

## COMMENT TROUVER DES SOURCES D'INFORMATION AU FORMAT RSS

Posted on 15 janvier 2021

**On entend souvent dire que les fils RSS sont morts et qu'ils ne servent plus à rien, qu'ils ont disparu de la surface du Web. Mais ceux qui disent cela ne savent peut-être pas bien les utiliser et les chercher, car les flux RSS sont bien présents à l'heure actuelle sur la Toile et peuvent s'avérer extrêmement intéressants à utiliser. Voici comment les identifier au mieux dans une démarche de veille.**



Les flux RSS ont été pendant longtemps faciles à trouver puisqu'ils étaient par défaut présents sur chaque blog. Une simple recherche dans Technorati, Blog Pulse ou Google Blogs permettait de repérer des blogs thématiques et de s'y abonner, mais, d'une part, le phénomène des blogs s'est tassé et d'autre part, beaucoup de ceux qu'on appelait à l'époque "blogs" se sont transformés en sites web traditionnels, mettant de côté les spécificités qui les en distinguaient initialement : présentation par ordre antéchronologique, possibilité de commentaires, trackbacks, blogroll et... flux RSS ! On aurait alors pu penser que les flux allaient disparaître, d'autant qu'à partir de 2015, les GAFAM les avaient presque totalement banni de leurs plateformes. Ce ne fût pourtant pas le cas et l'une des raisons de leur résistance est liée à la transformation d'un outil très populaire de création de blogs - Wordpress - en un CMS capable de créer tout type de sites web. Or Wordpress dispose nativement de très nombreux flux RSS ([plus d'une trentaine](#)) et il est également le CMS le plus utilisé dans le monde. Des chiffres publiés très récemment par le cabinet d'études W3techs indiquent qu'[il représente 40% des sites créés](#), loin devant Shopify (3%), Joomla (2,2 %) ou Drupal (- de 2%). Une information intéressante dans le contexte de la veille puisqu'elle nous assure qu'*a minima*, 40% des sites qu'on découvrira lors d'une phase de sourcing pourront être surveillés *via* leur flux RSS.

### Comment identifier les sites sous Wordpress ?

La démarche consiste dans un premier temps à identifier si le site que l'on souhaite surveiller a été bâti avec Wordpress. Rien de plus simple pour cela. Il suffit d'installer l'extension [Wappalyzer](#) et de cliquer dessus lorsqu'on est sur la page cible :



### *Détection du CMS de création du site Réacteur avec Wappalyzer*

Il ne reste plus ensuite qu'à repérer le flux dans la page et à l'ajouter à son agrégateur. Des extensions d'aide à la détection de flux RSS existent d'ailleurs afin d'aller plus vite :

Pour Chrome :

- [Extension abonnement RSS](#)

Pour Firefox :

- [WantMyRSS](#)

Il est également possible d'utiliser le modèle de structure de flux (cf. article sur le site [coderevolution.com](http://coderevolution.com) cité ci-dessus) pour les « recréer » de manière plus précise directement dans son agrégateur. Si par exemple vous souhaitez obtenir le flux de l'ensemble des commentaires déposés sur le site Réacteur, vous pouvez créer votre flux ainsi :

<https://www.reacteur.com/comments/feed>

AB

## <https://www.reacteur.com/comments/feed>

1 flux trouvé



### Commentaires pour Réacteur

Magazine de veille SEO et Moteurs de recherche

0 subscribers

0 article/semaine

M'abonner



Commentaires sur Rédiger une fiche produit efficace : votre plan d'action en 5...

4w



Commentaires sur les outils de création de champ lexicaux

4w

*Création d'un flux spécifique aux commentaires d'un blog Wordpress dans l'agrégateur Inoreader*

### Les annuaires de blogs ou de flux RSS

Ceci étant dit, il existe d'autres moyens d'identifier des flux RSS ou, *a minima*, les blogs susceptibles de les accueillir. Il s'agit des (antiques) annuaires qui n'ont pas totalement disparu et sont encore actifs sur une niche liée au référencement. Notre objectif sera ici d'obtenir soit :

- Une liste de blogs produisant du contenu sur une thématique spécifique ;
- Une liste de flux RSS produisant du contenu sur une thématique spécifique.

