

Créer des pages d'accueil optimisées et des fichiers Sitemap pour Google News

[Retour au sommaire de la lettre](#)

Le site Google News est très souvent un générateur de trafic important pour un site d'actualité. Il est donc important d'optimiser ses pages pour obtenir un référencement quantitatif et qualitatif de qualité... Voici deux façons d'y arriver : créer des pages "sommaire" attractives pour les robots du moteur et proposer à ce dernier un fichier Sitemap dédié à vos articles d'actualité. Explications...

Cet article regroupe différentes recommandations concernant le référencement d'un site au sein du moteur de recherche d'actualité Google News (<http://news.google.fr/>). Il s'attache particulièrement à décrire deux méthodes permettant l'indexation d'un maximum d'articles d'un site déjà éligible au système Google News.

En effet, le robot d'indexation de Google News ne parcourt pas les sites Web de la même façon que le robot du moteur de recherche Web Google. Alors que ce dernier est conçu pour trouver la totalité des pages d'un site en suivant ses liens, le robot de Google News n'a pour but que de trouver de nouveaux articles d'actualité mis en avant sur différentes pages d'accueil. Il se contente donc d'une profondeur de crawl d'une page, listant les articles directement disponibles depuis sa page cible.

De cette particularité se dégage la nécessité de mettre en ligne des pages (HTML ou sitemaps XML) dédiées aux robots de Google News, optimisant à la fois les contenus proposés et la fréquence de mise à jour de ceux-ci.

Les points importants du fonctionnement de Google News

Suite à différentes études portant sur la façon dont le robot d'indexation de Google News voit et référence les différentes pages d'un site, deux points se sont révélés importants dans la conception de ce type de page :

- La **fréquence de visite** du robot de Google News sur une page dépend de la fréquence de mise à jour de celle-ci. Ainsi, plus une page sera mise à jour de façon fréquente, plus Google News viendra voir les éventuels nouveaux articles et fera apparaître ceux-ci rapidement dans ses pages de résultats. *A contrario*, une page mise à jour de façon irrégulière ne sera pas visitée régulièrement par le robot de Google News et perdra ainsi de sa réactivité.

- La **pertinence d'une page** en fonction d'une actualité précise tient à l'utilisation, dans cette page, de mots clés précis relatifs à l'actualité en question. Les moyens d'optimisation d'un article ou d'une page pour ces mots clés sont les mêmes que ceux utilisés pour les moteurs de recherche classiques (titres, pertinence du texte, liens...) et doivent donc être considérés de la même façon.

Ces deux points permettent de concevoir une page d'accueil spécialement dédiée à l'indexation d'un site par Google News afin de garantir à la fois une meilleure réactivité du site par rapport au moteur, mais également une indexation optimale en fonction de mots clés précis. Sur ce dernier point, l'écriture générale et la présentation des articles sur un site est cruciale et devra également être adaptée en fonction des résultats souhaités.

1. Structure, contenu et mise en forme d'une page Sommaire

Le robot d'indexation de Google News n'a pour but que de trouver de nouveaux articles. Il se contente donc d'une profondeur de crawl d'une page, listant les articles directement disponibles depuis sa page cible. **Il convient donc qu'une page dédiée à Google News liste un maximum d'articles et mène directement vers le détail de ceux-ci.** Afin de garantir une meilleure réactivité de la part de Google News, cette page devra être mise à jour de façon très régulière, et si possible en temps réel suivant la publication des articles dans les différentes sections du site. Dans l'idéal, cette page doit diffuser l'actualité de toutes les rubriques déjà présentes sur Google News.

L'ajout d'une nouvelle page d'accueil sur le site ou d'un dossier spécial susceptible de contenir des articles non listés sur les autres pages d'accueil devra également se faire sur cette page dédiée à Google News. Chaque nouvel article publié sur le site doit automatiquement apparaître sur cette nouvelle page, si possible tout en haut de la liste des articles disponibles. La page Google News doit,

de fait, fonctionner à la façon d'un fil d'information permanent présentant le résumé de tous les articles du groupe et menant à chacun d'eux par le biais de liens définis de façon précise.

