

Anke Probst

Die Macht der XML-Sitemaps (Teil 2): SEO-Analysen und Monitoring mit Sitemaps

Im ersten Teil dieses Beitrags in der letzten Ausgabe hat Anke Probst bereits über die Macht von Sitemaps und wichtige Impulse für die Konzeption und das Set-up von Sitemaps berichtet. Vielleicht unterstützt auch bereits ein ähnliches Set-up Ihr individuelles Crawl- und Indexierungsmanagement und bildet damit eine solide Grundlage für eine gute SEO-Performance? Aber hier hört die Arbeit mit Sitemaps noch längst nicht auf. Im Gegenteil!

In diesem zweiten Teil erfahren Sie deshalb, wie Sie Ihre Sitemaps/Feeds für SEO- und Qualitätsanalysen verwenden und somit die ganze Power dahinter effizient nutzen können.

Ein wesentlicher Teil der SEO-Arbeit besteht bei großen Seiten aus kontinuierlichem (URL-) Monitoring und der Analyse technischer Fehler, deren Behebung sowie der Minimierung qualitativer Schwächen des Contents. Welche Analysegrundlage eignet sich hier besser als Sitemaps? Wenn Sie im Vorfeld bei der Konzeption die für Ihre Domain wichtigen Fragen beantworten konnten (vgl. erster Teil dieses Artikels) und aufgrund dessen ein Cluster von validen Sitemaps generiert haben, dann ist dies die beste Ausgangsbasis. Schließlich sollten diese Sitemaps nur gültige, crawl- UND indexierbare URLs enthalten und keine unwichtigen URLs, für die Sie weder Zeit noch Ressourcen verschwenden wollen. Eine Ausnahme bilden Feeds zur Deindexierung, falls Sie diese nutzen. Der Fokus einer SEO-Analyse, aber auch einer Qualitätsanalyse, liegt daher immer auf den in Sitemaps/Feeds enthaltenen URLs. Konzentrieren Sie sich daher bei der SEO-Analyse und im Monitoring zunächst auf genau diese – egal, ob Sie lieber manuell arbeiten oder sich automatisiert unterstützen lassen.

generiert. Der Clou dieser Applikation ist eine daran angedockte Engine, die alle Sitemaps nach exakt definierten Vorgaben prüft, URL-basierte Informationen speichert und an ein Monitoring schickt. In diesem letzten Prozess werden unterschiedliche Monitorings generiert und deren Daten in einem schlichten E-Mail-Format an SEOs und Entwickler versendet.

Die wichtigsten Grundgedanken dieser Applikation sind:

- » die Notwendigkeit, sich bei einer komplexen Plattform (Millionen von URLs mit vielen unterschiedlichen (De-)Indexierungslogiken) nur auf die rankingrelevanten URLs zu konzentrieren – und sich damit zunächst einen Überblick über die zu crawlenden und zu indexierenden URLs zu verschaffen;
- » die Notwendigkeit, dem Suchmaschinenindex automatisiert möglichst schnell neue URLs hinzuzufügen, Änderungen mitzuteilen und veraltete URLs entfernen zu lassen – insbesondere für den User, der eine gewisse Erwartungshaltung hat, wenn er z. B. sein XING-Userprofil ändert;
- » aber eben auch die Notwendigkeit, ebendiese URLs nach Seitentypen/Kategorien zu splitten, deren Logiken zu erfassen und vor allem aus technischer Sicht zu monitorieren, z. B. wo unter welchen Bedingungen Canonicals oder „noindex“ gesetzt werden, wann und wohin Redirects entstehen oder ein Statuscode 404/410 gesendet wird;

DIE AUTORIN



Anke Probst verantwortet als Senior-SEO-Managerin die SEO-Strategie der XING SE. Sie ist außerdem Mitglied im Expertenbeirat des Bundesverbands Digitale Wirtschaft (BVDW) e. V. und Co-Autorin des Buches „Der Online Marketing Manager“.

Grundgedanken für SEO-Analysen und Monitorings mit Sitemaps

Im ersten Teil dieses Beitrags in Ausgabe 53 wurde am Beispiel der Domain *xing.com* eine Sitemap-Applikation beschrieben, die unterschiedliche XML-Sitemaps und Feeds für bestimmte Inhalte und Seitentypen/Kategorien

