um das Abfragen verschiedener APIs für die Datenbeschaffung. In den folgenden Teilen werden die Daten aggregiert, visualisiert und in einem Report zusammengefasst.**

## Tipp

Um Ihnen das Abtippen des Codes zu ersparen, können Sie sich das Skript zu diesem Artikel unter folgendem Link auf GitHub herunterladen. Dennoch empfehle ich Ihnen, den Code selbst zu schreiben, um ein Gefühl für die Syntax zu bekommen. [einfach.st/r4seocode](https://einfach.st/r4seocode)

Das Google Data Studio (GDS) bietet eine einfache Möglichkeit, Daten aus dem Google-Kosmos – bspw. aus Google Analytics (GA) und der Google Search Console (GSC) – abzufragen, zu visualisieren und als Reports aufzubereiten. Problematisch wird es allerdings, wenn andere Datenquellen angebunden werden

sollen. Denkbar sind hier APIs anderer Anbieter wie Sistrich. In solchen Fällen muss der „Umweg“ über Google Spreadsheets und Google Apps Scripts gegangen werden, um die Abfragen selbst zu programmieren. Dies ist durchaus ein valides Vorgehen, das jedoch schnell an seine Grenzen stoßen kann. Denn meist sind Reports nur der erste Schritt. Sobald in diesen Auffälligkeiten erkennbar werden, schließt eine Analyse an, die eine umfassende Beschäftigung mit den zugrunde liegenden Daten erfordert. Derartige Analysen verlangen zumeist komplexe Datentransformationen und -aggregationen, für die Dashboard-Softwares, wie das GSA, mit ihren begrenzten Filtern- und Manipulationsoptionen naturgemäß nicht ausgelegt sind. Ganz abgesehen davon, dass der direkte Zugriff auf die Daten zur Weiterverarbeitung häufig gar nicht gegeben ist. Von großem Vorteil ist es daher, wenn Report- und Analysesoftware in einer Anwendung vereint sind. Diese Anforderungen kann R insofern erfüllen, als die Programmiersprache durch Packages so erweitert werden kann, dass mit ihr sowohl sehr individualisierte Reports erstellt als auch umfangreiche Analysen durchgeführt werden können.

Da Datenanalysen eine sehr komplexe Thematik sind, soll in diesem Artikel jedoch zunächst das Potenzial von R zur Report-Generierung vorgestellt werden. Denn im Grunde sind beide Aufgaben bis zu einem gewissen Grade sehr ähnlich. Zunächst werden Daten von verschiedenen Quellen abgefragt, dann aufbereitet und schließlich visualisiert. Bei Analysen ist dieser Prozess jedoch iterativ, ohne dass im Vorhinein klar ist, wie das Ergebnis resp. die Erkenntnis genau aussieht. Bei der Generierung von Reports hingegen ist der Ablauf des Datenprozesses vorgegeben. Sie sind daher kleine, standardisierte Analysen, die an einem bestimmten Punkt der Visualisierung aufhören. Das Ergebnisdokument ist somit immer gleich und jederzeit reproduzierbar. Daher eignen sie sich sehr gut, um exemplarisch die Möglichkeiten von R aufzuzeigen, und sind dabei aber gleichzeitig so vereinheitlicht, um durch den Leser selbst nachgebaut zu werden. Ich möchte Sie daher

gerne dazu auffordern, das folgende Skript selbst in RStudio umzusetzen.

## Hinweis

Der Autor ist sich bewusst, dass gerade für Einsteiger der Anfang etwas holprig sein kann. Aber auch bei erfahreneren Lesern können durchaus Fragen aufkommen. Um hier Hilfestellung zu bieten, die Antworten der Community zugänglich zu machen und den Austausch anzuregen, möchte er Sie ermutigen, in der extra hierfür gegründeten Facebook-Gruppe nachzufragen. Es gibt dabei keine „dummen“ Fragen. Vielmehr hätte der Autor sich in seiner Anfangszeit selbst ein solches Forum mit dem Fokus auf analytische Programmierung und Online-Marketing sehr gewünscht. [www.facebook.com/groups/ompyr/](http://www.facebook.com/groups/ompyr/)

Aber genug der langen Vorrede! In diesem Teil der Serie ist das Ziel, die Sessions aus GA sowie die Performance-Daten aus der GSC abzufragen, um die Packages für diese APIs besser kennenzulernen. Darüber hinaus werden Sie am Beispiel von Sistrix eine eigene API-Abfrage schreiben, um den Sichtbarkeitsindex für eine Domain zu ermitteln. Als Beispiel-Domain dient hier die Website dein-trueffel.de, ein Studentenprojekt an der Hochschule Darmstadt im Modul „Fortgeschrittene Suchmaschinenoptimierung“, das durch Jens Fauldrath und Stefan Keil betreut wird. Voraussetzung zum Nachvollziehen des Skripts sind eine aktuelle R-Installation ([www.r-project.org](http://www.r-project.org)) und RStudio ([www.rstudio.com](http://www.rstudio.com)).

## Anlegen eines neuen Projekts in RStudio

Zunächst wird ein neues Projekt angelegt (File → New Project ... → New Directory → New Project; Abb. 1). Benennung und Speicherort können Sie frei wählen. Der Einfachheit halber wird für diesen Beitrag das Projekt „report“ (Directory name)

unter `C:\Users\{USER}\Documents\` (im Eingabedialog durch `~` automatisch abgekürzt; Abb. 2) erzeugt.



Abbildung 1: Anlegen eines neuen Projekts



Abbildung 2: Speicherort und Benennung des Projekts

Das Projekt ist zunächst leer, im File-Explorer finden Sie nur eine RStudio-Projektdatei (`report.Rproj`), die ignoriert werden kann. Daher wird als erstes über `File → New File → R Script` (oder den darunter befindlichen Button) eine neue Datei erzeugt und unter der Benennung `api_calls.R` gespeichert.



Abbildung 3: (1) Neues Skript erstellen, (2) Projekt-Datei und Skript im File-Explorer, (3) Skript-Editor zeigt die erstellte Datei an

## Installation und Laden der benötigten Packages

In der soeben erstellen `api_calls.R`-Datei werden nun die benötigten Packages installiert bzw. – falls bereits vorhanden – aktualisiert und geladen (Abb. 4; 1). Die Google-Packages kennen Sie bereits aus dem Beitrag von Tobias Aubele. Das Package `lubridate` erweitert den Funktionsumfang von R für die komfortable Arbeit mit Datums- und Zeitwerten, `jsonlite` dient zum Abfragen und Verarbeiten von JSON-Dateien. Sobald Funktionen aus diesen Packages verwendet werden, wird explizit darauf hingewiesen. Die vorangestellten Rauten kommentieren den Code aus. Beim ersten Ausführen müssen Sie diese also entfernen. Markieren Sie dann den gesamten Code mit der Maus und führen Sie ihn durch Drücken der STRG- und ENTER-Tasten aus.



Abbildung 4: Installieren bzw. Aktualisieren und Laden der benötigten Packages

In der anschließenden `options()`-Funktion definieren Sie Ihre Google-Client-ID sowie Ihr -Secret (2). Diese können Sie sich unter `console.developers.google.com` generieren. Das Argument `scopes` definiert die Berechtigungen für die GA- bzw. GSC-API. `gar_auth()` dient schließlich der Authentifizierung (3). Führen Sie den Code das erste Mal aus, öffnet sich Ihr Browser, in dem Sie den Zugriff des Skripts bestätigen müssen. Ist dies erfolgt, finden Sie in Ihrem File-Explorer eine Authentifizierungsdatei (`sc.oauth`). Bei einem erneuten Ausführen des Skripts verwendet `gar_auth()` diese Datei, sodass Sie nicht erneut eine manuelle Authentifizierung durchführen müssen – praktisch, denn zukünftig soll das Skript ja automatisch laufen.

## Konfiguration wichtiger Variablen

Nun folgt ein Code-Abschnitt, in dem einige Variablen definiert werden (Abb. 5), die an verschiedenen Stellen des Skripts zum Einsatz kommen. Dass die Variablen hier großgeschrieben werden, ist eine Konvention des Autors, die anzeigt, dass die Variablen erst zu einem späteren Zeitpunkt genutzt werden – grundsätzlich könnten Sie die Variablen auch kleinschreiben. Dies ist somit die Stelle im Skript, an der Sie Ihre eigenen Daten eintragen müssen.



Abbildung 5: Definition globaler Variablen, für die die Daten von den APIs abgefragt werden sollen

Im Detail:

(1) Um die passende View-ID Ihrer GA-Datensicht zu ermitteln, können Sie die auskommentierte Funktion `ga_account_list()` `%>% View()` verwenden. Der Pipe-Operator `%>%` übergibt den von der API zurückgegebenen DataFrame mit den Account-Details an die Funktion `View()`, welche die Tabelle in einem Explorer statt auf der Console öffnet (Abb. 6). Falls der letzte Satz nur ein böhmisches Dorf für Sie war, seien Sie beruhigt, eine

ausführlichere Erklärung folgt im nächsten Artikel. Wichtig ist erst einmal nur, dass Sie Ihre View-ID definieren. Entnehmen Sie dazu die gewünschte View-ID aus der Spalte `viewId` und tragen Sie sie bei der Variable `GA_VIEW_ID` ein.

(2) Analog dazu verfahren Sie mit der `siteUrl` für die Variable `GSC_PROP`.

(3) Die Sistrix-Daten sollen später auf Domain-Ebene abgefragt werden, entsprechend tragen Sie bei `SISTRIX_DOMAIN` diese ein. Ihren Sistrix-API-Key für die Variable `SISTRIX_APY_KEY` finden Sie unter [app.sistrix.com/account/api](https://app.sistrix.com/account/api).

(4) `START_DATE` und `END_DATE` definieren den abzufragenden Berichtszeitraum. In der aktuellen Konfiguration umfasst dieser sechs Monate rückwirkend vom Beginn des aktuellen Monats bis zum heutigen Datum. `floor_date(today(), „month“)` `%m-% months(6)` bedient sich dreier Funktionen aus dem Package `lubridate`. `today()` gibt das heutige Datum zurück („2018-01-11“). `floor_date(„2018-01-11“, „month“)` rundet dann zum Beginn des aktuellen Monats („2018-01-01“). `%m-% months(6)` subtrahiert von diesem Datum sechs Monate („2018-07-01“). `%m-%` ist ein besonderer Operator aus dem Package, der das Subtrahieren von Monaten unter Berücksichtigung unterschiedlicher Tageszahlen im Monat und Schaltjahren unterstützt. Sie landen mit dieser Funktion also immer am ersten eines Monats.

Die definierten Variablen können Sie nun im Environment sehen (Abb. 7).



Abbildung 6: Data-Explorer mit den GA-Accounts



Abbildung 7: Environment oder Variable-Explorer listet alle zugewiesenen Variablen

# GA-Sessions abfragen

Wenn die Konfiguration abgeschlossen ist, kann mit der ersten API-Abfrage begonnen werden. Als Erstes sollen die Sessions, die aus den Channels Organic Search resp. Direct resultieren, aus GA abgerufen werden (Abb. 8). Um die Abfrage auf diese Dimensionen einzuschränken, werden mittels `dim_filter()` zwei Filter-Komponenten konfiguriert und in der Variablen `channel_filter_organic` bzw. `channel_filter_direct` gespeichert (1). Diese werden an eine weitere Funktion `filter_clause_ga4()` gegeben, welche intern die Filter in eine für die API verständliche Syntax übersetzt (2). Die Funktion `google_analytics()` führt schließlich die eigentliche Abfrage aus (3). Hier sehen Sie auch die Variablen für die GA-Datensicht sowie das Start- und Enddatum, die Sie zuvor konfiguriert haben. Führen Sie den Code aus, erscheint in Ihrem Environment die Variable `ga_sessions` (Abb. 9; 1), die Sie mit der Maus anklicken können und die einen DataFrame mit der API-Antwort enthält (2).



