

Comment optimiser ses recherches sur Google en manipulant les urls du moteur

[Retour au sommaire de la lettre](#)

Les pages de résultats de Google répondent toutes à une url bien précise, affichant de nombreux paramètres (exemple : http://www.google.fr/search?hl=fr &c2coff=1 &rlz=1B3GGGL_frFR246FR246 &q=knol &btnG=Rechercher &meta=lr%3Dlang_fr). Savez-vous qu'en "manipulant" les données présentées dans ces adresses, vous pouvez aller beaucoup plus vite dans vos recherches ? Voici quelques explications détaillées de la façon dont il faut s'y prendre...

Dans cet article, nous allons examiner comment optimiser les résultats renvoyés par votre moteur de recherche fétiche en manipulant l'URL qui s'inscrit dans la barre d'adresses de votre navigateur. De cette façon, nous allons pouvoir affiner les requêtes, accéder à des paramètres cachés et créer des favoris à partir des syntaxes de recherches que nous avons imaginées. Selon le même procédé, nous verrons qu'il existe les mêmes possibilités en utilisant Google Images et Google Video.

Adresses URL : explications et ressources utiles

Une URL (Acronyme de "Uniform Resource Locator" ou littéralement "Repère constant d'une ressource") est une dénomination permettant de localiser une ressource sur l'internet, et qui indique également la méthode pour y accéder.

- Les URL ne sont pas sensibles à la casse ;
- Elles ne doivent pas contenir d'espaces.

Notez qu'Internet Explorer sait automatiquement ajouter <http://> et [.com](http://www.com). Il vous suffit d'appuyer sur la combinaison de touches Ctrl + Entrée...

Par ailleurs, il existe beaucoup de services vous permettant de réduire une adresse URL longue en un énoncé plus court comme Tinyurl : <http://tinyurl.com/>. Afin de trouver d'autres services du même genre, lancez cette recherche dans Google : **"* mieux que tinyurl**.

Enfin, Google a récemment mis en place une fonctionnalité permettant de corriger une recherche portant sur une URL qui n'existe pas. Faites un test sur cette recherche : <http://google.biz>. Ce sera l'occasion de visiter Google Burundi...



Les principes de base

L'adresse URL qui s'inscrit quand vous effectuez une recherche dans Google répond à une syntaxe très précise que nous allons, dans un premier temps, apprendre à décoder...

Le chemin commence obligatoirement par l'expression **"/search?"** suivi par un ou plusieurs commutateurs séparés par une Esperluette ("Ampersand" ou **&**).

Tous les paramètres doivent être encodés. Dans ce cadre, vous devez tout d'abord :

- Remplacer les espaces par le signe + ;
- Remplacer chaque caractère non alphanumérique par sa valeur ASCII en hexadécimal. Dans ce dernier cas, le format utilisé est celui-ci : % suivi par les deux digits au format hexadécimal.

Les caractères spéciaux \$-_.+!*'(), et les caractères réservés ;/?:@=& peuvent être utilisés comme des valeurs à l'intérieur de la requête.

Par exemple, la recherche [google site:www.abondance.com](#) se traduira par cette valeur : **&q=google+site%3Awww.abondance.com.**

Vous pouvez vous amuser à comparer les résultats renvoyés par ces deux recherches [+fête](#) et [fête](#) et constater que le signe "Plus" est encodé de cette manière : **%2B.**

Une table des codes ASCII est accessible à partir de cette page : <http://www.commentcamarche.net/base/ascii.php3>.

Les principaux commutateurs

&pws=0 : désactive la personnalisation des résultats qui peut être dûe soit au fait que vous avez activé l'Historique Web soit que vous êtes connecté à votre compte Google.

&cat=gwd/Top : l'ajout de ce drapeau force directement une recherche dans l'annuaire Google et dans une sélection des sites les mieux notés pour la catégorie considérée.

&q= ou **&as_q=** : annonce les termes de recherche que vous avez spécifiés.

hl= : permet de préciser la langue d'interface qui sera utilisée.

