

# Atelier 5 : « Va donc chercher sur Google »

**S'informer via les moteurs de recherche**

**Brigitte Simonnot**

**Université de Lorraine, [Crem](#)**

brigitte.simonnot@univ-lorraine.fr

Séminaire ESPÉ DANE : Pépinière 4.0  
13 janvier 2016, Vandœuvre-lès-Nancy

# Atelier de réflexion

Les moteurs de recherche web sont insérés dans nos pratiques quotidiennes

Objectif :

réfléchir à la manière dont les moteurs de recherche du web peuvent être « invités » dans les pratiques d'apprentissage en contexte scolaire

# Atelier de réflexion

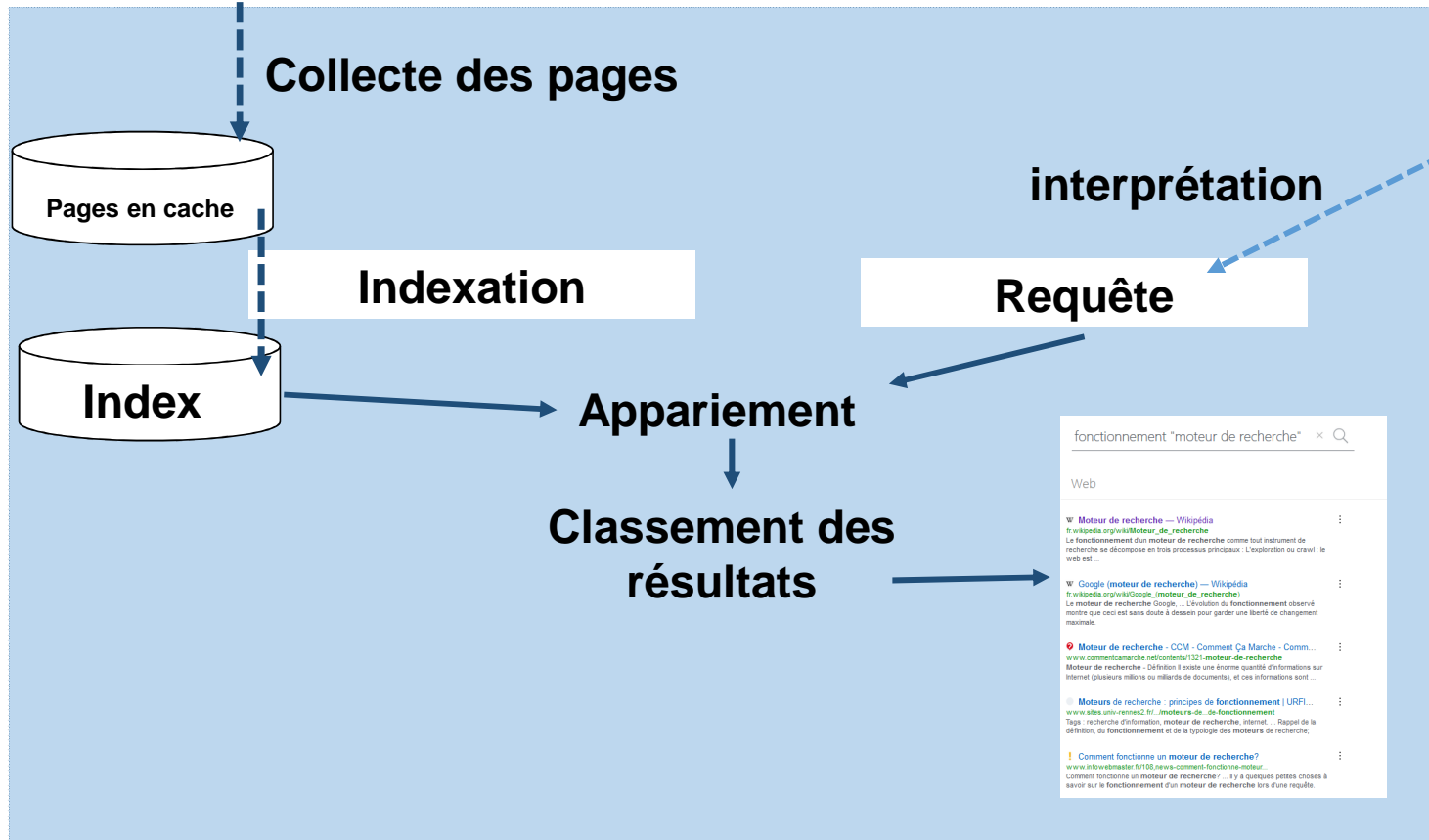
- Point sur les connaissances
  - Qu'est-ce qu'un moteur de recherche web ?
  - Quels moteurs ?
- Questions/réponses
- Intégrer une activité de recherche d'information en ligne en formation
  - Questions à se poser, objectifs, compétences développées