Abbildung 8: API-Abfrage der GA-Sessions



Abbildung 9: DataFrame mit den GA-Sessions

# GSC-Performance-Daten abfragen

Ähnlich funktioniert die anschließende Abfrage der GSC-API (Abb. 10). Von dieser werden verschiedenen Dimensions-Kombinationen abgefragt. (1) Die erste Abfrage ruft die Clicks, Impressions, CTR und Position für den searchType „web“ auf Tagesbasis ab. Auch hier sehen Sie die Verwendung der anfangs definierten Variablen. (2) Dann erfolgt die Abfrage der Metriken für die Kombination aus Datum und Suchanfrage sowie schließlich (3) für die Kombination aus Datum und URL. Auch hier finden Sie nach dem Ausführen der API-Calls die Antwort-Tabellen in Ihrem Environment.



Abbildung 10: API-Abfrage der GSC-Metriken für verschiedene Dimensionskombinationen

## Sistrix-Sichtbarkeitsindex abfragen

Zu guter Letzt müssen Sie noch selbst eine API-Abfrage schreiben, denn es gibt natürlich nicht für alle Datenquellen fertige Packages. Sistrix ist so ein Fall, der aber aufgrund der sehr einfachen Abfrage-Syntax ohne großen Aufwand umgesetzt werden kann. Die Abfrage ist nämlich nur eine URL, deren Query-Parameter die gewünschten Abfragewerte aufnehmen. Wird die URL aufgerufen, erhalten Sie als Antwort eine JSON-Datei. Diese URL bauen Sie mit der Funktion `paste0()` zusammen (Abb. 11; 1). Konkret verbindet die Funktion die in Klammern angegebenen Argumente. Diese bestehen aus den Bestandteilen der Basis-URL für die Abfrage des historischen Sichtbarkeitsindexes (SI), in die die zuvor definierten Variablen für Ihren API-Key sowie die abzufragende Domain eingefügt werden.

Anschließend fragen Sie die zusammengebauten URLs gegen die API an (2). Dazu verwenden Sie die Funktion `fromJSON()` aus dem Package `jsonlite`. Die Funktion führt einen GET-Request aus, nimmt die JSON-Antwort der API entgegen und überführt sie automatisch in eine Liste – ein R-Datentyp zum Speichern hierarchischer Daten. Die Antwort wird zunächst in die jeweiligen Variablen geschrieben. Klicken Sie eine der beiden Variablen im Environment an, sehen Sie, dass das JSON in eine Listen-Struktur überführt wurde (Abb. 12). Hier sehen Sie auch schon den „Weg“, den Sie innerhalb der Liste gehen müssen, um an die SI-Daten heranzukommen. Denn neben diesen erhalten Sie noch weitere Meta-Informationen, wie bspw. die verbrauchten Credits, zurück. Um die Tabelle mit den SI-Daten zu extrahieren, müssen Sie sich an den übergeordneten Listenpunkten entlanghangeln. Dies macht der Ausdruck `api_response_desktop$answer$sichtbarkeitsindex[[1]]` für den

Desktop-SI und überführt den DataFrame in die Variable `si_desktop` (Abb. 11; 3).



Abbildung 11: Konstruktion der Sistrix-API-Abfrage



Abbildung 12: Liste mit der Antwort der Sistrix-API

## Fazit und Ausblick

Damit haben Sie das Ziel dieses Beitrags erreicht. Sie haben ein reproduzierbares Skript geschrieben, das Sie jederzeit erneut ausführen können. Der Berichtszeitraum der abzufragenden Daten passt sich bei jeder erneuten Ausführung automatisch an. Als Resultat liegen Ihnen die GA-, GSC- und Sistrix-Daten in tabellarischer Form vor. Im nächsten Artikel dieser Reihe wird es dann um die Aufbereitung dieser Daten gehen. Aktuell liegen die Daten auf Tagesbasis vor. Für den Report sollen bspw. die GSC-Daten auf Query- resp. URL-Basis aggregiert werden, um somit Tabellen der Top-10-Suchphrasen bzw. URLs der Vorwoche zu erstellen.

---

## R4SEO: Automatisierte Reports mit R generieren (Teil 2)



### R4SEO: Automatisierte Reports mit R generieren (Teil 2) – websiteboosting.com

Im ersten Teil dieser Artikelserie in der letzten Ausgabe hat Ihnen Patrick Lürwer gezeigt, wie Sie mit R die Google-Analytics- und Search-Console- sowie SISTRIX-API abfragen

können. Im vorliegenden Teil werden Sie lernen, die Daten so aufzubereiten, dass sie zur Darstellung im Report geeignet...

**Im ersten Teil dieser Artikelserie in der letzten Ausgabe hat Ihnen Patrick Lürwer gezeigt, wie Sie mit R die Google-Analytics- und Search-Console- sowie SISTRIX-API abfragen können. Im vorliegenden Teil werden Sie lernen, die Daten so aufzubereiten, dass sie zur Darstellung im Report geeignet sind. Dazu werden Sie fehlende Datumsunkte in den DataFrames ergänzen. Sie werden den gleitenden Mittelwert für verschieden Metriken berechnen, um starke Schwankungen der Tagesdaten zu glätten. Außerdem erfahren Sie, wie Sie DataFrames transponieren, um sie von einem weiten in ein langes Format zu bringen.**

Im letzten Artikel haben Sie Daten aus drei unterschiedlichen Quellen abgefragt, sodass Ihnen diese nun in tabellarischer Form (DataFrames) in R vorliegen. Aber wie das so häufig ist, haben die APIs die Daten nicht so geliefert, wie sie für die Report-Erstellung benötigt werden. Beispielsweise liefert die Google-Search-Console-API (GSC) nur Datumsunkte zurück, an denen auch tatsächlich Impressions bzw. Clicks stattgefunden haben. Es fehlen folglich die Tage ohne Impressions. Außerdem liegen die Daten auf Tagesbasis vor und sind demzufolge unter Umständen sehr volatil. Sie wollen sie daher glätten, um Trends besser erkennen zu können. Es ist somit unabdingbar, dass Sie die Daten „aufräumen“ und in eine entsprechende Form bringen, um sie für die Visualisierung verwenden zu können. Im Englischen spricht man in dem Fall von *tidying* – daher auch die Benennung des Packages *tidyverse*, das Sie verwenden werden. Viele der darin enthaltenen Funktionen dienen allein dazu, Daten zunächst in die richtige Form zu bringen, um mit ihnen arbeiten zu können.

## **Neues Skript erstellen sowie Daten**

## und Packages laden

Bevor Sie mit dem Aufräumen beginnen, generieren Sie in Ihrem Projekt, das Sie im Zuge des letzten Artikels angelegt haben, ein neues R-Skript mit der Benennung *data\_tidying.R* und kopieren Sie den gesamten Inhalt des Skripts *api\_calls.R* hinein. Führen Sie das Skript aus, damit die aktuellen Daten von den APIs abgefragt werden und zur Weiterverarbeitung in DataFrames in Ihrem Environment vorliegen (Abb. 1).



Abbildung 1: Environment mit geladenen API-Daten

Im Anschluss installieren und laden Sie wieder einige Packages, die für die nachfolgenden Datentransformationen benötigt werden (Abb. 2). *tidyverse* (<https://www.tidyverse.org/>) ist eine sehr mächtige Sammlung von Packages für die Datenmanipulation. In diesem Package ist auch *ggplot* enthalten, welches Sie schon aus dem Artikel von Tobias Aubele kennen und das zum Plotten von Daten dient. Darüber hinaus erweitert es die R-Syntax um den Pipe-Operator (`%>%`), den Sie im vorherigen Artikel kurz verwendet haben. *zoo* und *padr* dienen dem Arbeiten mit Zeitreihen.

## Tipp

Eine sehr anschauliche Einleitung in die Syntax und gängige Funktionen des *tidyverse* finden Sie unter folgendem Link. Das dort beschriebene Package *dplyr* ist ein Bestandteil des *tidyverse* und dient der Datenmanipulation, wie Sie sie im Folgenden selbst durchführen werden: [speakerdeck.com/omaymas/data-manipulation-with-dplyr-first-steps](https://speakerdeck.com/omaymas/data-manipulation-with-dplyr-first-steps)



Abbildung 2: Installieren bzw. Aktualisieren und Laden der benötigten Packages

# Google-Search-Console-Daten vervollständigen

Wie bereits angesprochen, gibt die GSC-API nur Datumspunkte zurück, an denen auch tatsächlich mindestens eine Impression stattgefunden hat. Werden die Daten nur mit der Dimension *date* abgefragt, wie Sie es für die Variable *gsa\_dim\_date* getan haben, liegen meistens für alle Tage Daten vor, denn hier wird die gesamte Website betrachtet. Wird die Dimension *date* jedoch um die *query* erweitert, ist dies nicht mehr der Fall, denn nicht jedes Keyword wird an jedem Datum des Berichtszeitraums eine Impression generiert haben. Dies sehen Sie in Abbildung 3, in der die Anzahl der rankenden Keywords gezählt wurde. Um die Tabelle selbst zu reproduzieren, können Sie den Befehl, den Sie im Screenshot sehen, in der Console ausführen. Dadurch ist er kein Bestandteil des Skripts, sondern dient nur dazu, die Daten on-the-fly anzuzeigen. 2018-07-02 hat nur ein Keyword Impressions generiert, 2018-07-05 zwei. An den Tagen dazwischen gab es keine Rankings, sodass keine Datumspunkte von der API zurückgegeben wurden.



Abbildung 3: Zählen der Queries je Tag; fehlende Datumswerte an Tagen ohne Impressions

Wie geschrieben werden fehlende Datumspunkte bei Abfragen, die allein aus der Dimension *date* bestehen, recht selten vorkommen, da dies bedeuten würde, dass keine einzige Seite an diesem Tag eine Impression generiert hat. Sie müssen sich aber stets bewusst darüber sein, dass dieses Verhalten auftreten kann. Und dann möchten Sie dies auch bestimmt in Ihren Diagrammen sehen. Denn wenn an einem Tag keine Seite ein Ranking aufweisen kann, liegt definitiv etwas im Argen. Solche fehlenden Werte müssen Sie also explizit machen, denn standardmäßig werden beim Plotting Werte für fehlende Datumspunkte interpoliert. In Abbildung 4 sehen Sie einen exemplarischen Verlauf von Impressions je Tag im Monat, wobei die Anzahl der Impressions der Einfachheit halber konstant

ist. Für die Datumspunkte 5, 8, 12 / 13, 24 und 30 liegen keine Werte vor. Die blaue Linie (1) wurde ohne ergänzte Datumspunkte geplottet. Hier sehen Sie, dass die Linie zwischen den bestehenden Datumspunkten einfach weitergezogen wurde. Der fehlende Datumspunkt 5 wird durch Interpolation (vereinfacht: dem Mittelwert der beiden bekannten umliegenden Werte) berechnet. Bei der roten Linie (2) wurden die fehlenden Datumspunkte ergänzt, allerdings keine Ersatzwerte eingetragen. Dadurch bricht die Linie an diesen Punkten ab. Die fehlenden Werte liegen nun explizit in den Daten vor. Das Zielbild ist aber die grüne Linie (3). Bei dieser wurden die Datumspunkte ergänzt und die fehlenden Werte durch 0 aufgefüllt. Hier fällt somit bei der Betrachtung des Graphen sofort ins Auge, dass die Impressions auf 0 abgestürzt sind.