Cette valeur est source de découvertes car vous pouvez aussi afficher Google dans des interfaces linguistiques loufoques. Testez ces commutateurs : **hl=xx-hacker** - **hl=xx-bork** - **hl=xx-elmer** - **hl=xx-klingon** - **hl=xx-piglatin**.

Elles correspondent respectivement :

- À l'interface Google dédiée aux hackers ;
- Au chef suédois Bork (c'est un personnage de la série "Le Muppet Show") ;
- À Elmer Fudd du nom du chasseur de la série des "Bugs Bunny" ;
- Aux Klingons qui sont, dans la série "Star Trek", des extraterrestres originaires de la planète Kronos ;
- Au Pig Latin (appelé aussi backslang) qui est un genre d'argot utilisé originellement en Angleterre.

Ce type de syntaxe permet aussi de combiner différentes langues : **&lr=lang_en|lang_fr.**



lr= : restreint la recherche à la langue définie.

btnG= : correspond à l'utilisation du bouton Recherche Google.

btnI= : correspond à l'utilisation du bouton *J'ai de la chance*.

ie= : définit l'encodage des caractères en entrée (latin1, ut8, big5, etc.).

oe= : précise l'encodage des caractères en sortie.

q=cache: : affiche la page en cache.

q=related: : affiche les pages similaires.

&start= : indique le numéro de la page des résultats.

&meta=lr : dans le cas du moteur de Google France, le fait de cocher le bouton *Pages Francophones* provoque l'ajout de cette valeur : **&meta=lr%3Dlang_fr**.

&meta=cr : si vous cochez le bouton Pages : France, cette valeur va être alors ajoutée : **&cr=countryFR**.

Nous verrons un peu plus loin comment étendre ce type de recherche par pays ou par langue.

Ces commutateurs concernent directement la syntaxe des recherches :

&q + : tous les mots.

&as_epq= : cette expression exacte.

&as_oq= : au moins un des mots suivants.

&as_eq= : aucun des mots suivants.

Voici ceux qui relèvent des recherches avancées :

&lr= : permet de préciser la langue des pages qui apparaîtront dans les résultats. Les valeurs les plus courantes sont les suivantes :

Arabe : **lang_ar** ;

Chinois (Simplifié) : **lang_zh-CN** ;

Chinoise (Traditionnel) : **lang_zh-TW** ;

Tchécoslovaque : **lang_cs** ;

Danois : **lang_da** ;

Hollandais : **lang_nl** ;

Anglais : **lang_en** ;

Estonien : **lang_et** ;

Finois : **lang_fi** ;

Français : **lang_fr** ;

Allemand : **lang_de** ;

Grec : **lang_el** ;

Hébreu : **lang_iw** ;

Hongrois : **lang_hu** ;

Islandais : **lang_is** ;

Italien : **lang_it** ;

Japonais : **lang_ja** ;

Coréen : **lang_ko** ;

Lettonien : **lang_lv** ;

Lithuanien : **lang_lt** ;

Norvégien : **lang_no** ;

Portuguais : **lang_pt** ;

Polonais : **lang_pl** ;

Roumain : **lang_ro** ;

Russe : **lang_ru** ;

Espagnol : [lang_es](#) ;
Suédois : [lang_sv](#) ;
Turc : [lang_tr](#).

&as_filetype= : indique le format des fichiers.

&as_ft= : inclut (**i**) ou exclut (**e**) le format de fichier spécifié. Ce paramètre est à utiliser avec le précédent.

&as_qdr= : permet de rechercher les pages Web découvertes depuis une date précise. Quand on décompose cette URL, on voit que cette adresse contient ce commutateur : **&as_qdr=d**. Les valeurs possibles sont :

D : jour ;
W : semaine ;
M : mois ;
Y : année.

Par exemple, si vous spécifiez que votre requête porte sur les trois derniers mois, cette valeur sera ajoutée dans la barre d'adresses : **&as_qdr=m3**. Il est donc possible de définir des critères plus précis en utilisant ce type de syntaxe : **&as_qdr=d10** (les 10 derniers jours). Vous pouvez ensuite vous servir de la liste déroulante qui apparaîtra afin de changer le laps de temps défini.



