

Le référencement audio, une stratégie à prendre en compte dès maintenant

[Retour au sommaire de la lettre](#)

Google a lancé le mois dernier un service, baptisé GAUDI, de reconnaissance vocale à l'intérieur des vidéos. L'outil recherche ainsi, au sein de documents sonores, les occurrences des mots clés saisis au clavier par l'utilisateur... L'occasion était donc belle de faire un tour d'horizon des différentes technologies d'indexation disponibles aujourd'hui sur le marché pour l'audio et d'envisager leur impact sur le référencement. Un type de stratégie qu'il va falloir certainement mettre en œuvre très rapidement pour bon nombre de sites web orientés multimedia...

16 septembre 2008 : Google annonce le lancement de son premier portail consacré à la recherche de documents Audio, à l'adresse <http://labs.google.com/gaudi>. C'était l'occasion rêvée de faire le point sur les différentes technologies d'indexation audio et sur leur impact pour le référencement. Voici un tour d'horizon des différentes technologies disponibles à l'heure actuelle...

Google Audio Indexing (GAUDI)

En juillet 2008, Google avait déjà frappé un grand coup en pleine période électorale américaine en proposant un service baptisé "Speech Recognition", destiné à rechercher des mots clés dans les discours des politiques américains (voir le communiqué : <http://googleblog.blogspot.com/2008/07/in-their-own-words-political-videos.html>).

A l'époque, il s'agissait d'un "gadget" destiné aux possesseurs d'un compte Google et pouvant enrichir la recherche de vidéos Youtube. Ce module marquait la création de l'équipe "Google Speech Team". Le service s'est désormais généralisé sur un portail vidéo dédié, sans que l'utilisateur ait besoin d'installer un module complémentaire. Baptisé GAUDI (*Google Audio Indexing*), ce portail fonctionne comme un moteur de recherche vidéo classique... sauf que les mots clés recherchés se trouvent à l'intérieur même de la vidéo, dans sa bande-son. Pour le moment le moteur de recherche interroge exclusivement les "YouTube political channels".

The screenshot shows the Google Audio Indexing (GAUDI) interface. At the top, there is a search bar with the text "Search what the politicians are saying" and "france". Below the search bar, there is a list of search results for "france". The results include titles like "Barack Obama in Paris", "Harpers record disastrous for women", "Gov Romney Clinton Couldn't Even Get Elected In France", "Rudy Believes in School Choice", "Jack Layton and former Saskatchewan Premier Calvert", "Kucinich Campaign Update 1 29 08 All for One", "Rudy with Chris Matthews", and "Rudy on His Record of Cutting Taxes". To the right of the search results, there is a video player for "Barack Obama in Paris". The video player shows a video of Barack Obama and Nicolas Sarkozy. Below the video player, there is a search bar with the text "france" and a "Search inside this video" button. Below the search bar, there is a list of search results within the video, such as "...you should talk to the President of France he seems to have a good ..." and "...to strengthen the bilateral relationship between France and the united states a ...".

Comme souvent chez Google, l'interface est pensée en terme d'ergonomie et d'efficacité, et la présence de marqueurs temporels intégrés à la vidéo en rend l'utilisation immédiatement attractive.

Plusieurs informations intéressantes sont proposés par Google dans l'aide en ligne : <http://labs.google.com/gaudi/static/faq.html>

On y apprend ainsi que l'indexation audio est un étape supplémentaire dans le grand projet de Google : organiser l'information mondiale et la rendre universellement accessible de façon pratique. On s'en serait un peu douté... :-)

Dans ce cadre, les discours des hommes politiques ont été choisis comme terrain d'expérimentation, car des informations clés étaient transmises dans les élocutions. Mais comme l'affirme Google, le domaine des élections américaines n'est qu'une première étape. A terme, on peut imaginer un système qui se généralise à de nombreuses chaînes Youtube.

Sur le fonctionnement de la technologie et du classement des résultats, Google dit simplement qu'il utilise un outil de reconnaissance du langage parlé et que les résultats sont classés d'après le contenu audio, les méta données et la fraîcheur.