Qu'est-ce qu'un moteur web ?



Web

## Moteur - Opérations automatisées



Mise en mots



Analyse des résultats et accès aux pages

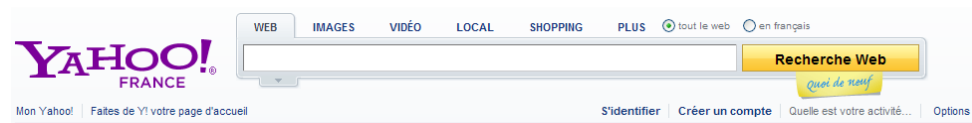
# Divers types de dispositifs d'accès à l'information en ligne

- Moteurs généralistes ou verticaux (= spécialisés dans un domaine ou un format de document)
- Annuaire (tombent en désuétude sauf dans des domaines spécialisés)
- Métamoteurs (pas d'index propre, interrogent un ou des autres moteurs)
- Systèmes de questions/réponses (ex : [Ask.com](http://Ask.com))
- Les médias sociaux ([Twitter](https://twitter.com), [Facebook](https://facebook.com), etc.)

# Un moteur de recherche, c'est

au moins 4 composants :

- un module de collecte des documents en ligne
- un module d'indexation
- un module d'appariement et de classement des résultats pour une requête
- une interface de recherche (interrogation, consultation des résultats)



# La collecte des pages

- Un ensemble de pages initiales
- Suivi des liens de ces pages et collecte des pages cibles par un « robot »
- Suivi des nouveaux liens trouvés dans les pages, etc.

## Remarques :

- La collecte n'est jamais exhaustive (**web invisible** = partie du web non accessibles aux moteurs)
- Chaque moteur a **sa** « recette » pour la collecte
- Passage périodique sur les pages déjà visitées (fréquence des mises à jour)
- Conservation des pages collectées (en « cache »)