The screenshot shows a Google search interface. The search bar contains 'site:abondance.com'. To the right of the search bar is a dropdown menu set to 'les 10 derniers jours' and a 'Rechercher' button. Below the search bar, it says 'Rechercher dans : Web Pages francophones Pages : France'. The search results are displayed under the heading 'Web' and show 'Résultats 1 - 10 sur un total d'environ 33 sur abondance.com durant le(s) les 10 derniers jours.' The first result is titled 'Abondance: SEOCamp, nouvelle association de référenceurs ...' and includes a snippet: 'Paru sur le Blog actualité d'Abondance le mercredi 12 décembre 2007, Auteur : Olivier Andrieu, Titre : Abondance: seocamp, nouvelle association de ...' and a link to 'actu.abondance.com/2007/12/seocamp-nouvelle-association-de.html - Il y a 9 heures - Pages similaires - À noter'.

&as_occt= : permet de définir l'emplacement des termes recherchés dans la page. Les valeurs possibles sont :

- Partout dans la page : **any** ;
- Dans le titre de la page : **title** ;
- Dans l'URL de la page : **URL** ;
- Dans les liens de la page : **links**.

- Par exemple : **&as_occt=title**.

&as_sitesearch= : permet de n'afficher que les pages d'un site ou d'un domaine.

&as_dt=e ou **&as_dt=i** : "Seulement" afficher les pages du site ou du domaine ou "Ne pas". À employer avec le commutateur **&as_sitesearch=**.

&as_rights= : définit les droits d'utilisation des pages des résultats. Les valeurs possibles sont les suivantes :

%28cc_publicdomain%7Ccc_attribute%7Ccc_sharealike%7Ccc_noncommercial%7Ccc_nonderived%29 : libres de droits d'utilisation et de distribution ;

%28cc_publicdomain%7Ccc_attribute%7Ccc_sharealike%7Ccc_nonderived%29.-

%28cc_noncommercial%29 : libres de droits d'utilisation et de distribution y compris à des fins commerciales ;

%28cc_publicdomain%7Ccc_attribute%7Ccc_sharealike%7Ccc_noncommercial%29.-

%28cc_nonderived%29 : libres de droits d'utilisation, de distribution ou de modification ;

%28cc_publicdomain%7Ccc_attribute%7Ccc_sharealike%29.-

%28cc_noncommercial%7Ccc_nonderived%29 : libres de droits d'utilisation, de distribution ou de modification y compris à des fins commerciales.

&safe= : active ou non le filtre "SafeSearch". Les valeurs possibles sont : **on** ou **off**.

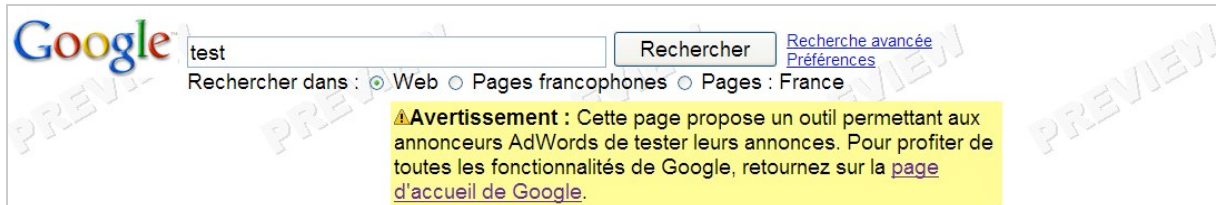
&num= : Nombre de résultats. Afin de n'afficher que les 5 premiers résultats, ajoutez cette valeur : **&num=5**.

&newwindow= : permet d'afficher les résultats dans une nouvelle page.

&as_rq= : trouve les pages similaires.

&as_lq= : trouve les liens qui pointent vers une page.

&adtest= (avec comme valeurs possibles **on** ou **off**) : autorise le test des annonces commerciales sans qu'une impression ne soit enregistrée. Les adresses URL ne seront pas cliquables. Il y aura un message d'avertissement vous signalant que "Cette page propose un outil permettant aux annonceurs AdWords de tester leurs annonces".



