Engineer

Kanonisk URL

den ene adressen en side oppgir som sin egen

Hva det er

En kanonisk URL, også kalt kanonisk lenke, canonical-tagg eller bare rel=canonical, er adressen en side utpeker som sin original når den svarer på flere adresser. Det er kopien siden ber søkemotorer og andre mottakere om å indeksere, vise og gi æren. Navnet kommer fra koden som bærer den, et `link`-element med `rel="canonical"`.

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

Duplikater oppstår oftere enn noen planlegger for. Én artikkel kan nås med og uten `www`, over `http` og `https`, med en sporingsparameter på slutten, sortert på en annen måte, som utskriftsvisning eller publisert på nytt hos en annen vert. RFC 6596, publisert i 2012, registrerte relasjonen og satte dens ene harde regel: målet må peke på innhold som enten er et duplikat av siden som viser til det, eller en større helhet som rommer hele den siden. Innenfor den regelen kan målet være siden selv, ligge på et annet vertsnavn eller bruke et annet skjema i adressen. HTML-standarden beskriver den som «preferred URL for the current document», den foretrukne URL-en for dokumentet. En fil uten HTML-hode, som en PDF, kan gi samme beskjed i en HTTP-header av typen `Link` med `rel="canonical"`.

En kanonisk URL er et hint og ikke en ordre. RFC-en lar programmer se bort fra en utpeking som er feilaktig, og dokumentasjonen til Google regner omdirigeringer og rel=canonical som sterke signaler, mens det å stå i en sitemap regnes som et svakt. Der en side ikke erklærer noe, velger søkemotoren en kopi selv.

Den står ved siden av tre verktøy som er lette å blande sammen med den. En permanent omdirigering fjerner duplikatet for mennesker så vel som maskiner, og det gjør den til det sterkeste valget når den ekstra adressen ikke har noen grunn til å være tilgjengelig. Et `noindex` -direktiv holder en side helt utenfor indeksen, og det er et annet utfall enn å slå den sammen med en annen side. Og hreflang knytter sammen sider som er oversettelser av hverandre. Oversettelser er likeverdige, og ingen av dem er et duplikat av en annen.

Bruk begrepet i en beslutning

Bruk «Hva er en kanonisk URL?» i en konkret beslutning. Vi kan knytte begrepet til systemet dere bygger eller kjøper. Et nettsted eller en app som AI-assistentene kundene dine bruker, kan finne, lese og handle ut fra, kontrollert ved hver utgivelse i stedet for lovet én gang.

Se tjenestene våre Ta kontakt →

Undersøk antakelsene bak ordet

Bruk forklaringen og det tilknyttede arbeidet til å formulere spørsmål for prosjektet. Vi kan undersøke antakelsene sammen før et verktøy velges. Bygde et flerspråklig Hugo-nettsted på 3 språk — 108 malfiler som betjener tre språk og fire formater

Bygde et flerspråklig Hugo-nettsted på 3 språk Ta kontakt →

Hvorfor det er viktig

Beslutningen en kanonisk URL tar, er hvilken adresse som blir den indekserte, og det eneste reelle valget står mellom å ta den bevisst og å la noen andre ta den. Søkemotorer straffer ikke så mye duplisert tekst som de slår den sammen: de beholder én kopi og bretter resten inn i den. Kopien de beholder, er den folk ser i søkeresultatene, den som lenkene til alle variantene krediteres, og den en søkerobot kommer tilbake til. Overlatt til søkemotoren kan den kopien være versjonen med en kampanjeparameter, testverten som lakk ut, eller en partners gjenpubliserte kopi i stedet for originalen.

Omdirigering eller kanonisk URL er for det meste et spørsmål om mennesker. Trenger ingen den andre adressen, omdirigerer du den permanent, og problemet forsvinner for alle klienter på en gang. RFC 6596 ber selv forfattere om å utelukke en permanent omdirigering før de griper til lenken. Må kopien fortsatt kunne nås, fordi en kampanjelenke peker på den, en utskriftsvisning bor der eller et annet nettsted bærer artikkelen for sitt eget publikum, er den kanoniske URL-en verktøyet. Den avgjør spørsmålet for indeksene og lar leserne være i fred.

Kostnaden ligger i konsekvensen og ikke i selve taggen. Én linje i hodet er billig. Hvert annet sted som navngir sidens adresse, må være enig med den: `og:url` som lenkeforhåndsvisninger bruker, identifikatorene i strukturerte sidedata,

sitemapen og nettstedets egne interne lenker. Når to av dem navngir én adresse og den tredje en annen, kommer siden med to påstander om seg selv, og en mottaker kan tro på hvilken som helst av dem. Det en kanonisk URL avgjør, er hvilken adresse all anseelse innholdet bygger opp, blir knyttet til.

Slik ser det ut i praksis

Dette nettstedet er et krevende tilfelle. Det samme innholdstreet publiserer sider i flere representasjoner. Treet bygges også for de fokuserte underdomenene.

