Engineer

Kanonisk URL

den enda adress som en sida säger är dess egen

Vad det är

En kanonisk URL är den adress som en sida med flera adresser pekar ut som sitt original: kopian som den ber sökmotorer och andra konsumenter att indexera, visa och tillgodoräkna. Namnet kommer från märkningen som bär uppgiften, ett `link` -element med `rel="canonical"` , och därför kallas samma sak också kanonisk länk, canonical-tagg eller bara rel=canonical. I vardagligt tal säger många svenska utvecklare helt enkelt canonical.

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

Dubbletter uppstår oftare än någon planerar för. En och samma artikel kan nås med och utan `www` , över `http` och `https` , med en spårningsparameter på slutet, sorterad på ett annat sätt, som utskriftsvy eller återpublicerad på en annan värd. RFC 6596, som publicerades 2012, registrerade relationen och gav den dess enda hårda regel: målet måste identifiera innehåll som antingen är en dubblett av den hänvisande sidan eller en övermängd av den, alltså innehåller allt den har, eventuellt också mer. Inom den regeln får målet vara sidan själv, ligga på ett annat värdnamn eller använda ett annat URI-schema. HTML-standarden beskriver det som ”preferred URL for the current document”. En fil som saknar HTML-huvud, som en PDF, kan skicka samma besked i HTTP-headern `Link` med `rel="canonical"` .

En kanonisk URL är en vink och ingen order. RFC 6596 låter program bortse från en utpekning som är felaktig, och dokumentationen från Google räknar omdirigeringar och rel=canonical som starka signaler och att sidan står i en webbplatskarta som en svag. Där en sida inte deklarerar något väljer sökmotorn en kopia själv.

Den står bredvid tre verktyg som är lätta att blanda ihop med den. En permanent omdirigering tar bort dubbletten för människor lika väl som för maskiner, och är därför det starkare valet så fort den extra adressen inte har något skäl att gå att nå. Direktivet `noindex` håller en sida helt utanför indexet, vilket är ett annat utfall än att slå ihop den med en annan sida. Och `hreflang` knyter ihop sidor som är översättningar av varandra; översättningar är likvärdiga, och ingen av dem är en dubblett av en annan.

Använd begreppet i ett beslut

Använd ”Vad är en kanonisk URL?” i ett konkret beslut. Vi kan knyta begreppet till systemet ni bygger eller köper. En webbplats eller app som de AI-assistenter dina kunder använder kan hitta, läsa och agera på, kontrollerad vid varje release i stället för utlovad en gång.

Se våra tjänster Hör av er →

Granska antagandena bakom ordet

Använd förklaringen och det länkade arbetet för att formulera frågor till projektet. Vi kan undersöka antagandena tillsammans innan ett verktyg väljs. Byggde en flerspråkig Hugo-webbplats på 3 språk — 108 mallfiler som täcker tre språk och fyra format

Byggde en flerspråkig Hugo-webbplats på 3 språk Hör av er →

Varför det spelar roll

Beslutet som en kanonisk URL fattar är vilken adress som blir den indexerade, och det enda verkliga valet står mellan att fatta det med avsikt och att låta någon annan fatta det. Sökmotorer straffar inte dubblerad text så mycket som de slår ihop den: de behåller en kopia och viker in resten i den. Kopian som behålls är den som människor ser i sökresultaten, den som länkar till varje variant tillgodoräknas och den som en sökrobot återvänder till. Lämnas valet åt sökmotorn kan den kopian bli versionen med en kampanjparameter, en stagingserver som läckte ut eller en partners återpublicering i stället för originalet.

Omdirigering eller kanonisk länk är mest en fråga om människor. Behöver ingen den andra adressen, så omdirigera den permanent, och problemet försvinner för varje klient på en gång; RFC 6596 ber själv författare att utesluta en permanent omdirigering innan de tar till länken. Måste kopian gå att nå, för att en kampanjlänk pekar på den, en utskriftslayout bor där eller en annan webbplats bär artikeln för sin egen publik, då är den kanoniska länken rätt verktyg, eftersom den avgör frågan för indexen och lämnar läsarna i fred.

Kostnaden ligger i konsekvensen, inte i taggen. En rad i sidhuvudet är billig. Varje annat ställe som namnger sidans adress måste hålla med den: `og:url` som länkförhandsvisningar använder, identifierarna i strukturerad data,

webbplatskartan och webbplatsens egna interna länkar. Om två av dem namnger en adress och den tredje en annan gör sidan två påståenden om sig själv, och en konsument kan tro på vilket som helst. Det en kanonisk URL avgör är vilken adress allt anseende som innehållet förtjänar knyts till.

Så ser det ut i praktiken

Den här webbplatsen är ett krävande exempel. Ur samma innehållsträd publiceras sidor i flera format. Trädet byggs också för de fokuserade underdomänerna.

