Engineer

Kanonische URL

die bevorzugte Adresse für einen Seiteninhalt

Was es ist

Eine Seite, die unter mehreren Adressen antwortet, kann eine davon als ihr Original benennen. Diese Adresse ist ihre kanonische URL: die Kopie, die Suchmaschinen und andere Abnehmer indexieren und anzeigen sollen und der sie die Anerkennung für den Inhalt zurechnen sollen. Der Name stammt aus dem Markup, das die Angabe trägt, einem `link`-Element mit `rel="canonical"`. Dieselbe Angabe heißt deshalb auch Canonical Link, Canonical Tag oder kurz rel=canonical.

```
<link rel="canonical" href="https://example.com/guide/" />
```

Duplikate entstehen häufiger als geplant. Ein Artikel ist mit und ohne `www`, über `http` und `https`, mit einem angehängten Trackingparameter, anders sortiert, als Druckansicht oder erneut auf einem zweiten Host erreichbar. RFC 6596 registrierte die Relation 2012 und legte ihre einzige verbindliche Regel fest: Das Ziel muss Inhalt bezeichnen, der entweder ein Duplikat der verweisenden Seite ist oder sie vollständig umfasst. Innerhalb dieser Regel darf das Ziel die Seite selbst sein, auf einem anderen Hostnamen liegen oder ein anderes Schema verwenden. Der HTML-Standard beschreibt es als „bevorzugte URL für das aktuelle Dokument“. Eine Datei ohne HTML-Kopf, etwa ein PDF, kann dieselbe Angabe als HTTP-Header `Link` mit `rel="canonical"` senden.

Eine kanonische Angabe ist ein Hinweis, kein Befehl. Der RFC erlaubt Anwendungen, eine unzulässige Zuordnung zu ignorieren. Googles Dokumentation stuft Weiterleitungen und rel=canonical als starke Signale ein, die Aufnahme in eine Sitemap dagegen als schwaches. Gibt eine Seite nichts an, wählt die Suchmaschine selbst eine Kopie.

Daneben stehen drei leicht damit verwechselte Werkzeuge. Eine permanente Weiterleitung beseitigt das Duplikat für Menschen ebenso wie für Maschinen. Sie ist deshalb die stärkere Wahl, wenn die zusätzliche Adresse nicht erreichbar bleiben muss. Eine Direktive `noindex` hält die Seite vollständig aus dem Index heraus. Das ist ein anderes Ergebnis, als sie mit einer anderen Seite zusammenzuführen. `hreflang` verbindet dagegen Seiten, die Übersetzungen voneinander sind. Übersetzungen sind Entsprechungen; keine ist ein Duplikat einer anderen.

Das Konzept für eine Entscheidung nutzen

Nutzen Sie „Was ist eine kanonische URL?“ für eine konkrete Entscheidung. Wir können helfen, das Konzept mit dem System in Verbindung zu bringen, das Sie entwickeln oder einkaufen. Eine Website oder App, die die KI-Assistenten Ihrer Kunden finden, lesen und bedienen können — bei jeder Veröffentlichung geprüft statt einmal versprochen.

Unsere Leistungen entdecken Kontakt aufnehmen →

Die Annahmen hinter dem Begriff prüfen

Nutzen Sie die Erklärung und die verlinkten Arbeiten, um Fragen für Ihr Projekt zu formulieren. Wir können die Annahmen gemeinsam mit Ihnen prüfen, bevor ein Werkzeug ausgewählt wird.
Hugo-Website mehrsprachig in 3 Sprachen aufgebaut — 108 Vorlagendateien für drei Sprachen und vier Formate

Hugo-Website mehrsprachig in 3 Sprachen aufgebaut Kontakt aufnehmen →

Warum es wichtig ist

Eine kanonische Angabe entscheidet, welche Adresse indexiert wird. Die eigentliche Wahl besteht darin, diese Entscheidung bewusst zu treffen oder sie jemand anderem zu überlassen. Suchmaschinen bestrafen doppelten Text weniger, als dass sie ihn zusammenführen: Sie behalten eine Kopie und ordnen den Rest ihr zu. Diese Kopie erscheint in den Ergebnissen, erhält die Anerkennung für Links auf sämtliche Varianten und wird vom Crawler erneut besucht. Bleibt die Entscheidung der Suchmaschine überlassen, kann die Kopie die Version mit Kampagnenparameter sein, ein versehentlich öffentlicher Staging-Host oder der Nachdruck eines Partners statt des Originals.

