

Document 1 : le fonctionnement des moteurs de recherche.

Pour retrouver une information dans les milliards de pages du Web, un moteur de recherche réalise trois grandes étapes : **explorer, indexer, classer**.

1- L'exploration

Les moteurs utilisent des **robots d'indexation**, aussi appelés **crawlers**.
Ce sont des programmes automatiques qui parcourent le Web en continu.

Ces robots :

1. **visitent automatiquement** des milliards de pages web ;
2. **récupèrent leur contenu** (texte, images, liens...) ;
3. **suivent les liens** présents dans la page pour découvrir d'autres pages.

Grâce à eux, le moteur construit une copie géante du Web, mise à jour en permanence.

2. L'indexation : ranger les informations

Une fois les pages explorées, le moteur fabrique un **index**.
Un index, c'est comme un **énorme dictionnaire inversé**.

Pour chaque mot, on stocke **la liste des pages où il apparaît** dans une base de données.

Exemples :

- **volcan** → page A, page C
- **danger** → page C

Ainsi, quand un utilisateur tape des mots, le moteur cherche dans sa base de données et sait **immédiatement** quelles pages correspondent.

3. Le classement : choisir les "meilleures" pages

Quand tu fais une recherche, le moteur :

1. repère les pages qui contiennent les mots demandés (grâce à l'index),
2. **les classe** pour afficher d'abord celles qui semblent les plus utiles.

Le classement dépend de plusieurs critères :

- **Pertinence** : est-ce que la page traite vraiment du sujet ?
- **Popularité** : combien d'autres sites font un lien vers elle ?
- **Ordre des mots** : une recherche "volcan actif" n'est pas traitée comme "actif volcan".
- **Fraîcheur** : la date de mise à jour peut compter pour certains sujets.

Le moteur affiche alors les résultats **en fonction de ces critères cumulés**.