

Effectuer de bonnes recherches avec Google

 frdl.ch/web/recherche-google

Table des matières

Le moteur de recherche Google est très souvent utilisé par les internautes, parfois sans même s'en rendre compte. C'est un outil aussi efficace que méconnu. Une recherche simple donne normalement de bons résultats – Google s'y emploie pour nous rendre dépendants. Mais on peut aller beaucoup plus loin avec certaines astuces très simples.

Pour un catalogue exhaustif des recherches Google, rien ne vaut la [documentation officielle](#) ou cette [documentation bien trapue](#). Je choisis ici de me limiter aux exemples concrets les plus utiles.

Il existe un [formulaire de recherche avancée](#), mais je le trouve peu commode à utiliser. Et, finalement, beaucoup moins clair que la *ligne de commande*. **Cette page est (et restera) en cours de rédaction!**

Le sujet vous intéresse? Vous souhaitez une formation pour vos employés? Des exercices pratiques? [Contactez-moi!](#)

Opérateurs logiques

AND (et)

La recherche la plus simple, c'est:

```
Nicolas Friedli
```

Ce qui signifie que Google propose des pages qui contiennent les mots Nicolas **et** Friedli. Je peux aussi préciser cela, en cherchant:

```
Nicolas+Friedli
```

L'opérateur est implicite. Si je ne dis rien, je recherche tous les termes demandés. Je peux aussi chercher, c'est synonyme:

```
Nicolas AND  
Friedli
```

OR (ou)

En changeant d'opérateur logique (qui n'est plus implicite), je peux rechercher:

```
Nicolas OR  
Friedli
```

Ou, dans une forme plus concise:

```
Nicolas|Friedli
```

Ces deux motifs de recherche me permettent de retrouver des pages qui contiennent ou un terme, ou l'autre terme, ou les deux. Le OR n'est pas exclusif.

- (négation)

Troisième opérateur, la négation. Je peux donc rechercher toutes les pages qui contiennent Friedli, mais pas Nicolas:

```
Friedli -  
Nicolas
```

C'est particulièrement utile quand une recherche est *contaminée*. Par exemple, DITA est un outil de documentation. Mais une recherche sur DITA proposera évidemment plein de résultats qui parlent de Dita von Teese. Je vais donc choisir:

```
DITA -  
Teese
```

C'est moins sexy, mais beaucoup plus efficace.

" " (recherche exacte)

L'utilisation de guillemets droits permet de signaler à Google que je veux trouver le motif exact. Dit autrement, les algorithmes de correction sont désactivés. Très utile si le motif de recherche est très proche d'un autre beaucoup plus connu ou si l'on souhaite retrouver une information erronée (faute d'orthographe).

Avec le motif suivant, Google corrige automatiquement (et partiellement) la recherche:

```
nicola  
friedli
```

Alors que

```
"nicola  
friedli"
```

ne permet plus, c'est logique, de me retrouver.

Attention, comme on cherche un motif exact, l'ordre des mots a évidemment son importance!

Dans la veille que j'effectue pour le projet Parpaillot.ch, j'utilise régulièrement la recherche suivante dans Google News:

```
"(église|paroisse) (protestante|réformée)"
```

Ce qui me permet de trouver 4 expressions exactes en évitant des sujets vagues qui utilisent deux des mots dans l'article!

site: (noms de domaine)

Il est possible de limiter la recherche à un seul site web. Je me souviens avoir écrit quelque chose pour ProtestInfo, donc:

```
Nicolas Friedli  
site:protestinfo.ch
```

Je peux aussi exclure un nom de domaine avec l'opérateur vu précédemment. Voici une recherche intéressante:

```
ProtestInfo -  
site:protestinfo.ch
```

Ce qui signifie que je recherche le terme ProtestInfo partout sauf sur le site protestinfo.ch.

Il n'est pas impératif d'utiliser un domaine complet. Je peux donc chercher les occurrences de mon nom sur les sites suisses¹⁾:

```
Nicolas Friedli  
site:.ch
```

Ou je peux travailler avec des sous-domaines. Très utile par exemple pour un blog hébergé chez wordpress.com:

```
site:arianebeldi.wordpress.com
```

À vous de décrypter ce que fait cette requête... Réponse: elle liste toutes les pages de l'excellent blog [Bloggo ergo cogito et sum....](#) Comme ce blog est hébergé en sous-domaine (arianebeldi) de wordpress.com, limitons-nous au sous-domaine en question.

filetype: (types de fichiers)

Une requête particulièrement utile: celle qui limite la recherche à certains types de fichiers. Par exemple:

```
Nicolas Friedli  
filetype:pdf
```

Encore un peu mieux, la recherche de mon nom dans les fichiers pdf hébergés sur des sites suisses:

```
Nicolas Friedli filetype:pdf  
site:.ch
```

Très utile pour retrouver quand on est mentionné dans des procès-verbaux et autres listes...

Exemples concrets

[Soumettez-moi](#) des exemples concrets de recherches qui vous sont utiles ou que vous n'arrivez pas à

formuler correctement. J'essaierai de vous répondre ici!

Réduire le nombre résultats pour un site web avec inurl

Si j'effectue une recherche limitée à l'excellent site www.ethikos.ch, j'obtiens des milliers de résultats. La requête, vous le savez maintenant, est:

```
site:ethikos.ch
```

En observant les résultats, on voit rapidement que certaines URL sont particulières, par exemple celles qui comprennent `category`. Je pourrais donc essayer:

```
site:ethikos.ch -  
site:ethikos.ch/category
```

Cela fonctionne, quelques centaines de pages en moins, mais c'est insuffisant. Essayons donc:

```
site:ethikos.ch -site:ethikos.ch/category -site:ethikos.ch/tag -  
site:ethikos.ch/page
```

Et là, je n'ai plus que les contenus, sans les taxonomies et les paginations. Comme WordPress a la mauvaise idée de dupliquer des contenus, il faut bien trouver des solutions.

On peut faire cela beaucoup plus élégamment, avec `inurl`, dont le nom semble explicite:

```
site:ethikos.ch -inurl:  
(category|tag|page)
```

Cet exemple utilise la notion de `site:`, la notion `inurl:`, la négation et le ou. Joli, non?

Détecter les hotlinks

Afin de savoir si des sites affichent des images hébergées sur votre serveur et volent votre bande passante, il suffit d'effectuer cette recherche²⁾ dans [Google Images](#):

```
inurl:frdl.ch -  
site:frdl.ch
```

Autrement dit: je recherche toutes les images (je suis sur Google Images) comportant mon nom de domaine dans l'URL mais qui ne se trouvent pas dans mon domaine.

Remarque: cette méthode permet aussi de détecter des liens, qui ne sont donc pas des images affichées sur le site en question. Soyez attentifs avant d'insulter le prétendu voleur.

web/recherche-google.txt · Dernière modification: 13.08.2015 à 17:11 par Nicolas Friedli