L'ensemble des actualités présentées sur la nouvelle page dédiée à Google News devra être présentée de façon à optimiser les mots clés des différents articles du site. A savoir :

- Le titre de l'article affiché à l'aide d'une balise <H2> et lié directement à l'article détaillé ;
- La date et l'heure de publication de l'article ;
- La rubrique initiale de publication de l'article, ou les thèmes associés à celui-ci ;
- Le chapeau ou résumé de l'article en question, lui aussi lié à la page de détail.

La présentation pourrait être la suivante :

L'Abbé Pierre est mort ce matin au Val de Grâce (France - 22/01/2006 à 08h52)
L'hôpital du Val de Grâce (Paris) a annoncé ce matin la mort de l'Abbé Pierre à la suite de son hospitalisation en fin de semaine dernière.

D'une manière générale, il convient de veiller à l'optimisation des titres et chapeaux des articles publiés sur les sites. Ces titres et textes doivent, au maximum, utiliser des termes clairs et évidents, susceptibles d'être à la fois utilisés par d'autres sites sur le Web et recherchés par les internautes sur les différents moteurs de recherche. On évitera ainsi de parler du "Président des Etats-Unis" et on préférera nommer explicitement "Georges Bush" dans les titres, chapeaux et URL des articles afin que l'article en question et les liens y menant soient correctement optimisés sur ce nom, plus susceptible d'être recherché par les internautes.

La technologie de Google News conçoit la page d'accueil du service (<http://news.google.fr>) en fonction des affinités de contenu observées sur les différentes sources de contenu du service. Ainsi, plusieurs articles provenant de plusieurs sites (Le Monde, Le Figaro, Libération, ...) et utilisant la même terminologie (*Georges W Bush* comme mot clé par exemple) seront regroupés sous un seul bloc et apparaîtront dans un seul élément de navigation sur la page d'accueil du service. Il convient donc de s'assurer de l'utilisation du même vocabulaire que les autres sites d'actualité afin de s'assurer une présence dans ces blocs présents en page d'accueil. Un article isolé, de par son sujet ou de par son vocabulaire et ses mots clés utilisés, ne remontera que rarement en page d'accueil de Google News.

Le nombre d'articles qui doit être présenté sur cette nouvelle page dépend à la fois de la fréquence de mise à jour de celle-ci - soit la fréquence d'apparition de nouveaux articles sur le site - et de la fréquence des visites du robot de Google News sur cette même page. La but étant bien entendu que Google News ne manque aucun des articles mis en avant sur la page, et donc que ceux-ci ne disparaissent pas trop rapidement de la page.

Au vu des fréquences de mise à jour initiale de Google News (visites allant de la demi-heure à l'heure suivant les pages d'accueil définies) et du nombre d'articles publiés en "heure de pointe" sur un site d'actualité classique, il est recommandé d'afficher, dans un premier temps, les cinquante derniers articles disponibles. Ce nombre sera bien entendu adapté en fonction de l'activité constatée sur le site et du rafraîchissement des pages de résultat de Google News.

Intégration dans Google News et dans le site

Google News se base sur une liste d'URL soumise manuellement au service pour parcourir un site d'actualité et en identifier les différents articles susceptibles d'être affichés sur le service.

Une fois le fil d'actualité spécifique à Google News conçu et mis en place sur le site, celui-ci sera soumis aux services de Google News France au moyen des formulaires mis à disposition sur le site de Google.

Il faut toutefois que la page destinée à Google News soit également accessible de façon naturelle aux visiteurs du site, même sans mise en avant excessive, afin que Google puisse s'assurer que le contenu soumis au robot est bien identique à celui mis à disposition des internautes. Cette page pourrait donc être ajoutée à la navigation générale du site à l'aide d'un lien "Toute l'actualité" présent dans le pied de page du site.

