

La rubrique du chercheur geek

Parlez-vous le Google ? (Partie 2)

Dans le dernier numéro, nous avons détaillé les fonctions élémentaires de Google. En pratique, poser la meilleure question revient à laisser à Google le moins de place possible pour l'interprétation de ce que l'on recherche. Pour cela, il faut être spécifique dans ce que l'on cherche (qui, quoi, combien), là où il faut rechercher (où, comment), et la période à rechercher (quand).

Où, comment, combien, quand (et bientôt pourquoi ?) : les opérateurs spécifiques

Il est possible d'affiner les recherches Google avec des opérateurs spécifiques. Le tableau ci-après regroupe les opérateurs les plus utiles :

Opérateur*	Description	Exemple
site:	La requête s'effectue sur les pages d'un site web spécifique	abduction site:lelibellio.com
filetype:	Filtre les résultats avec le type de fichier spécifié	épistémologie site:lelibellio.com filetype:pdf
intitle:	Exécute la recherche du terme sur les titres des pages indexées	intitle:épistémologie site:lelibellio.com filetype:pdf
allintitle:	Exécute la recherche <u>des</u> termes sur les titres des pages indexées	allintitle:épistémologie sciences de gestion
intext:	Exécute la recherche du terme sur le corps de texte des pages indexées	intext:épistémologie -intitle:épistémologie site:lelibellio.com filetype:pdf (= les pages en PDF ayant dans leur corps de texte "épistémologie" mais pas dans leur titre, sur le site lelibellio.com)
allintext:	Exécute la recherche <u>des</u> termes sur le corps de texte des pages indexées	allintext: star wars site:lelibellio.com
inurl:	Exécute la recherche du terme sur les adresses (les URL) des pages indexées	inurl:"vol.-7" site:lelibellio.com filetype:pdf
allinurl:	Exécute la recherche <u>des</u> termes sur les adresses (les URL) des pages indexées	allinurl:2011 egos site:lelibellio.com filetype:pdf
daterange:	Restreint les résultats selon la date d'indexation par Google. Problème : la date est à mettre au format du calendrier Julien**	"big data" daterange:2454466-2454831 Correspond à la recherche "big data" pour les pages indexées du 01/01/2008 au 31/12/2008
author:	Cet opérateur permet de chercher dans le champ « auteur » des métadonnées (spécifique à Google Scholar)	author:skywalker intitle:force

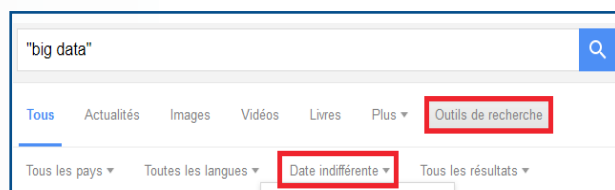
« La réponse à la grande question de la vie, de l'univers et de tout le reste est : 42 »

Maintenant maître dans l'art des recherches, il est temps de laisser place aux arcanes de Google. Utilisé pour effectuer des recherches, Google dispose aussi d'un certain nombre de fonctionnalités. Voici un florilège des fonctions disponibles, agrémenté de quelques exemples :

- Calculatrice (des fonctions élémentaires mais également incluant des fonctions avancées, des constantes physiques, etc.). La fiche Wikipédia de la calculatrice Google recense les fonctions et constantes disponibles (https://fr.wikipedia.org/wiki/Calculatrice_Google).
- Convertisseur ([convert 100 € to \$] renvoie 112.05 US\$)¹.
- Dictionnaire ([define:abstrus]).
- Horloge ([time now] ; [time:abu dhabi] ; [time:africa]).
- Prévisions météorologiques ([weather:saint petersburg, USA]).
- Cartes ([map:figueres]) ou trajets ([travel:lille naples]).
- Cours des actions ([stocks:apple]).
- Réponses aux questions ([évolution PIB France 2000] renvoie un graphique dynamique de l'évolution du PIB).
- Faire un tour sur lui-même ([do a barrel roll]).
- Connait même La réponse sur la vie, l'univers et le reste ([the answer to life the universe and everything]).

* Veillez à ne pas mettre d'espace entre l'opérateur, les deux-points et votre requête sinon Google interprétera cela comme un élément de la requête et non comme un opérateur.

** Bien qu'il existe des convertisseurs du format grégorien au format julien (par exemple celui de l'Observatoire naval des États-Unis, <http://aa.usno.navy.mil/data/docs/JulianDate.php>), l'utilisation de cette fonction reste malaisée. On lui préférera l'option de date présente dans les outils de recherche sur la page de recherche Google, plus facile et permettant de sélectionner des intervalles prédéfinis (1 heure, 1 semaine, 1 mois, 1 année) ou personnalisés.