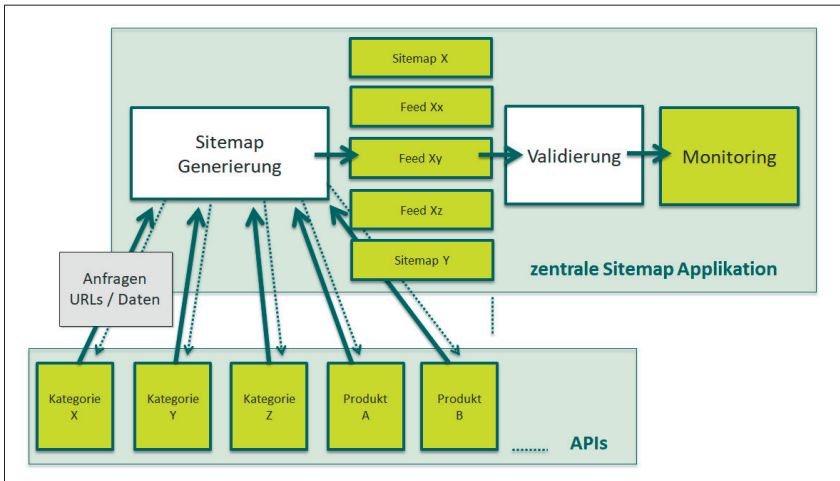


Abb. 1: Stark vereinfachte Darstellung einer Sitemap-Applikation inkl. Validierungs- und Monitoring-Engine

Eingereichte Sitemaps					
Sitemap	Typ	Eingereicht ↓	Zuletzt gelesen	Status	Gefundene URLs
/sitemap.xml. gz	Sitemap-Index	09.04.2018	10.01.2019	Erfolgreich	43.494

Abb. 2: Status einer gültigen Sitemap mit Anzahl eingereichter/gefundenen URLs

» sowie die Möglichkeit, wenn diese Logiken bereits regelmäßig geprüft werden, auch weitere SEO-relevante Kriterien zu definieren und ins Monitoring aufzunehmen.

Wenn man diese Grundgedanken auf die eigene Domain anwendet, lassen sich sicherlich einige Gemeinsamkeiten finden. Im Grunde sollte jede größere Webseite einen hohen Anspruch an ein sauberes Crawl- und Indexierungsmanagement haben, und jede Webseite hat ebenfalls unterschiedliche Logiken implementiert, die die Indexierung beeinflussen: Das Setzen eines Filters löst eine Parameter-URL aus, die ein Canonical auf die Kategorie erzeugt; dauerhaft ausverkaufte Produkte erzeugen einen Statuscode 404; vorübergehend ausverkaufte Produkte erhalten ein „noindex“; bei URL-Änderungen erfolgt ein 301-Redirect; um nur ein paar gängige Beispiele zu nennen. All dies sind Logiken, deren Implementierung regelmäßig überprüft werden sollte. Für den kontinuierlichen Prozess der Suchmaschinenoptimierung ist das aber nicht genug – auch Quali-

tätskriterien sollten ständig nachgehalten und bestenfalls sogar mit internen Daten verknüpft werden. Dieses Vorgehen kann Ihnen wertvolle Informationen zur Priorisierung Ihrer SEO-Tasks aufzeigen und ggf. Ihre Strategie beeinflussen. Beginnen Sie aber zunächst mit ein paar wichtigen Basics: dem Monitoring und der Analyse technischer Fehler mithilfe der Search Console.

Erste Anlaufstelle: Googles Search Console

Die Nutzung der Search Console in Verbindung mit Sitemaps/Feeds ermöglicht Ihnen vollautomatisch Einblicke und wertvolle Tipps. Insbesondere die neue Version der Search Console gibt Daten und Fehler aus, die sich konkret auf eingereichte Sitemaps und Feeds beziehen, und zwar wesentlich umfangreicher und kleinteiliger, als Sie es von der alten Version gewohnt waren. Unter „Sitemaps“ im Menüpunkt „Index“ einer Property finden Sie die Sitemap(s), die Sie für die jeweilige Property/das Verzeichnis eingereicht haben (Abbildung 2). Aber Achtung: In der Liste

TIPP

Übrigens: Seit geraumer Zeit gibt es die Indexing-API von Google: <https://google.com/B6Xdkr>. Zitat: „Currently, the Indexing API can only be used to crawl pages with either job posting or livestream structured data“ – wenn Sie solche Seitentypen haben, dann könnte die API sinnvoll sein. Sie scheint aber auch für andere Seitentypen zu funktionieren! Die API ermöglicht es, Google über neue, aktualisierte und zu entfernende Inhalte zu informieren. Google crawlt daraufhin die Seiten laut eigenen Aussagen schneller als Sitemaps (über Feeds wird keine Aussage getroffen). Das Testen bzw. die Nutzung macht sicherlich Sinn – vergessen Sie aber nicht, dass sie Sitemaps auch für andere Suchmaschinen bereitstellen sollten.

werden Ihnen zwar XML-Sitemaps und Index-Sitemaps angezeigt, jedoch keine Feeds. Eine Auflistung der Feeds finden Sie derzeit nur in der alten Version der Search Console (Abbildung 4).

