

PAPEROFFICE
ENTERPRISE./AI

DATASHEET · VASTLEGGEN · WEBSITEBRONNEN

Alles staat op uw website. Wie typt het opnieuw over?
Uw website wordt de kennis van uw agent.

Websitebronnen lezen de openbare HTTPS-pagina's van uw website volgens een schema. Elke pagina wordt een intern artikel in de kennisbank van uw keuze, en uw agents antwoorden daaruit aan de telefoon en in de websitechat.

IN ÉÉN OOGOPSLAG

Tot 2000 pagina's per bron, linkdiepte tot 6, een schema van elke 6 uur tot maandelijks. Gewijzigde pagina's worden opnieuw geschreven; verdwenen pagina's worden na twee runs buiten gebruik gesteld.

01 · WAAR HET OVER GAAT

Twee versies van hetzelfde antwoord. Nu bepaalt uw website de feiten.

Openingstijden, productpagina's en helpteksten staan al jaren op uw website. Ze voor de agent opnieuw overtypen kost tijd, en elke wijziging op de site maakt de kopie verouderd. Websitebronnen lezen de pagina's zelf en nemen wijzigingen mee bij de volgende run.

INVOER

Startadres met https, optioneel een sitemap en meer startadressen op dezelfde host

RESULTAAT

Eén intern artikel per pagina, met het paginaadres als bron en de herkende taal

SCHEMA

Elke 6 of 12 uur, dagelijks, wekelijks, elke 14 dagen of maandelijks

LEZERS

Gespreksagents aan de telefoon en in de websitechat, chatagents in Agent AI



Eén host, veel pagina's

Vanaf het startadres volgt de crawler links op dezelfde host, met of zonder www. Sitemaps uit robots.txt en een eigen sitemap breiden de lijst uit.

De reikwijdte bepalen

Maximaal aantal pagina's en linkdiepte per bron, plus pad-voorvoegsels, uitsluitpatronen, URL's met een querystring en lijstpaden zoals /tag.

JavaScript inbegrepen

Snel leest de HTML, Volledig rendert in een browser, Automatisch schakelt zelf over als een pagina leeg lijkt. Met een klikselector opent u bijvoorbeeld FAQ-accordeons.

Schone artikelen

Uitsluiting via CSS houdt gebieden zoals navigatie of de voettekst buiten het artikel. Desgewenst neemt het artikel een korte samenvatting uit de pagina over.

De regels van de website

robots.txt en noindex worden gerespecteerd. Wordt een geïndexeerde pagina later geblokkeerd, dan zet de crawler haar artikel buiten gebruik.

Wijzigingen, geen volledige herschrijvingen

De crawler herkent gewijzigde pagina's aan hun inhoud en laat ongewijzigde met rust. Verdwenen pagina's worden na twee runs zonder treffer buiten gebruik gesteld.

02 · ZO WERKT HET

Bron toevoegen, controleren, toewijzen aan de agent.

Een websitebron vraagt maar enkele velden. De preview toont wat de crawler vindt vóór de eerste run, en daarna toont het overzicht wat er met elke URL is gebeurd.

1

Bron toevoegen

Gebruik onder `Knowledge Base AI › Bronnen van de website` `Bron toevoegen` ; vul naam, startadres, doelkennisbank en vernieuwingsinterval in.

2

De preview controleren

De `Voorbeeld` toont robots.txt, sitemaps en kandidaten, zonder iets te indexeren. Stel daarna de reikwijdte en de crawler in.

3

Draaien en volgen

De eerste run start enkele minuten na het opslaan, en `Nu crawlen` plant nog een run. `Geïndexeerde pagina's` toont status en HTTP-code per URL.

4

Toewijzen aan de agent

Kies onder `Conversation AI › Gespreksagenten` de kennisbank op het tabblad `Kennis` . Vanaf dan antwoordt de agent uit de websiteartikelen.

TECHNISCHE GEGEVENS

KENMERK	DETAILS
Invoer	Openbare HTTPS-pagina's van één host, HTML-tekst
Reikwijdte per bron	Tot 2000 pagina's (standaard 200), linkdiepte tot 6 (standaard 3)
Bronnen per account	Tot 25
Crawler	Snel, Volledig of Automatisch, met klik- en wachtselectors
Uitvoer	Intern artikel per pagina met paginaadres, herkende taal en de tags website en host
Overzicht	`Geïndexeerde pagina's` met status, rendering, HTTP en laatste ophaling; `Synchronisatiegeschiedenis` per bron
Toegang en facturering	Webapp onder `Knowledge Base AI › Bronnen van de website` ; op basis van gebruik gefactureerd

03 · IN DE PRAKTIJK

Waar websitekennis bij telefoon en chat komt.

Handleidingen voor de websitechat

Een softwarebedrijf houdt zijn handleidingen onder /docs. Een websitebron met het pad-voorvoegsel /docs leest ze dagelijks, en de websitechat noemt in het antwoord de passende pagina.

Productpagina's voor de hotline

Een detailhandelaar laat zijn productpagina's wekelijks lezen. De gespreksagent aan de telefoon beantwoordt vragen over maten en materialen uit dezelfde teksten die klanten online zien.

Website plus uw eigen artikelen

Wat niet op de website staat, zoals een sluiting voor de feestdagen, voegt u toe als eigen artikel in dezelfde kennisbank. De agent leest beide samen.

VEELGESTELDE VRAGEN

Welke pagina's leest een websitebron?

Openbare HTTPS-pagina's van één host, vanaf het startadres en, desgewenst, via sitemaps en extra startadressen. Het leest de tekst van de HTML-pagina's, zonder login, cookies of inloggegevens.

Wat gebeurt er met mijn eigen wijzigingen in een websiteartikel?

Zodra de pagina verandert, schrijft de volgende run het artikel opnieuw. Aanvullingen horen daarom thuis in een eigen artikel in dezelfde kennisbank.

Hoe actueel is de kennis van de agent?

Zo actueel als de laatste run. U kiest het schema van elke 6 uur tot maandelijks, en Nu crawlen plant een extra run. Twee handmatige runs liggen 15 minuten uit elkaar.

Verschijnen de pagina's in het klantenportaal?

Websiteartikelen zijn intern: uw team en de agents van de kennisbank lezen ze, en ze blijven buiten het klantenportaal.

AAN DE SLAG

Maak uw website uw eerste bron.

Ga in de webapp naar Knowledge Base AI › Bronnen van de website, gebruik Bron toevoegen om het startadres in te voeren, controleer de preview en sla op. De eerste run start enkele minuten later.

paperoffice.ai/nl/ai-web-sources-webcrawler-kennisbronagent-antwoorden

MEER OVER DE MODULE

paperoffice.ai/nl/ai-web-sources-webcrawler-kennisbronagent-antwoorden

PRIJZEN

paperoffice.ai/en/pricing

CONTACT

paperoffice.ai/nl/contact