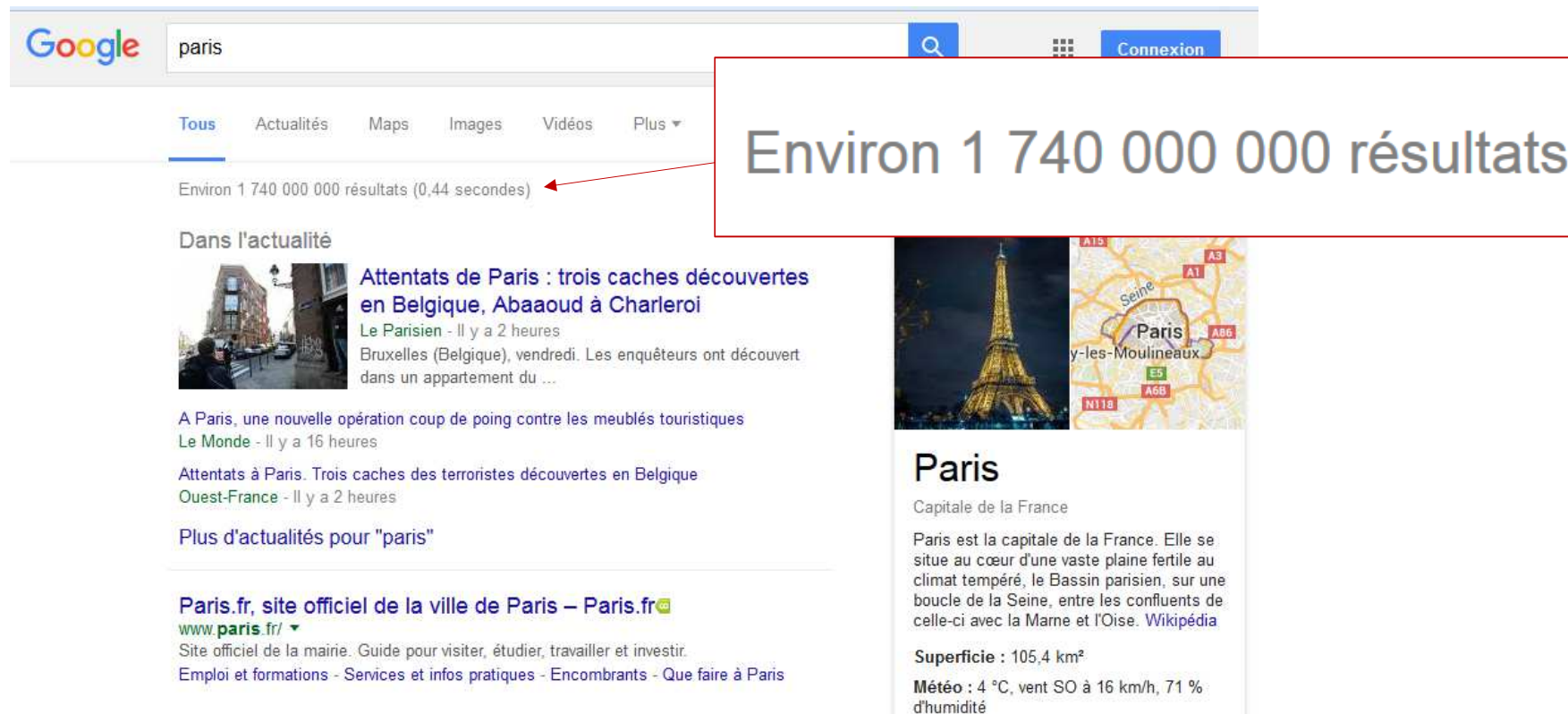
# L'indexation automatique

- « Extraction » des mots de la page (mot = suite de caractères entre deux séparateurs comme l'espace ou les signes de ponctuation)
- Problème : ambiguïtés du langage (homonymie, synonymie, etc.)

**Ex : paris**



# L'indispensable classement des résultats



Google paris

Tous Actualités Maps Images Vidéos Plus

Environ 1 740 000 000 résultats (0,44 secondes)

**Environ 1 740 000 000 résultats**

**Dans l'actualité**

**Attentats de Paris : trois caches découvertes en Belgique, Abaaoud à Charleroi**  
Le Parisien - Il y a 2 heures  
Bruxelles (Belgique), vendredi. Les enquêteurs ont découvert dans un appartement du ...

A Paris, une nouvelle opération coup de poing contre les meublés touristiques  
Le Monde - Il y a 16 heures

Attentats à Paris. Trois caches des terroristes découvertes en Belgique  
Ouest-France - Il y a 2 heures

Plus d'actualités pour "paris"

**Paris.fr, site officiel de la ville de Paris – Paris.fr**  
www.paris.fr/ ▾  
Site officiel de la mairie. Guide pour visiter, étudier, travailler et investir.  
Emploi et formations - Services et infos pratiques - Encombrants - Que faire à Paris

**Paris**  
Capitale de la France

Paris est la capitale de la France. Elle se situe au cœur d'une vaste plaine fertile au climat tempéré, le Bassin parisien, sur une boucle de la Seine, entre les confluent de celle-ci avec la Marne et l'Oise. Wikipédia

**Superficie** : 105,4 km<sup>2</sup>

**Météo** : 4 °C, vent SO à 16 km/h, 71 % d'humidité

# Classement des résultats

Sont privilégiées les pages

- dont le titre contient les mots de la requête
  - dont l'adresse (URL) contient le mot
  - où le mot revient souvent, surtout au début du texte
  - ... (+ de 500 critères)...
- + **pages souvent citées par d'autres sites** (pour Google)

**Critères de classement = « secret » des moteurs**  
(enjeux économiques importants)

# Quels moteurs ?

---

# Liste de moteurs de recherche

---

Cet article contient une **liste de moteurs de recherche** web ou d'entreprise.

