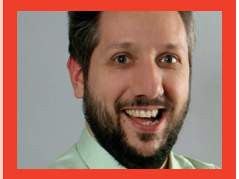
Fili Wiese

ALLES, WAS SIE FÜR DIE UMSTELLUNG AUF HTTPS WISSEN MÜSSEN (Teil 3/3)

In den ersten beiden Teilen dieses Leitfadens ging es um die Vorbereitung der Umstellung auf HTTPS und um die tatsächliche Realisierung auf der Server-Seite und in der Google Search Console. Dieser dritte und letzte Teil der umfassenden Serie des ehemaligen Googlers und Search-Experten Fili Wiese befasst sich mit dem Abschluss der Umstellung auf HTTPS: Messung der Wirkung, Identifizieren der bei der Migration aufgetretenen Probleme und Vorteile, die Sie durch die Umstellung Ihrer Website auf HTTPS erzielt haben.

DER AUTOR



Fili Wiese ist ein renommierter SEO-Spezialist und hat früher in leitender Funktion im Google Search Quality Team mitgearbeitet. Bei SearchBrothers.com geht er mit Erfolg gegen die Abstrafung von Websites durch Google-Penalties vor und bietet SEO-Consulting mit SEO-Audits und SEO-Workshops.

Abschluss der Umstellung auf HTTPS

Sie haben jetzt alles getan, damit Google die HTTPS-Version Ihrer Website bevorzugt und deren Umstellung auf HTTPS für den Googlebot möglichst klar verständlich ist. Je nach Größe der gecrawlten Website und des verfügbaren Crawl-Budgets kann es einige Wochen oder auch Monate dauern, bis der Großteil der Google-Search-Ergebnisse die Umstellung widerspiegelt.

Außerdem ist es wichtig zu beobachten, wie die Umstellung auf HTTPS in Google Search voranschreitet und ob neue Probleme auftreten.

Monitor Server Logs

Wenn etwas nicht so funktioniert, wie es soll, muss das Problem möglichst schnell behoben werden. Achten Sie darauf, dass ein Benachrichtigungssystem vorhanden ist, und prüfen Sie, ob 50x- oder 404-Anforderungen an die HTTP- oder HTTPS-Version der Website gestellt werden. Durchsuchen Sie die Server-Log-Files nach diesen Fehlern. Eine andere Möglichkeit ist, sofort bei Auftreten eines solchen Fehlers eine Nachricht per E-Mail oder an einen Slack-Kanal zu schicken. Arbeiten Sie mit dem IT-Team zusammen, damit genügend Personal bereitsteht, um kritische Fehler möglichst schnell zu

Foto: Imillian / thinkstockphotos.de

beheben. Auch wenn es sich bei markierten 404-Fehlern möglicherweise um einen falschen Alarm handelt, muss geprüft werden, ob er wiederholt auftritt und zu beheben ist.

```
84.25.65.243 - - [10/Oct/2016:13:55:36 -0700]
„GET /favicon.icof HTTP/1.1“ 200 326
„http://www.example.com/start.html“
„Mozilla/4.08 [en] (Win98; I ;Nav)“
2.5.45.7 - - [10/Oct/2016:13:55:45 -0700] „GET
/ HTTP/2“ 200 2956 „-“ „Mozilla/5.0 (iPad;
U; CPU OS 3_2_1 like Mac OS X; en-us) Apple-
WebKit/531.21.10 (KHTML, like Gecko) Mobi-
le/7B405“
64.233.191.102 - - [10/Oct/2016:13:55:49
-0700] „GET /home HTTP/1.1“ 200 43455 „-“
„Mozilla/5.0 (compatible; Googlebot/2.1;
+http://www.google.com/bot.html)“
```

Soweit möglich, verfolgen Sie mithilfe des Systems auch, welche Seiten der Googlebot indexiert und mit welcher Geschwindigkeit und welche Statuscodes er vorwiegend crawlt. Da die beste Quelle für die Gewinnung dieser Daten die Server-Log-Dateien sind, ist es möglicherweise eine gute Idee, einen ELK-Stack auf einem lokalen Server oder einen externen Log-Analyzer zu verwenden, beispielsweise Botify, Screaming Frog Log Analyser oder Splunk. Andernfalls laden Sie die Server-Log-Dateien in Google BigQuery und analysieren Sie die Daten mithilfe der BigQuery-Schnittstelle oder 360 Data Studio.