A noter qu'il existe déjà un groupe d'entraide consacré à l'indexation audio : <http://groups.google.com/group/google-audio-indexing>

Parmi les remarques et les bugs relevées par les utilisateurs, on peut relever une sorte de bêtisier : il arrive souvent que Google comprenne mal les mots employés par les orateurs !ce qui génère des lapsus assez amusants (voir <http://groups.google.com/group/google-audio-indexing/t/4d8708ec713ebadd>). La technologie n'est pas encore très au point...

Blinkx

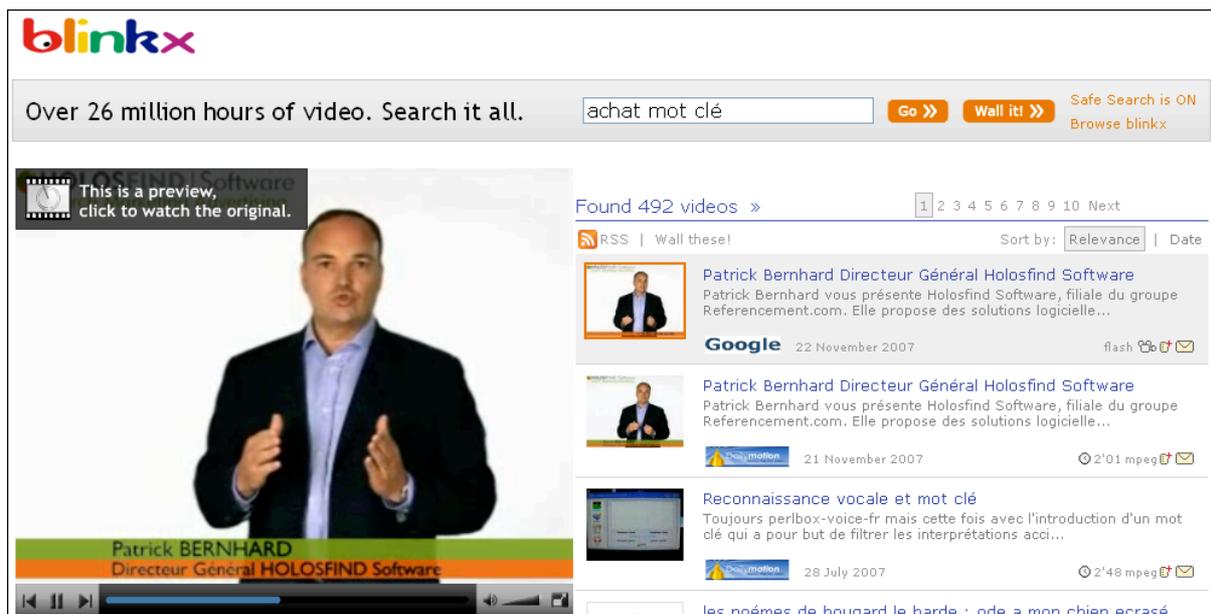
Google n'est pas le seul à avoir pensé à la reconnaissance vocale, loin de là : le portail vidéo Blinkx (<http://www.blinkx.com/>) s'y est intéressé dès 2004.

Récompensé en 2008 par le magazine Speech Technology, ce pionnier de la recherche vidéo utilise en effet sa propre technologie pour rechercher des expressions clés dans n'importe quelle langue (voir à ce sujet un article sur 01net : <http://www.01net.com/editorial/285297/blinkx.tv-le-moteur-de-recherche-qui-prend-tele-et-radio-aux-mots/>)

Par exemple la requête "référencement google" met en avant une vidéo où un artisan apporte son témoignage sur internet et sur la façon dont il s'est positionné dans Google.