## Sommaire [\[masquer\]](#)

- 1 Moteurs de recherche Web
  - 1.1 Solidaires et/ou écologiques
- 2 Moteurs de recherche d'entreprise
  - 2.1 Moteurs de recherche libre
- 3 Classement par spécialisation
  - 3.1 Moteurs de recherche Dark Web (Tor)
  - 3.2 Moteurs de recherche éducation et recherche
  - 3.3 Moteurs de recherche pour enfants
  - 3.4 Moteurs de recherche d'images
  - 3.5 Moteurs de recherche sur les sites de partage
  - 3.6 Moteurs de recherche RH
  - 3.7 Moteurs de recherche sémantique
  - 3.8 Moteurs en langage naturel
  - 3.9 Moteurs de recherche assurant la confidentialité des recherches
  - 3.10 Métamoteurs
  - 3.11 Annuaire de moteurs
- 4 Notes et références

## Moteurs de recherche Web [\[ modifier \]](#) [modifier le code \]](#)

---

- [AltaVista](#) (*multilingue*)
- [Ask](#) (*multilingue*)
- [Baidu](#) (*chinois*)
- [Bing](#) (*multilingue*)
- [Blekko](#) (*anglais*)
- [Dazoo FR](#) (*français*)
- [DuckDuckGo](#) (*multilingue*)
- [Exalead](#) (*français*)
- [Gigablast](#) (*anglais*)
- [Google](#) (*multilingue*)
- [LeMoteur](#) (*français*), qui remplace *Voila*
- [Lycos](#) (*multilingue*)
- [Mozbot](#) (*multilingue*)
- [Panguso](#) (*chinois*)
- [Premsgo](#) (*français*)
- [Qwant](#) (*multilingue*)
- [Sohu](#) (*chinois*)
- [Soso](#) (*chinois*)
- [Swisscows](#) (*suisse*)
- [Naver](#) (*Coréen*)
- [V9](#) (*chinois, multilingue*)
- [Wopa!](#) (*multilingue*)
- [Yahoo!](#) (*multilingue*)
- [Yandex](#) (*russe, anglais*)

## Moteurs de recherche éducation et recherche [\[ modifier \]](#) [modifier le code \]](#)

---

- [BASE](#) (développé dans le cadre du projet [Open Archives Initiative](#))
- [CiteSeerX](#) : bibliothèque numérique pour les articles scientifiques et académiques
- [Google Scholar](#)
- [ScientificCommons](#)
- [PubGene](#) : moteur de recherche biomédical
- [PubMed](#) : recherche bibliographique biomédicale
- [MyScienceWork](#) : moteur de recherche multidisciplinaire d'articles scientifiques en Open Access
- [FreeFullPDF](#) : publications scientifiques en format pdf (articles, thèses, brevets)

Source : [https://fr.wikipedia.org/wiki/Liste\\_de\\_moteurs\\_de\\_recherche](https://fr.wikipedia.org/wiki/Liste_de_moteurs_de_recherche)

# Paradoxe

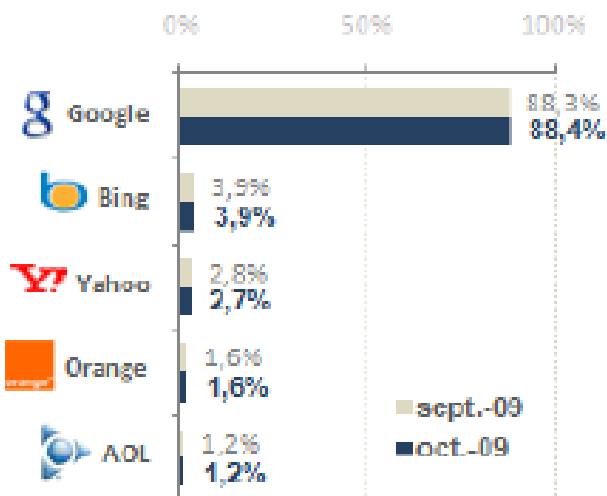
Le Web propose une multitude d'informations hétérogènes  
(monde numérique des publications en ligne très vaste)