Überwachung der Google Search Console

Um die Indexnummern der HTTP- und der HTTPS-Version zu überwachen, behalten Sie alle vorgelegten XML-Sitemaps im Auge. Zusammen mit dem in der Google Search Console gemeldeten Google-Index-Status zeigt das Ergebnis an, wie viele der URLs auf der HTTP-Version de-indexiert und auf die HTTPS-Version umgestellt werden.

Überprüfen Sie auf der Übersicht der Crawl-Statistik für die entsprechenden HTTP- und HTTPS-Versionen auch die Angaben zur Crawl-Geschwindigkeit und vergewissern Sie sich, dass die Infrastruktur schnell auf den Googlebot reagiert. Wenn nötig und falls durchführbar, denken Sie darüber nach, für die nächsten paar Monate zusätzliche und stärkere Server-Instanzen mit schnelleren Netzwerkverbindungen bereitzustellen, solange der Googlebot die HTTP- und HTTPS-Versionen der Website erneut crawlt. Ob dies zu einem schnelleren Crawlen der Website führt, lässt sich nur an den Crawl-Statistiken und Server-Log-Dateien erkennen.

Prüfen Sie in der Google Search Console den AMP- und Rich-Cards-Überblick sowie die Übersicht der strukturierten Daten, um zu sehen, ob Indexierungsprobleme wie beispielsweise Fehler angezeigt werden oder Probleme mit der Geschwindigkeit, mit der die Daten gefunden werden. Falls ja, beheben Sie die Schwierigkeiten möglichst schnell.

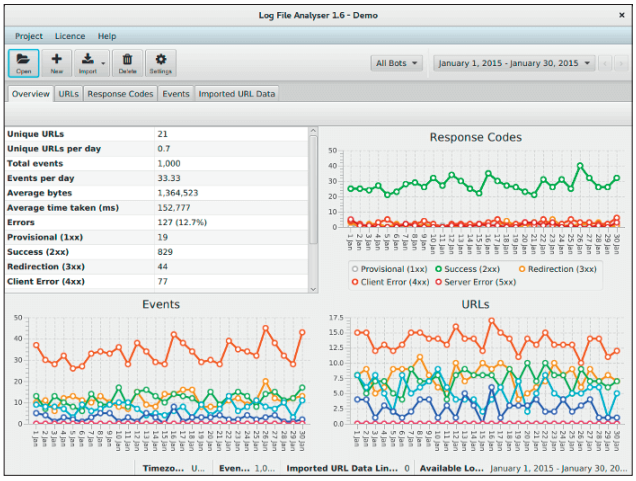


Abb. 1: Beispiel: Screaming Frog Log Analyser mit Daten

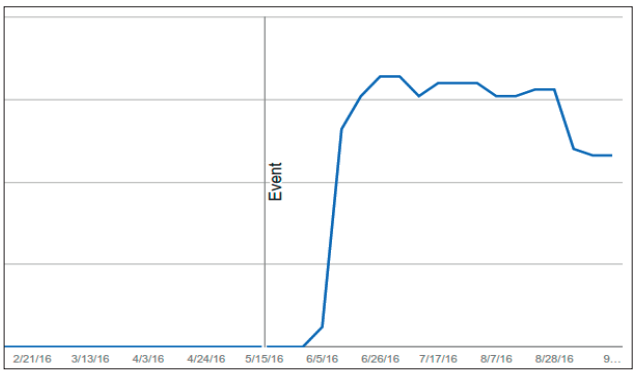


Abb. 2: Beispiel: Index-Status einer neuen HTTPS-Website in der Google Search Console