Weiterleitung oder kanonische Angabe ist vor allem eine Frage der Menschen. Braucht niemand die zweite Adresse, leiten Sie sie dauerhaft weiter. Das Problem verschwindet dann für alle Clients zugleich. RFC 6596 selbst fordert Autoren auf, eine permanente Weiterleitung auszuschließen, bevor sie zum Link greifen. Muss die Kopie erreichbar bleiben, weil ein Kampagnenlink darauf zeigt, dort eine Druckansicht liegt oder eine zweite Website den Artikel für ihr eigenes Publikum veröffentlicht, ist die kanonische Angabe das Werkzeug. Sie klärt die Frage für Indizes und lässt die Leser unbehelligt.

Die Kosten liegen in der Einheitlichkeit, nicht im Tag. Eine Zeile im Kopf ist günstig. Jede andere Stelle, die die Adresse der Seite nennt, muss damit übereinstimmen: das `og:url` für Linkvorschauen, die Kennungen in strukturierten Daten, die Sitemap und die eigenen internen Links der Website. Nennen zwei davon eine Adresse und eine weitere Stelle eine andere, macht die Seite zwei Aussagen über sich selbst. Ein Abnehmer kann jeder davon glauben. Die kanonische Angabe legt fest, welcher Adresse die vom Inhalt erworbene Reputation zugerechnet wird.

Wie es in der Praxis aussieht

Diese Website stellt anspruchsvolle Anforderungen daran. Derselbe Inhaltsbaum veröffentlicht Seiten in mehreren Darstellungsformen. Der Baum wird auch für die spezialisierten Subdomains gebaut.

Die Vorlagen klären Duplikate mit einer Regel: Eine Seite ist nur dann ihr eigenes Original, wenn sie für die Adresse geschrieben wurde, unter der sie ausgeliefert wird. Auf der Hauptwebsite gilt das für alles, also nennt sich jede Seite selbst. Für eine spezialisierte Subdomain wurde nur eine eigene Startseite geschrieben. Diese nennt sich deshalb selbst, während gespiegelte Arbeiten, gemeinsame Seiten und jede vom Generator erstellte Liste auf die Kopie unter demselben Pfad der Hauptwebsite verweisen. Die Voreinstellung geht in die sichere Richtung: Alles zeigt zur Hauptwebsite, sofern es nichts anderes angibt. Eine Seite, die eine eigene Identität verdient, muss sie im Front Matter ausdrücklich anfordern, wo diese Entscheidung sichtbar ist. Eine neue Seitenart erbt damit die vorsichtige Antwort, statt sich stillschweigend als Original auszugeben.

Der Wert wird in einer Vorlage berechnet und überall dort gelesen, wo er erscheint: von `rel="canonical"`, `og:url` und `twitter:url` sowie von `@id` und `url` des Knotens für die strukturierten Daten der Seite. Sitemap und Sprachalternativen stellen dieselbe Frage als Ja oder Nein. Sie werden nur ausgegeben, wenn die kanonische Adresse die eigene Adresse der Seite ist. Eine Adresse zur Indexierung vorzuschlagen und gleichzeitig eine andere als Original zu benennen ist ein Widerspruch, auf den ein Crawler reagieren kann, indem er keiner glaubt. Eine Kopie auf einem spezialisierten Host wird weiterhin vollständig gerendert, und ihr Sprachumschalter funktioniert weiter. Zurückgehalten wird die Aufforderung zur Indexierung, niemals der Inhalt. Der Markdown-Zwilling wird auf der HTML-Seite mit `rel="alternate"` und seinem Medientyp angekündigt. Der HTML-Standard versteht das als dasselbe Dokument in einem anderen Format statt als konkurrierende Seite.

Die Ausgabeprüfung der Hauptwebsite verlangt auf normalen HTML-Seiten genau eine absolute, selbstreferenzierende kanonische Adresse und prüft interne Seitenlinks auf fehlende abschließende Schrägstriche. Eigenständige ausgelieferte Dokumente und die Social Card folgen eigenen Regeln; die Prüfungen der spezialisierten Hosts prüfen die Unterscheidung zwischen Spiegeln und Originalen.

