Engineer

Kanonisk URL

den ene adresse en side kalder sin egen

Hvad det er

Når en side svarer på flere adresser, kan den udpege én af dem som sin original, og den udpegede adresse er sidens kanoniske URL: den kopi, siden beder søgemaskiner og andre modtagere om at indeksere, vise og give æren. Navnet kommer fra den markup, der bærer den, et `link` -element med `rel="canonical"` , og derfor hører man også kanonisk link, canonical-tag eller bare rel=canonical.

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

Dubletter opstår oftere, end nogen planlægger. Den samme artikel kan nås med og uden `www` , over `http` og `https` , med en sporingsparameter i enden, sorteret på en anden måde, som printvisning eller genudgivet på et andet domæne. RFC 6596 fra 2012 registrerede relationen og gav den én fast regel: målet skal pege på indhold, der enten er en dublet af den henvisende side eller et supersæt af den. Inden for den regel må målet være siden selv, ligge på et andet værtsnavn eller bruge et andet skema. HTML-standarden kalder det dokumentets foretrukne URL. En fil uden HTML-head, for eksempel en PDF, kan sende samme besked som en HTTP-header af typen `Link` med `rel="canonical"` .

En kanonisk URL er et hint og ikke en ordre. RFC'en giver programmer lov til at se bort fra en udpegning, der er forkert, og Googles dokumentation regner omdirigeringer og rel=canonical for stærke signaler, mens en plads i et sitemap er et svagt signal. Hvor en side ikke angiver noget, vælger søgemaskinen selv en kopi.

Den står ved siden af tre værktøjer, som er lette at forveksle med den. En permanent omdirigering fjerner dubletten for mennesker såvel som maskiner og er derfor det stærkere valg, når den ekstra adresse ikke behøver at kunne nås. Et

`noindex` -direktiv holder en side helt ude af indekset, og det er noget andet end at lægge den sammen med en anden side. Og `hreflang` binder sider sammen, der er oversættelser af hinanden; oversættelser er ligeværdige, og ingen af dem er en dublet af en anden.

Brug begrebet i en beslutning

Brug »Hvad er en kanonisk URL?« i en konkret beslutning. Vi kan knytte begrebet til det system, I bygger eller køber. Et site eller en app, som de AI-assistenter, jeres kunder bruger, kan finde, læse og handle på, kontrolleret ved hver udgivelse i stedet for lovet én gang.

Se vores ydelser Kontakt os →

Undersøg antagelserne bag ordet

Brug forklaringen og det tilknyttede arbejde til at formulere spørgsmål til jeres projekt. Vi kan undersøge antagelserne sammen, før et værktøj vælges. Har bygget et Hugo-site på 3 sprog — 108 skabelonfiler, der betjener tre sprog og fire formater

Har bygget et Hugo-site på 3 sprog Kontakt os →

Hvorfor det betyder noget

Den beslutning, en kanonisk URL træffer, er, hvilken adresse der bliver den indekserede, og det eneste reelle valg står mellem at træffe den bevidst og at lade en anden træffe den. Søgemaskiner straffer ikke så meget dubleret tekst, som de slår den sammen: de beholder én kopi og folder resten ind i den. Den kopi, de beholder, er den, folk ser i søgeresultaterne, den, som links til alle varianterne bliver krediteret, og den, en crawler vender tilbage til. Overlader man det til søgemaskinen, kan det blive versionen med en kampagneparameter, et testmiljø, der er sluppet ud, eller en partners genudgivne kopi i stedet for originalen.

Omdirigering kontra kanonisk URL er mest et spørgsmål om mennesker. Har ingen brug for den anden adresse, så omdiriger permanent, og problemet forsvinder for alle klienter på én gang; RFC 6596 anbefaler selv at undersøge, om en omdirigering kunne klare opgaven. Skal kopien kunne nås, fordi et kampagnelink peger på den, et printlayout bor der, eller et andet site bringer artiklen for sit eget publikum, så er den kanoniske URL værktøjet, fordi den afgør spørgsmålet for indeksene og lader læserne være.

Omkostningen ligger i sammenhængen og ikke i selve tagget. Én linje i head er billig. Alle andre steder, der nævner sidens adresse, skal være enige med den: den `og:url` , som linkforhåndsvisninger bruger, identifikatorerne i strukturerede data, sitemappet og sitets egne interne links. Nævner to af dem én adresse og

det tredje en anden, påstår siden to ting om sig selv, og en modtager kan tro på hvilken som helst af dem. Det, den kanoniske URL afgør, er, hvilken adresse den anseelse, indholdet opnår, bliver knyttet til.

Sådan ser det ud i praksis

Dette site er et krævende tilfælde. Det samme indholdstræ udgiver sider i flere repræsentationer. Træet bygges også til de fokuserede subdomæner.