The screenshot shows the Blinkx search interface. At the top, the Blinkx logo is displayed. Below it, a search bar contains the text "référencement google". To the right of the search bar are buttons for "Go", "Wall it", and "Safe Search is ON". Below the search bar, a video player is visible, showing a man speaking. The video title is "Jean Ménard à Nueil-les-Aubiers". The video description reads: "En moins de dix minutes, un superbe cours d'e-business par Jean Ménard, réparateur et vendeur en cycles et motoculture à Nueil-les-Aubiers. Avec humour, il raconte comment il a...". To the right of the video player, there are search results for "référencement google". The first result is "Jean Ménard à Nueil-les-Aubiers" with a duration of 12'24 mpeg. The second result is "sillaz parle moi" with a duration of 3'36 flash. The third result is "Commentaire du 08/07/2007, la guerre Mac - PC est terminée..." with a duration of 32'17 mpeg.

Autre exemple, la requête "achat mot clé" fait ressortir une vidéo d'explication en français sur le search marketing.



Blinkx propose donc une technologie qui a fait ses preuves et qui semble assez efficace... D'ailleurs, la société a noué des partenariats avec de nombreux diffuseurs tels que CNN ou la BBC et également avec Microsoft (voir article sur Abondance : <http://actu.abondance.com/2006-41/microsoft-blinkx.php>)

Le point fort de Blinkx est sans contestation sa capacité à interpréter différentes langues (Blinkx gère l'anglais, l'allemand, le français et l'espagnol) et à trouver des vidéos basées non pas sur un mot clé mais sur une thématique, en explorant le champ sémantique.

Podscope / TVEyes

Podscope, disponible à l'adresse <http://www.podscope.com/> s'intéresse plus particulièrement aux flux audio présents dans les podcasts.

Lancé en 2005 par TVEyes, il a fait preuve de son efficacité et a été retenu par AOL pour son portail radio <http://music.aol.com/radioguide/podcasting> ainsi que par la société Evoca (<http://www.evoca.com>) consacrée à l'animation audio de sites web (technique dite de *phone-to-web*)

A noter que TVEyes connaît bien son affaire puisque cette société indexe les émissions radio et TV anglophones depuis 1999. (voir interview de David Ives de TVEyes : http://www.masternewmedia.org/audio_search/audio_search_engines/Podscope_and_TVEyes_search_audio_video_content_20051005.htm)

Podscope se présente comme un moteur de recherche classique; une liste d'émissions radio renfermant le mot clé est proposée à l'internaute, avec des marqueurs temporels. Par exemple sur la requête "spielberg" on obtient une sélection d'émission où les projets du cinéaste sont évoqués.

The screenshot shows the Podscope search results for the query "spielberg". At the top, there is a search bar with the text "podscope" and "spielberg" entered, and a "Search" button. Below the search bar, a green banner displays "Podscope Search Results for 'spielberg'" and "Sort Order: score, date". The results are organized into two main sections:

- NPR: 09-28-2007 Movies (8 Hits)**
Source: NPR: Movies
Stories: 1) Coming Soon, to a Theater Near You: Naked People 2) 'The Kingdom' Mirrors the Headlines 3) 'The Darjeeling Limited': Self-Discovery, by Rail 4) Will Paramount and DreamWorks Really Split? 5) Filming 'The Jane Austen Book Club'
A timeline on the left shows time markers: 00:14:01, 00:13:31, 00:13:37, 00:14:01, and 00:15:34. Below the timeline are buttons for "SITE", "LINK", "PLAY", and "XHL".
- Thursday 14/02/2008 (7 Hits)**
Source: Asia Pacific Podcast
AUST: Review of government support for automotive industry | ETIMOR: Emotions run high at Alfredo Reinado's funeral | AUSTRALIA: PM prepares to travel to East Timor | ETIMOR: Reinado's petitioners fear being targeted | INDONESIA: Demonstrators call for dissolution of Papuan Council | CHINA:...
A timeline on the left shows time markers: 00:19:24, 00:19:30, 00:20:35, and 00:21:07. Below the timeline are buttons for "SITE", "LINK", "PLAY", and "XHL".