Wo es schiefgeht

Suchergebnisse zeigen eine Adresse, die niemand gewählt hat. Ein Trackingparameter, ein ungewollt erreichbarer Host oder ein Spiegel erscheint als gelistete URL. Die Seite nennt entweder keine kanonische Adresse oder nennt eine, während ihr Sitemapeintrag, `og:url` und interne Links etwas anderes nennen. Die Suchmaschine gewichtet die widersprüchlichen Signale und entscheidet. Geben Sie auf jeder Seite eine absolute kanonische Adresse aus, beim Original eine selbstreferenzierende. Leiten Sie die Adressen in Sitemap, Vorschauen und strukturierten Daten aus demselben Wert ab.

Die zweite Seite einer Reihe verschwindet aus dem Index. Jede Seite einer paginierten Liste oder eines aufgeteilten Artikels nennt die erste Seite als kanonisch. Die erste Seite ist weder ein Duplikat noch eine vollständige Obermenge der zweiten. Was nur auf späteren Seiten steht, kann deshalb als Duplikat behandelt und übergangen werden. RFC 6596 warnt ausdrücklich davor. Lassen Sie jede Seite einer Reihe sich selbst nennen oder verweisen Sie alle auf eine einzige Gesamtansicht, die tatsächlich alles enthält.

Ein in robots.txt gesperrtes Duplikat bleibt gelistet, oder das Original verschwindet mit. Die Adresse eines Duplikats in robots.txt zu sperren hindert den Crawler daran, die Seite einschließlich ihrer kanonischen Angabe zu lesen. Google weist darauf hin, dass eine gesperrte URL trotzdem ohne ihren Inhalt indexiert werden kann. Duplikate mit `noindex` zu markieren entfernt sie vollständig aus der Suche, statt ihre erworbene Reputation dem Original zuzuschreiben. Lassen Sie Duplikate crawlbar und überlassen Sie die Zusammenführung einer kanonischen Angabe oder Weiterleitung.

Die angegebene kanonische Adresse wird ignoriert. Das Ziel leitet dauerhaft woandershin weiter, nennt selbst eine andere kanonische Adresse, liefert 404 oder konkurriert mit einem zweiten kanonischen Element auf derselben Seite. Auch ein Skript kann den Link nach Empfang des HTML umschreiben. Jede dieser Situationen macht die Zuordnung unzulässig. Der RFC erlaubt dem Abnehmer

dann, nach eigenem Ermessen zu entscheiden. Verweisen Sie auf die endgültige Adresse, die 200 zurückgibt, liefern Sie ein Element im ausgelieferten HTML-Kopf aus und lassen Sie es danach von keinem Skript ändern.

Weiterführende Lektüre

- RFC 6596: Die kanonische Linkrelation — die Registrierung, die Regel für Duplikate oder Obermengen und die zu vermeidenden Zuordnungen.
- HTML-Standard: Linktyp „canonical“ — wie HTML das Schlüsselwort definiert und wo es stehen darf.
- RFC 8288: Verknüpfungen im Web — der HTTP-Header `Link`, der eine kanonische Angabe für Dateien ohne HTML-Kopf übermittelt.
- Google Search Central: Doppelte URLs zusammenführen — wie eine Suchmaschine Weiterleitungen, kanonische Angaben und Sitemaps gegeneinander gewichtet.

Glossar

Für Ihre tatsächlichen Rahmenbedingungen auswählen

Beschreiben Sie den Arbeitsablauf, die vorhandenen Schnittstellen und die Nutzer, die die Veränderung benötigen. Wir können eine prüfbare Lieferung definieren und die Übergabe so gestalten, dass Ihr Team die gelieferten Bestandteile warten kann.

Preise ansehen Kontakt aufnehmen →

Die Begriffe praktisch nutzbar machen

Fragen Sie, wie „Was ist eine kanonische URL?“ auf Ihre Organisation zutrifft. Ein konkretes Beispiel aus Ihrem bestehenden System reicht aus, um das Gespräch zu beginnen.

Unsere Leistungen entdecken

welcome@engineer.company +45 50 10 61 84 Kopenhagen, Dänemark

Nehmen Sie diese Seite mit

Kontaktkarte

Mit Sorgfalt und Liebe gestaltet von D.

CVR 45969703 · Tagensvej 105C, 3. th, 2200 København N

Copyright © 2025 – heute · Engineer ApS · Alle Rechte vorbehalten.

Gedruckt von https://engineer.company/de/glossary/canonical/