

# Moteurs de recherche

## SNT thème Web

Frédéric Junier<sup>1</sup>

Lycée du Parc, Lyon

---

1. <https://frederic-junier.org/>

# Plan



Analyse critique des résultats fournis par un moteur de recherches

## Quel est le rôle d'un navigateur ? Citer trois navigateurs connus.

- Un navigateur est un logiciel qui joue le rôle de client Web. Il permet de demander à un serveur une page Web identifiée par son adresse URL. Les échanges entre le serveur et le client suivent le protocole HTTP. La réponse du serveur est une page écrite en langage HTML qui est interprétée puis affichée par le navigateur.
- *Safari, Firefox, Edge, Chrome, Opera* sont les navigateurs les plus connus.

## Quels usages différents permettent de distinguer moteur de recherche et navigateur Web ?

- Un navigateur ou *browser* est le logiciel permettant de naviguer sur le Web, réseau de documents reliés par des liens hypertextes.  
Il permet d'accéder directement à une page Web en saisissant son URL dans la barre d'adresse.

  <https://interstices.info/un-moteur-de-recherche-pour-le-meilleur-et-pour-le-pire/>

- Il n'existe pas de carte du Web. Un moteur de recherche est un logiciel s'exécutant sur un serveur Web permettant une recherche par mots clefs dans l'ensemble des pages Web qu'il a indexées. On interroge un moteur de recherche depuis son navigateur en remplissant un champ de formulaire contenu dans une page Web. *Google search*, *Bing*, *Qwant* sont les moteurs de recherche les plus connus.

## Quel est l'origine du nom Google ?

Le nom Google provient du mot *googol* introduit par Edward Kasner pour désigner le nombre  $10^{100}$  (un 1 suivi de 100 zéros). Il fait référence au nombre gigantesque de pages indexées par les robots d'indexation de Google.

Quelles sont les tâches principales réalisées par un moteur de recherche (1/3) ?

Animation du manuel Delagrave / Cliquez

## Quelles sont les tâches principales réalisées par un moteur de recherche (2/3) ?

Voici la description simplifiée donnée par Google pour son moteur de recherche sur

<https://support.google.com/webmasters/answer/9128586>.

Cliquez

## Quelles sont les tâches principales réalisées par un moteur de recherche (3/3) ?

Voici une description simplifiée. Pour une description plus complète voir cette **infographie**.

Phase de précalcul :

- **Exploration** : Des robots parcourent le Web en suivant les liens et repèrent les mots clefs.
- **Indexation** : Les pages analysées sont rangées en fonctions des mots clefs contenus.
- **Classement de notoriété** : Algorithme PageRank.

Phase de recherche :

- **Recherche par mots clefs** : Le moteur recherche les pages indexées par les mots clefs de la requêtes.
- **Tri et classement** : Plusieurs algorithmes interviennent dont fonctionnement reste secret (PageRank, filtrage par langue, localisation géographique, historique de recherche, terminal



## Quel est le rôle de l'algorithme PageRank de Google search ?

L'algorithme PageRank de Google search classe les pages par notoriété/popularité. C'est le principe d'un vote démocratique qui se mord la queue : chaque page transmet une fraction de son PageRank vers les pages liées par ses liens sortants et reçoit des fractions de PageRank par ses liens entrants. Le PageRank est déterminé de façon itérative : un surfeur aléatoire parcourt les pages de lien en lien un certain nombre de fois. Le nombre de passages par page au bout d'un grand nombre d'itérations permet d'estimer le PageRank. Si une page n'a pas de lien sortant et 15 fois sur 100 environ, le surfeur se téléporte aléatoirement vers une autre page du Web ce qui l'empêche de rester bloqué ou de tourner en rond.

## Quelle requête faut-il écrire dans un moteur de recherche pour obtenir des pages traitant du tennis mais pas du tennis de table ?

Il faut écrire `tennis -"tennis de table"`.



Il ne faut pas mettre d'espace après l'opérateur `—`.

Pour affiner ses recherches sur Google, voir la page :

<https://support.google.com/websearch/answer/2466433?hl=fr>

## Débusquer les fake news ?

Parmi les cinq articles figurant sur cette page <https://epreuves.pix.fr/sourceinfo/fredericf2.html>, pour détecter si l'information est plutôt fiable ou non en fonction de sa source, on peut utiliser le moteur de recherche Decodex du journal Le Monde.



**DÉCODEX**

Entrez l'adresse (URL) d'une page Web ci-dessous ou le nom d'un site **pour savoir si la source de l'information (c'est-à-dire celui ou celle qui la diffuse) est plutôt fiable ou non**

**En Bref**

Un site qui permet à n'importe quel internaute de créer un article fictif et de le partager sur les réseaux sociaux.

**Notre avis :**

**!** Attention, il s'agit d'un site satirique ou parodique qui n'a pas vocation à diffuser de vraies informations. A lire au second degré.

## Quelle est la signification de l'acronyme SEO ?

L'acronyme SEO signifie Search Engine Optimization ou optimisation du référencement naturel en français. On désigne ainsi l'ensemble des techniques permettant d'améliorer le référencement naturel d'une page par les moteurs de recherche.

On trouvera des conseils de Google pour le SEO sur la page <https://support.google.com/webmasters/answer/7451184?hl=fr>