Wichtige Analysemöglichkeiten mit der Search Console

1. Validierung der eingereichten Sitemaps: Konnte die Sitemap erfolgreich verarbeitet werden? Zunächst gilt es, zu prüfen, ob der Statuscode der Sitemaps korrekt ist und die Sitemap erreichbar ist. Google zeigt Ihnen hier außerdem, wann die Sitemap zuletzt gelesen wurde (Abbildung 2). Dies kann ein wichtiger Hinweis sein, falls neuer Content noch nicht indexiert wurde.
2. Indexierungsstatus: Wurden alle eingereichten URLs erfolgreich indexiert? Verschaffen Sie sich unbedingt einen Überblick darüber, wie viele URLs in einer Sitemap übergeben werden – und ob diese Anzahl zum Indexierungsstand der Kategorie/der Property passt, auf die sich die Sitemap bezieht. Falls nicht und Sie sicher sind, dass die Sitemap korrekt konzipiert und generiert wurde, dann müssen Sie sich auf die Suche nach der

TIPP

Denken Sie auch an den Tipp im ersten Teil, die Sitemap/den Feed einer Kategorie immer ins entsprechende Verzeichnis hochzuladen – dieser kleine Trick kommt Ihnen jetzt für die Auswertung in der Search Console zugute, vor allem, wenn Sie viele Verzeichnisse und innerhalb dieser Verzeichnisse viele unterschiedliche Sitemaps haben. Sie profitieren jetzt von besserer Übersichtlichkeit dank Auflistung aller zur Property gehörigen Sitemaps.

Ursache begeben. Und nicht selten kann daraus ein Quickwin resultieren! Die Anzahl der URLs in der Sitemap finden Sie direkt neben dem Status unter „gefundene URLs“ (Abbildung 2). Hier im Beispiel werden über die Sitemap 43.494 URLs übermittelt. Klicken Sie rechts neben der Sitemap auf das kleine Balkendiagramm-Symbol und Sie gelangen direkt zur Indexabdeckung, um den Wert mit dem tatsächlichen Stand der Indexierung zu vergleichen (Abbildung 3). Alternativ finden Sie in der alten Version der Search Console eine anschauliche Grafik über eingereichte und indexierte URLs (Abbildung 4). In einer perfekten Welt stimmt der Wert für gültige, also indexierte URLs exakt mit dem Wert der übermittelten/gefundenen URLs überein – aber keine Sorge, das passiert wirklich selten. Wesentlich wahrscheinlicher ist es, dass sich die Werte geringfügig voneinander unterscheiden, denn der Zeitpunkt der Sitemap-Generierung weicht oft deutlich vom Zeitpunkt des Crawlens ab, sodass sich möglicherweise in der Zwischenzeit für URLs je nach implementierter Logik der Statuscode ändert oder ein Canonical/noindex angesteuert wird. In diesem Fall würde Google hier einen Hinweis unter „Ausgeschlossene URLs“ oder „Fehler“ geben. Im schlimmeren Fall liegt dann tatsächlich ein technischer Fehler vor. Werfen Sie daher in regelmäßigen