Skabelonerne afgør dubletterne med én regel: en side er kun sin egen original, hvis den blev skrevet til den adresse, den serveres fra. På hovedsitet gælder det alt, så hver side udpeger sig selv. Et fokuseret subdomæne har skrevet sin egen forside og intet andet, så den side udpeger sig selv, mens det spejlede arbejde, de fælles sider og alle de lister, generatoren bygger, peger på hovedsitets kopi på samme sti. Standardvalget går i den sikre retning: alt peger hjem, medmindre det siger andet, og en side, der fortjener sin egen identitet, må bede om den i front matter, hvor anmodningen er synlig. En ny slags side arver dermed det forsigtige svar i stedet for i stilhed at give sig ud for at være en original.

Værdien beregnes i én skabelon og læses alle de steder, den optræder: af `rel="canonical"`, af `og:url` og `twitter:url` og af `@id` og `url` i sidens node med strukturerede data. Sitemappet og sprogalternativerne stiller samme spørgsmål som ja eller nej og udskrives kun, når den kanoniske URL er sidens egen adresse, for at bede om at få én adresse indekseret og samtidig kalde en anden for originalen er en modsigelse, som en crawler kan løse ved at stole på ingen af dem. En kopi på et fokuseret subdomæne vises stadig i sin helhed, og sprogvælgeren virker; det, den holder tilbage, er anmodningen om at blive indekseret, aldrig indholdet. Markdown-tvillingen annonceres fra HTML-siden med `rel="alternate"` og sin medietype, hvilket HTML-standarden læser som det samme dokument i et andet format og ikke som en konkurrerende side.

Tjekket af hovedsitets output kræver én absolut, selvrefererende kanonisk URL på almindelige HTML-sider og tjekker interne sidelinks for manglende afsluttende skråstreger. Selvstændige dokumenter til andre værter og det sociale kort følger deres egne regler; tjek af de fokuserede værter kontrollerer skellet mellem spejle og originaler.

Hvor det går galt

Søgeresultaterne viser en adresse, ingen har valgt. En sporingsparameter, et forkert værtsnavn eller et spejl dukker op som den viste URL. Siden angiver enten ingen kanonisk URL, eller den angiver én, mens dens sitemap-post, `og:url` og interne links nævner noget andet, så søgemaskinen vejer blandede signaler og vælger selv. Sæt én absolut kanonisk URL på hver side, også en selvrefererende på originalen, og afled adresserne i sitemap, forhåndsvisning og strukturerede data af netop den værdi.

Side to i en serie forsvinder fra indekset. Hver side i en pagineret liste eller en opdelt artikel udpeger side ét som sin kanoniske URL. Side ét er hverken en dublet eller et supersæt af side to, så det, der kun står på de senere sider, kan blive behandlet som dublet og ignoreret, hvilket RFC 6596 advarer om ved navn. Lad hver side i en serie udpege sig selv, eller lad dem alle pege på én samlet side, som virkelig rummer det hele.

En dublet, der er spærret i robots.txt, bliver ved med at stå i søgeresultaterne, eller originalen forsvinder sammen med den. Spærrer man en dublets adresse i robots.txt, kan crawleren ikke læse siden og dermed heller ikke dens kanoniske URL, og Google gør opmærksom på, at en spærret URL stadig kan blive indekseret uden sit indhold. Mærker man dubletter `noindex`, fjernes de helt fra søgning i stedet for at lægge det, de har opnået, over på originalen. Lad dubletter være tilgængelige for crawleren, og lad en kanonisk URL eller en omdirigering stå for sammenlægningen.

Den angivne kanoniske URL bliver ignoreret. Målet omdirigerer permanent videre, udpeger selv en anden kanonisk URL, svarer med 404 eller konkurrerer med et andet canonical-element på samme side; et script kan også skrive linket om, efter at HTML'en er ankommet. Hver af dem gør udpegningen forkert, og RFC'en lader en modtager falde tilbage på sin egen vurdering. Peg på den endelige adresse, der svarer med 200, sæt ét element i head i den HTML, serveren sender, og lad intet script ændre det bagefter.

Læs mere

- RFC 6596: The Canonical Link Relation — registreringen, reglen om dublet eller supersæt og de udpegninger, man skal undgå.
- HTML Standard: link type "canonical" — hvordan HTML definerer nøgleordet, og hvor det må stå.

- RFC 8288: Web Linking — HTTP-headeren `Link`, der bærer en kanonisk URL for filer uden HTML-head.
- Google Search Central: consolidate duplicate URLs — hvordan én søgemaskine vejer omdirigeringer, kanoniske URL'er og sitemaps op mod hinanden.

Ordliste

Vælg ud fra jeres faktiske vilkår

Beskriv arbejdsgangen, de eksisterende grænseflader og de brugere, der har brug for ændringen. Vi kan definere en testbar leverance og en overdragelse, der gør teamet i stand til at vedligeholde løsningen.

Se, hvad det koster Kontakt os →

Gør begreberne praktiske

Spørg, hvordan »Hvad er en kanonisk URL?« gælder i jeres organisation. Et konkret eksempel fra jeres nuværende system er nok til at begynde samtalen.

Se vores ydelser

welcome@engineer.company +45 50 10 61 84 København, Danmark

Tag denne side med dig

Visitkort

Lavet med kærlighed af D.

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

Copyright © 2025 – nu · Engineer ApS · Alle rettigheder forbeholdes.

Udskrevet fra https://engineer.company/da/glossary/canonical/