Prüfen Sie im Crawlfehler-Überblick für die relevanten HTTP- und HTTPS-Versionen in der Google Search Console, ob Serverfehler (50x), Not-Found-Fehler (404), Verbindungsprobleme und Soft-404-Fehler für PCs, Feature Phones und/oder Smartphones angezeigt werden. Liegen diese Fehler vor, insbesondere bei der HTTPS-Version, beheben Sie sie möglichst schnell. Bei markierten 404-Fehlern kann es sich um einen falschen Alarm handeln, aber vergewissern Sie sich, dass der Fehler nicht unerwartet ist; prüfen Sie, ob er erneut auftritt und behoben werden muss.

Last, but not least prüfen Sie die Search Analytics in der Google Search Console. In Teil 2 dieses Leitfadens habe ich beschrieben, wie Sie verschiedene Sätze erstellen, um die diversen Properties zu gruppieren, unter anderem einen Satz mit allen für diese Content-Umstellung relevanten Eigenschaften der HTTP- und HTTPS-Versionen. Jetzt kommt der Zeitpunkt, wo sich die investierten Mühen auszahlen.

Durch die Zusammenlegung aller relevanten HTTP- und HTTPS-Eigenschaften kann man in Google Search Analytics die Gesamtwirkung der Umstellung auf HTTPS auf das Google-Ranking sehen, so wie Google selbst sie bewertet.

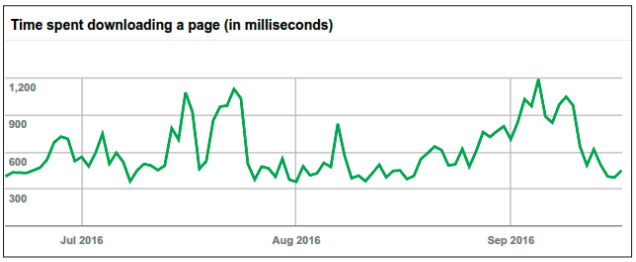


Abb. 3: Beispiel: So schnell lädt der Googlebot eine Website herunter

Es ist nämlich so: Bei der Umstellung von Content von HTTP auf HTTPS zeichnet die Google Search Console die auf eine Property entfallende Suchanalytik ausschließlich für die jeweilige Property auf, und sie kann auch nur in dieser Property geteilt werden. Das führt dazu, dass die HTTP-Property sämtlichen Traffic und alle Rankings verliert und dass die HTTPS-Version sie gewinnt. Möglicherweise muss man jedoch mehr als 90 Tage warten, bevor man einen guten Überblick über die Auswirkung bekommt, und die Search Analytics in der Google Search Console stehen nur für die letzten 90 Tage zur Verfügung.

Der kombinierte Satz ermöglicht einen Überblick über alle Rankings und über die Gesamtauswirkung der Umstellung auf HTTPS. Diesen Überblick können Sie über Search Analytics API auch herunterladen und/oder Sie klicken den Download-Button der Übersichtsseite von Search Analytics.

Eingehende Links aktualisieren

Die Umstellung Ihrer Website auf HTTPS ist abgeschlossen und jetzt soll der Rest der Welt davon erfahren. Ist die betreffende Website Teil eines größeren Website-Netzwerks der gleichen Organisation, gehen Sie zu jeder Website in diesem (internen) Netzwerk und ändern Sie alle Referenzen zur Website auf die HTTPS-Version um. Um alle Link-Referenzen zu finden und auf die HTTPS-Version zu aktualisieren, müssen Sie möglicherweise mit verschiedenen Teams in der Organisation sprechen und/oder die Datenbanken durchsuchen.