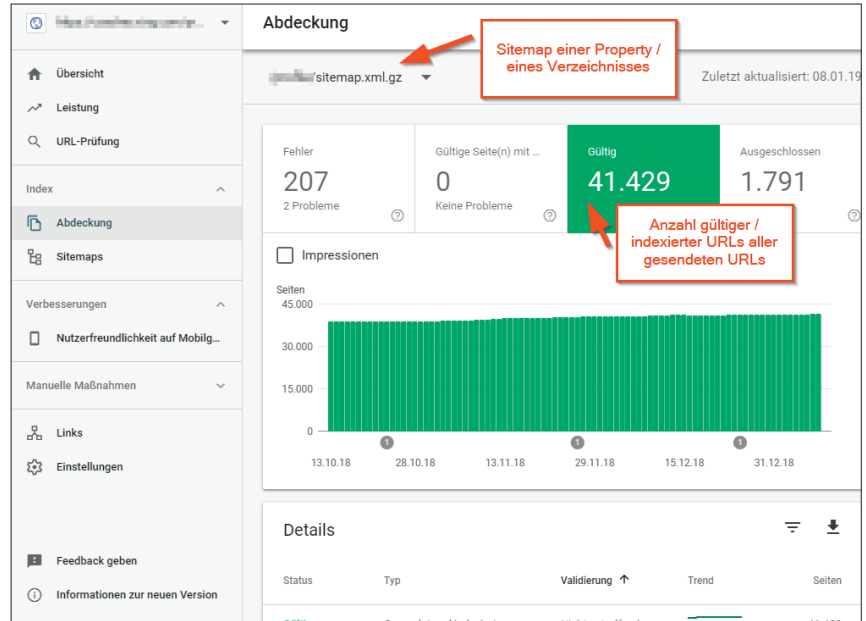


Abb. 3: Gültige, also indexierte URLs einer spezifischen Kategorie-Sitemap

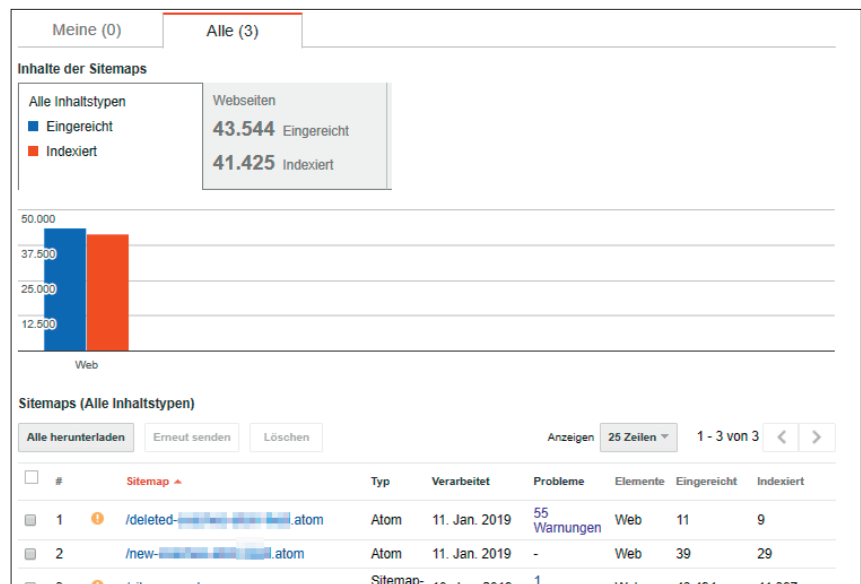


Abb. 4: Liste mit XML-Sitemaps und Atom-Feeds einer Property/eines Verzeichnisses mit Angabe der eingereichten und indexierten URLs sowie Fehlermeldungen/Warnungen

Abständen einen Blick auf die URLs unter „Ausgeschlossene“ und „Fehler“. Die hier angegebenen Fehlermeldungen können vielfältig sein (Abbildung 5), besonderes Augenmerk sollten Sie jedoch auf folgende Hinweise legen:

- » Gesendete URL nicht gefunden (404)/Soft-404-Fehler/Crawling-Anomalie/Serverfehler (5xx): Beim Crawl dieser URLs sind Server-Fehler oder (Soft-)404-Fehler aufgetreten. Je nach Anzahl und Häufigkeit dieser Fehlermeldung

muss unbedingt nach der Ursache geforscht werden.

- » Gesendete URL mit „noindex“ gekennzeichnet: Wenn URLs mit der Meta-Tag-Anweisung „noindex“ in der Sitemap übermittelt werden, lohnt sich ein genauer Blick auf diese URLs. Liegt ein Bug vor?
- » Gesendete URL durch robots.txt-Datei blockiert: Anscheinend wurden durch einen falschen Eintrag in der robots.txt-Datei URLs vom Crawl ausgeschlossen, die eigentlich crawlbar sein sollten.

Fehler	Gesendete URL hat Crawling-Problem	Nicht gestartet	178
Fehler	Gesendete URL nicht gefunden (404)	Nicht gestartet	29
Fehler	Gesendete URL ist ein Soft 404-Fehler	Nicht zutreffend	0
Fehler	Serverfehler (5xx)	Nicht zutreffend	0
Ausgeschlossen	Gecrawlt – zurzeit nicht indexiert	Nicht zutreffend	1.051
Ausgeschlossen	Gefunden – zurzeit nicht indexiert	Nicht zutreffend	603
Ausgeschlossen	Duplikat – eingereichte URL nicht als kanonisch festgelegt	Nicht zutreffend	137

Abb. 5: Beispiele von Hinweisen in der Search Console zu „ausgeschlossenen URLs“ und „Fehlern“

Indexierung zulässig?	Ja
Vom Nutzer angegebene kanonische URL	https://www.xing.com/companies/5vor12gmbh
Von Google ausgewählte kanonische URL	https://www.xing.com/companies
	PRÜFEN

Abb. 6: Beispiel einer nicht-indexierten URL, für die Google selbst eine kanonische URL bestimmt

» Duplikat – eingereichte URL nicht als kanonisch festgelegt/alternative Seite mit richtigem kanonischem Tag: In der Sitemap stehen URLs, die per Canonical auf URLs mit stark ähnlichen Inhalten verweisen. Hier ist entweder das Canonical oder aber die Sitemap falsch konfiguriert.

Eine vollständige Liste finden Sie unter <https://goo.gl/YPSPjB>. Tipp: Viele Daten sind auch über die API der Search Console verfügbar: <https://goo.gl/oK6bER>.

Beispiel aus der Praxis

Ein Online-Shop verkauft Sneaker, Stiefeletten und Sandalen und hat spezifische Sitemaps für alle drei Kategorien. In jeder Kategorie gibt es 200 Produkte – die Anzahl der übermittelten URLs beträgt pro Sitemap also 200. Von jeweils 200 in der Sitemap vorhandenen URLs sind jeweils ca. 195 URLs indexiert – das ist ein realistischer Wert. Nur von der Sitemap sandalen.xml sind lediglich 35 URLs indexiert. Wie kann das sein?