*mais*

les moteurs que nous utilisons sont très peu nombreux  
(portes étroites)

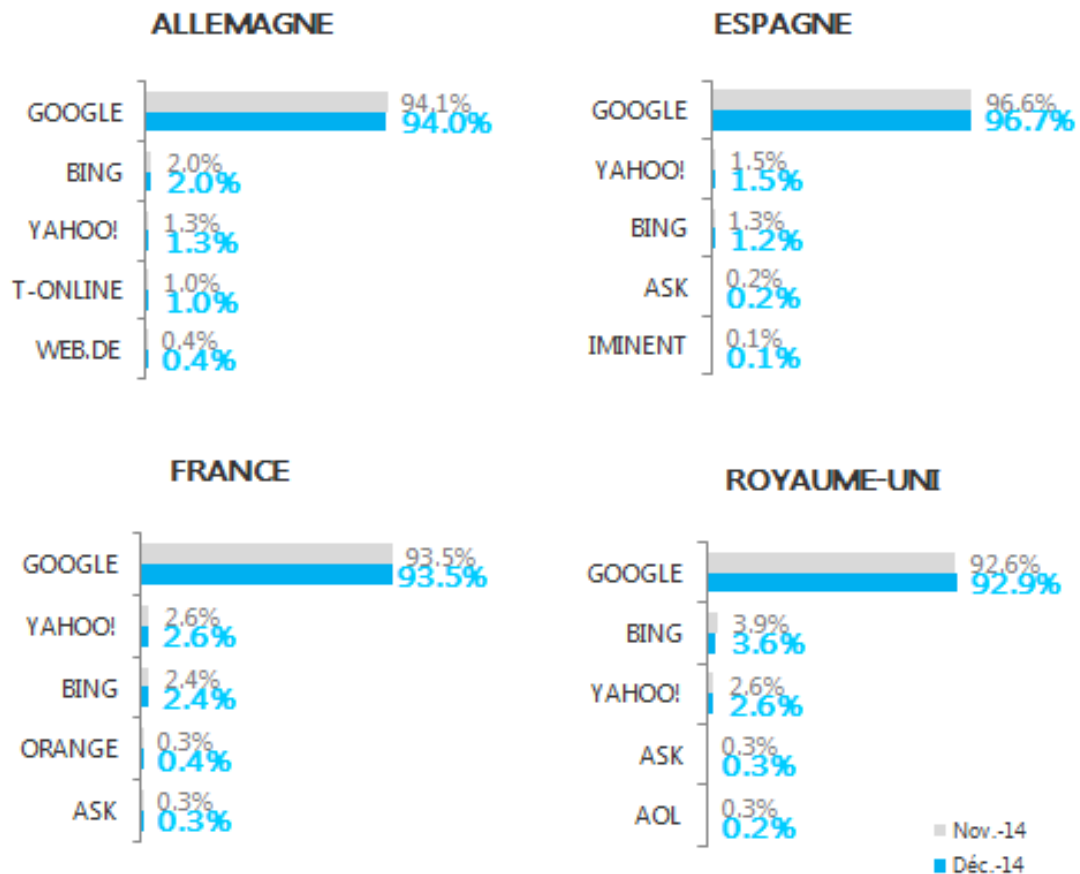
Pour mémoire, en 2009 :

(France)



### TOP 5 des MOTEURS DE RECHERCHE en part de visites web

Indicateur moyen par site



# Tous les moteurs ne sont pas équivalents

Certains ont un 5<sup>ème</sup> module : captation des « traces » des internautes

Les « traces » permettent de « personnaliser » les annonces publicitaires (l'historique du navigateur aussi, ou la géolocalisation)

Respect de la vie privée des internautes ?