Malene løser duplikatene med én regel: en side er sin egen original bare hvis den ble skrevet for adressen den serveres fra. På hovednettstedet gjelder det alt, så hver side navngir seg selv. Et fokusert underdomene skrev sin egen forside og ingenting annet. Den siden navngir seg selv, mens det speilede arbeidet, de delte sidene og hver liste generatoren bygger, peker på kopien med samme sti på hovednettstedet. Standardvalget går i den trygge retningen: alt peker hjem med mindre det sier noe annet, og en side som fortjener sin egen identitet, må be om den i front matter, der bønnen er synlig. En ny type side arver derfor det forsiktige svaret i stedet for i det stille å gi seg ut for å være en original.

Verdien regnes ut i én mal og leses overalt der den dukker opp: av `rel="canonical"`, av `og:url` og `twitter:url`, og av `@id` og `url` i sidens node i de strukturerte sidedataene. Sitemapen og språkalternativene stiller det samme spørsmålet som ja eller nei og skrives bare ut når den kanoniske URL-en er sidens egen adresse. Å be om at én adresse indekseres og samtidig navngi en annen som original er en selvmotsigelse en søkerobot kan løse ved å stole på ingen av dem. En kopi på et fokusert underdomene vises likevel i sin helhet, og språkvelgeren virker fortsatt. Det kopien holder tilbake, er forespørselen om å bli indeksert, aldri innholdet. Markdown-tvillingen annonseres fra HTML-siden med `rel="alternate"` og medietypen sin, noe HTML-standarden leser som det samme dokumentet i et annet format og ikke som en rivaliserende side.

Sjekken av hovednettstedets resultat krever én absolutt, selvrefererende kanonisk URL på vanlige HTML-sider og sjekker interne sidelenker for manglende avsluttende skråstreker. Selvstendige dokumenter til andre verter og det sosiale kortet følger egne regler; sjekker av de fokuserte vertene kontrollerer skillet mellom speil og originaler.

Hvor det går galt

Søkeresultatene viser en adresse ingen har valgt. En sporingsparameter, en vert på avveie eller et speil dukker opp som den listede URL-en. Siden erklærer enten ingen kanonisk URL, eller den erklærer én mens oppføringen i sitemapen, `og:url` og de interne lenkene navngir noe annet. Da veier søkemotoren blandede signaler og velger selv. Skriv ut én absolutt kanonisk URL på hver side, også en selvrefererende på originalen, og avled adressene i sitemapen, forhåndsvisningen og de strukturerte sidedataene fra den samme verdien.

Side to i en serie faller ut av indeksen. Hver side i en paginert liste eller en oppdelt artikkel oppgir side én som sin kanoniske URL. Side én er verken et duplikat av side to eller en større helhet som rommer den, så det som bare står på de senere sidene, kan bli behandlet som duplikat og oversett. RFC 6596 advarer mot akkurat dette. La hver side i en serie navngi seg selv, eller la alle sidene peke på én samleside som virkelig inneholder alt.

Et duplikat sperret i robots.txt blir stående i indeksen, eller originalen forsvinner med det. Å sperre adressen til et duplikat i robots.txt hindrer søkeroboten i å lese siden, også den kanoniske lenken, og Google påpeker at en sperret URL likevel kan bli indeksert uten innholdet sitt. Å merke duplikater med `noindex` fjerner dem helt fra søk i stedet for å samle det de har opparbeidet, i originalen. La duplikatene være åpne for søkeroboter, og la en kanonisk URL eller en omdirigering gjøre sammenslåingen.

Den oppgitte kanoniske URL-en blir oversett. Målet omdirigerer permanent til et annet sted, oppgir en annen kanonisk URL selv, svarer 404 eller konkurrerer med et andre kanonisk element på samme side. Et skript kan også skrive om lenken etter at HTML-en har kommet fram. Hver av disse gjør utpekingen feilaktig, og RFC-en lar mottakeren falle tilbake på sitt eget skjønn. Pek på den endelige adressen som svarer 200, skriv ut ett element i HTML-hodet slik det serveres, og la ingen skript endre det etterpå.

Les mer

- RFC 6596: The Canonical Link Relation — registreringen, regelen om at målet må være et duplikat eller romme hele siden, og utpekingene du bør unngå.
- HTML Standard: link type «canonical» — hvordan HTML definerer nøkkelordet, og hvor det kan stå.

- RFC 8288: Web Linking — HTTP-headeren `Link`, som bærer en kanonisk URL for filer uten HTML-hode.
- Google Search Central: consolidate duplicate URLs — hvordan én søkemotor veier omdirigeringer, kanoniske lenker og sitemaps mot hverandre.

Ordliste

Velg ut fra deres faktiske rammer

Beskriv arbeidsflyten, eksisterende grensesnitt og brukerne som trenger endringen. Vi kan definere en testbar leveranse og en overlevering som gjør teamet i stand til å vedlikeholde løsningen.

Se hva det koster Ta kontakt →

Gjør begrepene praktiske

Spør hvordan «Hva er en kanonisk URL?» gjelder i organisasjonen deres. Et konkret eksempel fra dagens system er nok til å starte samtalen.

Se tjenestene våre

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

Ta med deg denne siden

Visittkort

Laget med kjærlighet av D.

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

Opphavsrett © 2025 – nå · Engineer ApS · Alle rettigheter forbeholdt.

Skrevet ut fra https://engineer.company/nb/glossary/canonical/