



# Up2news.tv, le moteur de recherche audiovisuel

Aurélie Vathonne

Up2news.tv est un moteur de recherche permettant d'effectuer des recherches textuelles sur un contenu audiovisuel.

Ce service innovant, apparu dès 2004, a été relancé récemment à l'occasion de l'élargissement de la couverture aux principales sources audiovisuelles françaises.

Ainsi, le système up2news.tv permet d'effectuer une recherche par mot-clé sur le contenu prononcé à l'antenne de 20 chaînes et 8 radios françaises, incluant toutes les chaînes hertziennes et la plupart des chaînes de la TNT.

Le service fonctionne théoriquement avec les principaux navigateurs (Internet Explorer, Firefox, Safari) mais lors de nos tests, l'utilisation de Firefox a été impossible dans notre configuration réseau, même après avoir téléchargé le plug-in spécifique. Nous avons donc effectué les tests à partir d'Internet Explorer. Les fichiers vidéos étant lisibles au format Windows Media Player, le système fonctionne d'ailleurs de façon beaucoup plus fluide en utilisant Internet Explorer, même lorsque l'utilisation de Firefox est par ailleurs rendue possible.

Une fois connecté au service, l'utilisateur découvre une interface assez simple, comportant essentiellement à gauche les vignettes correspondant aux différentes sources et à droite le lecteur de fichiers vidéos.

Une boîte de recherche – très "googlienne" – située en haut et au centre de l'écran, permet d'entrer un ou plusieurs mots ou expressions (à saisir sans guillemets). Nous avons effectué des tests sur des morceaux de phrase assez longs, tels que "*mieux vaut se tourner vers les valeurs sûres*", qui fonctionnent sans problème.

Les mots ou expressions peuvent être combinés à l'aide des opérateurs booléens ET, OU (forme française), qu'il faut entrer impérativement en majuscules. Le corpus de recherche étant un flux continu d'information, on considère comme unité de "document" tout extrait de 2 minutes. Autrement dit, l'opérateur ET permet de rechercher deux mots ou deux expressions prononcés dans un intervalle de 2 minutes. La recherche s'effectue par défaut sur les 7 derniers jours.

Une fois la requête lancée, le nombre et la liste des résultats s'affiche par ordre anté-chronologique dans une colonne à

gauche : il s'agit de tous les extraits, d'une durée d'une minute, contenant les mots ou expressions saisis.

Pour chaque réponse de la liste, up2news restitue :

- une vignette correspondant soit à la première image de l'extrait (pour les réponses issues de la télévision), soit à la source (pour les réponses issues de la radio) ;
- la date et l'heure de diffusion de l'extrait ;
- un extrait du script (texte) ;
- le logo de la source (chaîne de télévision ou radio).

L'extrait textuel du script est un élément précieux, qui permet d'écartier certaines réponses visiblement hors-sujet.



## Bases

28.09.2009

En cliquant sur l'un des résultats par le biais de la vignette, l'utilisateur charge l'extrait sélectionné dans le lecteur à droite de l'écran.

Si le temps de recherche à l'étape précédente est très rapide, celui correspondant au chargement de l'extrait vidéo peut être beaucoup plus long, voire se bloquer dans le pire des cas, mais un accès haut débit professionnel couplé à l'utilisation d'Internet Explorer diminue ce risque de blocage.

A droite de la fenêtre de lecture, on retrouvera des détails sur l'extrait qu'on a sélectionné (notamment le nom de la chaîne, la date et l'heure de diffusion), mais également des informations supplémentaires telles que le nom de l'émission d'où est tiré l'extrait, le nom de son présentateur, la durée de l'émission et sa thématique.

Le lecteur procède ensuite à la recherche de la citation et l'on voit visuellement la progression de cette opération à l'écran, puis la lecture s'enclenche automatiquement.

Des flèches "précédent" et "suivant", situées de part et d'autre de la barre de lecture, permettent de visualiser la tranche de une minute appartenant au même programme qui suit ou précède l'extrait initial.

Pour des raisons de droits de rediffusion, cette fonctionnalité ne peut être répétée plus de 2

fois, afin de ne pas autoriser la visualisation intégrale d'un programme, fut-ce par tranche d'une minute.

Il est également possible de télécharger l'extrait vidéo sur son disque dur. Mais ce n'est pas tout : l'utilisateur a la possibilité de créer une ou plusieurs alertes, pour être prévenu par e-mail dès que le ou les mots mis sous surveillance sont prononcés à l'antenne. Nous avons testé cette fonctionnalité qui relève quasiment du temps réel, à quelques minutes près.

Le service reposant sur les technologies de reconnaissance vocale, sa fiabilité est déterminée en grande partie par la qualité de l'environnement sonore du programme audiovisuel : la reconnaissance sera excellente si le flux est prononcé en studio par un journaliste, tandis qu'il sera médiocre s'il existe un bruit de fond ou si l'interlocuteur articule mal, possède un fort accent, etc.

Cette approximation se constate parfois dans les extraits de scripts, qui relèvent alors du charabia : "...ste dentition les combinaisons chanel qui s'est déroulé vous voyez sous forme de joël façades américain...". Mais globalement, le taux de reconnaissance oscille autour des 80 %.

De plus, le système doit être alimenté régulièrement par les nouveaux noms qui font

l'actualité, et qu'il doit apprendre à reconnaître. Cette mise à jour, qui est effectuée très régulièrement, concerne exclusivement les noms propres, notamment les noms de personnalités, y compris dans leurs diminutifs (exemple : *Sarko* pour *Sarkozy*).

Enfin, un petit graphique en bas à droite de l'écran donne accès à des analyses quantitatives, qui sont effectuées sur un mois glissant.

Les tarifs sont établis en fonction du nombre d'utilisateurs, des fonctionnalités offertes et du nombre de chaînes sélectionnées.

Toutefois, des ordres de grandeur de prix nous ont été délivrés :

- accès illimité : environ 500 € HT/mois pour un utilisateur :

- paiement à l'acte : 10 à 20 € HT/requête

Malgré les imperfections

inhérentes à la nature du service, Up2news.tv nous a séduits.

Il rend accessible sur le Web, et à un large public, un contenu informationnel qui était difficilement exploitable auparavant, faute d'indexation.

A l'instar des agrégateurs de presse écrite, qui permettent depuis des années de rechercher simultanément dans des milliers de publications, Up2news.tv étend le champ de la recherche au contenu audiovisuel, et permet aux professionnels de l'information d'élargir leurs sources.

### Informations

[www.up2news.com](http://www.up2news.com)

55, rue Aristide BRIAND

92309 Levallois Cedex

tél. 01 71 11 32 11

Fax. 01 71 11 32 12

[info@up2news.com](mailto:info@up2news.com)

contact : Jean-Dominique

Quien

### Un futur concurrent gratuit ?

On s'intéressera aussi de près au service similaire Voxlead, proposé depuis quelque temps en version bêta dans les laboratoires du moteur de recherche Exalead.

L'interface, très ergonomique, propose une boîte de recherche ainsi qu'un nuage de tags permettant d'embrasser rapidement tout ce qui fait l'actualité : personnalités, thèmes... Le service fonctionne pour le moment en français, anglais, chinois et arabe. Les sources disponibles sont encore peu nombreuses, mais contiennent notamment France24, i-télé, LCI, BFM.tv...

<http://voxleadnews.labs.exalead.com/>