Das Auffüllen ist im Übrigen nicht nur beim Plotting wichtig, sondern auch bei der Aggregation der Daten auf einer höheren Zeitebene. Soll beispielsweise der Durchschnitt der Impressions auf Monatsbasis berechnet werden, verfälschen fehlende Werte das Ergebnis. Angenommen, jeder Datumspunkt im Diagramm zeigt 10 Impressions an, wäre der Mittelwert für die blaue Linie 10 ( $10 \text{ Impressions} \cdot 25 \text{ Tage} / 25 \text{ Tage}$ ), für die grüne Linie hingegen 8,06 ( $(10 \text{ Impressions} \cdot 15 \text{ Tage} + 0 \text{ Impressions} \cdot 6 \text{ Tage}) / 31 \text{ Tage}$ ). Dies müssen Sie daher immer berücksichtigen, wenn Sie mit den Rohdaten aus der GSC arbeiten.



Abbildung 4: Fehlende Datumspunkte und Werte auffüllen, um Verfälschungen des Graphen zu vermeiden

Zurück zum Skript: Nachdem Sie die Packages geladen haben, fügen Sie den Code-Abschnitt, wie in Abbildung 5 zu sehen, ein. Der DataFrame *gsa\_dim\_date*, der die Performance-Daten auf Tagesbasis enthält, wird an die Funktion *pad()* aus dem Package *padr* gegeben. Die Funktion erkennt automatisch die Datumsspalte und füllt die fehlenden Datumspunkte auf. Über die Argumente *start\_val* und *end\_val* definieren Sie die untere und obere Grenze des aufzufüllenden Berichtszeitraums. Die

hierfür verwendeten Variablen `START_DATE` und `END_DATE` sind erneut jene, die Sie bereits bei den API-Abfragen genutzt haben. Nach diesem Schritt liegen alle Datumspunkte vor, allerdings enthalten die Spalten `clicks` und `impressions` noch keine Werte für diese Punkte. Diese geben Sie durch die anschließende Funktion `replace_na()` aus dem Package `tidyverse` an. Der bearbeitete DataFrame wird wieder in die Variable `gsa_dim_date` zurückgeschrieben.



Abbildung 5: Fehlende Datumspunkte ergänzen und Metriken mit 0 auffüllen

In Abbildung 6 sehen Sie die einzelnen Schritte noch einmal im Detail anhand eines exemplarischen DataFrames mit Daten für eine Woche. (1) sind die Rohdaten, bei denen drei Tage fehlen. Bei (2) wurden mittels `pad()` die fehlenden Datumspunkte ergänzt, die Werte für die Metriken fehlen aber noch (`NA`; not available). In (3) wurden diese Metriken via `replace_na()` mit dem Wert 0 aufgefüllt.



Abbildung 6: Exemplarischer DataFrame mit aufgefüllten Werten

## Gleitenden Mittelwert der Impressions und Clicks berechnen

Für die Performance-Daten liegt Ihnen nun der gesamte Berichtszeitraum im DataFrame `gsa_dim_date` auf Tagesbasis vor. Bei der Visualisierung sorgt dies jedoch dafür, dass der Graph stark „zappeln“ wird, wenn bspw. die Werte für die Impressions starken Schwankungen unterliegen. In Abbildung 7 (1) sehen Sie dies für exemplarische Daten eines Jahres. Hier einen Trend zu erkennen, ist somit sehr schwer, da er im Rauschen des Graphen untergehen kann. Ein gängiges Mittel, um den Graphen zu glätten, ist daher, den gleitenden Mittelwert für ein Zeitfenster von sieben Tagen zu berechnen (2).

# Hinweis

Wie Sie die Daten visualisieren, erfahren Sie in einem späteren Teil. Hier dienen die Plots zunächst nur der Veranschaulichung, warum die Daten wie beschrieben transformiert werden.



Abbildung 7: Plotting der Rohdaten versus gleitenden Mittelwert

Sie müssen sich dazu vorstellen, dass von einer gegebenen Zeitreihe fortschreitend sieben Tage betrachtet werden und für diese der Mittelwert gebildet wird. In Abbildung 8 sehen Sie dies beispielhaft für ein Zeitfenster von drei Tagen. Zunächst wird der Mittelwert für die Tage 1 bis 3 berechnet, dann für die Tage 2 bis 4 usw.



Abbildung 8: Exemplarische Berechnung des gleitenden Mittelwerts für ein Zeitfenster von drei Tagen

Den gleitenden Mittelwert berechnen Sie mit dem Code-Abschnitt, den Sie in Abbildung 9 sehen. Der DataFrame *gsa\_dim\_date* wird an die Funktion *mutate()* gegeben, die eine bestehende Spalte bearbeiten oder eine neue Spalte hinzufügen kann. In diesem Fall werden die Spalten *rollmean\_clicks* und *rollmean\_impressions* hinzugefügt. Die Werte der beiden Spalten werden mit der Funktion *rollmean()* aus dem Package *zoo* berechnet. Das erste Argument *x* nimmt die Spalte auf, auf deren Basis der gleitende Mittelwert berechnet werden soll. Das Argument *k* definiert das Zeitfenster – hier sieben Tage. Mit *fill* kann der Wert angegeben werden, der in die Zellen geschrieben werden soll, für die aufgrund der Reduktion der Daten keine Werte mehr vorliegen. Das sehen Sie ebenfalls in Abbildung 8. Für die Datumspunkte 1 und 2 kann kein Mittelwert berechnet werden, da das Zeitfenster ansonsten links über den Rand des Datumsbereichs hinauslaufen würde. Entsprechend werden diese beiden Datumspunkte mit *NAs* aufgefüllt. Das Argument *align* steuert dieses Verhalten. Es zeigt an, ob der

berechnete Mittelwert am rechten Rand des Berichtszeitraums endet und demnach links mit *NAs* aufgefüllt werden soll oder ob am Datumspunkt 1 mit dem Mittelwert begonnen wird und somit die *NAs* am rechten Rand eingetragen werden.

Für die exemplarischen Daten der Abbildung 8 sähe die finale Tabelle wie in Abbildung 10 dargestellt aus.



Abbildung 9: Berechnung des gleitenden Mittelwerts für Clicks und Impressions



Abbildung 10: Beispieltabelle mit berechnetem gleitendem Mittelwert

## Umwandlung des DataFrames von einem weiten in ein langes Format

Damit folgt der letzte Schritt zum Aufräumen der Performance-Daten. Ihr DataFrame sieht zum jetzigen Zeitpunkt wie in Abbildung 11 aus. Man spricht in diesem Fall von einer weiten Tabelle, denn je Datumspunkt liegen die einzelnen Metriken (*clicks*, *impressions*, *ctr* etc.) als Spalten vor. Für die Visualisierung mit *ggplot* müssen die Daten jedoch in einem langen Format vorliegen. Konkret bedeutet dies, dass die Benennungen der bisherigen Spalten die Werte einer neuen Spalte bilden. Die Tabelle wird quasi gedreht, wobei die Datumsspalte als Ankerpunkt dient. Das umgekehrte Verfahren, die Überführung einer langen in eine weite Tabelle, kennen Sie wahrscheinlich aus Excel, wenn Sie dort eine Pivot-Tabelle erstellen.

Das war jetzt sehr abstrakt: Eine schematische Darstellung mit nur vier Metrik-Spalten sehen Sie daher in Abbildung 12.



Abbildung 11: DataFrame `gsc_dim_date` mit den ergänzten Spalten für die gleitenden Mittelwerte



Abbildung 12: Umwandlung einer weiten in eine lange Tabelle mittels `gather()`

Ihren `gsa_dim_date` DataFrame überführen Sie mit dem Code aus Abbildung 13. `select()` dient dazu, nur die Spalten `date`, `rollmean_clicks` und `rollmean_impressions` beizubehalten. Die restlichen Spalten `clicks`, `impressions`, `ctr` und `position` löschen Sie somit aus dem DataFrame, da sie in der späteren Visualisierung nicht genutzt werden. Die Funktion `gather()` transponiert dann den DataFrame. Mit dem Argument `key` geben Sie den Namen der neuen Spalte an, in die die bisherigen Spaltenbenennungen als Werte überführt werden. Das Gleiche machen Sie mit dem Argument `value` für die neue Spalte, die die Werte der jeweiligen Metrik-Spalten aufnimmt. Damit haben Sie das Aufräumen der GSC-Daten erledigt. Ihr DataFrame sieht nun wie in Abbildung 14 aus. Die DataFrames mit den Queries (`gsa_dim_date_query`) resp. den URLs (`gsa_dim_date_page`) werden Sie im nächsten Teil dieser Reihe aufbereiten.



Abbildung 13: Überführen des weiten DataFrames `gsa_dim_date` in ein langes Format



Abbildung 14: Die ersten Zeilen des langen DataFrames `gsa_dim_date`

## Google-Analytics-Daten vervollständigen

Analog zu den Performance-Daten aus der GSC müssen Sie noch die Google-Analytics-Daten (GA) aufräumen. Auch bei diesen beginnen Sie damit, fehlende Datumswerte zu ergänzen. Insbesondere bei den GA-Daten ist dies eine reine Vorsichtsmaßnahme, da Sie diese für die Channels *Organic Search* und *Direct* abgefragt haben. Dass an einem Tag keine organischen oder direkten Sessions auf Ihrer Website erfolgen, ist höchst unwahrscheinlich – kann aber nichtsdestotrotz bei

sehr kleinen bzw. neuen Websites vorkommen. Um sich auf den später generierten Report verlassen zu können, empfiehlt es sich daher, immer für eine „saubere“ Datenbasis zu sorgen.

Fügen Sie den Code aus Abbildung 15 in Ihr Skript ein. Der Unterschied zum Code für die GCS-Daten aus Abbildung 5 besteht hier in dem zusätzlichen Argument *group* der Funktion *pad()*. Sie erinnern sich: Diese ergänzt fehlende Datumspunkte. Die Besonderheit des GA-DataFrames besteht darin, dass dieser bereits im langen Format von der API zurückkommt (Abb. 16). Er besteht nur aus den drei Spalten *date*, *channelGrouping* und *sessions*. Die Spalte *channelGrouping* ist vergleichbar mit Spalte *metric* im DataFrame *gsa\_dim\_date*. Sie qualifiziert die nebenstehenden Werte in der Spalte *sessions* (vergleichbar mit *value* im GSC-DataFrame) nach den Groups *Direct* und *Organic Search*. Das Argument *group* ist daher nötig, um die fehlenden Datumspunkte individuell für die beiden Ausprägungen *Direct* und *Organic Search* zu ermitteln. In der Abbildung 16 sehen Sie bspw., dass am 2018-07-01 zwar eine *Direct*-Session erfolgt ist, für *Organic Search* jedoch kein Wert vorliegt. Für 2018-07-02 verhält es sich genau andersherum. Für Ersteres muss somit der *Organic-Search*-, für Letzteres der *Direct*-Wert ergänzt werden (Abb. 17).