Ce qui exclut les services qui agrègent eux-mêmes des flux et permettent ensuite de rechercher des articles par mots-clés sans donner accès à leur liste de flux. En effet, même s'ils peuvent être utiles, ils sont chronophages dans une phase de sourcing car ils impliquent de remonter à chaque fois au blog pour en déduire s'il est pertinent ou non sur notre thématique.

**Blogarama**

Blog Directory Submit Your Blog Blogging Jobs Contact Us

Search

## "fuel Cell" Blog Posts

**Quality Stocks News Breaks – MGX Minerals Inc. (CSE: XMG) (FKT: 1MG) (OTCQB: MGXMF) Subsidiary Receives Notice Of Publication For Metal-Air Fuel Cell Patent**  
Monaker Group, Inc. · 14:39 07 Feb 2018  
Diversified resource and technology company MGX Minerals Inc. (CSE: XMG) (FKT: 1MG) (OTCQB: MGXMF) this morning announced that its wholly-owned subsidiary, ZincNyx Energy Solutions, Inc., h... [Read More](#)

**Hydrogen And Fuel Cell History In Videos**  
Hydrogen Cars Fuel C... · 18:17 27 Sep 2017  
Now we are talking about Hydrogen and Fuel Cell FUTURE: "China targets 1 million hydrogen fuel cell electric vehicles by 2030" — Nasdaq HydrogenCarsNow.com and this... [Read More](#)

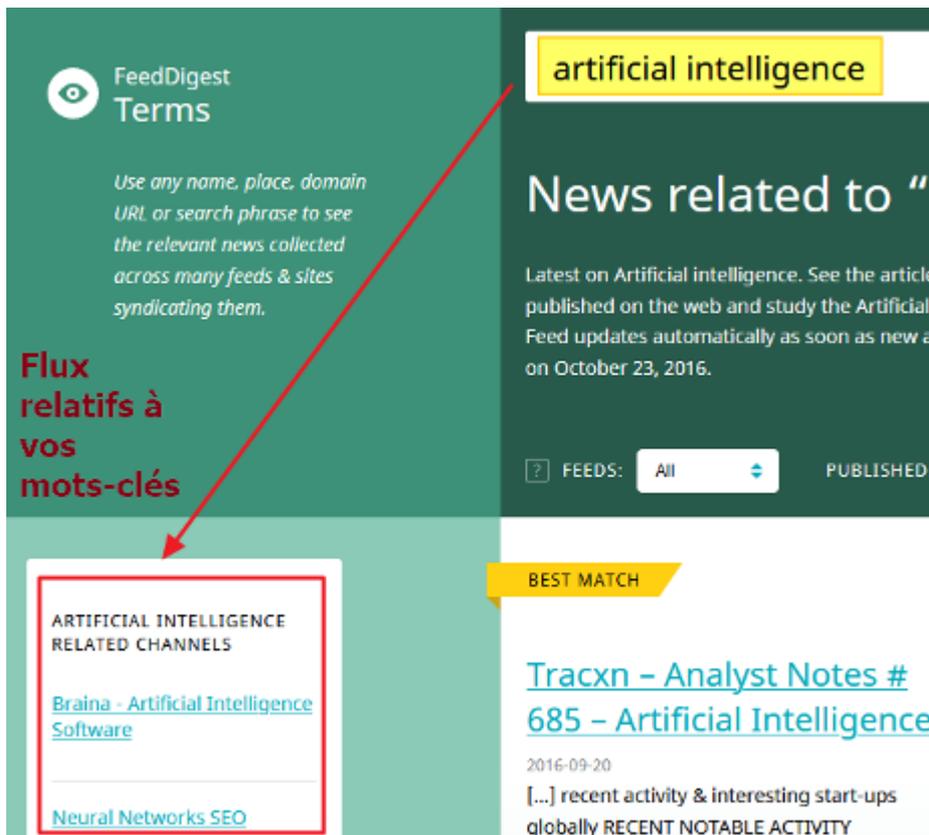
*Blogarama, un agrégateur générique sans annuaire de sources*