¹ Les crochets représentent ce qui est saisi exactement dans le moteur de recherche ; ces crochets ne doivent pas être saisis dans la requête

Limites et alternative

Nous l'avons vu, une des clés d'une recherche efficace est de laisser le moins de place possible à l'interprétation par le moteur de recherche. Malgré cela, Google interprétera tout de même la requête afin de proposer ce qu'il considère comme des résultats plus pertinents plutôt que de renvoyer exactement ce que vous souhaitez. Si de manière générale le fait de transformer la requête sert l'utilisateur (par exemple la correction des fautes de frappe par exemple), le fait de ne pas avoir le choix de désactiver cette assistance pose question. Le flou dans les transformations faites sur les requêtes (correction orthographique, dictionnaires de mots interdits, lemmatiseur, etc.) et la différence de résultats pour une même requête servent de rappel au fait que Google n'est pas qu'une interface de requête sur une immense base de données mais surtout une entreprise monétisant les données à sa disposition. Sans rentrer dans le débat de l'utilisation des données personnelles par une entreprise privée, il ne faut pas perdre de vue que les résultats proposés par Google ne sont pas absolus en termes de pertinence et d'exhaustivité. Au contraire, le profilage effectué par Google tend à renvoyer des résultats similaires à ce que Google considère comme vos attentes et à atténuer la variété des résultats disponibles. Lors d'une requête, Google ne vous renvoie pas les meilleurs résultats mais ce qu'il considère comme vos résultats.



Le canard contre-attaque



Faute d'apporter une solution aux limites de Google, il est possible d'utiliser un moteur de recherche alternatif. Un exemple est donné par le moteur de recherche DuckDuckGo qui s'appuie sur les résultats de Google en supprimant les mouchards, cookies et autres outils de collectes de données personnelles (malheureusement, ce moteur ne propose pas d'alternative à Google Scholar). DuckDuckGo propose les mêmes fonctions de recherche que Google (opérateurs logiques, distributivité, opérateurs spécifiques) qui sont disponibles sur la page d'aide du moteur de recherche (<https://duck.co/help/results/syntax>).

En conclusion de cette rubrique, nous vous présentons quelques fonctions spéciales de ce moteur, inutiles et donc parfaitement indispensables :

- !bang permet d'accéder à d'autres sites depuis la page de recherche de DuckDuckGo : [!gsc] permet d'envoyer directement une requête sur Google Scholar (ex. : [!gsc author:dumez]), [!cnrtl] envoie une requête au moteur du centre national de ressources textuelles et lexicales du CNRS, etc. Il en existe des milliers d'autres accessibles en saisissant [!] avec les premières lettres du site web que vous souhaitez utiliser pour voir si les résultats suggérés correspondent à ce que vous désirez.
- Calculer le nombre de jours entre deux dates (ex. : [number of days between 11/01/2013 and 31/12/2015] renvoie 1084).
- Convertir un nombre arabe en nombre romain et inversement (ex. : [roman numeral MCCCXXXVII] renvoie le nombre 1337 ; [numeral roman 314] renvoie CCCXIV).
- Générer un QR Code (ex. : [QR http://lelibellio.com/wp-content/uploads/2013/01/Le-Libellio-d-Volume-11-numéro-4-Hiver-2015-10-ANS-2.pdf]).



*** Pour aller plus loin ***

- Le concept de bulle filtrante expliqué par *DuckDuckGo: Escape your Search Engine's filter bubble* - <http://dontbubble.us/>
- Un article d'Élisabeth Noël publié dans le *Bulletin des bibliothèques de France* qui discute l'hégémonie de Google et sa « pertinence » algorithmique. Paru en 2005, soit quelques mois après la mise en ligne de Google Scholar, cet article soulève des questionnements qui sont encore d'actualité, notamment celui de réfléchir à l'offre d'un portail standardisé, gratuit et convivial d'accès à la connaissance : Noël Élisabeth (2005) "Google Scholar", *Bulletin des bibliothèques de France* [en ligne], n° 4, <http://bbf.enssib.fr/consulter/bbf-2005-04-0043-009>. ISSN 1292-8399.

Geoffrey Leuridan
IAE Lille, LEM UMR CNRS 9221

Cécile Chamaret
Université Paris Sorbonne Abu Dhabi