Abbildung 15: Fehlende Datumspunkte ergänzen und Sessions mit 0 auffüllen



Abbildung 16: Langer DataFrame *ga\_sessions* mit teilweise fehlenden Kombinationen aus *date* und *channelGrouping*



Abbildung 17: DataFrame *ga\_sessions* mit ergänzten Kombinationen aus *date* und *channelGrouping*

## Gleitenden Mittelwert der GA-

# Sessions berechnen

Im Anschluss errechnen Sie den gleitenden Mittelwert für die Sessions (Abb. 18). Auch hier müssen Sie dies separat für die beiden Channels machen. Daher erfolgt in der zweiten Zeile ein `group_by()` der Spalte `channelGrouping`, sodass die nachfolgenden Funktion `rollmean()` separat auf diese beiden Gruppen angewendet wird.



Abbildung 18: Berechnung des gleitenden Mittelwerts der Sessions je channelGrouping

## Fazit

Gratulation! Wenn Sie es bis hierher geschafft haben, sind Sie schon ein gutes Stück in die Gefilde der Datentransformation vorgedrungen. Sie haben das Wehklagen eines jeden Daten-Analysten nachvollziehen gelernt, dass Analysen zu 80 % aus Datenbereinigung und nur zu 20 % aus der eigentlich spannenden Analysetätigkeit bestehen.

Darüber hinaus haben Sie erfahren, wie Sie fehlende Datumswerte in Zeitreihen vervollständigen, um Verfälschungen bei Datenaggregationen zu vermeiden. Außerdem haben Sie den gleitenden Mittelwert verschiedener Metriken berechnet, um Graphen zu glätten und somit Trends besser erkennbar zu machen. Last but not least haben Sie ein mächtiges Werkzeug an die Hand bekommen, um Tabellen zu transponieren und sie so in ein handhabbares langes Format zu bringen.

Lesen Sie im dritten Teil in der nächsten Ausgabe, wie Sie Ihr SEO-Reporting mit R automatisieren können. So lassen sich z. B. mit wenig Aufwand die Top-SEO-Seiten nach Klicks und die Top-Suchanfragen aus der Google Search Console ermitteln und für Berichte in speziellen DataFrames zusammenfassen.

---

# Strukturierte Daten: Eine Bestandsaufnahme und ein Ausblick



## Strukturierte Daten: Eine Bestandsaufnahme und ein Ausblick – websiteboosting.com

Strukturierten Daten ermöglichen Webseiten-Betreibern, Inhalte für Suchmaschinen und deren Crawler mit einer semantischen Bedeutung auszuweisen und somit verständlicher zu machen. Der Beitrag von Matthäus Michalik und Lennart Leienbach beleuchtet, was strukturierte Daten sind, was sie können, wie... **Strukturierten Daten ermöglichen Webseiten-Betreibern, Inhalte für Suchmaschinen und deren Crawler mit einer semantischen Bedeutung auszuweisen und somit verständlicher zu machen. Der Beitrag von Matthäus Michalik und Lennart Leienbach beleuchtet, was strukturierte Daten sind, was sie können, wie man sie am besten verwendet, was sie mit SEO zu tun haben und welche Bedeutung ihnen in Zukunft noch zukommen könnte.**

Der Begriff „strukturierte Daten“ ist zunächst einmal ein sehr allgemeiner. Strukturiert liegen Daten – egal in welchem Kontext – immer dann vor, wenn sie einem bestimmten Muster, einem Format folgen, wenn sie also nach einem bestimmten Schema organisiert sind. Das Wort „Schema“ wird uns genau deshalb gleich noch einmal begegnen.

Eine bestimmte Struktur schafft Ordnung, stellt eine Interpretierbarkeit her, und genau an dieser Stelle kommt das

Internet ins Spiel – genauer gesagt: Websites und Suchmaschinen. Letztere dienen dem Zweck, einem Internetnutzer Ergebnisse zu präsentieren, die seiner Intention bei einer bestimmten Recherche möglichst genau entsprechen.

Dabei haben Suchmaschinen es in der schier unendlich großen Welt des World Wide Web mit Unmengen an Informationen zu tun. Wörter, Bilder, technische Details, Mail-Adressen, Preise – dies und vieles, vieles mehr sind Daten. Die Herausforderung besteht für Google, Yahoo und Co. darin, diese Daten wahrzunehmen, zu sortieren, in gewisser Weise „zu verstehen“ (eigentlich: zu interpretieren) und dann zum geeigneten Zeitpunkt als Suchergebnis auszugeben.

Betreiber von Websites (zum Beispiel Online-Shops) haben die Möglichkeit, den Suchmaschinen bei ihrem Job behilflich zu sein. Man muss ihnen die Interpretationsleistung nicht ganz allein überlassen. Diese Hilfestellung danken sie dem Webmaster mitunter durch aussagekräftigere Suchergebnisse, die dem Nutzer noch besser zeigen, was ein Website-Betreiber ihm auf einer bestimmten Seite bietet.

## **Das gemeinsame Projekt „Schema.org“**

2011 haben sich die weltweit bedeutendsten Suchmaschinen zusammengetan und die Initiative Schema.org ins Leben gerufen. Heute stehen Google, Yahoo, Bing (Microsoft) und Yandex hinter dem Open-Source-Projekt. Dessen Ziel ist die einheitliche Auszeichnung von Website-Inhalten mit dem Zweck, besser für die Suchmaschinen interpretierbar zu sein.

Schema.org stellt gewissermaßen ein gemeinsames Vokabular, eine Ontologie dar. Es werden in dem Verzeichnis, das jeder auf [www.schema.org](http://www.schema.org) einsehen kann, nach und nach alle erdenklichen Entitäten („things“), Handlungen und Relationen katalogisiert. Verkaufe ich beispielsweise in meinem Online-Shop Bücher, so kann ich die einzelnen Produktseiten jeweils so kennzeichnen, dass Suchmaschinen wissen: „Hier geht es um

[Book]!“

Die Entität [Book] steht normalerweise für konkrete literarische Werke mit ganz spezifischen Eigenschaften. Diese Eigenschaften kann ich den Suchmaschinen in einem zweiten Schritt näherbringen. Dabei fülle ich mehr oder weniger umfangreich (je nach meinen Ressourcen und Absichten) eine Maske aus „Properties“ aus; jedes Property steht für eine bestimmte Eigenschaft. So liegt es etwa nahe, unter anderem Buchtitel, Verfasser, Seitenzahl, Sprache, ISBN-Nummer und/oder Preis zu definieren.



Abbildung 1: Eintrag für „Book“ bei Schema.org

## Zum Vokabular fehlt noch die Grammatik

Mit Schema.org haben die „Big Player“ unter den Search Engines also eine Art stetig wachsendes Lexikon geschaffen mit Einträgen, die für das Internet besonders relevant sind. Damit aus einzelnen Vokabeln aber eine Sprache wird, braucht es außerdem Grammatik. Sie lässt es mit ihren Regeln zu, einzelne Wörter so miteinander zu verbinden, dass etwas darüber etwas ausgesagt werden kann. Genau diese Grammatik stellen die verschiedenen Auszeichnungssprachen („Annotationen“) zur Verfügung: JSON-LD, HTML Microdata und RDFa.

Jede Grammatik funktioniert dabei unterschiedlich, aber sie alle bieten mir als Website-Betreiber die Möglichkeit, meine gesamte Homepage und auch einzelne Seiten darin mithilfe strukturierter Daten auszuzeichnen. Was die Vor- und Nachteile der Auszeichnungssprachen sind und was sich derzeit aus welchen Gründen als die brauchbarste erwiesen hat, wird an späterer Stelle näher beleuchtet. Wichtig ist zunächst, dass eine Suchmaschine nun jede Menge Dinge über mich und meine Seiten weiß.

# Die Früchte der Auszeichnungssarbeit

Die Informationen, die ich der Suchmaschine leicht verdaulich übermittelt habe, sind in einigen Fällen besonders nützlich für sie – und damit auch für mich. Sie kann nämlich nun im Brustton der Überzeugung noch mehr über mein Suchergebnis sagen. So kann es mitunter dazu kommen, dass mein Suchergebnis auf einer Suchergebnisseite nicht nur als gewöhnliches Snippet aus Title, URL und Meta-Description ausgegeben wird, sondern als ein um bestimmte Informationen angereichertes Snippet, nämlich als „Rich Result“ (ehemals auch „Rich Snippet“ genannt).



Abbildung 2: Google-Desktop-Suche nach „der insasse“ mit normalen und Rich Results

Abbildung 2 zeigt einen Screenshot einer Google-Suche nach „der insasse“, dem neuen Roman von Sebastian Fitzek. Ganz oben ist ein gewöhnliches Snippet zu sehen (das von Amazon.de allerdings überhaupt nicht ordentlich gestaltet wurde), während im Fall der letzten drei Ergebnisse Rich Results geboten werden. Das Rich Result von Buecher.de bietet anstelle der eigentlichen URL eine sogenannte Breadcrumb, die es dem Suchenden leichter machen soll, die Einordnung der Seite auf der Gesamt-Homepage nachzuvollziehen. Wirklich gelungen ist das hier allerdings nicht, denn was ein Thriller, den ich überall auf der Welt in deutscher Sprache online bestellen kann und der auch nicht in Berlin spielt, nun mit „Europa > Deutschland > Berlin“ zu tun hat, weiß wohl nur Buecher.de.

Besser macht es da schon Weltbild.de, wo nicht nur eine vernünftige Meta-Description formuliert wurde, sondern die Breadcrumb tatsächlich klar macht: Das Buch ist offenbar als ein „packender Thriller“ im Shop eingeordnet. Dazu ist das Snippet auch noch mit Bewertungssternen angereichert, die sich aus den Käufer-Rezensionen auf der Seite ergeben. Thalia.de wiederum verzichtet bei seinem Ergebnis zwar auf eine

Breadcrumb-Navigation, gibt dem Suchenden aber dank Bewertungssternen ebenfalls einen ersten Hinweis zur Qualität des Buches. Thalia.de trumpft aber zusätzlich noch mit der Angabe des Preises sowie dem Lagerstatus auf.

Bin ich als Kunde kurz angebunden und möchte nur wissen, wo ich das Buch auf jeden Fall bestellen kann und was es mich kostet, dann werde ich wohl zuerst hier klicken – die Click-Through-Rate (CTR) von Thalia.de könnte also durchaus besser aussehen als die von Weltbild.de, obwohl Letztere zwei Positionen weiter oben rankt. An diesem Beispiel lässt sich gut ablesen, dass strukturierte Daten in einer Hinsicht eher nicht als SEO-Instrument taugen: Durch sie werde ich nicht unmittelbar meine Rankings verbessern können. Indirekt aber vielleicht doch, denn eine bessere CTR wird von Google durchaus wahrgenommen und unter Umständen über kurz oder lang auch mit einer besseren Platzierung belohnt.

Breadcrumb, Bewertungen, Preis, Verfügbarkeit – all dies speist sich aus strukturierten Daten. Um eines aber klar zu sagen: Dass ich strukturierte Daten anbiete, garantiert nicht, dass ich auch Rich Results erhalte. Aber anders herum ist garantiert, dass es ein Rich Result ohne strukturierte Daten für mich nicht geben wird. Eine Form von Rich Results, die vor allem im Mobile-Search-Bereich Bedeutung haben, sind sogenannte „Carousels“. Auf dem Screenshot von Abbildung 3 sieht man beispielsweise, wie Chefkoch.de seine Rezepte für „spinatlasagne“ um Bilder als strukturierte Daten angereichert hat.



Abbildung 3: Google-Mobile-Suche nach „spinatlasagne“ mit Carousel

Das belohnt Google auf dem Smartphone damit, dass der Nutzer die unterschiedlichen Rezept-Vorschaubilder durchswipen kann, bis er eines sieht, das ihn besonders anspricht. Die Chance, dass Chefkoch.de hier anstelle eines normalen, weiter unten stehenden Ergebnisses angeklickt wird, ist ziemlich hoch.