2. Créer un fichier Sitemaps pour Google News

Depuis l'été 2007, Google offre la possibilité aux organismes de presse de faciliter leur inclusion dans Google News par le biais du protocole *Google News Sitemaps* (<http://googlewebmastercentral.blogspot.com/2006/11/introducing-sitemaps-for-google-news.html>). Ce protocole, hérité du protocole Sitemaps classique (<http://www.sitemaps.org/>) mis en place par Google en juin 2005, permet de lister les derniers articles publiés sur un site et de soumettre ceux-ci de façon plus automatique aux robots de Google News.

Un seul pré-requis est indispensable pour la mise en place du protocole News Sitemaps sur un site : Celui-ci doit déjà faire partie des sites authentifiés comme source de contenu pour Google News. Les sites non authentifiés doivent passer par une soumission manuelle de leur contenu et une approbation humaine de leurs articles. Une fois cette approbation faite, l'option Google News Sitemaps sera alors disponible pour eux également dans l'interface Google Webmaster Tools associée à leur compte.

La structure d'un fichier News Sitemaps est très proche à celle d'un fichier Sitemaps classique. Il s'agit toujours d'un fichier XML listant l'ensemble des URL dont les contenus doivent être indexés par Google News. Toutefois, au vu de la nature particulière des documents fournis au moteur à l'aide de ce service, Google a pris soin d'ajouter quelques balises spécifiques destinées à renforcer la sémantique des pages référencées.

La structure de base d'un fichier News Sitemaps est la suivante :

```
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9"
  xmlns:news="http://www.google.com/schemas/sitemap-news/0.9">
  <url>

    <loc>http://mynewssite.com/article100.html</loc>
    <news:news>
      <news:publication_date>2006-08-
17T03:19:00Z</news:publication_date>
      <news:keywords>Sports, Baseball</news:keywords>

    </news:news>
  </url>
</urlset>
```

La déclaration d'en-tête est immuable et confirme la conformation aux protocoles Sitemaps classique et News Sitemaps pour le document en cours. Le reste du fichier est ensuite composé d'une suite de balises <url> contenant chacune la description précise des URL soumises. Les différentes balises qui décrivent les pages sont les suivantes :

Balise		Usage
<loc>	obligatoire	Définit l'URL de la page à indexer au sein de Google News.
<news:news>	<i>optionnelle</i>	Définit le début des META Données propres à Google News et destinées à qualifier l'article soumis.
<news:publication_date>	<i>optionnelle</i>	Contient la date de publication de l'article soumis, au format W3C (YYYY-MM-DDThh:mmTZD).
<news:keywords>	<i>optionnelle</i>	Contient une liste de mots clés, séparés par des virgules, qualifiant l'article soumis. Ces mots clés sont destinés à aider Google News à déterminer le sujet réel de l'article

		en question et à positionner celui-ci dans les bonnes catégories d'actualité. Les noms des rubriques Google News peuvent être utilisés dans cette balise.
<news:stock_tickers>	<i>optionnelle</i>	Contient le code MNEMO de l'entreprise citée dans l'article, ou les codes des indices cités, dans le cadre d'informations économiques, boursières ou financières.

Le contenu d'un fichier News Sitemaps

Comme un fichier Sitemaps classique, le fichier News Sitemaps possède des limitations quant au nombre d'URL soumises (un maximal de 1 000) et quant au poids total du fichier (10 Mo).

Le système News Sitemaps possède une contrainte quant au contenu proposé. Chaque fichier News Sitemaps est associé pour Google News à une source d'information précise. Cette source correspond au nom du site ou du journal cité dans l'interface générale de Google News aux côtés du titre d'un article :



Les articles listés dans un fichier News Sitemaps ne peuvent appartenir qu'à une source d'information unique. Si un site web est déjà reconnu au sein de Google News comme faisant partie de deux sources distinctes (sur deux versions locales de Google News par exemple), il faudra alors concevoir un fichier News Sitemaps pour chacune des sources en question et déclarer ces deux fichiers de manière distincte au sein de l'interface Google Webmaster Tools.