Désactiver certains filtres

Il existe deux filtres visant à limiter l'affichage de contenus trop semblables :

- Activer ou désactiver le filtrage des résultats pour un même nom de domaine : Google n'affichera que deux résultats pour un même site même si ce dernier contient plus de deux pages en rapport avec la requête (notion de *clustering*) ;
- Activer ou désactiver le filtrage des résumés : si deux pages n'ont pas le même contenu mais correspondent toutes deux à une requête donnée et présentent le même résumé ou "Snippet" (les deux lignes qui apparaissent sous le titre du résultat), Google n'affichera qu'une seule des deux pages. Bien entendu, ce cas est plutôt rare...

Voici les valeurs possibles pour ces filtres :

&filter=1 : les deux filtres sont activés ;

&filter=0 : les deux filtres sont désactivés ;

Par défaut, Google n'affiche pas les pages dont il juge le contenu trop similaires aux résultats déjà affichés. Ajoutez ce commutateur afin d'afficher tous les résultats. Cela revient donc à cliquer sur le lien *Relancer la recherche en incluant les pages ignorées* qui apparaît en fin des pages de résultats.

Pour limiter les résultats aux pages les plus pertinentes (total : 151), Google a ignoré certaines pages à contenu similaire.
Si vous le souhaitez, vous pouvez [relancer la recherche en incluant les pages ignorées](#).

Les pages affichées font partie d'un index de moindre qualité car Google les considère comme peu importantes. C'est dû au fait qu'elles présentent trop de similarité avec d'autres pages de votre site ou que leur contenu est jugé trop "court". Si, néanmoins, une page que vous estimez comme importante apparaît dans cette "Sandbox" d'un genre particulier, étoffez son contenu et optimiser les déclarations figurant dans les balises Meta.

&filter=p : seul le second filtre est activé ;

&filter=s : seul le premier filtre est activé.

Testez les résultats renvoyés par ces requêtes : [pagerank](#) ou [test site:abondance.com](#) en ajoutant ou non dans la barre d'adresses le commutateur "**&filter=0**".

Localisation géographique

L'adresse <http://www.google.com/webhp> évite la redirection automatique vers la version pour mobiles de Google Recherche sur le Web ;

L'adresse <http://www.google.com/ncr> revient à cliquer, à partir de la page d'accueil de Google, sur le lien *Google.com in English*. Vous serez dans la version globale de Google mais pas sur la version localisée USA.

Afin d'accéder à Google USA (et d'éviter les redirections automatiques), vous pouvez utiliser un Proxy en ligne (indiquez simplement cette adresse : <http://google.com>) ou saisir ce type d'adresse : <http://www.google.com/webhp?hl=en&gl=us>. Nous forçons ici à la fois l'utilisation de l'anglais comme langue d'interface ainsi que l'utilisation de la version américaine de Google.

Une autre manière consiste à utiliser cette syntaxe : <http://www.google.com/intl/us> (pour les États-Unis) ou celle-ci : <http://www.google.de/intl/de/> (pour l'Allemagne).

Il existe d'autres commutateurs mais qui concernent cette fois-ci les pages de résultats :

&cr= : nom du pays. Par exemple : **&cr=countryIT** si vous ne souhaitez rechercher que les pages hébergées sur des serveurs en Italie. Vous pouvez aussi utiliser ce type de syntaxe : **&gl=it**. Les résultats seront les mêmes à la différence près que vous conservez les avantages liés à votre version locale (comme la fonctionnalité des OneBox).

Les codes des pays et des territoires sont visibles à partir de cette adresse : http://www.google.com/apis/adwords/developer/adwords_api_countries.html.



Créer des favoris de vos syntaxes de recherche

La syntaxe de départ est donc celle-ci : <http://www.google.com/search?hl=fr&q=>.