In vielen Fällen finden sich Links zu der umgestellten Website auch in der E-Mail-Signatur jedes Mitarbeiters, auf der offiziellen Profilseite der Website und/oder Organisation in den sozialen Netzen (wie Twitter, Facebook, Pinterest etc.) und den Unternehmensprofilseiten auf LinkedIn und/oder den Gelben Seiten und/oder Wikipedia und regionalen Firmenverzeichnissen, PPC-Kampagnen und -Anzeigen, Kampagnen in den sozialen Medien, in Newsletter-Software und/oder Mailing-Listen, Werbetexten für Direktvermarktung, Videos, Analytik-Software, externer Web-Tracking- und Reporting-Software, Google-My-Business-Eintragungen, Visitenkarten, externen Beurteilungsplattformen und so weiter. Natürlich müssen Sie nicht all diese Links gleich heute updaten, aber führen Sie sofort realisierbare Änderungen so schnell wie möglich durch und machen einen Zeitplan für die anderen

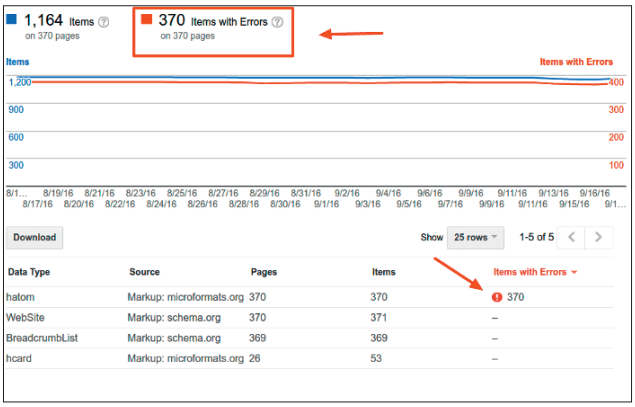


Abb. 4: Beispiel: Überblick über strukturierte Daten mit Fehleranzeige in der Google Search Console

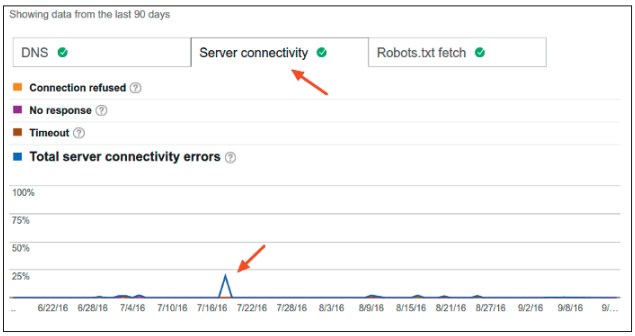


Abb. 5: Beispiel: In der Google Search Console gemeldete Verbindungsfehler

Links, damit auch diese in naher Zukunft aktualisiert werden.

Wenn Sie die Social-Counts aufzeichnen müssen, lesen Sie den Artikel von Michael King.

Und schließlich gehen Sie zu dem „Links to Your Site"-Überblick in der Google Search Console, um zu sehen, woher die wichtigsten Links auf Ihre Website stammen – zum Beispiel von Hauptgeschäftspartnern, Nachrichtenmeldungen oder Domains, die häufig zu der Website verlinken. Kontaktieren Sie diese Websites und weisen Sie sie auf die Umstellung Ihrer Website auf HTTPS hin. Vielleicht nutzen Sie auch gleich diese großartige Gelegenheit, um den Gesprächspartnern mitzuteilen, was Ihr Team für deren Unternehmen tun kann – möglicherweise ergibt sich eine Zusammenarbeit.

Content-Security-Policy

Um Probleme mit Mixed Content zu vermeiden, sollte Ihre Website auf jeden Fall eine Content-Security-Policy haben. Dieses Sicherheitskonzept schränkt das Laden bestimmter Ressourcen ein und verhindert XSS-Angriffe. In dem vorliegenden Leitfaden befassen wir uns aber nur mit der Funktion „upgrade-insecure-requests".

