MOTEURS DE RECHERCHE

Table of contents

- MOTEURS DE RECHERCHE
- Introduction
- Moteurs de recherche
    - Définition
    - Moteurs de recherche génériques
    - Moteurs qui respectent votre vie privée
    - Moteurs et réseaux spécialisés
        - Académique
        - Wikipedia
        - Moteurs de recherche de baladodiffusions/podcasts
        - Moteurs de recherche dans les médias sociaux
        - Moteurs verticaux
        - Moteur de recherche d'objets connectés à Internet (Internet of things)
        - Moteur de recherche pour enfants
- Méta-moteurs
    - Définition
        - Exemples
- Autres outils de recherche
    - Annuaires/répertoires
        - Définition
    - Exemples
    - Répertoires sélectifs
    - Archives Internet
    - Les réseaux sociaux
        - Les réseaux généraux
        - Les réseaux professionnels
        - Les réseaux de questions/réponses
        - Les réseaux de social bookmarking
        - Les réseaux de microblogging

Introduction

Google est devenu le moteur de recherche par défaut de la majorité des internautes. Si Google demeure un incontournable, pour réaliser une veille efficace, il est nécessaire de combiner différents outils de recherche pour explorer de façon combinatoire, le **Web visible ou surfacique** (indexé par les moteurs de recherche), le **Web invisible ou profond** (non indexé par les moteurs de recherche), et le **Web social**.

Les vidéos suivantes en anglais vous expliquent comment fonctionne la recherche sur Google et ce qu'est le Web invisible ou "Deep Web".

Une requête Google c'est quoi? Le "Deep Web" c'est quoi?

Search increases choice at the expense of serendipity, the experience of discovering beneficial stories that we weren't intentionally seeking. When we can easily choose the news that interests us, we may miss stories that didn't appear interesting but that help us make unexpected and useful connections. [1] Une bonne veille est une *veille éveillée* telle que la définie Béatrice Foenix-Riou :

- *Parce qu'«éveillé» signifie « qui ne dort pas», et qu'il est impératif de ne pas s'endormir avec les premiers résultats de Google, en croyant qu'il a cherché systématiquement ce qu'on lui demandait et en se contentant des premières pages de résultats, obtenues après avoir saisi quelques mots-clés;*
- *Parce qu'«éveillé» signifie aussi « ouvert à», et qui est essentiel d'être attentif aux alternatives qui se présentent, un moteur étant incapable de répondre à toutes les questions;*
- *Parce qu'il est important enfin de rester « en état d'éveil » – en état de veille–, et que la vision des différents outils et modèles qui se sont succédé sur Internet permet d'appréhender les évolutions de la recherche d'information sur le Web et de les anticiper.»* [2]

Moteurs de recherche

Définition

Un moteur de recherche est une application web permettant de retrouver des ressources (pages web, articles de forums Usenet, images, vidéo, fichiers, etc.) associées à des mots quelconques. Certains sites web offrent un moteur de recherche comme principale fonctionnalité ; on appelle alors moteur de recherche le site lui-même (Google Video par exemple est un moteur de recherche vidéo). Instrument de recherche sur le web constitué de « robots », encore appelés bots, spiders, crawlers ou agents qui parcourent les sites à intervalles réguliers et de façon automatique (sans intervention humaine, ce qui les distingue des annuaires) pour découvrir de nouvelles adresses (URL). Ils suivent les liens hypertextes (qui relient les pages les unes aux autres) rencontrés sur chaque page atteinte. Chaque page identifiée est alors indexée dans une base de données, accessible ensuite par les internautes à partir de mots-clés. (Source : **Wikipedia** consultée le 16.02.2014) Vous trouverez ci-dessous une liste en développement d'outils alternatifs à ***Google**.

Moteurs de recherche génériques

- **Bing** (Microsoft)
- **Exalead** (Dassault System)
- **Yahoo!**

Astuce : fouillez les moteurs de recherche dans différentes langues (ex: Google.de, Google.es, Bing.de, etc.). Les résultats obtenus sont souvent totalement différents. La traduction de la page peut se faire facilement avec Google Translate.

Moteurs qui respectent votre vie privée

- **DuckDuckGo**
- **Startpage** (résultats Google + protection de la vie privée ixquick)
- **Qwant**
- **Gibiru**

- **Yippy**
- **Yacy** - Moteur décentralisé en P2P

Moteurs et réseaux spécialisés

Académique

- **Google Scholar** - Engin de recherche académique. Comme pour Google, essayez aussi les extensions d'autres pays (ex. : .fr, .es, .de, etc.)
- **Science Direct** (Elsevier) - Engin de recherche académique
- **Microsoft Academic Search** - Engin de recherche académique
- **Academia** - Une plateforme de partage où les chercheurs partagent leurs travaux
- **PeerLibrary** - Moteur de recherche de publication et d'auteurs
- **Base Bielefeld Academic Search Engine** - Engin de recherche académique
- **Academic Torrents** - Un nouveau joueur qui grossit vite et qui repose sur un de partage distribué pour les chercheurs et par les chercheurs
- **Wikidata**
- **Science Research**
- **Science.gov** (Méta moteur qui contient Science.gov, Scitopia.org, WorldWideScience.org et E-Print Network)
- **WorldWideScience**

