



데이터시트 · 캡처 · 웹사이트 출처

내용은 모두 웹사이트에 있습니다. 누가 또 옮겨 적고 있습니까? 웹사이트가 에이전트의 지식이 됩니다.

웹사이트 출처는 일정에 따라 웹사이트의 공개 HTTPS 페이지를 읽습니다. 각 페이지는 선택한 지식 베이스의 내부 기사가 되며, 에이전트는 이를 바탕으로 전화와 웹사이트 채팅에서 답합니다.

한눈에 보기

출처당 최대 2000개 페이지, 링크 깊이 최대 6, 6시간마다부터 월 단위까지의 일정. 변경된 페이지는 다시 작성되고, 사라진 페이지는 두 번의 실행 후 보관 처리됩니다.

01 · 개요

같은 답변이 두 가지 버전으로 존재합니다. 이제 웹사이트가 사실을 정합니다.

영업시간, 제품 페이지, 도움말 글은 오래전부터 웹사이트에 있습니다. 에이전트를 위해 다시 옮겨 적는 데는 시간이 걸리고, 사이트의 모든 변경은 사본을 낚게 만듭니다. 웹사이트 출처는 페이지를 직접 읽고, 다음 실행 때 변경 사항을 함께 가져옵니다.

입력

https 시작 URL, 선택 사항으로 사이트맵과 같은 호스트의 추가 시작 URL

결과

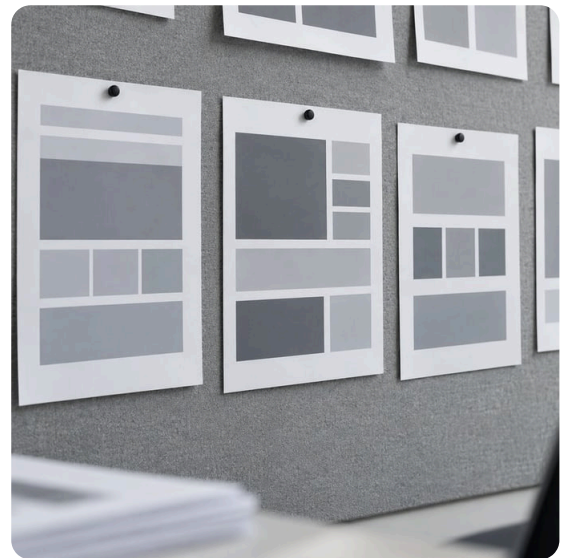
페이지마다 하나의 내부 기사, 페이지 URL을 출처로, 인식된 언어와 함께

일정

6시간 또는 12시간마다, 매일, 매주, 14일마다 또는 매월

독자

전화와 웹사이트 채팅의 대화 에이전트, Agent AI의 채팅 에이전트



한 호스트, 여러 페이지

시작 URL에서 크롤러는 같은 호스트의 링크를 따라갑니다(www 유무와 관계없이). robots.txt의 사이트맵과 직접 지정한 사이트맵이 목록을 확장합니다.

범위 설정

출처별 최대 페이지 수와 링크 깊이, 경로 접두사, 제외 패턴, 쿼리 문자열이 있는 URL, /tag 같은 목록 경로를 설정할 수 있습니다.

JavaScript 포함

Fast는 HTML을 읽고, Full은 브라우저에서 렌더링하며, Auto는 페이지가 비어 보이면 스스로 전환합니다. 클릭 선택자로 예를 들어 FAQ 아코디언을 열 수 있습니다.

깨끗한 기사

CSS 제외 설정으로 탐색 메뉴나 바닥글 같은 영역은 기사에서 빠집니다. 원하면 기사가 페이지에서 짧은 요약물 가져옵니다.

웹사이트의 규칙

robots.txt와 noindex를 존중합니다. 색인된 페이지가 나중에 차단되면 크롤러는 그 기사를 보관 처리합니다.

전체 재작성이 아닌 변경만

크롤러는 내용으로 변경된 페이지를 찾고, 변경되지 않은 페이지는 그대로 둡니다. 두 번의 실행 동안 조회되지 않은 페이지는 보관 처리됩니다.

02 · 작동 방식

출처를 추가하고, 확인하고, 에이전트에 배정하십시오.

웹사이트 출처에는 몇 개의 필드만 필요합니다. 미리 보기는 첫 실행 전에 크롤러가 찾는 내용을 보여 주며, 이후 카탈로그에는 각 URL에 무슨 일이 있었는지 나타냅니다.