Mallarna löser dubbletterna med en enda regel: en sida är sitt eget original bara om den skrevs för den adress den serveras från. På huvudwebbplatsen gäller det allt, så varje sida namnger sig själv. En fokuserad underdomän har skrivit sin egen startsida och inget annat, så den sidan namnger sig själv, medan det speglade arbetet, de delade sidorna och varje lista som generatorn bygger pekar på huvudwebbplatsens kopia på samma sökväg. Standardläget går åt det säkra hållet: allt pekar hem om inget annat sägs, och en sida som förtjänar en egen identitet måste be om den i sin front matter, där begäran syns. En ny sorts sida ärver därför det försiktiga svaret i stället för att i tysthet utge sig för att vara ett original.

Värdet räknas ut i en enda mall och läses överallt där det förekommer: av `rel="canonical"`, av `og:url` och `twitter:url`, och av `@id` och `url` i sidans nod för strukturerad data. Webbplatskartan och språkalternativen ställer samma fråga i ja- eller nej-form och skrivs bara ut när den kanoniska adressen är sidans egen, eftersom att be om att en adress ska indexeras och samtidigt peka ut en annan som original är en motsägelse som en sökrobot kan lösa genom att lita på ingen av dem. En kopia på en fokuserad värd renderas ändå i sin helhet, och dess språkväljare fungerar; det den håller inne är begäran om att bli indexerad, aldrig innehållet. Markdown-tvillingen annonseras från HTML-sidan med `rel="alternate"` och sin mediatyp, vilket HTML-standarden läser som samma dokument i ett annat format och inte som en konkurrerande sida.

Kontrollen av huvudwebbplatsens resultat kräver en absolut, självrefererande kanonisk URL på vanliga HTML-sidor och kontrollerar interna sidlänkar för saknade avslutande snedstreck. Fristående dokument för andra värdar och det sociala kortet följer egna regler; kontroller av de fokuserade värdarna verifierar skillnaden mellan speglar och original.

Var det går fel

Sökresultaten visar en adress som ingen valde. En spårningsparameter, en vilsen värd eller en spegel dyker upp som den listade URL:en. Sidan deklarerar antingen ingen kanonisk URL, eller deklarerar en medan dess post i webbplatskartan, `og:url` och de interna länkarna namnger något annat, så sökmotorn väger blandade signaler och väljer själv. Skriv ut en absolut kanonisk URL på varje sida, också en självrefererande på originalet, och härled adresserna i webbplatskartan, förhandsvisningen och den strukturerade datan från samma värde.

Sida två i en serie faller ur indexet. Varje sida i en paginerad lista eller en uppdelad artikel pekar ut sida ett som sin kanoniska URL. Sida ett är varken en dubblett eller en övermängd av sida två, så allt som bara finns på senare sidor kan behandlas som dubblett och ignoreras, en risk som RFC 6596 varnar för uttryckligen. Låt varje sida i en serie namnge sig själv, eller peka alla sidor mot en enda visa-allt-sida som verkligen innehåller allt.

En dubblett som blockerats i robots.txt ligger kvar i sökresultaten, eller så försvinner originalet med den. Att blockera en dubbletts adress i robots.txt hindrar sökroboten från att läsa sidan, också dess kanoniska länk, och Google påpekar att en blockerad URL ändå kan indexeras utan sitt innehåll. Att märka dubbletter med `noindex` tar bort dem helt ur sökningen i stället för att föra över det de har förtjänat till originalet. Låt dubbletterna vara åtkomliga för sökrobotar och låt en kanonisk URL eller en omdirigering sköta sammanslagningen.

Den deklarerade kanoniska URL:en ignoreras. Målet omdirigerar permanent någon annanstans, anger en annan adress som sin egen kanoniska URL, svarar 404 eller konkurrerar med ett andra canonical-element på samma sida; ett skript kan också skriva om länken efter att HTML-koden har kommit fram. Vart och ett av dessa fall gör utpekningen felaktig, och RFC 6596 låter en konsument falla tillbaka på sitt eget omdöme. Peka på den slutliga adressen som svarar 200, skriv ut ett enda element i huvudet på den HTML som serveras, och låt inget skript ändra det efteråt.

Läs mer

- RFC 6596: The Canonical Link Relation — registreringen, regeln om dubblett eller övermängd och de utpekningar som ska undvikas.

- HTML Standard: link type ”canonical” — hur HTML definierar nyckelordet och var det får stå.
- RFC 8288: Web Linking — HTTP-headern `Link`, som bär en kanonisk URL för filer utan HTML-huvud.
- Google Search Central: consolidate duplicate URLs — hur en sökmotor väger omdirigeringar, kanoniska URL:er och webbplatskartor mot varandra.

Ordlista

Välj utifrån era faktiska förutsättningar

Beskriv arbetsflödet, befintliga gränssnitt och användarna som behöver förändringen. Vi kan definiera en testbar leverans och en överlämning som gör att teamet kan underhålla lösningen.

Se vad det kostar Hör av er →

Gör begreppen praktiska

Fråga hur ”Vad är en kanonisk URL?” gäller i er organisation. Ett konkret exempel från ert nuvarande system räcker för att börja samtalet.

Se våra tjänster

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

Ta med dig den här sidan

Visitkort

Omsorgsfullt skapad med kärlek av D.

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

Upphovsrätt © 2025 – nu · Engineer ApS · Alla rättigheter förbehållna.

Utskrivet från https://engineer.company/sv/glossary/canonical/