Droit à l'oubli (Cf. arrêt de la Cour de justice de l'Union européenne, 2014 et la CNIL)



# Des alternatives aux Gafa\*

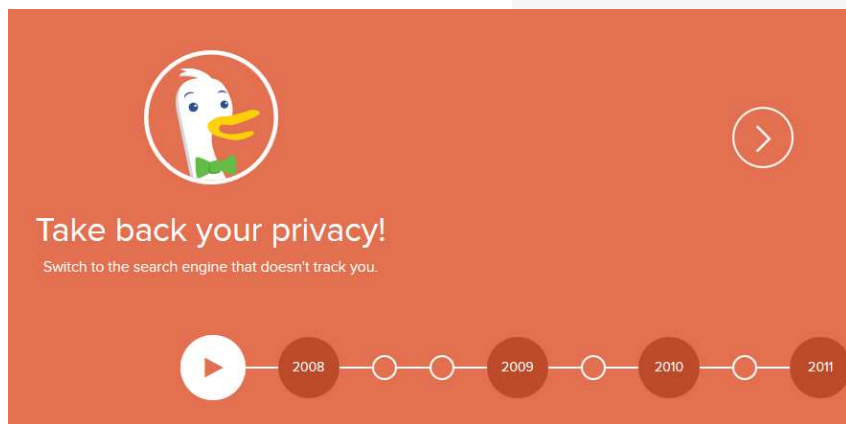
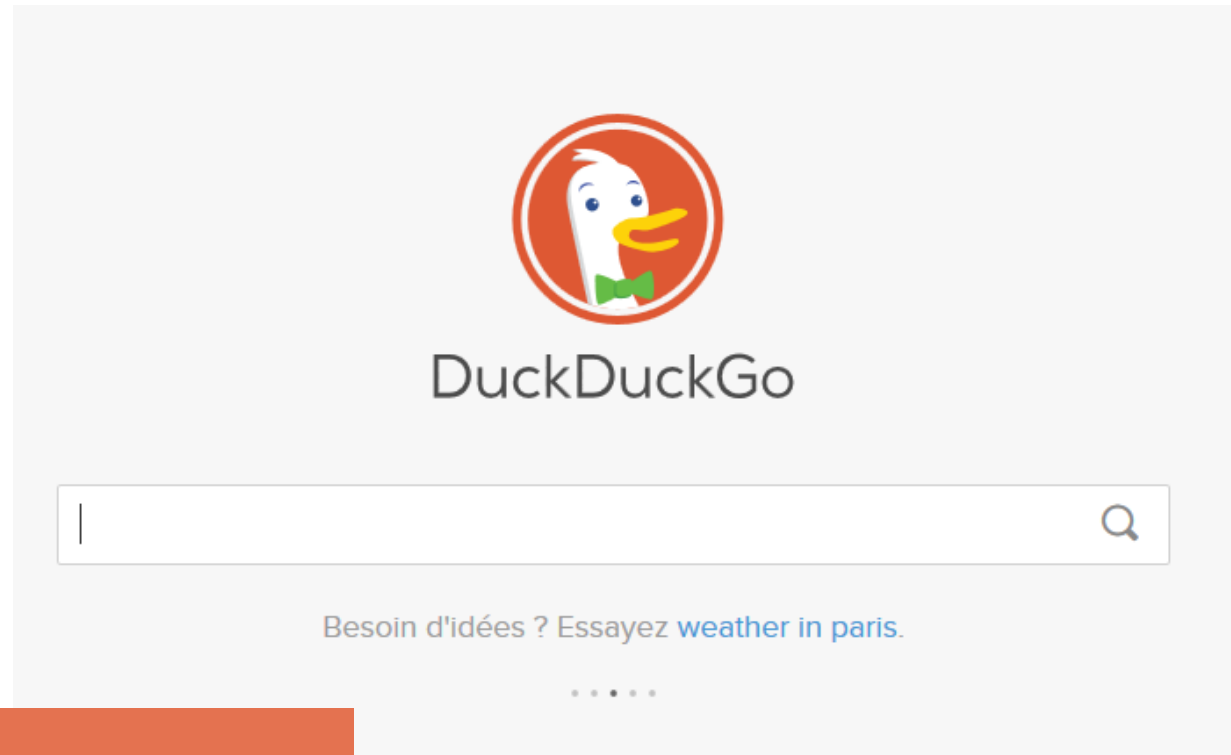
BOTTE DE FOIN