Die Search Console liefert die Antwort: Die 165 nicht-indexierten URLs werden unter „Ausgeschlossen“ mit dem Hinweis „Duplikat – eingereichte URL nicht als kanonisch festgelegt“ ausgegeben. Klicken Sie hier einfach

auf den Hinweis, dann auf eine einzelne URL und auf „URL prüfen“. Die Search Console ruft die Daten nun direkt aus dem Index ab und gibt einen ähnlichen wie in Abbildung 6 dargestellten Hinweis.

Google hat hier autonom festgelegt, dass die URL ein Duplikat darstellt, und eine andere URL als kanonische URL bestimmt. Ein Vergleich dieser URLs offenbart den Bug: Lediglich für die Sandalen einer bestimmten Marke wird die Produktbeschreibung angezeigt, nicht aber für die anderen Marken. Dadurch unterscheiden sich die Seiten ohne Textinhalte kaum voneinander, sodass Google eine URL davon als primäre/kanonische Version auswählt – und alle anderen von der Indexierung ausschließt. Ärgerlich, wenn dies nicht bemerkt wird!

Tipp: Vergleichen Sie die Daten eventuell mit den Daten in der alten Search Console. Erfahrungsgemäß werden hier auch Fehlermeldungen für Bilder-URLs ausgegeben, die den Content-URLs in der Sitemap beigelegt waren. In der neuen Search Console scheinen sich die angegebenen Hinweise vor allem auf die Content-URLs zu beziehen. Das ist aktuell ein wenig umständlich, heißt aber nicht, dass Google die neue Version nicht noch nachbessert.

TIPP

Mit dem Screaming Frog lassen sich auch (Bilder-)Sitemaps nach bestimmten Parametern erstellen – je nach Bedarf auch zur De-Indexierung mit Statuscodes 3xx, 4xx oder auch „noindex“. Dies kann enorm hilfreich sein, wenn Sie kurzfristig für einen gewissen Zeitraum eine neue Sitemap brauchen, um z. B. eine URL-Änderung mit 301-Redirects mithilfe einer Sitemap zu unterstützen.

„Manuelle“ SEO-Analysen und Qualitätsanalysen mit dem Screaming Frog

Wer sich tiefer mit den eigenen Daten beschäftigen will und mehr Informationen zu Analyse Zwecken benötigt, für den sind Crawling-Tools die bessere Wahl. Der Screaming Frog (<https://www.screamingfrog.co.uk/>) als günstige und mächtige Crawling-Lösung eignet sich hervorragend zu Sitemap-Analysen – und damit indirekt auch zur SEO-Analyse.

Wählen Sie hierfür den Modus „List“. Der Upload-Button bietet Ihnen dann die Möglichkeit, sowohl einen Sitemap-Index als auch eine einzelne Sitemap auszuwählen. Geben Sie hier einfach die URL der zu prüfenden Sitemap ein, dies erspart Ihnen lästiges und zeitaufwendiges Downloaden und Extrahieren der URLs, um sie schließlich hochladen zu können. Der Screaming Frog ist in der Lage, auch gepackte Verzeichnisse mit umfangreichen Sitemap-Indizes zu verarbeiten.

Für das Crawling von URLs aus Sitemaps werden dieselben Informationen ausgegeben wie beim eher gängigen Crawl von Webseiten im Spider-Modus. Interessant für die Fehleranalyse sind hier, ähnlich wie in der Search Console, u. a. die folgenden Daten, um herauszufinden, ob wirklich nur gültige URLs übermittelt werden:

- » Statuscodes
- » Indexability, Directives
- » Canonicals

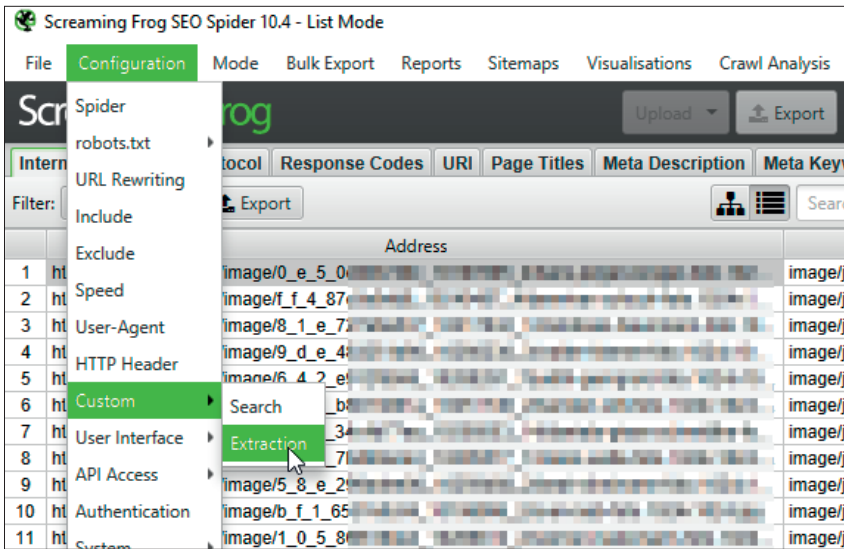


Abb. 7: Hier finden Sie die Custom Extraction im Screaming Frog

Der entscheidende Vorteil des Screaming Frog ist jedoch, dass pro URL alle Informationen gebündelt ausgegeben werden – und zwar nicht nur die gängigen technischen Fehler, die auch in der Search Console auffindbar wären, sondern noch jede Menge wei-