At the bottom right of each section, there are links for "DiggIt!" and "Del.icio.us".

En théorie, la technologie peut s'appliquer à de nombreux domaines, et pas seulement aux flux audio. Ainsi en avril 2008, TVEyes a annoncé que Podscope pouvait traiter les vidéos, ce qui va ouvrir de vastes possibilités en matière d'indexation. Il est fort possible que la technologie soit bientôt utilisée sur un portail vidéo, qui deviendrait alors un bon concurrent de Blinkx et Youtube.

L'avenir du référencement audio

Nous venons de voir 3 technologies intéressantes en terme de reconnaissance audio et de transcription textuelle. Il en existe d'autres... Que faut-il en retirer pour le référencement ? Faudra-t'il bientôt optimiser une vidéo comme on le fait pour un contenu textuel ?

Pour le moment il n'existe aucun guide permettant de faire de l'optimisation audio mais quelques tests sur les outils vont permettre de faire ressortir des principes de base.

Tout d'abord, il est certain que la qualité de la bande son et l'absence de "bruit de fond" seront primordiales : plus le son est intelligible et la voix reconnaissable (imaginez un système de reconnaissance vocale tentant de comprendre ce que dit une personne située à côté d'un marteau-piqueur...), meilleures seront les chances de voir le contenu analysé avec efficacité... Pensez-y au moment de l'enregistrement de vos vidéos...

1. Occurrence de mots clés

GAUDI est clairement basé sur les occurrences de mots clés dans la vidéo, comme le montre l'étude de plusieurs résultats de recherche.

La requête "auto industry" affiche d'abord un résultat qui comprend 4 occurrences de l'expression (donc 4 marqueurs), puis un résultat avec 3 occurrences du mot clé, et ainsi de suite, jusqu'à afficher des mots clés avec "industry" et "auto" séparés. Il est clair que Google favorise tout simplement les résultats qui renferment le plus d'occurrence de mots clés.

The screenshot shows a Google Audio Indexing search for "auto industry". The search bar contains "auto industry" and the search button is labeled "Search videos". Below the search bar, there are several video results listed under "All Politicians | McCain | Obama". The top result is "Barack on the Economy in Michigan" with a thumbnail showing Barack Obama in a white SUV. To the right of the search results is a video player for the top result, titled "Barack on the Economy in Michigan". The video player shows a play button and a progress bar. Below the video player, there is a search bar with "auto industry" and a "Search inside this video" button. Below that, there are four video segments with their start times (all "min 0") and snippets of text containing the search term "auto industry".

Le premier résultat contient 4 occurrences de mots clés

Il n'est pas pour autant vérifié que la vidéo renfermant le plus d'occurrence de mots clés soit placée systématiquement en première position : pour la requête "nuclear weapon", le 1er résultat renferme 4 occurrences de mots clés, alors que le 6e résultat renferme pas moins de 9 occurrences.

En fait c'est tout simplement l'ancienneté qui détermine l'affichage des résultats. La vidéo la plus récente est placée en première position (et pour le moment ce système de classement n'est pas paramétrable alors qu'on peut le faire dans Youtube).

The screenshot shows a Google Audio Indexing search for "nuclear weapon". The search bar contains "nuclear weapon" and the search button is labeled "Search videos". Below the search bar, there are several video results listed under "All Politicians | McCain | Obama". The top result is "Joe Biden Wooster Ohio" with a thumbnail showing Joe Biden at a campaign event. To the right of the search results is a video player for the top result, titled "Joe Biden Wooster Ohio". The video player shows a play button and a progress bar. Below the video player, there is a search bar with "nuclear weapon" and a "Search inside this video" button. Below that, there are four video segments with their start times (min 3, min 4, min 6, min 6) and snippets of text containing the search term "nuclear weapon".

Sur ce test, le nombre d'occurrence n'est pas lié au classement

2. Thématique du contenu

La technologie utilisée par Blinkx est un peu plus subtile que celle de Google : en effet ce moteur ne prend pas seulement en compte les occurrences de mots clés mais évalue la thématique de la vidéo et la présence de mots clés connexes.