---

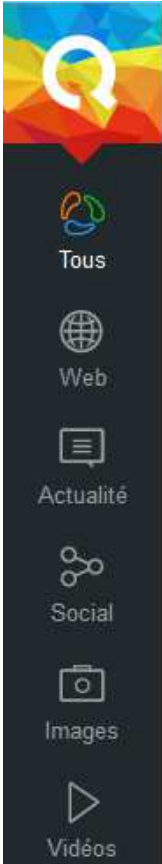
aiguille »

Crédit image : O. Ertzscheid ([affordance.typepad.com/](http://affordance.typepad.com/))

\*Gafa = Google, Amazon, Facebook, Apple (les multinationales du numérique et du web)



<https://duckduckgo.com/>



Sécurité, confidentialité et respect de la vie privée. [En savoir plus >>](#)



Que recherchez-vous ?



LES TENDANCES DU JOUR

FAVORIS

BLOG

<http://www.qwant.com>



Sécurité, confidentialité et respect de la vie privée. [En savoir plus >>](#)



Que recherches-tu ?



<http://www.qwantjunior.com>

Vos questions ?

# Questions de la salle (1/3)

- **Comment utiliser de manière plus efficace un moteur web ?**

Consulter l'aide en ligne des moteurs

- Aide Google : <https://support.google.com/websearch/answer/134479?hl=fr>
- Aide Bing : <http://help.bing.microsoft.com/#apex/18/en-us/10002/-1> (en anglais)  
ou <http://www.microsoft.com/france/bing/>
- Aide DuckDuckGo : <https://duck.co/help/results/syntax> (en anglais)

## Questions de la salle (2/3)

- **Comment utiliser de manière plus efficace un moteur web ?**

Utiliser des opérateurs dans les requêtes :

- Les guillemets pour une expression exacte ("moteur de recherche")
- Le signe moins devant un mot pour exclure (moteur -google)
- Rechercher uniquement sur un site (moteur site:education.gouv.fr)

# Questions de la salle

- **Et les méta-moteurs ?**

Un méta-moteur fonctionne comme un moteur de recherche mais il n'a pas son propre index : il interroge un ou plusieurs autres moteurs et fait une synthèse des résultats. Exemple : <https://ixquick.com/fra/>





# Intégrer une activité de recherche d'information en ligne en formation

Questions à se poser

Objectifs

Compétences développées

# Chercher soi-même l'information

Valeurs mises en avant par les usagers des moteurs

- Sensation d'autonomie (vs dépendance vis-à-vis de l'outil)
- Réussite perçue, performance (vs manque d'évaluation)
- Plaisir : variété des informations trouvées, stimulation liée à la découverte des résultats, interface qui plait
- Réputation du moteur (vs pouvoir)
- Sécurité, conformité : à la marge

[étudiants de 1<sup>ère</sup> année de licence, enquête par questionnaire Simonnot, 2009]

# Quel scénario d'usage en éducation

- Recherche sur un sujet prescrit vs sujet libre
- Temporalité dans l'activité principale : en amont d'une réalisation, pendant ou en aval ?
- Pendant les cours ou en dehors ?
- Quel accompagnement est prévu (y en a-t-il besoin) ?
- L'occasion d'aborder une présentation des moteurs (Gafa) ?
- Quelles consignes donner ?
- Comment aider à citer les sources (éviter le plagiat) ?

# Les enjeux

## Acquérir une culture de l'information

- Connaître les sources d'information
- Savoir
  - reconnaître un besoin d'information
  - formuler une requête et affiner
  - repérer et qualifier l'information
  - utiliser l'information et la partager

Dans les établissements secondaires, les professeurs documentalistes peuvent aider à organiser des séquences de formation