



FICHE TECHNIQUE · CAPTURER · SOURCES WEB

Tout est sur votre site. Qui le retape ? Votre site devient le savoir de votre agent.

Les sources web lisent selon un calendrier les pages HTTPS publiques de votre site. Chaque page devient un article interne dans la base de connaissances de votre choix, et vos agents y puisent leurs réponses au téléphone et dans le chat du site.

EN BREF

Jusqu'à 2000 pages par source, profondeur de liens jusqu'à 6, une planification de toutes les 6 heures à mensuelle. Les pages modifiées sont réécrites, les pages disparues sont retirées après deux exécutions.

01 · DE QUOI IL S'AGIT

Deux versions de la même réponse. Désormais, votre site fait foi.

Horaires d'ouverture, pages produits et textes d'aide figurent sur votre site depuis longtemps. Les retaper pour l'agent prend du temps, et chaque modification du site rend la copie obsolète. Les sources web lisent les pages elles-mêmes et intègrent les modifications lors de la prochaine exécution.

ENTRÉE

URL de départ en https, plan du site facultatif et autres URL de départ du même hôte

RÉSULTAT

Un article interne par page, avec l'URL de la page comme source et la langue détectée

PLANIFICATION

Toutes les 6 ou 12 heures, quotidienne, hebdomadaire, toutes les 14 jours ou mensuelle

LECTEURS

Agents de conversation au téléphone et dans le chat du site, agents de chat dans Agent AI

**Un hôte, de nombreuses pages**

À partir de l'URL de départ, le robot suit les liens du même hôte, avec ou sans www. Les plans du site indiqués dans robots.txt et un plan du site propre complètent la liste.

Définir le périmètre

Nombre maximal de pages et profondeur de liens par source, ainsi que des préfixes de chemin, des motifs d'exclusion, des URL avec paramètre de requête et des chemins de liste comme /tag.

JavaScript inclus

Rapide lit le HTML, Complet affiche la page dans un navigateur, Auto bascule seul lorsqu'une page semble vide. Un sélecteur de clic ouvre par exemple les accordéons de FAQ.

Articles propres

L'exclusion CSS écarte de l'article les zones comme la navigation ou le pied de page. Si vous le souhaitez, l'article reprend un court résumé de la page.

Les règles du site

robots.txt et noindex sont respectés. Si une page indexée est ensuite bloquée, le robot retire son article.

Des modifications, pas des réécritures complètes

Le robot repère les pages modifiées à leur contenu et laisse les autres intactes. Les pages disparues sont retirées après deux exécutions sans correspondance.

02 · FONCTIONNEMENT

Ajoutez une source, vérifiez-la, attribuez-la à l'agent.

Une source web ne demande que quelques champs. L'aperçu montre ce que le robot trouve avant la première exécution, puis le catalogue indique ce qui s'est passé pour chaque URL.

1

Ajouter une source

Sous **Knowledge Base AI** > **Sources de sites web**, utilisez **Ajouter une source** : saisissez le nom, l'URL de départ, la base de connaissances cible et l'intervalle d'actualisation.

2

Contrôler l'aperçu

Aperçu affiche robots.txt, les plans du site et les candidats, sans rien indexer. Définissez ensuite le périmètre et le robot.

3

Exécuter et suivre

La première exécution démarre quelques minutes après l'enregistrement, et **Explorer maintenant** planifie une autre exécution. **Pages indexées** affiche le statut et le code HTTP pour chaque URL.

4

Attribuer à l'agent

Sous **Conversation AI** > **Agents de conversation**, choisissez la base de connaissances dans l'onglet **Connaissances**. Dès lors, l'agent répond à partir des articles du site web.

CARACTÉRISTIQUES TECHNIQUES

FONCTIONNALITÉ	DÉTAILS
Entrée	Pages HTTPS publiques d'un seul hôte, texte HTML
Périmètre par source	Jusqu'à 2000 pages (par défaut 200), profondeur de liens jusqu'à 6 (par défaut 3)
Sources par compte	Jusqu'à 25
Robot	Rapide, Complet ou Auto, avec sélecteurs de clic et d'attente
Sortie	Article interne par page, avec URL de la page, langue détectée et les étiquettes website et host
Vue d'ensemble	Pages indexées avec statut, rendu, HTTP et dernière lecture ; Historique de synchronisation par source
Accès et facturation	Application web sous Knowledge Base AI > Sources de sites web ; facturé selon l'usage

03 · EN PRATIQUE

Là où le savoir du site atteint le téléphone et le chat.

Tutoriels pour le chat du site

Une société de logiciels tient ses tutoriels sous /docs. Une source web avec le préfixe de chemin /docs les lit chaque jour, et le chat du site cite la page correspondante dans sa réponse.

Pages produits pour la ligne d'assistance

Un détaillant fait lire ses pages produits chaque semaine. L'agent de conversation au téléphone répond aux questions sur les tailles et les matières à partir des mêmes textes que les clients voient en ligne.

Site web et vos propres articles

Tout ce qui ne figure pas sur le site, comme une fermeture pour les vacances, vous l'ajoutez comme article propre dans la même base de connaissances. L'agent lit les deux ensemble.

QUESTIONS FRÉQUENTES

Quelles pages une source web lit-elle ?

Les pages HTTPS publiques d'un seul hôte, à partir de l'URL de départ et, si vous le souhaitez, via les plans du site et d'autres URL de départ. Il lit le texte des pages HTML, sans connexion, cookies ni identifiants.

Que deviennent mes propres modifications d'un article web ?

Lorsque la page change, la prochaine exécution réécrit l'article. Les ajouts doivent donc figurer dans un article distinct, dans la même base de connaissances.

Les connaissances de l'agent sont-elles à jour ?

Elles sont aussi actuelles que la dernière exécution. Vous choisissez la planification, de toutes les 6 heures à mensuelle, et **Explorer maintenant** ajoute une exécution supplémentaire. Deux exécutions manuelles sont espacées de 15 minutes.

Les pages apparaissent-elles dans le portail client ?

Les articles web sont internes : votre équipe et les agents de la base de connaissances les lisent, et ils restent hors du portail client.

COMMENCER

Faites de votre site votre première source.

Dans l'application web, allez sous **Knowledge Base AI > Sources de sites web**, utilisez **Ajouter une source** pour saisir l'URL de départ, contrôlez l'aperçu et enregistrez. La première exécution démarre quelques minutes plus tard.

paperoffice.ai/fr/ai-web-sources-agent-source-connaissances-analyseur-site



EN SAVOIR PLUS SUR LE MODULE

paperoffice.ai/fr/ai-web-sources-agent-source-connaissances-analyseur-site

TARIFS

paperoffice.ai/fr/tarifs

CONTACT

paperoffice.ai/fr/contact