

# 1 Problématique

Pour naviguer sur internet, il existe principalement deux solutions :

- en tapant l'*URL* (*Uniform Resource Locator*) exacte du site web dans la barre d'adresse du navigateur ; exemple : <https://lyceejaydebeaufort.fr/>,
- en utilisant un *moteur de recherche* qui trouve les sites en rapport avec un thème donné.

Comment fonctionne un moteur de recherche ?

## 2 Différents moteurs

### À retenir

Un moteur de recherche est un outil qui permet de chercher des contenus (documents, vidéos, images...) sur le web, à partir de mots clés.

#### Activité 1 :

1. Lire l'article suivant <https://tinyurl.com/yya3bj9x> et répondre aux questions :
  - (a) Quel est le moteur le plus utilisé dans le monde ?
  - (b) En France, quel pourcentage de parts de marché représente ce moteur (tout appareil confondu) ?
  - (c) Citer deux autres moteurs de recherche.
2. Se rendre sur les pages des deux moteurs suivants :
  - <https://duckduckgo.com>
  - <https://qwant.com>
3. Quel principe semble défendre ces deux moteurs de recherche ?

*Qwant* est un moteur de recherche français mis en ligne en 2013. L'extension du moteur de recherche pour le navigateur Mozilla Firefox fait partie de la liste des logiciels libres préconisés par l'État français. Les navigateurs (Firefox, Chrome...) utilisent un moteur de recherche par défaut. Il est possible de modifier ce choix.

#### Activité 2 : Depuis les préférences du navigateur :

1. trouver le moteur de recherche utilisé par défaut,
2. changer de moteur de recherche.

## 3 Indexation

### 3.1 Définition

#### À retenir

Les moteurs analysent, identifient et organisent les documents sur le web pour faciliter une recherche ultérieure.

**Activité 3 :**

1. Ouvrir un onglet pour les trois moteurs de recherche suivant :
  - <https://duckduckgo.com>
  - <https://qwant.com>
  - <https://google.com>
2. Effectuer la même recherche dans les trois moteurs, par exemple « chaussure de sport ».
3. Les résultats obtenus sont-ils identiques ?
4. Pour quelle raison les premiers résultats contiennent le mot *Annonce* ?

### 3.2 Algorithme d'indexation

Le *PageRank* est un algorithme d'analyse des pages web, utilisé par le moteur *Google*. Le principe de base est d'attribuer à chaque page web un score proportionnel au nombre de fois qu'un internaute visite la page quand il clique aléatoirement sur les hyperliens.

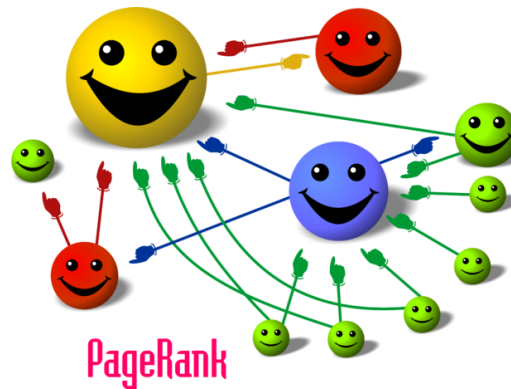


FIGURE 1 – Illustration du pagerank

Plus une page a des hyperliens qui pointent vers elle, plus son score sera élevé et plus elle sera présente dans les premiers résultats de recherche.