1

출처 추가

Knowledge Base AI > 웹사이트 소스 아래에서 소스 추가 를 사용하십시오. 이름, 시작 URL, 대상 지식 베이스, 갱신 간격을 입력합니다.

2

미리 보기 확인

미리보기 에는 아무것도 색인하지 않고 robots.txt, 사이트맵, 후보가 표시됩니다. 그다음 범위와 크롤러를 설정하십시오.

3

실행과 모니터링

첫 실행은 저장한 뒤 몇 분 후에 시작되며, 크롤링 시작 는 추가 실행을 예약합니다. 인덱싱된 페이지 에는 URL별 상태와 HTTP 코드가 표시됩니다.

4

에이전트에 배정

Conversation AI > 대화 에이전트 아래에서 지식 탭의 지식 베이스를 선택하십시오. 그 순간부터 에이전트는 웹사이트 기사를 바탕으로 답합니다.

기술 데이터

기능	세부 정보
입력	한 호스트의 공개 HTTPS 페이지, HTML 텍스트
출처별 범위	최대 2000개 페이지(기본 200), 링크 깊이 최대 6(기본 3)
계정당 출처	최대 25개
크롤러	Fast, Full 또는 Auto, 클릭 및 대기 선택자 포함
출력	페이지 URL, 인식된 언어, 웹사이트와 호스트 태그가 있는 페이지별 내부 기사
개요	인덱싱된 페이지 : 상태, 렌더링, HTTP, 마지막 가져오기 표시. 출처별 동기화 기록
접근과 요금	Knowledge Base AI > 웹사이트 소스 아래의 웹 앱. 사용량에 따라 청구됩니다

03 · 실제 활용

웹사이트 지식이 전화와 채팅에 닿는 곳입니다.

웹사이트 채팅용 사용 방법

소프트웨어 회사는 사용 방법을 /docs 아래에 둡니다. 경로 접두사 /docs를 가진 웹사이트 출처는 이를 매일 읽고, 웹사이트 채팅은 답변에서 일치하는 페이지를 알려 줍니다.

하라인용 제품 페이지

소매업체는 제품 페이지를 매우 읽게 합니다. 전화의 대화 에이전트는 고객이 온라인에서 보는 것과 같은 글로 사이즈와 소재에 대한 질문에 답합니다.

웹사이트와 자체 기사

휴일 휴무처럼 웹사이트에 없는 내용은 같은 지식 베이스에 자체 기사로 추가합니다. 에이전트는 둘을 함께 읽습니다.

자주 묻는 질문

웹사이트 출처는 어떤 페이지를 읽습니까?

시작 URL에서 한 호스트의 공개 HTTPS 페이지를 읽으며, 원하면 사이트맵과 추가 시작 URL을 통해 확장합니다. HTML 페이지의 텍스트를 읽으며, 로그인, 쿠키, 자격 증명은 사용하지 않습니다.

웹사이트 기사를 직접 편집하면 어떻게 됩니까?

페이지가 변경되면 다음 실행이 기사를 다시 작성합니다. 따라서 추가 내용은 같은 지식 베이스의 별도 기사에 넣으십시오.

에이전트의 지식은 얼마나 최신입니까?

마지막 실행만큼 최신입니다. 6시간마다부터 월 단위까지 일정을 고를 수 있고, 크롤링 시작 버튼을 누르면 추가 실행이 예약됩니다. 수동 실행 두 번은 15분 간격입니다.

이 페이지들이 고객 포털에 나타납니까?

웹사이트 기사는 내부용입니다. 팀과 지식 베이스의 에이전트가 읽으며, 고객 포털에는 나타나지 않습니다.

시작하기

웹사이트를 첫 출처로 만드십시오.

웹 앱에서 Knowledge Base AI > 웹사이트 소스 로 가서 소스 추가 를 사용해 시작 URL을 입력하고, 미리 보기를 확인한 다음 저장하십시오. 첫 실행은 몇 분 뒤에 시작됩니다.

paperoffice.ai/ko/ai-web-sources-websaiteu-keuolreo-jisig-boseu-eijeonteu



모듈 자세히 보기

paperoffice.ai/ko/ai-web-sources-websaiteu-keuolreo-jisig-boseu-eijeonteu

요금

paperoffice.ai/en/pricing

문의

paperoffice.ai/ko/munui