Nicht zu verwechseln sind die Rich Results übrigens mit „Featured Snippets“, bei denen Google einen größeren Content-Ausschnitt von einer Seite in einem prominenten Fenster ganz oben anzeigt. Auf Featured Snippets, die dem Suchenden ohne jeden Website-Besuch möglichst alle Fragen direkt beantworten sollen, haben strukturierte Daten keinen Einfluss.



Abbildung 4: Google-Desktop-Suche nach „spinatlasagne rezept“ mit Featured Snippet

Dass Google bei der Desktop-Suche nach „spinatlasagne rezept“ ein Featured Snippet aus der Rezeptseite von Essen-und-trinken.de anbietet (Abbildung 4), liegt also allein daran, dass die Suchmaschine den Inhalt der Seite als denjenigen bewertet hat, der die Suchanfrage am besten und prägnantesten beantwortet. Google sagt: „Du willst also Spinatlasagne backen und brauchst ein Rezept? Ich habe hier neulich eine Seite gecrawlt, die ein prima Rezept bietet, das du direkt auf der Suchergebnisseite lesen kannst. Du musst gar nicht mehr auf die Seite gehen!“

Weitere mögliche Arten der Anreicherung von Snippets zu Rich Results mittels strukturierter Daten zeigt Google in seiner Search Gallery: [einfach.st/seagal](http://einfach.st/seagal). Bevor man sich nun näher damit beschäftigt, wie strukturierte Daten technisch am besten zu implementieren sind und welche zukünftigen Potenziale in ihnen schlummern könnten, bleibt zunächst das Zwischenfazit: Rein theoretisch kann man den Suchmaschinen nie zu viele (korrekte!) Informationen über sich, seine Produkte oder Dienstleistungen liefern. Da allerdings die Ressourcen beim Integrationsprozess naturgemäß endlich sind, sollte man versuchen, die für sich wesentlichen Aspekte zu bestimmen und als „structured data“ anzubieten.

## Wie lassen sich strukturierte Daten

# implementieren?

Unterschieden werden kann zunächst einmal zwischen Spezifikationen für die Qualität der Daten und Spezifikationen für die technische Umsetzung. Zu den inhaltlichen Qualitätsanforderungen zählt, dass nur einzigartige Inhalte ausgezeichnet werden dürfen. Mit anderen Worten: Nur von mir selbst oder durch meine Nutzer erstellte Inhalte sollten ausgezeichnet werden. Des Weiteren müssen die ausgezeichneten Inhalte für Nutzer sichtbar sein. Wird bspw. ein Produkt ausgezeichnet, sollte dieses auch tatsächlich auf der Seite zu sehen sein. Ebenso dürfen keine irrelevanten oder irreführenden Inhalte ausgezeichnet werden. Darunter fallen etwa gefälschte Rezensionen oder Inhalte, die nichts mit dem Hauptinhalt einer Seite zu tun haben.

Aus technischer Sicht ist es natürlich eine Grundvoraussetzung, dass alle Inhalte, die in den Suchergebnissen gelistet werden sollen, für Suchmaschinen und ihre Bots aufrufbar sind. Das heißt, der Zugang sollte nicht durch Einträge in der robots.txt oder mittels anderer Methoden blockiert werden. Für die Implementierung sollten zudem JSON-LD, RDFa oder Microdata verwendet werden. Um etwaige Fehler und Doppel-Auszeichnungen zu vermeiden, ist zu empfehlen, sich auf eines der drei Formate zu beschränken. Die drei Annotationsformate unterscheiden sich dabei wie folgt:

**JSON-LD** (JavaScript Object Notation for Linked Data) ist die jüngste Methode, strukturierte Daten auszuzeichnen. Zudem ist es das Format, das Google regelmäßig empfiehlt und am breitesten unterstützt. Einer der größten Vorteile von JSON-LD ist zudem, dass es – im Gegensatz zu anderen Formaten – unabhängig vom restlichen Quellcode einer Internetseite eingebunden werden kann. Wie unter anderem auch die Meta-Description wird auch JSON-LD im <head>-Bereich (dem „Kopf“) eines HTML-Dokumentes implementiert. Die dadurch gewonnene Unabhängigkeit vom tatsächlichen Content der Seite ermöglicht

die einfache Anpassbarkeit des Seiten-Templates sowie der strukturierten Daten – und zwar ganz unabhängig voneinander.

Die Implementierung mittels **RDFa** basiert auf einem etwas anderen Ansatz: Anstatt die strukturierten Daten wie bei JSON-LD in einem (auch für Menschen) leicht lesbaren Block zusammenzufassen, werden sie über den gesamten Inhalt einer Seite verteilt. Bei RDFa handelt es sich um „HTML5 Extensions“. Bestehende HTML-Elemente werden erweitert und mit zusätzlichen Informationen versehen, etwa zur Auszeichnung strukturierter Daten.

Auch bei **Microdata** kommen HTML-Erweiterungen zum Einsatz. Microdata gilt jedoch als veraltet und wird nicht mehr aktiv weiterentwickelt. Dennoch setzen nach wie vor einige Websites das Format ein. Es schadet deshalb nicht, schon einmal davon gehört zu haben.

## **Welche Form der Auszeichnung sollte genutzt werden?**

Für den Fall, dass bisher noch keine strukturierten Daten auf einer Website eingebunden sind, empfiehlt sich eindeutig die Nutzung von JSON-LD. Aufgrund der Unterstützung von Google, Bing (seit Juli 2018) und Baidu ist man für drei der großen vier Suchmaschinen somit gewappnet. Einzig die russische Suchmaschine Yandex unterstützt bislang nur Microdata sowie RDFa. Sollten bereits Inhalte mithilfe von Microdata oder RDFa ausgezeichnet worden sein, kann man dies selbstverständlich beibehalten. Es sollte jedoch regelmäßig überprüft werden, wie lange die Suchmaschinen das Format in Zukunft noch unterstützen.

## **Weitere Anforderungen:**

Für einige der Markups gelten zusätzliche Anforderungen, die bei der Implementierung berücksichtigt werden sollten. Diese speziellen Richtlinien werden wie oben schon erwähnt im

„Feature Guide“ im Google-Developers-Portal (<http://einfach.st/seagal>) benannt.

## **Häufige Fehler bei strukturierten Daten**

Bei der Vielzahl an verschiedenen Richtlinien können bestimmte Fehler in den Markups recht schnell vorkommen. Selbst Googles Test-Tool für strukturierte Daten kann nicht immer hilfreich sein. Einige dieser schnell zu übersehenden, aber leicht zu vermeidenden Fehler sind die folgenden:

### **Auszeichnen von Bewertungen oder Rezensionen von Drittanbietern**

Auch wenn die Verwendung von Widgets von Drittanbietern zur Einbettung von Bewertungen im Allgemeinen kein Problem darstellt, gibt es in Bezug auf strukturierte Daten einige Punkte zu beachten. Besonders wichtig ist es, dass sich das Review-Markup immer auf den Hauptinhalt einer Seite bezieht. Falsch wäre es z. B., wenn die allgemeine Bewertung eines Shops (per Widget) seitenweit angezeigt und dadurch auch für jede einzelne Produktseite ausgezeichnet würde. Da die Bewertung sich auf das Unternehmen selbst, nicht jedoch auf die einzelnen Produkte bezieht, würde Google dies als Spam ansehen. Zudem müssen Bewertungen und Rezensionen, die nicht von der eigenen Website stammen, immer einen Link zur Quelle enthalten. Benutzern und Searchbots muss gleichermaßen gezeigt werden, woher die Bewertungen stammen.

### **Inhalte, die für den Benutzer nicht sichtbar sind**

Ausgezeichnete Inhalte müssen für Desktop- und Mobile-Nutzer auch grundsätzlich sichtbar sein. Der Nutzer muss immer klar erkennen können, wo sich die Inhalte auf einer bestimmten

Seite befinden, die durch ein Rich Result in den Suchergebnissen angeteasert wurden. Allerdings erlaubt es Google ausdrücklich, wenn aus Usability-Gründen die entsprechenden Inhalte auf der mobilen Seite zunächst in einem Tab „versteckt“ sind und erst nach einem Klick durch den Nutzer sichtbar werden. Voraussetzung ist natürlich auch hier, dass der Suchmaschinen-Bot den „Tabbed Content“ auch auslesen kann.

## **Irreführende oder ungeeignete Inhalte**

Es sollte immer darauf geachtet werden, dass die ausgezeichneten Inhalte auch korrekt sind. „Katzenspielzeug, das Ihre Kätzchen lieben werden“ ist für Google bspw. kein gültiger Produktname. Besser wäre „Kratz- und Kletterbaum in Schwarz für Katzen“. Ein weiteres Negativbeispiel bestünde darin, eine Pauschalreise inklusive Flug, Hotel und Tickets für ein Konzert der Foo Fighters als Veranstaltung (Event) auszuzeichnen. Die Kennzeichnung des Konzertes selbst als Veranstaltung entspricht jedoch den Richtlinien.

## **Haben strukturierte Daten eine SEO-Relevanz?**

Die Frage, ob strukturierte Daten für die Suchmaschinenoptimierung bedeutsam sind, lässt sich zugleich bejahen und verneinen. Einen direkten Einfluss auf die Ranking-Position einer Seite in den Suchergebnissen haben strukturierte Daten meist nicht. Google und Co. bewerten die Inhalte einer Internetseite nicht anders, wenn diese ausgezeichnet wurden. In den meisten Fällen erhöhen allerdings die durch die strukturierten Daten möglich gemachten Rich Results die Click-Through-Rate (CTR, auch Klickrate) eines Ergebnisses. Suchergebnisse mit gelben Bewertungssternen etwa setzen sich allein schon visuell von normalen Snippets ab. Wurde obendrein eine gute Bewertung

erreicht, ist die Wahrscheinlichkeit dafür, dass Nutzer auf das Ergebnis klicken, weitaus höher als bei Ergebnissen ohne Bewertungssterne. Dadurch verbessern sich oftmals die generellen Nutzer-Metriken einer Seite. Indirekt können strukturierte Daten somit durchaus einen Einfluss auf SEO haben.

Eine Besonderheit stellen außerdem Rich-Result-Ergebnisse für Artikel („Article“), Rezepte („Recipe“) und einige weitere „Search Features“ dar: Werden eigene Inhalte korrekt ausgezeichnet, haben sie die Chance, in besonderen Boxen wie dem weiter oben vorgestellten „Rezept-Carousel“ ausgespielt zu werden. Solche Boxen befinden sich meist sehr weit oben auf der ersten Ergebnisseite – und weit oben zu sein, ist das Ziel von Suchmaschinenoptimierung.

## Tools für strukturierte Daten

Für strukturierte Daten gibt es mittlerweile zahlreiche Tools, die einem Webmaster beim Implementieren oder auch Überprüfen der eigenen Auszeichnungen helfen können.

Hat man gerade erst damit begonnen, sich mit dem Thema „strukturierte Daten“ zu befassen, kann Googles „Codelab“ (<http://einfach.st/codelab2>) dabei helfen, den Einstieg schnell zu meistern. Anhand praktischer Beispiele wird man hier Schritt für Schritt an das Thema herangeführt. Derzeit ist diese Hilfe allerdings nur auf Englisch verfügbar.