La seule difficulté consiste à ce que vos paramètres URL ne soient pas interprétés comme faisant partie de votre recherche. Pour être plus clair : si vous définissez un favori qui vous permet de trouver facilement des fichiers MP3 avec Google, vous devez toujours avoir le choix du nom de l'artiste ou du groupe que vous allez spécifier. À titre d'exemple, testez les résultats renvoyés par ces deux requêtes :

[http://www.google.com/search?hl=en&q=intitle%3A%22index.of%22+\(mp3%7Cavi%7Cmov\)++beatles+-html+-htm+-php+-aspx+-cf&btnG=Search](http://www.google.com/search?hl=en&q=intitle%3A%22index.of%22+(mp3%7Cavi%7Cmov)++beatles+-html+-htm+-php+-aspx+-cf&btnG=Search) ;
[http://www.google.com/search?hl=en&q=intitle%3A%22index.of%22+\(mp3%7Cavi%7Cmov\)+-html+-htm+-php+-aspx+-cf&btnG=Search](http://www.google.com/search?hl=en&q=intitle%3A%22index.of%22+(mp3%7Cavi%7Cmov)+-html+-htm+-php+-aspx+-cf&btnG=Search).

La seconde ne fonctionne donc pas... Voici une manière de contourner le problème en utilisant Mozilla Firefox :

- Cliquez sur *Marque-pages - Organiser les marque-pages...*
- Cliquez sur le bouton *Nouveau marque-page*.
- Saisissez un nom.

Dans la zone de texte Adresse Web, copiez ce code :

<http://www.google.com/search?hl=en&q=intitle%3A%22index.of%22+%28mp3%29+%s+-html+-htm+-php+-jsp&btnG=Google+Search&meta=>.

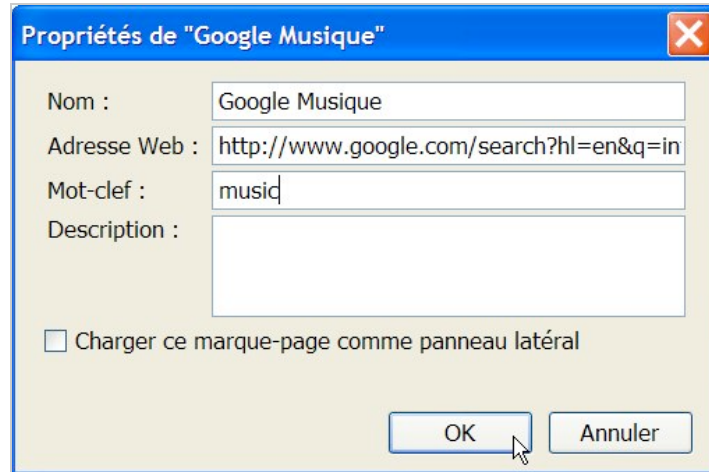
Vous pouvez l'adapter en fonction de vos préférences. Voici, par exemple, une syntaxe de recherche qui est plus efficace :

[intitle:index.of last modified parent directory \(mp3|wma|ogg\) beatles -htm -html -php -asp -e-mp3s.eu -fast-mp3.com -franceradio.net -mp3-network.net.](#)

Indiquez un mot-clé comme par exemple celui-ci : [music](#).

Laissez éventuellement le champ *Description* vide puis cliquez sur *OK*.

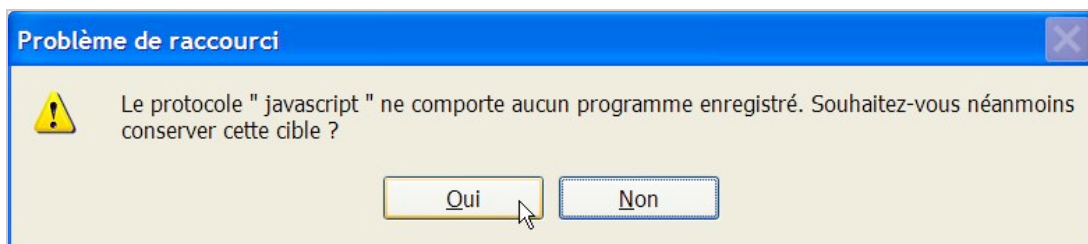
Dans la barre d'adresses de Firefox, saisissez, par exemple, cette recherche : [music beatles](#).



