robots.txt: comprendre et bien configurer son fichier

SEO · https://academy.drgeek.com/docs/avance/seo/robots-txt-comprendre-et-bien-configurer-son-fichier/

Le fichier **robots.txt** permet de contrôler ce que les moteurs de recherche peuvent explorer sur votre site. Bien configuré, il évite d'indexer des pages inutiles et améliore la qualité du référencement.

En bref

- Le robots.txt indique aux moteurs ce qu'ils peuvent explorer
- Il ne bloque pas totalement l'accès, mais guide les robots
- Une mauvaise configuration peut nuire au référencement

Qu'est-ce que le fichier robots.txt?

Le fichier robots.txt est un fichier texte placé à la racine de votre site (exemple: **https://votresite.ch/robots.txt**).

Il sert à donner des instructions aux robots des moteurs de recherche comme Google pour leur indiquer quelles pages ou sections doivent être explorées... ou ignorées.

À quoi sert le robots.txt?

- éviter l'indexation de pages inutiles (admin, tests...)
- optimiser le crawl des moteurs de recherche
- protéger certaines parties du site

Important

Le robots.txt ne protège pas réellement vos données. Il indique seulement aux robots de ne pas explorer certaines pages, mais celles-ci peuvent rester accessibles.

Exemple de robots.txt pour WordPress

```
User-agent: *
Disallow: /wp-admin/
Allow: /wp-admin/admin-ajax.php

Sitemap: https://votresite.ch/sitemap.xml
```

Explication

- **User-agent: *** → règle pour tous les robots
- **Disallow** → interdit l'exploration
- **Allow** → autorise un fichier spécifique
- **Sitemap** → indique où trouver votre plan du site

Où trouver ou modifier le robots.txt?

- via un plugin SEO (ex: AIOSEO)
- via votre hébergement (FTP / gestionnaire de fichiers)

Erreurs fréquentes à éviter

- bloquer tout le site par erreur (Disallow: /)
- empêcher Google d'accéder aux ressources importantes
- oublier d'ajouter le sitemap

Astuce utile

Utilisez Google Search Console pour tester votre fichier robots.txt et vérifier qu'il fonctionne correctement.

Pour conclure

Le fichier robots.txt est un outil simple mais puissant pour guider les moteurs de recherche. Une configuration correcte permet d'améliorer la visibilité et d'éviter des erreurs d'indexation.