Ces fichiers ne sont toutefois pas destinés à lister l'intégralité des actualités d'un site, mais uniquement les derniers articles qui devraient apparaître dans les pages de résultats de Google News. Dans l'idéal, ils doivent donc contenir l'ensemble des URL qui devraient être indexées depuis le dernier passage du robot de Google News sur ce fichier. Il est donc préférable de lister dans les fichiers News Sitemaps un nombre limité de pages (les 500 à 800 dernières actualités par exemple) mais de garder ces fichiers le plus à jour possible afin de pouvoir informer Google News au plus vite de la présence de nouveaux articles.

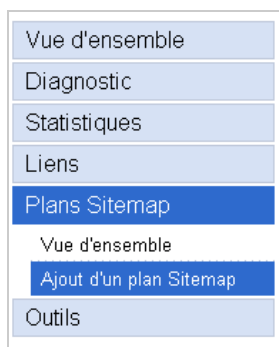
Dernier point, les fichiers News Sitemaps doivent être mis à jour en direct avec les derniers articles publiés sur le site afin de fournir les informations les plus pertinentes possibles au moteur d'actualité de Google.

Soumission d'un Sitemaps Google News

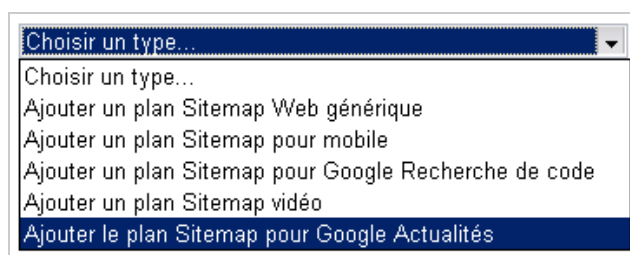
Comme les fichiers Sitemaps classiques, les fichiers Sitemap pour Google Actualités se déclarent auprès de Google via l'interface Google Webmaster Tools (<http://www.google.com/webmasters/tools/>). Comme dit précédemment, l'option de déclaration de ces fichiers n'est disponible que si un site est déjà reconnu comme source d'information pour le service.

Si ce pré-requis est rempli, la déclaration du fichier se fait de la manière suivante :

1. Dans l'interface Google Webmaster Tools, se rendre dans la section *Plan Sitemap* correspondant au site dont dépend le fichier Sitemap pour Google Actualités et sélectionner la sous-rubrique *Ajout d'un plan Sitemap* :



2. Dans la liste déroulante des types de Sitemaps disponible, sélectionner l'option *Ajouter le plan Sitemap pour Google Actualités* :



3. Les options d'inclusion s'affichent alors :

1. J'ai créé un plan Sitemap pour Google Actualités dans l'un des formats pris en charge. ?

2. J'ai transféré mon plan Sitemap vers le répertoire de plus haut niveau auquel j'avais accès.

3. L'URL de mon plan Sitemap est :
http://www.autojournal.fr/
Exemple : sitemap.xml

4. Libellé de publication :

En créant un plan Sitemap pour Google Actualités, vous demandez à Google d'inclure des extraits de vos articles dans Google Actualités. Pour plus d'informations, [cliquez ici](#).

4. Compléter l'URL du fichier Sitemap pour Google Actualités (*L'URL de mon plan Sitemap est :*) avec le nom effectif du fichier Sitemap pour Google Actualités disponible sur le site ;

5. Sélectionner dans la liste déroulante *Libellé de publication* le nom de la source qui apparaîtra dans Google News pour l'ensemble des URL disponibles dans le fichier Sitemap pour Google Actualités.

Une fois ces renseignements fournis, le fichier Sitemap pour Google Actualités apparaît dans le tableau récapitulatif des Sitemaps dans l'interface de Google. Ce tableau permettra de connaître la fréquence de visite du robot de Google News sur le fichier Sitemaps et la présence d'éventuelles erreurs dans le contenu ou le codage de ce dernier.

François Houste

Directeur SEO - LSF Interactive (<http://www.lsfinteractive.fr/>)

Search Engine Feng Shui (<http://www.search-engine-feng-shui.com/>)

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

<http://abonnes.abondance.com/blogpro/2008/03/creer-des-pages-d-accueil-optimisees-et.html>