

David Botella (Orange) : "Un moteur de recherche est une somme de challenges techniques à relever..."

[Retour au sommaire de la lettre](#)

Domaine :	Recherche	Référencement
Niveau :	Pour tous	Avancé

David Botella est directeur des projets marketing du moteur de recherche d'Orange, déployé sur le portail de la marque ainsi que sur les sites Voila et Lemoteur.fr. Il fait un point, au travers de cette interview pour notre lettre, sur les projets actuels et à venir de cet outil, ainsi que sur les partenaires actuels d'Orange dans la recherche d'information sur le Web...

David Botella s'occupe depuis plusieurs années des projets de "search" chez France Telecom puis Orange. Nous lui avons posé plusieurs questions sur son métier et sur les projets d'Orange en termes de moteurs de recherche sur le Web. Voici ses réponses...

- David Botella, merci de répondre à nos questions. Pouvez-vous présenter votre parcours professionnel et expliquer à nos lecteurs votre rôle dans le cadre du moteur de recherche d'Orange ?

Bonjour. Je suis entré chez France Telecom en 2000 lors de la mise en place de la régie publicitaire Internet de France Telecom. En 2002, nous avons lancé la première offre de liens sponsorisés sur les listes-réponses de Wanadoo et Voila et quand la commercialisation a été confiée à Overture (racheté par la suite par Yahoo!), j'ai pris en charge la direction de projets marketing du moteur. Avec le développement du marché des liens sponsorisés, il a été possible de redonner de l'ambition avec des moyens bien plus conséquents. Mon rôle est de définir avec les équipes techniques les sujets prioritaires d'amélioration de la pertinence du moteur et de développement de son audience.



- Aujourd'hui, où en est le moteur de recherche d'Orange / Voila / Lemoteur.fr ? Pouvez-vous nous le présenter en chiffres par exemple ?

Selon les derniers chiffres de Nielsen NetRatings, nous recevons aujourd'hui 6,5 millions de personnes par mois, soit 1 internaute sur 5. Au niveau du moteur lui-même, nous exploitons plus d'un milliard de documents francophones, mais nous en connaissons beaucoup plus. Je pourrais dire que nous en avons 4, 8 ou 10 milliards trouvées sur internet, mais le chiffre le plus gros n'est pas le plus représentatif. Pour moi, l'idéal est d'avoir à n'indexer que le web "utile" pour nos internautes, c'est-à-dire les pages où ils trouvent réponse à leur question. Et bien sûr de le proposer dans les premières réponses ce qui représente le challenge quotidien de nos ingénieurs.

- La technologie "France" est-elle toujours maintenue en interne (technologie "KE") ? Combien de personnes travaillent-elles aujourd'hui dessus ? Le moteur est-il "internalisé" à 100% ? Son mode de fonctionnement est-il similaire à celui de Google, Bing ou autres leaders du marché ?

Oui, la technologie est entièrement opérée en interne pour la partie France. En nombre de personnes qui travaillent dessus, c'est environ 70 personnes, en partant des documentalistes jusqu'à l'hébergement et l'exploitation. Nous avons un large éventail de compétences...

Pour l'externalisation, nous avons souvent essayé, mais il est toujours apparu que la solution la plus opérationnelle et économiquement efficiente était de garder la compétence en interne. Externaliser est intéressant quand vous partagez des coûts de R&D ou d'hébergement avec d'autres acteurs de taille équivalente (ou supérieure) ce qui n'existe pas à ma connaissance à notre niveau. Nous utilisons par contre des bibliothèques issues de la communauté du Libre, notre plateforme étant entièrement sous Linux.

Son fonctionnement est similaire dans les grandes lignes à tout moteur (mais qui connaît vraiment le fonctionnement exact de Google ou Bing ? ;-). Mais en fait, il faut plus voir l'activité comme une somme de challenges techniques à relever avec des moyens spécifiques à mettre en oeuvre pour les résoudre. Comment faire par exemple pour mettre à jour efficacement notre connaissance d'un site qui comporte plusieurs millions de pages (un forum par exemple) quand le site n'accepte pas plus d'une visite de notre robot par seconde ? Nous avons beaucoup de "petits" problèmes de ce genre à résoudre ! Notre approche ciblée du web francophone nous permet d'étudier des solutions *ad hoc* moins coûteuses qu'une approche globale du web mondial complet.

