

La vérité sur l'indexation dans Google News

Les éditeurs considèrent qu'il existe un grand flou dans la façon dont le géant de la recherche sur le Net procède au classement des actualités.

Conséquence de ce manque de transparence, l'actuel « Big Boss » de cette division très importante pour le groupe américain, l'ancien journaliste de l'agence Reuters, Josh Cohen, a enfin décidé de détailler le fonctionnement du système.

Le blog officiel de Google News, relaye désormais une page Web intitulée « Mythes et Réalités » ou « Truths and Myths » dans la version originale. Cet article est supposé répondre aux nombreuses questions que se posent les éditeurs de la presse en ligne.

Les « mythes »...

Premier mythe, le fait de placer une image dans un article n'améliore pas le classement d'un article dans Google News. Bien entendu, une bonne image améliore les chances de diffusion de cette dernière sur le site de Google News, mais cela n'a pas d'impact sur le « ranking pur » du papier.

Ensuite, de nombreux éditeurs pensent que l'heure de la publication d'article peut avoir un impact sur le classement des articles dans la page actualité du moteur de recherche. Là aussi il s'agit d'un mythe d'après Google.

A ce sujet, Google explique que si ces équipes cherchent effectivement les articles les plus récents, cela n'a pas d'impact sur le classement d'un article et cela même si le document est publié à l'instant T ou les équipes de Google News cherchent les informations les plus pertinentes. En réalité, les algorithmes de Google prennent en compte plusieurs éléments afin de choisir le meilleur article parmi une série de papier traitant d'un même sujet.

Par exemple, deux sites publient une actualité relativement similaire, la priorité est donnée à celui qui est à l'origine de l'histoire. Un statut qui est déterminé par Google. Enfin, si une source essaye de tricher en modifiant constamment son article alors ce dernier est signalé au système d'indexation.

Autre idée reçue, les éditeurs pensent qu'il n'est pas possible de savoir pourquoi un article n'est pas référencé dans Google. Le moteur de recherche rappelle que cela n'est pas vrai dans la mesure où il propose un outil à destination des Webmasters afin d'analyser la publication et l'indexation des news. Google explique également que si un site est inscrit sur Google News, l'utilisateur peut se créer un compte et ainsi analyser les erreurs.

La création d'une sitemap (carte du site) n'aide pas à l'indexation contrairement à ce que pensent les éditeurs. Cela peut aider les équipes de Google à trouver une actualité et dans une certaine mesure s'ils ne peuvent pas trouver l'information, ils ne peuvent pas la classer. Mais d'après Google, cela n'a pas d'impact sur la position d'un papier sur la page de News.

Pour Google, l'idée de réaliser un sitemap est bonne malgré tout. Car cela offre un meilleur contrôle et permet à ses équipes d'identifier les derniers papiers plus rapidement. Enfin dernier

avantage, cela permet de spécifier des meta-données pour chaque article. Par exemple, sa date de publication, ou les mots clés qu'il contient.

Enfin, il existe un dernier mythe présenté par Google selon lequel l'utilisation d'AdSense favorise l'indexation sur Google News. Cela est faux affirme Google. « *Nous voulons garder notre objectivité et ce principe ne l'est absolument pas* ».

Les « réalités »...

Pour ce qui est des réalités de l'indexation sur Google News, la firme de Mountain View est moins bavarde. Premier point, elle confirme que la mise à jour d'un article après sa publication provoque des problèmes au niveau de Google News. Il vaut donc mieux vérifier à deux fois avant de procéder à la mise en ligne d'un article.

Pourquoi ? Tout bonnement parce que le système de recherche des articles, le Google Crawler, ne visite les URL qu'une seule fois. Notons que Google pense pouvoir mettre en place un système plus évolué prochainement, capable de re-indexer les URL et ainsi prendre en compte les modifications apportées à un article.

Les articles qui ne contiennent que de la vidéo et des images ne seront pas classés dans Google News. « *C'est vrai* », affirme Cohen sur le blog du moteur de recherche.

Les articles qui contiennent des données multimédias sont inclus dans la recherche, mais s'ils ne contiennent pas de texte ils ne sont pas retenus.

La refonte d'un site peut avoir un effet sur le classement des News dans Google. C'est vrai, mais il faut que les changements soient importants précise Google. Il est donc préférable d'anticiper sur un changement de design ou de nom en contactant les équipes de Google.