Für Anfänger und auch Fortgeschrittene gleichermaßen geeignet ist der „Schema Markup Generator (JSON-LD)“ von Technicalseo.com (<http://einfach.st/markupgen>). Er deckt alle gängigen Markups ab und valider Quellcode kann mit wenigen Klicks erstellt werden.

Zu den essenziellen Tools für das Testen strukturierter Daten gehört Googles „Test-Tool für strukturierte Daten“ (<http://einfach.st/markupgen2>). Damit lassen sich einzelne

Quellcode-Schnipsel oder auch ganze URLs überprüfen.

Kein Tool im engeren Sinne, aber durchaus auch praktisch ist die Website [Jsonld.com](https://jsonld.com). Frei nach dem Motto „Steal our JSON-LD“ (zu Deutsch: „Stiehl unser JSON-LD“) werden hier für verschiedenste Anwendungsfälle gut aufbereitete Quellcode-Beispiele präsentiert.

Für denjenigen, der des Öfteren verschiedene URLs auf strukturierte Daten hin überprüfen möchte, ist das folgende JavaScript-Lesezeichen für den Browser eine gute Wahl. Dazu muss im Browser zunächst ein neues Lesezeichen angelegt werden. Anstelle einer URL trägt man jedoch folgenden Code ein:

```
javascript:(function(){  
window.open('https://search.google.com/structured-data/testing  
-tool/u/0/#url='+encodeURIComponent(location.href))})();
```

Surft man anschließend auf eine Seite, die man hinsichtlich strukturierter Daten überprüfen möchte, reicht ein Klick auf das Lesezeichen. Sogleich öffnet sich in einem neuen Browser-Tab die entsprechende URL in Googles „Test-Tool für strukturierte Daten“.

Zuletzt sei auch auf die zahlreichen Plug-ins hingewiesen, die es für viele CMS (Content-Management-Systeme) gibt, etwa für WordPress. Sie sind mal mehr, mal weniger umfangreich, erledigen ihren Job jedoch in den meisten Fällen zuverlässig und vereinfachen das Implementieren strukturierter Daten. Es ist daher definitiv lohnenswert, einmal einen Blick in den Plug-in-Store des jeweiligen CMS zu werfen.

## Was kommt als Nächstes?

Seit der Einführung strukturierter Daten kamen immer wieder neue Auszeichnungsmöglichkeiten hinzu. Was sich aktuell in einer Art Beta-Phase befindet und in Zukunft eventuell bei [Schema.org](https://schema.org) einmal als Standard hinzugefügt werden könnte,

findet man auf [pending.schema.org](https://pending.schema.org). So hat Google im August 2018 bspw. die Auszeichnung für „Speakable Content“ aufgenommen und unterstützt sie seitdem: [einfach.st/speak](https://einfach.st/speak). Dabei handelt es sich um Inhalte, die gut von Voice-Assistenten und Smart Speakern vorgelesen werden können. Derzeit ist die Auszeichnung „speakable“ ausschließlich für Publisher vorgesehen, aber es ist davon auszugehen, dass sie in Zukunft für alle Websites genutzt werden kann.

Es ist sehr wahrscheinlich, dass in den kommenden Jahren immer wieder neue Auszeichnungsmöglichkeiten von Google unterstützt werden. Des Weiteren hat das Einbinden strukturierter Daten nicht mehr nur allein einen Effekt auf die klassischen Suchergebnisse bei der Google-Suche, wie am Beispiel von „Speakable Content“ ersichtlich ist. Vielmehr ist es Websites dadurch möglich, sich auch in anderen Search-Bereichen zu positionieren und somit den eigenen Traffic zu steigern.

---

## **Title-Geddon – war da was bei Ihnen?**

### **R-Leuchtungen! Teil 9**



### **R-Leuchtungen! Teil 9: Title-Geddon – war da was bei Ihnen? – [websiteboosting.com](https://websiteboosting.com)**

In den Ausgaben bis 54 bis 58 der Website Boosting konnten Sie in der Serie „R4SE0“ von Patrick Lürwer nachvollziehen, wie man die kostenlose Software R verwendet, was sie leistet und

wie man sie nutzbringend für die eigene Arbeit für SEO bzw. die Aufklärung im Online-Marketing einsetzen kann. R...

In den Ausgaben bis 54 bis 58 der Website Boosting konnten Sie in der Serie „R4SEO“ von Patrick Lürwer nachvollziehen, wie man die kostenlose Software R verwendet, was sie leistet und wie man sie nutzbringend für die eigene Arbeit für SEO bzw. die Aufklärung im Online-Marketing einsetzen kann. R wurde ja ursprünglich für Statistik entwickelt. Wer das bisher als Entschuldigung verwendet hat, es deswegen links liegen zu lassen, dem sei versichert, dass er damit komplett falschliegt. R kann für den Einsatz im Unternehmen, allen Bereichen voran „Online“, sehr viel mehr leisten, als statistische Berechnungen durchzuführen. Genau genommen ist es ein wirklich nützliches Helferlein bei allen Aufgaben im Umgang mit größeren Datenmengen, mit Daten, die man erst in eine gewisse Struktur bringen muss, und bei der automatischen oder halb automatischen Datenbeschaffung aus praktisch fast allen Quellen aus dem Web! Für die interessierten Einsteiger, aber auch für alle, die nach der Serie von Patrick Lürwer „R-Blut“ gelect haben, startete in der Ausgabe 62 die neue anwendungsorientierte Serie „R-Leuchtungen“. Sie werden in jeder Ausgabe erfahren, wie Sie ohne Programmierkenntnisse jeweils ein definiertes und in der Online-Praxis häufiger auftretendes Problem rund um das Thema Daten und Auswertungen lösen können. Und keine Sorge, die kleinen Hilfe-Tutorials nehmen Sie Schritt für Schritt an der Hand, sodass Sie auch als Neuling von der Power von R profitieren können. Was hält Sie also ab, das einfach mal auszuprobieren? Die einzelnen Schritte müssen Sie übrigens nicht im Detail verstanden haben. Um an die hilfreichen Daten für ein besseres Ranking zu kommen, müssen Sie im Prinzip nur nachmachen, was Sie hier beschrieben finden. Für dieses Beispiel brauchen Sie übrigens keinerlei kostenpflichtige Tools.

# Worum geht es diesmal?

Um den 21.08.2021 herum hat Google das sogenannte Title-Update ausgerollt. Dabei wurde in den Suchergebnissen bei nicht wenigen Seiten der im jeweiligen Quellcode hinterlegte Title durch einen von Google generierten Title ersetzt. Das führte zu einigem Aufruhr bei den aufmerksamen Sitebetreibern, weil zum Teil wohl weniger passende Title erzeugt wurden und man zum anderen einen gewissen Kontrollverlust verspürte.

Einige Tools, so z. B. die SISTRIX-Toolbox, zeigen bereits für einzelne Domains an, ob und welche Title getauscht wurden. Ebenso gibt es bereits Browser-Plug-ins, die Veränderungen direkt in den Suchergebnissen sichtbar machen. Das alles sind aber nur punktuelle Hinweise und bisher bekommt man kaum weiter verarbeitbare Daten geliefert bzw. kann die Ergebnisse nicht downloaden für Checklisten und Handlungsanweisungen.

Die Kernfrage ist, ob sich bei (nahezu) gleichbleibender Position eines Keywords die CTR, also die Klickrate, signifikant derart verändert hat, dass sich der Grund beim Title-Tausch durch Google vermuten lässt. Bleibt z. B. vor und nach dem Update die Position 3 stabil stehen, aber die Klickrate sinkt von 15 % auf 5 %, dann könnte mit hoher Wahrscheinlichkeit das veränderte Ergebnissnippet dafür verantwortlich sein, da sich sonst ja nichts verändert hat. Würde die Position steigen oder fallen, würde sich in jedem Fall bekanntlich auch immer die CTR ändern. Insofern gilt es, Rankings zu filtern, deren Position sich nicht stark verändert hat, jedenfalls nicht über eine normale Schwankungsbreite hinaus. Natürlich tritt auch bei der CTR eine gewisse Schwankung auf.

Die Aufgabe lautet also: Finde alle Rankings, deren Position gleich geblieben ist UND deren CTR sich deutlich verändert hat. Dabei sollten alle Suchphrasen, die als Besonderheit den Domainnamen (Brandbegriffe) enthalten, ignoriert werden.

Das Skript in der vorliegenden Ausgabe nimmt sich dieses Problems an und ist wie folgt strukturiert:

1. Abholen aller Rankingdaten für zwei definierbare Zeitfenster (vor und nach dem Update)
2. Definieren gewisser Schwankungstoleranzen für die CTR
3. Definieren gewisser Schwankungstoleranzen für die Position
4. Angabe der Brand-Suchbegriffe (Domainname)
5. Filtern aller ausgeschlossenen Fälle von Punkt 2 bis 4.
6. Ausgabe der übrig gebliebenen Daten in ein CSV-File zur weiteren Verwendung

Um tatsächlich umfassend mit den Rohdaten direkt von Google arbeiten zu können, muss ein sog. (kostenloses) Entwicklertoken von Google vorliegen (siehe den Infokasten am Rand).

## Ihr API-Account bei Google

Um automatisiert bzw. computergesteuert Daten von Google-Tools holen zu können, brauchen Sie dort einen „Developer-Account“. Für diesen bekommen Sie eine von Google erzeugte Mailadresse und ein File mit der Endung .json für die Authentifizierung Ihres Computers bzw. für R, um Daten zu holen. Die Mailadresse dient dazu, einen Nutzer für die Search Console anlegen zu können, und mit dem File kann sich eine Maschine gegenüber Google als berechtigt ausweisen. Wie Sie einen solchen Account anlegen können, haben wir in der Ausgabe 65 ausführlich beschrieben. Falls Sie den Account da angelegt haben, können Sie jetzt erneut darauf zugreifen und müssen nichts weiter tun.

Haben Sie das übersprungen, müssen und sollten Sie das (Schritt 1 und 2 im oben genannten Beitrag) unbedingt nachholen. Bereits in den vergangenen Ausgaben hatten wir

Ihnen immer wieder Skripts an die Hand gegeben, mit denen Sie wirklich sehr nützliche Auswertungen mit echten Google-Daten ganz einfach mit wenigen Klicks erzeugen können. Ohne diesen Account bekommen Sie leider keine ausführlichen Daten. Und ohne diese Daten können Sie natürlich auch keine eigenen und detaillierten Hinweise auf Optimierungsmöglichkeiten für Ihr Ranking erzeugen.



Abbildung 1: An einigen Stellen im Code („# TODO“) müssen einfache Anpassungen vorgenommen werden

## Schritt 1: Das R-Skript downloaden und die Parameter eingeben

Holen Sie sich unter [einfach.st/rcode6](http://einfach.st/rcode6) das Skript für diese Ausgabe 70 bzw. kopieren Sie von dort ganz einfach den Code in Ihre Zwischenablage. Starten Sie R-Studio als „Benutzerhülle“ für R. Mit File/New File erhalten Sie links oben ein neues, noch leeres Fenster für den R-Code. Kopieren Sie ihn von der Webseite direkt in dieses Fenster.