Sous Internet Explorer, il est possible d'utiliser un code Javascript :

- Créez un favori dans Internet Explorer (le nom et la page choisie n'ont pas d'importance).
- Ouvrez le volet des favoris.
- Cliquez avec le bouton droit de la souris sur le favori que vous venez d'ajouter puis sur le sous-menu *Propriétés*.
- Remplacez la mention visible dans la zone de texte URL par celle-ci :
`javascript:var i=prompt('Termes recherchés :','');if (i){this.location='http://google.com/search?q=intitle:%22index.of%22 '+i+' (mp3|mp4|flac|ogg) -htm -html -php -asp -jsp -cfm';}`

Une boîte de dialogue va vous signaler deux fois d'affilée que "*Le protocole 'Javascript' ne comporte aucun programme enregistré. Souhaitez-vous néanmoins conserver cette cible ?*" Cliquez sur les boutons *Oui*, *OK* et *Oui*.



Par ailleurs et dès que vous cliquerez sur votre favori, la barre d'informations va signaler que "*Ce site Web utilise une fenêtre scriptée pour vous demander des informations*". Afin de désactiver ce message d'avertissement, faites cette manipulation :

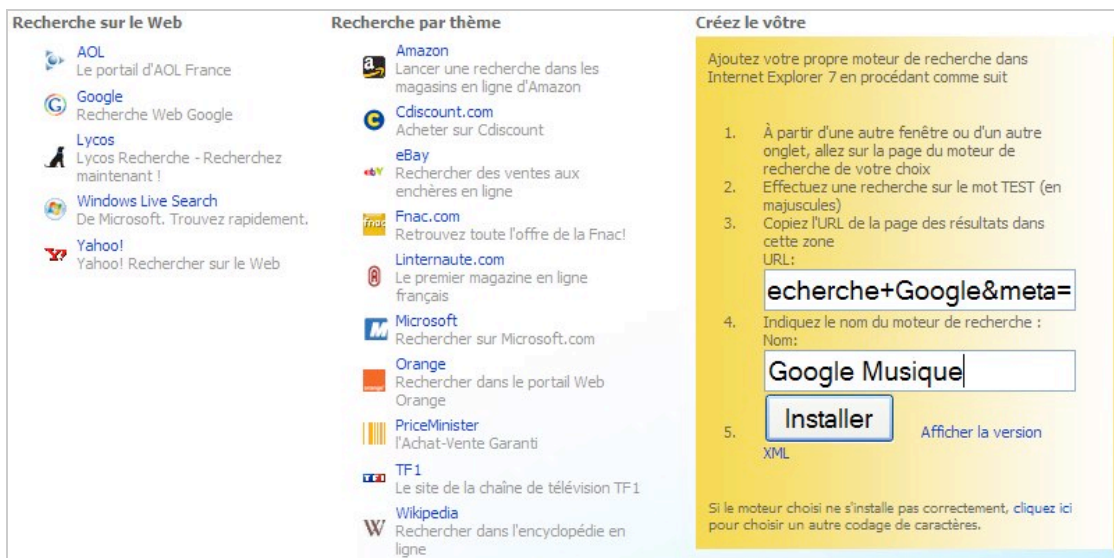
- Cliquez sur *Outils - Options Internet*.
- Cliquez sur l'onglet *Sécurité* puis le bouton *Personnaliser le niveau...*
- Sous la mention *Autoriser les sites Web à demander des informations à l'aide de fenêtres scriptées*, cochez le bouton radio *Activé*.
- Validez pour le reste...

Cette page propose également des Bookmarklets qui sont souvent intéressants :

<http://www.imilly.com/bm.htm>.

Mais il y a encore une autre solution :

- Dans Internet Explorer, cliquez sur la petite flèche placée à droite de la zone de recherche puis cliquez sur le lien *Rechercher encore des moteurs de recherche*.
 - Dans la rubrique *Créez le vôtre*, saisissez l'adresse URL de votre recherche telle qu'elle apparaît quand vous procédez à l'ajout du mot-clé **TEST** (en majuscules). Il suffit de faire une recherche dans Google avec l'ensemble des commutateurs que vous voulez. Par exemple : [intitle:index.of last modified parent directory \(mp3|wma|ogg\) TEST -htm -html -php -asp -e-mp3s.eu -fast-mp3.com -franceradio.net -mp3-network.net](http://intitle:index.of last modified parent directory (mp3|wma|ogg) TEST -htm -html -php -asp -e-mp3s.eu -fast-mp3.com -franceradio.net -mp3-network.net).
 - Copiez l'adresse URL puis définissez un nom pour votre moteur.
 - Cliquez sur le bouton *Installer*.
- Le tour est joué !