- Le partenaire "monde" du moteur est-il toujours Ask.com ? Pourquoi ce choix ?

Nous ne faisons pas l'anglophone et cela nous permet de proposer des réponses aux internautes qui souhaitent en avoir sur Orange. Les requêtes restent assez ciblées et peu pointues pour le côté anglophone. Le moteur Ask nous permet d'y répondre de façon efficace et intéressante. C'est aussi une façon de continuer à proposer des alternatives à d'autres moteurs que tout le monde connaît déjà ;-)

- Qui est également votre prestataire en termes de liens sponsorisés dans les SERPs des différents portails de "search" (Orange.fr, lemoteur.fr, voila.fr) ?

C'est Google qui commercialise et nous fournit les liens sponsorisés sur nos listes réponses, comme sur tous les portails européens d'Orange (Pologne exceptée). Nous étudions aussi une offre de mise en avant à destination des marques qui souhaitent mieux se présenter lorsque l'internaute les cherche explicitement.

- Quelles sont les dernières innovations intégrées sur le moteur ?

Nous avons sorti récemment un système de questions / réponses sur les données géographiques françaises basé sur les informations de Wikipedia et qui est presque incollable sur ces sujets, c'est impressionnant ! Demandez lui le nombre d'habitants en France (http://www.lemoteur.fr/S/voila?profil=lemoteur&bhv=web_fr&rtype=kw&rdata=nombre%20d'habitants%20en%20france), le maire d'Heiligenstein (http://www.lemoteur.fr/S/voila?profil=lemoteur&bhv=web_fr&rtype=kw&rdata=qui%20est%20le%20maire%20d'Heiligenstein), le président de la région Rhone Alpes (http://www.lemoteur.fr/S/voila?profil=lemoteur&bhv=web_fr&rtype=kw&rdata=president%20de%20la%20r%20E9gion%20Rhone%20Alpes), et le moteur répond directement. Nous avons aussi sortis précédemment dans la même logique d'interprétation de la requête, des moteurs thématiques exploités sur Voila et lemoteur.fr qui "comprennent" quand vous cherchez une recette de cuisine par exemple et qui vont pousser les résultats issus des sites de la thématique cuisine en premier.

Il y a d'autres évolutions primordiales qui ne sont pas forcément visibles ou matérialisée directement sur la liste-réponse. Il est parfois nécessaire d'aller chercher des méthodes très pointues pour maintenir une qualité de service optimale. Ces efforts ne sont pas directement vus par l'internaute, même si le résultat final se trouve dans la pertinence des réponses.

- Quels sont les projets à court et moyen termes sur ce moteur ?

La prochaine sortie est le moteur video qui est déjà disponible en version beta sur lemoteur.fr et qui indexe plus de 300 000 vidéos. Nous sommes en cours d'intégration de nouveaux partenaires afin de couvrir un maximum de sujets. Une prochaine version devrait sortir dans les semaines à venir sur orange.fr. Notre positionnement est de cibler les contenus à la fois de qualité et plebiscités par nos internautes. Nous ferons aussi un lien vers les programmes TV, les programmes disponibles en VOD et sur web Tv pour nos abonnés. Si vous cherchez "michael jackson", vous aurez ainsi les clips, les documentaires d'archives et les émissions

programmées sur lui qui passent à la TV dans les jours suivants. Une fois de plus, notre objectif est d'avoir les vidéos qui vont intéresser les internautes au quotidien, et en français.

A moyen terme, nous pensons que l'interprétation de la requête de l'internaute et une présentation d'une liste réponse adaptée à celle-ci sont vraiment un axe à approfondir. On n'a pas à présenter la même réponse à la requête "pôle emploi" que "telecharger driver carte graphique" ou même "pb avec driver de carte graphique". La difficulté étant d'arriver à faire supporter la rupture du modèle de liste-réponse standardisé actuel. Nous pensons que c'est un axe sur lequel il devrait être pourtant simple de se différencier, avec une mise en page agréable, riche, synthétique et qui se charge instantanément, car cela reste un pré-requis pour les internautes.

- Quels sont les champs d'action à améliorer en urgence sur le moteur selon vous ?