Stellen Sie sicher, dass R mit dem richtigen Arbeitsverzeichnis (Working Directory, kurz WD) arbeitet bzw. legen Sie das Verzeichnis auf Ihrem Computer fest, das für diese Session verwendet werden soll. Dazu klicken Sie sich einfach rechts unten in dem Fenster wie in Abbildung 2 unter den beiden Ziffern 1 gezeigt zu dem gewünschten Verzeichnis durch. Achtung: In diesem Verzeichnis muss dann auch die Datei (oder eine Kopie davon) von Google liegen, die man zu Authentifizierung des Computers braucht. Sie endet auf .json und enthält vorne die bei Google angegebene Mailadresse, gefolgt von einer Zahlen-/Buchstabenreihenfolge. Sie sieht in etwa so aus:

```
mein-accountname-cr4436ad9828.json
```

Stimmt das angezeigte Verzeichnis, klickt man einfach unter

dem Zahnradsymbol „More“ auf „Set As Working Directory“. Im Fenster links daneben wird dies nun bestätigt mit:

```
> setwd("~/was/sie/ausgewaehlt/haben")
```

Wie immer müssen für R dann noch bestimmte Funktionsbibliotheken (Librarys) installiert bzw. aufgerufen werden. Nach dem erstmaligen Installieren mit dem Befehl

```
install.packages("libraryname")
```

behält R diese Bibliothek in ihrem Dateispeicherbereich. Sie muss dann nicht mehr installiert werden. Der Aufruf einer Bibliothek, um damit arbeiten zu können, bzw. das Laden in den aktiven Arbeitsspeicher muss hingegen nach jedem Start von R vollzogen werden. Der Befehl dazu lautet:

```
library(libraryname)
```

Es gilt zu beachten, dass hier keine Anführungszeichen wie beim Installieren verwendet werden. Um den Aufruf einer Bibliothek in den Arbeitsspeicher muss man sich nicht aktiv kümmern, weil er in der Regel immer am Anfang im Skript steht und daher sowieso immer mit ausgeführt wird. Wer die R-Leuchtungen der letzten Ausgaben aktiv angewendet hat, hat bereits alle nötigen Bibliotheken installiert und kann den Code daher später einfach durchlaufen lassen. Bibliotheken kann man übrigens auch leicht über den Menüpunkt „Tools“ und dort „Install Packages“ einzeln installieren. R sucht sich die Speicherorte im Web von alleine.



Abbildung 2: So ist R-Studio strukturell aufgeteilt

Die Anpassungen im Skript sind dort nochmals dokumentiert (alle Zeilen mit einem führenden „#“ stellen Kommentarzeilen dar und sind kein Code).

Sie müssen an sechs Stellen Eingaben/Änderungen vornehmen. Diese sind alle mit „#TODO“ gekennzeichnet.

Zunächst müssen Sie den genauen Namen Ihrer Authentifizierungsdatei (die mit .json am Ende) angeben. Überschreiben Sie einfach die Vorgabe entsprechend:

```
KEY_FILE <- "name-der-authentifizierungsdatei.json"
```

Anschließend müssen Sie Ihren Domainnamen noch hinterlegen bzw. überschreiben:

```
GSC_PROP <- https://www.domainname.com/
```

**Optional:** Möchten Sie Suchanfragen nach Ihrem Domainnamen/Brand ignorieren? Dann überschreiben Sie einfach ebenfalls die Vorbelegungen. Brauchen Sie mehr Zeilen, kopieren Sie einfach z. B. die Zeile 2 und fügen Sie diese mit einem überschriebenen Suchwort wieder ein. Diese müssen nur jeweils durch Anführungszeichen und ein abschließendes Komma getrennt werden:

```
BRAND_QUERIES <- c("brandbegriff-1",  
  "brandbegriff-2",  
  "brandbegriff-3",  
  "brandbegriff-n")
```

**Optional:** Nun müssen bzw. können Sie noch die beiden Zeiträume für den Vergleich definieren. Vorbelegt ist:

```
START_DATE_PRE <- ymd("2021-08-10")  
END_DATE_PRE <- ymd("2021-08-16")
```

```
START_DATE_POST <- ymd("2021-09-01")  
END_DATE_POST <- ymd("2021-09-06")
```

PRE steht für „vor dem Update“ und POST entsprechend für einen Zeitraum danach. Das Startdatum ist im Beispiel hier so gewählt, dass sicherheitshalber ein Zeitraum einige Tage vor dem Title-Update fixiert wurde. Bei Bedarf können Sie beide Zeiträume entsprechend anpassen, sollten z. B. neue Updates das nötig erscheinen lassen.

# Nicht nur für Title-Updates!

Mit dem vorliegenden Skript können Sie übrigens jede Art von Klickveränderungen Ihrer Suchphrasen „überwachen“ bzw. prüfen. Da Sie den Start- und Endzeitraum des Vergleichs frei wählen können, lässt sich damit prinzipiell auch jeder Zeitraum mit jedem anderen vergleichen! Damit haben Sie für jede Art von Sonderanalyse ein recht machtvolles kleines Skript an der Hand, das Ihnen in wenigen Sekunden datengestützte Antworten liefern kann.

**Optional:** Legen Sie einen Wert fest, um den die Position schwanken darf, ohne gleich als echte Positionsveränderung zu gelten. Da Google Kommawerte bzw. die durchschnittliche Position angibt, schwankt diese eigentlich fast immer um einige Prozentpunkte hinter dem Komma (das in R mit einem Punkt gesetzt wird). Würde man diese Schwankung nicht berücksichtigen, gäbe es für jedes Keyword eine Positionsveränderung, weil bereits ein Ansteigen von Position 2,3 auf 2,4 als Änderung gelten würde. Vorbelegt ist der Wert 0,9, aber Sie können diesen Ihren Bedürfnissen entsprechend verändern. Das macht dann Sinn, wenn Sie zu viele oder zu wenige Daten in der Auswertung haben. Da dies von Domain zu Domain schwankt, kann man keine generellen, verbindlichen Werte angeben. Hier müssen Sie ggf. etwas experimentieren.

```
POSITION_FLUCTUATION <- 0.9
```

**Optional:** Ebenso macht es Sinn, einen natürlichen Korridor für Schwankungen im CTR-Wert zu berücksichtigen. Hier wurde 5 % vorbelegt. Das bedeutet, dass eine Schwankung unter oder über 5 % noch nicht als nennenswerte Veränderung gesehen wird und die Daten weggefiltert werden, die sich innerhalb dieser Breite befinden. Sollten Sie zu viele oder zu wenig Daten bekommen, versuchen Sie einfach, das Skript mit veränderten Werten noch mal laufen zu lassen.

```
CTR_FLUCTUATION <- 5
```

Das war auch schon alles, was Sie hinterlegen bzw. anpassen müssen. Klicken Sie nun einfach auf „Source“ rechts oben im linken oberen Fenster (Abbildung 3). Dies löst das Abarbeiten des gesamten Skripts auf einen Rutsch aus.



Abbildung 3: Ein Mausklick auf „Source“ startet das komplette Skript

Das Durchlaufen kann je nach Datenumfang etwas dauern, da die Daten von der Search Console im Batchbetrieb für jeden angeforderten Tag (über das Zeitintervall oben definiert) abgeholt werden (Abbildung 4). Anschließend werden alle empfangenen Daten automatisch ohne weiteres Zutun vom Skript verarbeitet.



Abbildung 4: So bzw. so ähnlich sieht das laufende Skript im Consolenfenster aus

Am Ende des Skripts sorgt noch die letzte Codezeile dafür, dass die fertigen Daten in einer CSV namens „ctr\_vergleich.csv“ im Arbeitsverzeichnis gespeichert werden:

```
write_csv2(gsc_data, "ctr_vergleich.csv")
```

Selbstverständlich können Sie hier auch einen eigenen Dateinamen vergeben, indem Sie die Vorgabe „ctr\_vergleich.csv“ im Skript einfach nach Ihren Wünschen verändern.

## **Schritt 2: Sehen Sie sich die Auswertungen an und arbeiten Sie damit**

Wenn Sie die erzeugte Datei öffnen, finden Sie alle URLs, die in dem von Ihnen definierten Zeitraum bei gleichbleibender Position (innerhalb der definierten Schwankungsbreite) eine CTR-Veränderung über der Prozentzahl haben, die Sie definiert haben. Das ist der Startpunkt für weitere Analysen. Durch die Verwendung der SEO-Tools für Excel kann man die Tabelle

problemlos mit weiteren Metriken anreichern, wie z. B. den aktuellen Title und den Inhalt der H1 dazu holen. Dem Vernehmen nach wird beim Austausch der Title von Google häufig auf den Inhalt der H1 zurückgegriffen, sofern er nicht identisch mit dem Title oder wenig aussagekräftig ist.

Beachten Sie bitte, dass Sie von der Google Search Console deutlich mehr Daten bekommen als mit vielen anderen Tools. Es ist daher nicht unwahrscheinlich, dass Sie auch für nur wenige Impressions oder Klicks Datensätze im Ergebnis bekommen. Hier sind die Änderungen der CTR natürlich statistisch nicht aussagekräftig bzw. belastbar. Sie sollten Sie daher in Excel am besten wegfiltern (Zahlenfilter/größer als).



Abbildung 5: Formatiert, nach CTR-Differenz sortiert und nach Klicks gefiltert



Abbildung 6: Einige Tools wie z. B. SISTRIX zeigen veränderte Titles in den Suchsnippets für zwei Datumsangaben an

Bei den meisten Analysen, die wir testhalber vorgenommen haben, hat sich das „Title-Geddon“, wie das Update im Netz zum Teil genannt wurde, übrigens nicht besonders stark ausgewirkt. Es gab zwar möglicherweise viele Title, die tatsächlich getauscht wurden, in eine stark messbare CTR-Veränderung mündete dies allerdings nicht. Die Fälle, die wir extrahieren konnten, zeigten bei einer manuellen Einzelfallprüfung, dass man sicherlich mit der getauschten Zeile gut leben kann.

Trotzdem gibt es im Web zum Teil heftige Beschwerden von Webmaster, dass die Algorithmen alles deutlich verschlimmert hätten. Es ist noch zu früh, stabile Aussagen dazu zu machen, vor allem weil Google angekündigt hat, nochmals nachbessern zu wollen. Trotzdem kann es natürlich nichts schaden, sich einmal in Ruhe anzusehen, ob die Performance Ihrer Suchergebnisse möglicherweise gelitten hat. Hier hilft Ihnen das vorliegende Skript!

---

# 101 Rechtsfehler im Online-Marketing – Folge II



## “101 Rechtsfehler im Online-Marketing – Teil 2 ” – [websiteboosting.com](http://websiteboosting.com)

Im Online-Marketing kann man bekanntlich viele teure Fehler machen. Für viele Marketer besonders mysteriös sind rechtliche Mängel. In den folgenden Heften wollen wir den häufigsten Fehlern auf den Grund gehen und jeweils kurz darstellen, was man alles falsch machen kann – und wie es besser geht. ...

Im Online-Marketing kann man bekanntlich viele teure Fehler machen. Für viele Marketer besonders mysteriös sind rechtliche Mängel. In den folgenden Heften wollen wir den häufigsten Fehlern auf den Grund gehen und jeweils kurz darstellen, was man alles falsch machen kann – und wie es besser geht. In Folge 2 geht es um Social Media und Display-Werbung.

## Unzureichende Sensibilisierung von Mitarbeitern für Social Media

Ein Hauptproblem bei Social-Media-Auftritten kleiner und mittlerer Unternehmen ist die unzureichende Schulung der Kolleg:innen, die die Accounts mitbetreuen. Es würde schon helfen, wenn den Mitarbeiter:innen bewusst wäre, dass Facebook & Co keine Spielwiese sind, auf der man Dinge ausprobieren kann, die für die Website „zu gefährlich“ erscheinen.