Google Images

Google Images se prête à toutes sortes de manipulation d'URL. Nous vous présentons simplement ici celles qui diffèrent de Google Recherche sur le Web :

&imgtype=news : Actualités ;

&imgtype=face : Visages ;



&imgsz= : Taille des images.

Dans ce dernier cas, les valeurs possibles sont les suivantes :

Icon : petites ;

Small : moyennes ;

Xxlarge : grandes ;

Huge : images de très grandes tailles.

La valeur "**Small**" regroupe des tailles intermédiaires : **medium|large|xlarge**. Cela correspond approximativement à ces dimensions : 100X100 - 240X320 - 300X400 - 600X800

&imgc= : renvoyer uniquement les images en :

Mono : Noir et Blanc ;

Gray : Niveaux de gris ;
Color : Couleur.

Le commutateur **&as_filetype=** vous permet de définir l'extension de fichier possible... Dans la pratique vous pouvez spécifier toutes sortes d'extensions : tiff, par exemple. Voici une autre suggestion de requête : [paysage filetype:pcx](#).

Google Video

Voici une petite astuce pour terminer. Il arrive souvent que vous ayez envie de partager une vidéo avec quelqu'un de votre entourage mais en faisant un lien à partir d'un moment précis à l'intérieur du fichier. Inutile dans ce cas de lui infliger toutes les scènes qui encadrent l'extrait qui va l'intéresser. Le principe consiste alors à modifier l'adresse URL du clip en ajoutant son minutage exact. Pour cela, ajoutez à l'adresse URL, le signe **dièse** suivi de la durée exprimée en **h, m, s** (ou m, s pour les vidéos plus courtes).

À titre d'exemple, comparez les résultats renvoyés par ces deux liens :

<http://video.google.fr/videoplay?docid=7294604022161146201&q=google&total=4562257&start=0&num=10&so=0&type=search&plindex=0> ;

<http://video.google.fr/videoplay?docid=7294604022161146201&q=google&total=4562257&start=0&num=10&so=0&type=search&plindex=0#1h15m42s>.

Tracer des graphiques en utilisant des adresses URL

Vous l'aurez compris : Google est une sorte de magicien en ce qui concerne l'utilisation des adresses URL. Alors pourquoi s'arrêter en si bon chemin ? La société de Mountain View a récemment créé une API appelée "Google Chart API".

Le principe est très simple : ouvrez votre navigateur puis copiez ce type de lien :

<http://chart.apis.google.com/chart?cht=p3&chd=s:Uf9a&chs=200x100&chl=A|B|C|D>.

Appuyez sur la touche *Entrée* pour afficher immédiatement le graphique que vous venez de créer ! Oui, c'est tout ! Vous êtes limité à 50 000 affichages par jour et par site. Ne croyez pas que ce soit beaucoup : en imaginant que vous avez créé 50 graphiques et que votre site enregistre plus de 1 000 visiteurs dans la journée, il y aura forcément un moment à partir duquel vos œuvres ne seront plus visibles...

Signalons enfin que cet API est couramment utilisée dans Google Finance :

- Accédez à ce service en saisissant cette URL : <http://finance.google.com/finance>.

- Cliquez avec le bouton droit de la souris sur un des graphiques présents puis sur *Propriétés*.

L'adresse URL visible vous donnera des indications sur le type de syntaxe qui est utilisée. Afin de mieux comprendre comment s'imbriquent les différents éléments, plongez-vous dans la documentation pour les développeurs qui est accessible à partir de cette adresse :

<http://code.google.com/apis/chart>.

Bonnes recherches !

Article écrit par un rédacteur free-lance pour la lettre "Recherche & Référencement"

Réagissez à cet article sur le blog des abonnés d'Abondance :

<http://abonnes.abondance.com/blogpro/2007/12/dcembre-2007-comment-optimiser-ses.html>