terer Werte. Diese können die Analyse relevanter URLs unter technischen, aber auch qualitativen Gesichtspunkten erheblich vereinfachen: z. B. die Analyse komplexer hreflang-Set-ups oder rel="prev"/„next“-Implementierungen.

Impulse für individuelle Qualitätsanalysen

Die Qualität der Webseite spielt eine große Rolle für SEO – nicht nur für die Suchmaschine, sondern vor allem für den User. Die in den vorherigen Kapiteln genannten technischen Faktoren müssen zwar kontinuierlich überwacht und ggf. auftretende Fehler behoben werden, letztendlich zählen aber auch inhaltliche Faktoren wie Contentqualität/-länge für das Erfüllen der Userintention sowie das Auffinden und die Analyse thematisch ähnlicher Inhalte. Sie wollen den Nutzer schließlich nicht in seiner Erwartungshaltung enttäuschen, ihn sogar bestenfalls zum „Wiederholungstäter“ machen – und Ihr Unternehmen schlussendlich dank treuer Kunden oder Leser zur Marke aufbauen. Auch Suchmaschinen werden dies zu schätzen wissen und mit guten Rankings belohnen!



ISBN 978-3-8362-6250-1

ISBN 978-3-8362-6351-7

ISBN 978-3-8362-5963-7

Wir bringen Ihr Online-Marketing auf Hochtouren!

Unsere Bücher zeigen, wie schnelles und flexibles Digitalmarketing geht. Erfahren Sie, wie Sie den Tag Manager für Tracking und Analyse nutzen. Lernen Sie, wie Sie Ihre Website mit gutem SEO nach vorne bringen. Und rasen Sie dem Mitbewerber einfach davon!

Der Screaming Frog bringt für solche inhaltlichen Qualitätsanalysen bereits ein Set an Funktionen mit: beispielsweise das Auslesen der Meta-Description bzw. des Seitentitels und deren Länge, womit Sie Duplikate aufstöbern oder die Zeichenanzahl optimal anpassen können; verbesserungswürdige User-Experience aufgrund fehlender Bilder im Content usw. Für individuellere Analysen bietet sich die mächtige Funktion der „Custom Extraction“ an (Abbildung 7). Die Custom Extraction ermöglicht es, genau die Daten während eines Crawls zu scrapen, die Sie interessieren: z. B. bestimmte Absätze oder die generelle Länge von Produktbeschreibungen, Anzahl von Kundenbewertungen je Produkt/-variante, Anzahl von Listings oder Bildern – Elemente also, die nach Ihrer Definition eine qualitativ hochwertige Seite enthalten sollte. Denken Sie ebenso an die Überprüfung von Structured Data Markup wie JSON-LD, Open-Graph-Daten oder Tracking-Codes.

Nutzen Sie wieder Ihre Sitemaps als Ausgangsbasis. Wenn Sie verschiedene Sitemaps/Feeds nach Seitentyp definiert haben, umso besser! Definieren Sie für jeden Seitentyp/jedes Template spezifische Qualitätskriterien (anhand eigener Einschätzung, aber vielleicht auch anhand der besser rankenden Wettbewerber) und nutzen Sie die Custom Extraction, um diese Informationen zu scrapen, indem Sie den Screaming Frog die dazu gehörige Sitemap verarbeiten lassen (Abbildung 8). Eine ausführliche Anleitung finden Sie unter <https://goo.gl/SrP9XE>.

Wertvolle Erkenntnisse solcher Analysen können sein:

- » Wie hochwertig sind die Seiten einer bestimmten Kategorie im Vergleich zum Wettbewerb?
- » Wie schneiden verschiedene Kategorien Ihrer Domain ab, wenn Sie sie nach Qualitätskriterien miteinander vergleichen? Spiegelt sich dies auch in den Suchergebnissen wider?