Mit „upgrade-insecure-requests" werden interne Linkverweise auf Assets und/oder andere interne Seiten von einer HTTP-Anfrage auf eine HTTPS-Anfrage im Browser aktualisiert. Im Beispiel unten werden die Link-Referenzen im Quellcode:

```
<a href="http://www.example.com/">example</a>
<img src="http://www.example.com/image1.png" />
```

vom Browser automatisch wie folgt aktualisiert:

```
<a href="https://www.example.com/">example</a>
<img src="https://www.example.com/image1.png" />
```

Um diese Funktion in Apache mit .htaccess zu aktivieren, müssen Sie nur den folgenden HTTP-Header hinzufügen:

Header-Satz Content-Security-Policy „upgrade-insecure-requests“

Oder verwenden Sie den folgenden Code in der HTML-Quelle der HTTPS-Version:

```
<meta http-equiv="Content-Security-Policy"
content="upgrade-insecure-requests" />
```

Damit werden alle Links zwangsweise auf HTTPS aktualisiert und möglicherweise arbeiten dann einige Teile der Website nicht mehr richtig. Prüfen Sie unbedingt, ob die Website im Hinblick auf Optik und Funktion normal erscheint.

HSTS

HSTS steht für HTTP Strict Transport Security und wird verwendet, um redundante Weiterleitungen im Browser für nur unter HTTPS laufende Websites zu verhindern.

Um zu verstehen, was das bedeutet, fragen Sie sich (oder Ihre Nutzer): „Wie oft tippt jemand das Protokoll in die Adresszeile eines Browsers, wenn er den Domainnamen eingibt?“

Die Antwort lautet vermutlich: „Praktisch nie“. Das Problem dabei ist, dass die Browser standardmäßig so eingestellt sind, dass sie bei der Eingabe eines Domainnamens automatisch zu HTTP gehen. Tippt der Nutzer also den Namen der Domain (ohne Protokoll) ein, fragt der Browser die HTTP-Version an, etwa *http://www.Beispiel.com*. Dann – vorausgesetzt, die Weiterleitungsregeln sind korrekt eingestellt – sendet der Server dem Browser eine Antwort 301 mit einer neuen Adresse, also *https://www.Beispiel.com/*. Jetzt muss der Browser den gesamten Vorgang noch mal ausführen und die nächste Anfrage an die HTTPS-Version schicken, bevor Inhalt heruntergeladen und dem Nutzer angezeigt werden kann.

Abb. 6: Das fehlende „s“ bei HTTP

So ist HSTS entstanden. Mit HSTS kann eine Website, die ausschließlich und vollständig unter HTTPS läuft, ihren Domainnamen zum „Preload“ durch den Browser in die Liste eintragen lassen. Dahinter steht folgendes Konzept: Die Browser-Teams führen eine hart codierte Liste mit einer langen Reihe von Domainamen, die alle seit Kurzem unter HTTPS laufen. Erscheint der in die Adressleiste eingegebene Name auf dieser Liste, umgeht der Browser die Serververbindungen und leitet den Nutzer mithilfe einer 307 (interne Weiterleitung) direkt auf die HTTPS-Version um, ohne vorher die HTTP-Version anzufragen. Die Wirkung ist wie bei einer 301-Weiterleitung, wird aber vom Browser und nicht vom Server generiert, sodass die HTTPS-Version der Website jetzt im Browser viel schneller geladen wird (wenige Millisekunden statt 100+ Millisekunden) und redundante HTTP-Weiterleitungen vermieden werden.