Les trois derniers outils utilisent le logiciel de recherche Explorit Research Accelerator de **Deep Web Technologies**

Wikipedia

- **Wikiwix** moteur de recherche pour Wikipédia

Moteurs de recherche de baladodiffusions/podcasts

- **Stitcher**
- **Listen Notes**

Moteurs de recherche dans les médias sociaux

- **tagboard**
- **Talkwalker**
- **Twitter recherche avancée**

Moteurs verticaux

- **HuriSearch** - Droits de l'homme
- **Chemindustry** - Chimie
- **Health On the Net Foundation** - Santé (patients et professionnels)
- **EurekaSanté** - Santé (patients)
- **Healthfinder** - Santé (patients)
- **Information hospitalière** - Santé (professionnels)

- **Biomalin** - Biologique
- **FishBase** - Pêche
- **Closer Look Search** - Moteurs verticaux sur mesure

Moteur de recherche d'objets connectés à Internet (Internet of things)

- **Shodan** - Premier moteur pour trouver des appareils connectés à Internet

Moteur de recherche pour enfants

- **Qwant Junior** un moteur de recherche pour les enfants qui protège la vie privée (**source**).

Méta-moteurs

Définition

Un métamoteur (ou méta-moteur) ou un méta-chercheur est un moteur de recherche qui puise ses informations à travers plusieurs moteurs de recherche généralistes. De manière plus précise, le métamoteur envoie ses requêtes à plusieurs moteurs de recherche et retourne les résultats de chacun d'eux. Le métamoteur permet aux utilisateurs de n'entrer le sujet de leur recherche qu'une seule fois tout en accédant aux réponses de plusieurs moteurs de recherche différents. (Source : **Wikipedia** consultée le 16.02.2014)

Exemples

- **Metacrawler**
- **Mamma**
- **Manhack.net**

Méta-moteur qui garantit le respect de votre vie privée

- **Ixquick** (recherche approfondie universelle, recherche internationale, filtrage)

Autres outils de recherche

Annuaires/répertoires

Définition

Un annuaire web, répertoire web, annuaire Internet ou répertoire Internet est un site Web proposant une liste classée de sites Web. (Source : **Wikipedia** consultée le 16.02.2014)

Exemples

- **Open Directory Project** (dmoz)
- **Search Engine Colossus** - Répertoire d'outils de recherche spécifiques à un pays ou une région
- **The World Press** - Répertoire de plus de 5200 journaux à travers 192 pays
- **ABYZ News Links** - Répertoire de journaux et de guides des médias à travers le monde
- **Les annuaires** - Un annuaire d'annuaires et de moteurs de recherche européens

- **Podcast Directory** - Balladodiffusions/Podcasts
- **Directory of open access journals** - Répertoire de journaux académiques en accès libre
- **Big List of Websites** - Recherche de sites par titre, méta données et mots clés (+5M de sites)

Répertoires sélectifs

- **The Environment Directory** - Environnement
- **Agrimondial** - Agriculture
- **The Food World** - Agriculture
- **Thèses** - Thèses et écrits académiques français
- **Signets** - Liens sélectionnés par des bibliothèques universitaires françaises
- **CiSMeF** - Catalogue et Index des Sites Médicaux de langue Française

Archives Internet

- **Internet Archive**
- **The Way Back Machine** - Captures d'écran des sites Web dans le temps

Les réseaux sociaux

Les réseaux généraux

Noms	Origines	Influence
Facebook	États-Unis	Monde
QZone	Chine	Chine
V Kontakte	Russie	Russie
Odnoklassniki	Russie	Russie
Cloob	Iran	Iran
Draugiem	Lettonie	Lettonie

Les réseaux professionnels

Noms	Origines	Influence
LinkedIn	États-Unis	Monde
Viadeo	France	Europe et pays émergents
Xing	Allemagne	Europe
Ushi	Chine	Chine

Les réseaux de questions/réponses

Noms	Origines	influence	Langue
Quora	États-Unis	Amérique du Nord/Europe	Anglais
Baidu Knows	Chine	Chine	Chinois

Liste Wikipedia des sites de questions/réponses

Les réseaux de social bookmarking

Noms	Origines
Delicious	États-Unis
Diigo	États-Unis
Reddit	États-Unis
Digg	États-Unis
Stars	États-Unis

Les réseaux de microblogging

Noms	Origines	Influence
Twitter	États-Unis	Monde
Identi.ca	Canada	Monde des geeks
Weibo	Chine	Chine

1 Zuckerman Ethan, 2013, *Rewire - Digital Cosmopolitans in the Age of Connection* p. 105

2 Foenix-Riou Béatrice, 2011, ***Recherche éveillée sur Internet : mode d'emploi*** p.7-8

N'hésitez pas à me contacter pour me faire part de vos commentaires ou pour me suggérer d'autres sources : regis.barondeau -at- mac.com (remplacer -at- par @). Pour lire d'autres fiches allez sur **Outils**