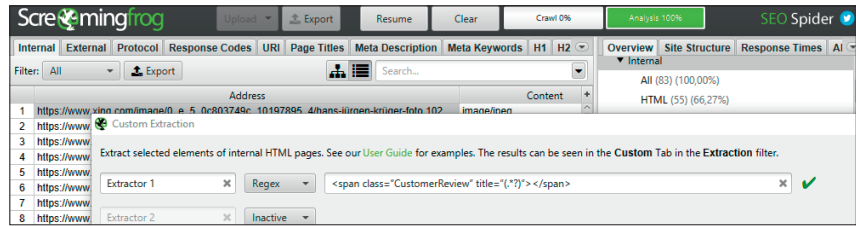


Abb. 8: Beispiel einer Custom Extraction, die den Titel einer Kundenbewertung auslesen würde

» Oder gibt es gar Seiten, die aufgrund mangelnder Qualität nicht von Suchmaschinen indiziert werden?

Mit diesen Daten könnten Sie außerdem Qualitätsindizes pro URL vergeben, die Sie dann mit relevanten internen Informationen wie Bot-Zugriffe, Rankings, Bounce-Rates, Produktmargen und Conversion-Rates anreichern. So erhalten Sie eine priorisierte URL-Liste, nach der Sie gezielt Ihre Webseite optimieren oder aber Logiken zur Deindexierung/Konsolidierung von Inhalten entwickeln können. Dies kann z. B. für minderwertige Seiten empfehlenswert sein, die Crawlingbudget verschwenden, aber aufgrund mangelnder Qualität nicht indiziert werden. Nutzen Sie hierfür die „URL-Prüfung“ in der Search Console oder die „info:URL“-Abfrage direkt in Google, um den Indexierungsstatus zu prüfen.

Eigene Sitemap-Validierung und Monitoring

Natürlich kann man sich jederzeit externer Tools bedienen und diese mit internen Daten aus BI-Systemen, der Webanalyse usw. verknüpfen. Für kreative Analysen bei großen Domains ist es jedoch möglich, dass die Search-Console-API nicht ausreicht – vielleicht möchte man sich aber auch nicht davon abhängig machen. So manch einer ist bei der Nutzung der Search Console auch über Ungereimtheiten gestolpert. Der Screaming Frog ist für Ad-hoc-Analysen sicherlich äußerst hilfreich, aber weniger geeignet für kontinuierliche automatisierte Monitorings – andere Tools am Markt sind vielleicht auf Dauer zu teuer oder bringen nicht die benötigten Funktionen mit. Gründe für die Entwicklung eigener Validierungs- und

Monitoring-Mechanismen gibt es viele! Empfehlenswert ist, zunächst mit einer Konzipierungsphase zu starten, das Konzept mit „Bordmitteln“ wie der Search Console, dem Screaming Frog etc. zu testen und sich schließlich am Markt nach der perfekten Lösung umzusehen.

So könnte eine eigene Lösung aussehen

Wie eingangs beschrieben, liegen für die Domain xing.com u. a. aufgrund der Domaingröße und vieler unterschiedlicher (De-)Indexierungs-Logiken gravierende Gründe für ein automatisiertes technisches Monitoring vor. Nach der Konzeption von Sitemaps und Feeds, die diese Logiken widerspiegeln, wurde schnell klar, dass der Aufwand für ein regelmäßiges manuelles Monitoring zu groß ist und die Search Console den Anforderungen nicht genügt. Zudem rechtfertigt die hohe Menge an SEO-Traffic die nötigen Entwicklungsressourcen zur Umsetzung einer eigenen Lösung. Da es in der Vergangenheit bei der Generierung von Sitemaps immer wieder zu Problemen kam und diese zukünftig vermieden werden sollten, war ein Validierungsmechanismus geplant – dieser wird als logische Konsequenz für die Überprüfung weiterer SEO-relevanter Kriterien stetig ausgebaut. Welche Gründe auch immer in Ihrem Fall für die Entwicklung einer eigenen Lösung sprechen, machen Sie sich um folgende Themen Gedanken:

Grundlegende Validierung

Ein Script sollte regelmäßig prüfen, ob die Sitemap-URLs und Feeds erreichbar sind. Zusätzlich wird der Zeitpunkt der letzten Aktualisierung gespeichert sowie die Anzahl der darin enthaltenen

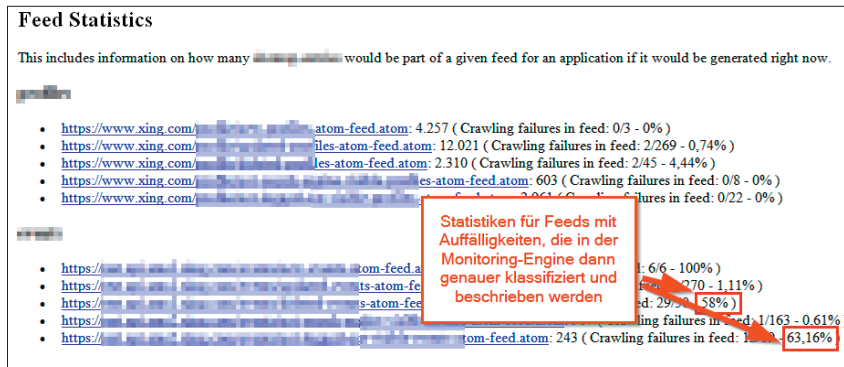


Abb. 9: Auszug von Feed-Statistiken, die erste Hinweise auf bestehende Fehler liefern

URLs. Diese werden pro Sitemap/Feed historisch festgehalten, um deren Entwicklung nachvollziehen zu können, etwaige Fehler in der Datenüberlieferung der APIs (falls genutzt) schnell zu bemerken und mit dem tatsächlichen Indexierungsstand vergleichen zu können.