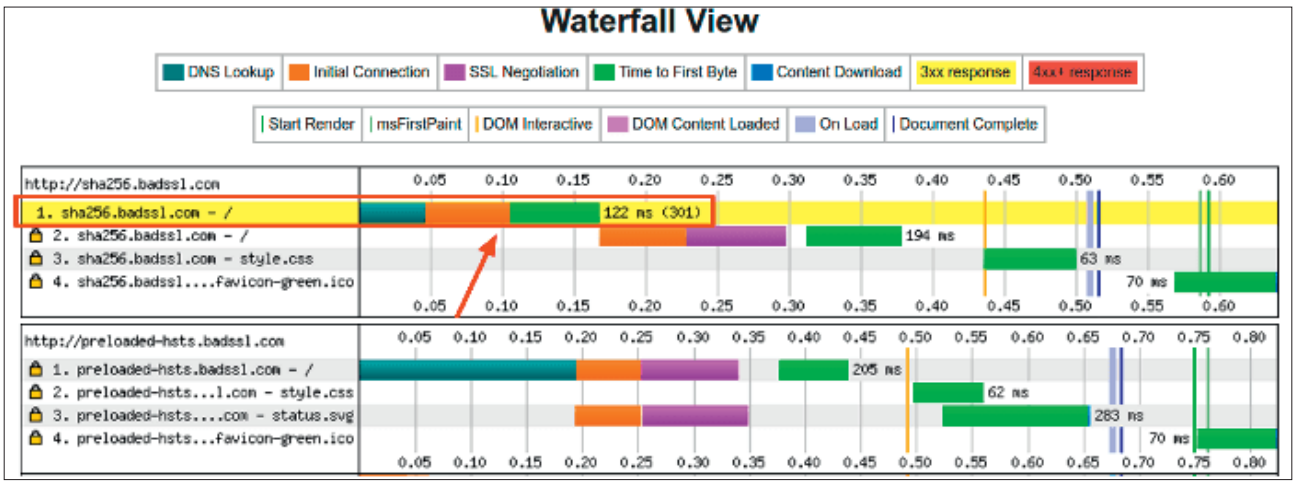


Abb. 7: Wasserfall-Diagramme für Websites mit HSTS preloaded und ohne HSTS preloaded

Website zur Eintragung auf der Preloading-Liste einreichen

Um zum HSTS-Preloading in Browsern zugelassen zu werden, muss die Website folgende Bedingungen erfüllen:

» Verwendung eines gültigen SSL-Zertifikats
» Weiterleitung aller HTTP-Anfragen zur HTTPS-Version
» Alle Hostnamen einschließlich aller Subdomains müssen Inhalt von der HTTPS-Version liefern
» Die „nackte" Domain muss den HSTS-Header liefern, selbst wenn dieser weiterleitet

Hier ist ein Beispiel für die meistbenutzten gültigen HSTS-Header:

```
Strict-Transport-Security: max-age=63072000; includeSubDomains; preload
```

Als bewährte Praxis wird empfohlen, den HSTS-Header generell mit jeder Anfrage auf der HTTPS-Version mitzuliefern. Bei Apache kann das in der .htaccess-Datei im HTTPS-Wurzelverzeichnis erfolgen:

```
<IfModule mod_headers.c>
Header set Strict-Transport-Security „max-age=63072000; includeSubDomains; preload"
</IfModule>
```

Sind alle Bedingungen erfüllt, testen Sie die Website und reichen Sie den Domainnamen zur Eintragung auf der Preloading-Liste ein.

Achtung!

Bevor Sie den Domainnamen bei den Browsern zur Eintragung in die HSTS-Preloading-Liste einreichen, beachten Sie den folgenden Warnhinweis: Ist ein Name einmal auf der Liste, bleibt er auch dort. Nach Einreichen und Zulassung zum HSTS-Preloading wird der Name der Domain allen neuen und zukünftigen Browser-Updates hinzugefügt.