Nous savons que nous sommes très perfectibles sur les requêtes à plusieurs mots (plus de 3-4 mots), alors qu'il devrait être plus simple d'y répondre. Nous travaillons, bien sûr, sur le sujet, et visons des évolutions en ligne d'ici la fin du semestre. De plus, le nombre croissant de moteurs spécialisés que nous opérons (vidéos, actualités, l'annuaire avec le 118712, et avec les contenus de nos portails) va nous aider à formuler des réponses directes qui ne sont pas de simples URL mais des réponses enrichies. Nous ne nous demandons plus comment remonter des sites web qui proposent d'écouter le dernier album d'un artiste, mais comment le proposer directement en écoute sur le moteur.

- Finalement, pourquoi maintenir une technologie propre à Orange et ne pas travailler avec Google ou Bing, voire Exalead, par exemple ?

Nous ne voyons pas l'intérêt d'une simple liste réponse aux couleurs d'Orange ou de Voila signée par Google ou Bing qui proposerait quelque chose de moins bien. Pour Exalead ou d'autres fournisseurs de technologies de recherche, il n'y a pas forcément une équation économique viable derrière. Ne serait-ce que sur la partie hébergement, notre plateforme est "sur-mesure" et elle est parfaitement adaptée à notre tenue en charge. Ce sont chez nous des équipes dédiées qui exploitent et maintiennent la plateforme, qui est de plus capitalisée avec le portail mobile et d'autres services pour lesquels nous assurons aussi un service de recherche (ex: l'annuaire 118712 sur Orange ou la recherche chez d'autres sites du groupe comme www.tendance-sante.fr). De façon moins matérielle, développer des compétences en recherche d'information me semble stratégique pour tout acteur qui veut perdurer sur Internet, même en dehors d'un moteur de recherche "web" car l'information en est bien la matière première.

- Y a-t-il encore de la place pour un acteur "indépendant" dans le domaine du search en France et dans le monde ?

A mon avis, la question est de savoir comment un acteur "indépendant" avec une solution très innovante peut se financer tout en restant "indépendant". Car les moyens techniques à mettre en oeuvre sont conséquents, et le deviennent de plus en plus dès qu'il faut faire face au succès. Le succès arrivant, il faut monétiser et donc soit être aussi indépendant sur ses revenus en les générant avec sa propre régie, soit se tourner vers le quasi-monopole actuel... Et avoir comme seule source de revenus le leader de son marché, ça complique vite les choses pour bien maîtriser sa croissance.

- En quoi un moteur de recherche peut-il aujourd'hui se démarquer des "gros" comme Google, Yahoo! ou Bing ?

On peut se démarquer en ayant plus de proximité avec les internautes et leurs besoins au quotidien, et sans forcément faire de l'intrusion dans leur vie privée. Notre idée est d'inscrire l'utilisation du moteur de recherche dans l'action quotidienne en regroupant en un seul endroit différentes sources. Si vous avez besoin de renouveler votre passeport, on peut imaginer une page qui vous fournit l'adresse où se rendre avec les horaires d'ouverture (avec le 118712), les formulaires à imprimer et à remplir (via le moteur web), etc. Cela peut venir compléter d'autres opportunités de différenciation sur le mobile car rien de très différent d'un portage du web internet vers une interface mobile existe. Et cela représente une croissance d'usage importante dans les mois à venir.

- Pour finir, une question que je ne vous aurais pas posée ? :-)

Oui : "Comment suivre au mieux les prochaines améliorations et services du moteur Orange ?"

Je sais que nous ne communiquons pas beaucoup coté *Search* sur Orange sur nos ambitions et réalisations. Je vous invite à suivre le twitter de notre équipe (<http://twitter.com/lemoteur>) où nous communiquerons les nouveautés désormais, de façon plus soutenue. Nous communiquons aussi sur dreamorange (<http://www.orange-innovation.tv/dreamorange/?p=Portal/searching&JContentOid=6&docType=13>).

Merci, David Botella, pour vos réponses...

Interview réalisée par mail par Olivier Andrieu, éditeur du site Abondance.com, en mars 2010.

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

<http://blog-abonnes.abondance.com/2010/03/david-botella-orange-un-moteur-de.html>