URL-Monitoring zur Logik-Prüfung

Ein URL-Monitoring sollte enthalten sein, um die implementierten Logiken regelmäßig zu prüfen. Dies kann über

einen eigenen Crawler abgearbeitet werden oder man bedient sich der APIs externer Tools.

Monitoring

Alle erhobenen Daten fließen in ein Dashboard ein. Wichtig ist, dass die Daten nicht nur historisch einsehbar sind, sondern auch zu Analyse Zwecken weiterverarbeitet werden können. Denken Sie außerdem an automatisierte Reports (Abbildung 9), z. B. wenn bestimmte Schwellenwerte bei der Fehleranzahl überschritten werden.

Fazit

Für große Domains machen solche Monitorings viel Sinn, sowohl aus technischer Sicht, aber auch, um die Qualität der Inhalte ständig automatisiert zu prüfen. Nur so kann gewährleistet werden, dass Suchmaschinen UND Nutzer die Webseite würdigen und dass positive Nutzersignale sowie letztendlich auch Umsatz generiert werden.

Jetzt liegt es an Ihnen, aktiv zu werden! Überlegen Sie genau, was Ihre Domain braucht, welche Erkenntnisse Sie gewinnen wollen, welche Logiken und Implementierungen Sie dafür regelmäßig prüfen müssen und mit welchen Daten Sie diese Informationen anreichern können, um erfolgreich und nachhaltig damit arbeiten zu können. Viel Spaß und Erfolg dabei! ¶

IMPRESSUM

Herausgeber & Chefredakteur (verantwortlich):

Mario Fischer
E-Mail: redaktion@websiteboosting.com

Autoren dieser Ausgabe:

Thorsten Abrahamczik, Tobias Aubele, Dr. Martin Bahr, Claudia Heinze, Fabian Jaeckert, Mirko Lauer, Patrick Lürwer, Benjamin O´Daniel, Anke Probst, Dr. Philipp Spreer, Matthias Süß, Patricia Unfried

Anzeigenleitung:

Markus Lutz
E-Mail: anzeigenleitung@websiteboosting.com

Art Direction:

Kai Neugebauer,
E-Mail: layout@websiteboosting.com

Layout/Produktion:

Kai Neugebauer

Lektorat:

Ursula Wenke, www.lektorat-wenke.de

Fotos & Illustrationen:

Website Boosting / GettyImages
Titel: Website Boosting

Druck:

Vogel Druck und Medienservice GmbH
Leibnizstr. 5, 97204 Höchberg

Vertrieb:

PressUp GmbH
Postfach 70 13 11
22013 Hamburg
E-Mail: websiteboosting@pressup.de

Abonnement:

Website Boosting Aboservice
PressUp GmbH
Postfach 70 13 11
22013 Hamburg
Tel. 040 / 38 6666 - 342
Fax: 040 / 38 6666 - 299
E-Mail: websiteboosting@pressup.de

Erscheinungsweise: 6 x jährlich

Bezugspreis: Einzelheft: 9,80€

Bezugspreis Inland jährlich 51,00€ inkl. Versand

Bezugspreis Ausland jährlich 63,00€
inkl. Versand

Studenten im Inland erhalten gegen Vorlage einer Immatrikulationsbescheinigung einen Preisvorteil von 20%.

Verlagsleitung:

Michael Müßig
Tel: +49 931 / 26 038 04,
verlag@websiteboosting.com

Anschrift des Verlages

Hotspot Verlag GmbH
Obere Landwehr 4a, 97082 Würzburg
Tel: + 49 931 / 26 038 04
Fax: +49 931 / 26 038 05
E-Mail: verlag@hotspotverlag.de
www.hotspotverlag.de

Geschäftsführung:

Kai Neugebauer

Die Inhaber- und Beteiligungsverhältnisse lauten wie folgt: Gesellschafter zu 100% ist die Webvalue Holding GmbH

ISSN: 2191-6241

Für unverlangt eingereichte Texte und Daten kann keine Haftung übernommen werden. Sämtliche Veröffentlichungen in Website Boosting erfolgen ohne Berücksichtigung eines eventuellen Patentschutzes. Markennamen werden ohne Gewährleistung einer freien Verwendung benutzt. Trotz sorgfältiger Recherche kann für die Richtigkeit des Inhalts keine Haftung übernommen werden. Namentlich gekennzeichnete Artikel geben nicht unbedingt die Meinung der Redaktion wieder.