Wird der Domainname auf Antrag des Seitenbetreibers wieder entfernt, kann es Monate dauern, bevor die Browser die Änderung vollzogen haben, und Jahre, bevor die meisten Browser-Nutzer ihren Browser auf die neueste Version aktualisiert haben (denken Sie nur daran, wie lange IE6 nach dem Ende des Supports durch Microsoft noch verwendet wurde). Während dieser Zeit können Browser, die noch die Preload-Liste mit dem Domainnamen verwenden, nicht auf die URLs der unsicheren HTTP-Version zugreifen und diese auf die HTTPS-Version aktualisieren.

Überlegen Sie gut, bevor Sie Ihre Entscheidung treffen, denn sie wirkt langfristig und kann nicht so einfach rückgängig gemacht werden. In Anbetracht der höheren Ladegeschwindigkeit der Website ist es jedoch meistens richtig, diesen Schritt zu tun. Definieren Sie für den Anfang eine kurze „max-age", d. h. die Zeitspanne, für die die Seite in Zukunft verschlüsselt erreichbar sein wird, und prüfen Sie, ob dies für Ihre Nutzer einen Unterschied macht. Erhöhen Sie die Max-age-Zeitspanne nach und nach auf die oben empfohlene Zahl, und wenn alles gut aussieht, bringen Sie die Sache zum Abschluss und reichen den Domainnamen zur Aufnahme in die HSTS-Preloading-Liste ein.

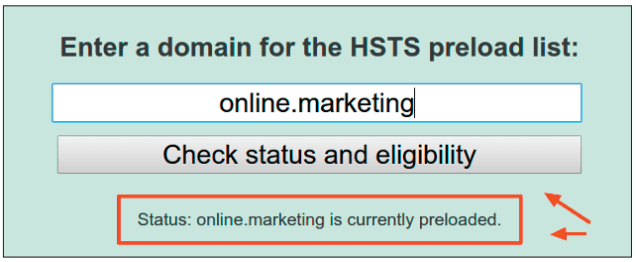


Abb. 8: Für HSTS Preloading prüfen und einreichen

HTTP/2 und Resource Hints

Sie haben es geschafft: Ihre Website läuft problemlos unter HTTPS. Jetzt können Sie endlich von einem der größten Pluspunkte von HTTPS profitieren: Sie können HTTP/2 verwenden. HTTP/2 ist ein Upgrade des älteren HTTP/1.1-Protokolls und bietet eine Menge an Verbesserungen, unter anderem höhere Geschwindigkeit. Fast alle modernen Browser unterstützen HTTP/2, aber nur, wenn die Website unter HTTPS läuft. Da dies bei Ihrer Website nun der Fall ist, können Sie die Server-Software auf HTTP/2 upgraden und „Preload Resource Hints" verwenden, damit Ihre Website bei der Kommunikation mit einem Client-Browser intelligent reagiert und Websites schneller lädt.

Schlussfolgerung

Die Umstellung auf HTTPS ist eine große technische Herausforderung und sollte nicht unterschätzt werden. Gleich wie gründlich man plant und wie sorgfältig man die Umstellung durchführt, vorübergehende Fluktuationen bei der Sichtbarkeit für die organische Suche sind kaum zu vermeiden.

Zudem kann jede kleine Schwierigkeit in der Übergangsphase weitreichende Auswirkungen haben, die die Sichtbarkeit Ihrer Website für die Suchmaschinen weiter verschlechtern. Solange der Prozess andauert, müssen andere wichtige Updates der Website gestoppt werden, um die Suchmaschinen nicht noch mehr zu verwirren.

Nicht jede Website muss auf HTTPS umgestellt werden. Tim Berners-Lee's Aussage „HTTPS überall ist schädlich" ist durchaus berechtigt. Urteilt man jedoch aus der Erfahrung als Benutzer und aus dem Blickwinkel eines SEOs, kann es sich ein Websitenbetreiber heute nicht mehr leisten, HTTPS zu ignorieren. Ich hoffe, dieser Leitfaden wird Ihnen bei der Umstellung behilflich sein. ¶