## **Zu spät bei Account-Namen**

Es ist für Unternehmen um ein Vielfaches günstiger, sich die einschlägigen Aliase bei neuen sozialen Netzwerken jeweils zu reservieren oder durch entsprechende Dienstleister reservieren zu lassen, als später mühsam die Accounts kaufen oder einklagen zu müssen. Dies gilt insbesondere für Start-ups und Unternehmen ohne starke Marken.

## **Facebook, Twitter und Co: fehlendes Impressum**

Auch bei Twitter und Facebook muss ein Impressum vorhanden sein, wenn die Social-Media-Plattformen geschäftsmäßig genutzt werden. In Konzernen ist es wichtig, sich genau zu überlegen, welche Gesellschaft Betreiberin der Social-Media-Profile sein soll.

## **Fehlende Datenschutzinformation für Facebook-Page**

Jede professionelle Facebook-Seite braucht eine Datenschutzinformation. Aus dieser muss sich ergeben, dass die Seite aus datenschutzrechtlicher Sicht gemeinsam mit Facebook betrieben wird. Außerdem sollte darin beschrieben sein, wie die Nutzer:innen ihre Rechte – etwa auf Auskunft und Löschung – geltend machen können.

## **Fehlende Einwilligung bei Mitarbeiterfotos**

Wer Angestellte auf Fotos oder in Videos auf Social-Media-Präsenzen präsentieren möchte, sollte vorher schriftlich die Einwilligung einholen oder gleich eine Art Model-Release-Vertrag mit den jeweiligen Beschäftigten schließen. Wichtig

ist, dass denjenigen, die das nicht möchten, kein Nachteil entsteht.

## **Firmenevents auf Facebook & Co**

Wer vorhat, Firmenevents – wenn diese denn irgendwann einmal wieder stattfinden können – in sozialen Netzwerken zu bebildern, muss dafür sorgen, dass die Abgebildeten davon wenigstens wissen. Am besten ist es, schon bei der Einladung darauf hinzuweisen und auch am Veranstaltungsort entsprechende Aushänge zu machen. Wer sich nicht auf Facebook sehen möchte, hat ein Widerspruchsrecht. Wichtig ist in jedem Falle: Augenmaß bei der Auswahl der Bilder.

## **Selbstbewertungen**

Nur weil die Wahrscheinlichkeit, entdeckt zu werden, gering scheint, heißt das nicht, dass Unternehmen sich selbst positiv bewerten sollten. Es ist als Schleichwerbung wettbewerbswidrig. Auch eine Incentivierung positiver Bewertungen durch eigene Mitarbeiter:innen sollte unterbleiben. Im Übrigen kommt insbesondere systematische Selbstbeweihräucherung meist doch heraus, sei es über ausgeschiedene Kollegen, verärgerte Agenturen oder deren Mitarbeiter.

## **Rabatt gegen Bewertungen**

Die Rechtsprechung sieht es außerordentlich kritisch, wenn Kunden Vorteile für die Abgabe einer Bewertung geboten werden. Evident rechtswidrig ist das Versprechen eines Rabatts für eine positive Bewertung. Vertragliche Vereinbarungen, etwa in AGB, negative Bewertungen zu unterlassen, sind unwirksam.

# **Werbung mit Preisen in Social Media ohne Pflichthinweis**

Die Preisangabenverordnung gilt natürlich auch auf Facebook und Instagram. Wer dort gegenüber Privatleuten mit Preisen wirbt, muss mitteilen, dass die Preise inklusive Mehrwertsteuer sind und ob gegebenenfalls Versandkosten hinzukommen.

## **Anzeigen auf Facebook ohne Pflichtangaben**

Auch Werbung in sozialen Netzwerken ist Werbung. Gibt es spezielle branchenspezifische Regeln für die Ausgestaltung von Werbung, müssen diese auch auf Twitter und Facebook eingehalten werden.

## **Laxer Umgang mit beleidigenden Kommentaren**

Unternehmen sind für Nutzerkommentare verantwortlich, sobald sie Kenntnis davon haben. Enthalten etwa die Kommentare von YouTube-Videos Beleidigungen oder falsche Behauptungen, sind Unterlassungsansprüche der Betroffenen die Folge. Zensurvorfürworf hin oder her: Rechtswidrige Kommentare müssen gelöscht werden, sobald der Betreiber der Seite davon Kenntnis erlangt.

## **Verwendung von User-Fotos ohne ausreichende Rechte**

Es ist unzulässig, Fotos von Nutzer:innen zu verwenden, ohne dass diese damit einverstanden sind. Die Fotograf:innen haben stets das Recht zu bestimmen, ob und wo ihre Fotos

veröffentlicht werden. Im Zweifel braucht es also eine ausdrückliche Erklärung von User:innen, wenn deren Fotos in Social-Media-Präsenzen von Unternehmen eingesetzt werden sollen.

## **Influencer-Marketing ohne Werbekennzeichnung**

Jede nicht ohne Weiteres als solche erkennbare Werbung muss gekennzeichnet werden. In Verträgen mit Influencern müssen Unternehmen auf Transparenz drängen, um als Auftraggeber nicht selbst verklagt zu werden.

## **Unscheinbare Werbekennzeichnung von Influencern**

Die Rechtsprechung ist streng und verlangt eine deutliche Kennzeichnung von Influencer-Marketing als Werbung. Ein unscheinbares #ad am Ende einer Wolke aus Hashtags genügt nicht. Besser ist eine deutliches #Werbung oder #sponsored.

## **Anbieten von Glücksspielen**

Die Grenze zwischen erlaubtem Gewinnspiel und verbotenem Glücksspiel sind teilweise fließend. Man muss also darauf achten, dass das Marketing keine Spiele veranstaltet, bei denen die Spieler einen finanziellen Einsatz leisten müssen, um teilnehmen zu können. Einen Like oder Kommentar zu hinterlassen, zählt allerdings nicht als geldwerter Einsatz, der ein Gewinnspiel zu einem illegalen Glücksspiel macht.

## **Unzureichende Teilnahmebedingungen**

## **bei Gewinnspielen**

Es gibt keinen Zwang, für jedes Gewinnspiel Teilnahmebedingungen vorzusehen. Allerdings empfiehlt sich das häufig, um die gesetzlichen Pflichtangaben unterzubringen. Diese sollten dann alle wesentlichen Angaben zu dem Gewinnspiel enthalten.

## **Verwendung von Gewinnspieldaten zu Werbezwecken**

Die Tatsache, dass jemand an Ihrem Gewinnspiel teilgenommen hat, rechtfertigt nicht, dieser Person Werbung zuzusenden. Die Nutzung von Daten zu Werbezwecken setzt in aller Regel eine Einwilligung des Betroffenen voraus. Ob man die Teilnahme an einem Gewinnspiel von der Erteilung einer Werbeeinwilligung abhängig machen darf (Kopplung), wird unterschiedlich bewertet. Die Datenschutzbehörden sind eher kritisch.

## **Werbung in Direktnachricht ohne Einwilligung**

Für die Direktwerbung in sozialen Netzwerken gelten die gleichen Regeln wie für die Werbung per E-Mail. Es ist unzulässig, werbende Direktnachrichten zu versenden, ohne dass der Empfänger zuvor eingewilligt hat.

## **Custom Audience ohne Nutzereinwilligung**

Für das Re-Targeting von Nutzern über soziale Netzwerke braucht man in der Regel eine Einwilligung der Nutzer. Dies gilt nach Ansicht der Datenschutzbehörden auch für den Einsatz des Facebook-Pixels.

# **Direkteinbindung von Social-Plugins auf der Website**

Datenschützer halten die Einbindung des Like-Buttons und anderer Plug-ins für rechtswidrig, weil eine Übermittlung personenbezogener Daten unmittelbar an das soziale Netzwerk erfolge, ohne dass der Nutzer eingewilligt habe. Drittinhalte sollten wo möglich mit datensparsamen Varianten integriert werden.

# **Anreicherung von CRM-Daten um Informationen aus Social Media**

Sollen Daten aus dem eigenen Customer-Relationship-Management-System mit Daten zusammengeführt werden, die über Social-Media-Monitoring erhoben wurden, setzt dies praktisch immer die Einwilligung des Betroffenen voraus. Fehlt diese, liegt ein Datenschutzverstoß vor.

# **Getarnte Displaywerbung**

Beachten Sie das Gebot der Trennung von Werbung und redaktionellem Inhalt. Wenn sich nicht aus dem Banner unmittelbar ergibt, dass es Werbung ist, muss das Werbemittel als Anzeige gekennzeichnet sein.

# **Datenerhebung in interaktiven Bannern**

Werden personenbezogene Daten über interaktive Banner erhoben, ohne dass der Betroffene dies erkennen kann, ist dies ein Datenschutzverstoß. Jedenfalls bedarf es einer Datenschutzhinweisung. Vorsicht ist geboten bei zu vielen Pflichtfeldern – es gilt das Gebot der Datenminimierung.

# **B2C-Bannerwerbung mit Nettopreisen**

Werden in Bannern Preise für bestimmte Produkte angegeben, muss es sich dabei um Gesamtpreise handeln. Es ist ein Verstoß gegen die Preisangabenverordnung und damit wettbewerbswidrig, wenn zu den angegebenen Preisen noch weitere Kosten hinzutreten.

## **Umgehung von Pop-up-Blockern**

Werden Pop-ups technisch so ausgestaltet, dass sie herkömmliche Pop-up-Blocker umgehen, ist dies eine unzumutbare Belästigung.

## **Gekaufte Advertorials**

Werbung ist stets von redaktionellen Inhalten zu trennen. Wird Unternehmen die Möglichkeit gegeben, sich im Rahmen eines Advertorials werblich darzustellen, muss die Werbung als Anzeige gekennzeichnet werden.

## **Fake-Bewertungen**

Es ist zwar üblich, aber klar wettbewerbswidrig, vermeintlich objektive Produktbewertungen in Online-Portalen abzugeben. Wenn Mitarbeiter einer PR-Agentur Produkte von Kunden in Online-Portalen lobend erwähnen oder der Verlag auf Amazon eine geschönte Rezension eines eigenen Buches verfasst, ist dies als Schleichwerbung wettbewerbswidrig.

## **Fehlende Einwilligung beim Targeting**

Werden personenbezogene Daten für Targeting-Maßnahmen verwendet, bedarf es häufig einer Einwilligung, jedenfalls aber einer transparenten Information des Nutzers.

# **Targeting: fehlendes Opt-out**

Auch wenn keine Einwilligung erforderlich sein sollte (etwa bei einfachem Re-Targeting), muss dem Nutzer die Möglichkeit gegeben werden, der Profilbildung (auch unter Pseudonym) zu widersprechen. Es bedarf also einer Opt-out-Möglichkeit. In der Praxis geschieht dies in der Regel durch Setzen eines Opt-out-Cookies.

## **Location Based Messages ohne Einwilligung des Nutzers**

Wenn Sie mit mobilen Nachrichten werben, müssen Sie die Einwilligung des Empfängers haben – sowohl für die Zusendung von Werbung als auch für die Nutzung der Standortdaten.

## **Kundenprofile im Mobile Targeting**

Die Bildung umfangreicher Nutzerprofile bedarf nach der Rechtsprechung einer Einwilligung der Betroffenen. Dies gilt jedenfalls, wenn Daten, die auf der Website erhoben werden, mit Klardaten des Betroffenen zusammengeführt werden. Werden Nutzerprofile etwa bei der mobilen Werbung mit Log-in-Daten des Nutzers abgeglichen, um noch mehr Informationen über einen konkreten User zu erhalten, ist das ein Datenschutzverstoß, wenn keine Einwilligung des Nutzers vorliegt.