Par exemple, la première vidéo affichée sur la requête "auto industry" est un reportage sur l'industrie automobile en Thaïlande, et on y trouve des expressions comme "trucks manufacturing", "car manufacture", "automobile industry", qui sont fortement liées à la requête.

The screenshot shows the Blinkx search interface. At the top, it says "Over 26 million hours of video. Search it all." with a search bar containing "auto industry" and buttons for "Go" and "Wall it!". Below the search bar, there's a video player showing a car on a street. To the right of the video player, there's a list of search results. The first result is "Thailand: The Future of the Auto Industry?" from NYJP, dated 27 March 2008. The second result is "China's thriving auto industry -- 22 April 2008" from YouTube, dated 22 April 2008. The third result is "Auto Industry Outlook" from Kiplinger.com, dated 1 May 2007. The fourth result is "Global Coverage: The auto industry" from Reuters, dated 27 February 2007. The fifth result is "Auto industry hopes to secure \$50 billion in govt. loans" from NECN, dated 27 February 2007.

C'est donc le champ sémantique qui doit être privilégié pour un contenu audio de qualité, du moins pour les moteurs "intelligents" qui sont capables de reconnaître de nombreux mots clés connexes.

On peut imaginer qu'à terme, les vidéos renfermant du spam de mots clés risquent d'être moins bien notées, voire sanctionnées par rapport à des vidéos utilisant le fameux langage naturel cher à Google.

3. Optimisation du discours

Il est tout d'abord évident qu'un mot clé sera d'autant mieux pris en compte qu'il est facilement intelligible dans la vidéo. Si un mot ou une expression est mal comprise, il y a peu de chances d'obtenir des positions !

On ne connaît pas le fonctionnement des logiciels d'analyse, mais il est probable qu'il faut privilégier des expressions bien articulées, avec la prononciation adéquate, et pourquoi pas, un changement de ton permettant de mettre en relief telle ou telle expression. La gestion du phrasé aura sans doute le même effet que les balises <h1> ou dans un texte web...

Ce qui est sûr c'est qu'il faudra s'adapter au fonctionnement des logiciels de reconnaissance vocale. Les figures de style et jeux de mots risquent de ne pas être bien compris par les robots analyseurs, et que se passera-t'il si un orateur possède un accent prononcé ?

4. L'internaute aura-t'il le dernier mot ?

Avant de se focaliser sur l'aspect purement technique, il faut aussi penser aux internautes. Ce qui fait le succès d'une vidéo dans Youtube ce n'est pas sa qualité technique mais plutôt son originalité et la façon dont elle interpelle l'internaute.

On peut très bien imaginer un classement basé sur le comportement des internautes qui viendrait compléter l'optimisation technique d'une vidéo. Des portails comme Dailymotion ont déjà mis en place un système de vote, qui permet aux internautes d'intervenir eux-même dans le classement.



Dans ce cadre il ne faut pas imaginer l'indexation audio comme un nouveau système de classement à part entière mais plutôt comme un outil permettant de classer et identifier les vidéos. Ce sera ensuite l'internaute qui prendra le relais.

Conclusion

En définitive il est vraisemblable que l'on va se retrouver avec la même approche que pour une page web : proposer un contenu compréhensible pour les moteurs, optimisé en terme de mots clés et accrocheur pour l'internaute.

Les méthodes ne changeront pas mais la modification du support fera intervenir de nouveaux spécialistes, tels que des ingénieurs sons, des comédiens, des réalisateurs... Le référencement a encore de beaux jours devant lui !

Guillaume Thavaud, cellule de veille (veille@brioude-internet.fr) de la société Brioude-Internet Référencement (<http://www.referencement-2000.com/>)

Réagissez à cet article sur le blog des abonnés d'Abondance :
<http://abonnes.abondance.com/blogpro/2008/10/le-rfrencement-audio-une-strategie.html>