Commençons ce tour d'horizon par un comparatif des annuaires de blogs :

Nom	Type	Requêtage / navigation	Remarques	Intérêt pour le sourcing
<b>Wingee</b> <a href="http://www.wingee.com/">http://www.wingee.com/</a>	Annuaire de flux RSS classés par catégories	Moteur interne ou navigation par catégories	Corpus anglo-saxon. Flux peu nombreux.	□
<b>1001rss</b> <a href="https://1001rss.com/sites/">https://1001rss.com/sites/</a>	Répertoire et agrégateur générique de flux RSS	Moteur de recherche interne limité. Utiliser Google : site: <a href="https://1001rss.com/sites/">https://1001rss.com/sites/</a> « <b>mot-clé</b> »	Corpus de sources France. Pas de recherche ou navigation par catégories.	□□
<b>Best of the Web Blogs</b> <a href="https://blogs.botw.org">https://blogs.botw.org</a>	Annuaire de flux RSS classés par catégories	Moteur interne inefficace ou navigation par catégories Utiliser Google : site: <a href="https://blogs.botw.org">https://blogs.botw.org</a> « <b>mot-clé</b> »	Corpus anglo-saxon.	□□□

<b>FeedDigest</b> <a href="http://terms.feeditdigest.com/">http://terms.feeditdigest.com/</a>	Répertoire et agrégateur générique de flux RSS	Moteur interne Très complet. La recherche par mots-clés remonte des flux thématiques (voir « Related channels » en bas à gauche des résultats).	Corpus anglo-saxon. Génère un flux RSS par requête.	□□□□
<b>Bloggernity</b> <a href="http://www.bloggernity.com/">http://www.bloggernity.com/</a>	Annuaire de flux RSS classés par catégories	Moteur interne ou navigation par catégories. La recherche par mots-clés remonte les résultats par ordre chronologique. Donc démarrer la consultation par la dernière page...	Beaucoup de sources blogs. Interface désuète.	□□□
<b>Bloglovin'</b> <a href="https://www.bloglovin.com/">https://www.bloglovin.com/</a>	Annuaire et agrégateur de flux RSS	Moteur interne ou navigation par catégories ( <a href="https://www.bloglovin.com/blogs">https://www.bloglovin.com/blogs</a> )	Rechercher un mot-clé ou une expression dans le moteur puis choisir l'onglet « Blogs » pour accéder aux sources.	□□□
<b>Atlas des flux RSS</b> <a href="http://atlasflux.saynete.net/">http://atlasflux.saynete.net/</a>	Annuaire de flux RSS classés par catégories	Moteur interne ou navigation par catégories	Corpus France. Possibilité d'importer tous les flux d'une catégorie dans un fichier OPML, XML ou CSV. Mises à jour régulières.	□□□
<b>Blogs Collection</b> <a href="https://www.blogs-collection.com/">https://www.blogs-collection.com/</a>	Annuaire de flux RSS classés par catégories	Navigation par catégories. Moteur interne limité. Utiliser Google : site: <a href="https://www.blogs-collection.com/tag/*">https://www.blogs-collection.com/tag/*</a> "mot-clé"	Corpus anglo-saxon Peu ergonomique.	□□

Si ces annuaires ont le mérite d'exister et de continuer à être mis à jour, ils s'avèrent globalement limités, tant dans la manière d'y rechercher des sources (requêtage et navigation), que dans leur graphisme souvent très daté.



Page de résultats de FeedDigest avec flux thématiques ("Related channels")

En pratique, notre choix se portera sur FeedDigest, pour son importante base de flux, ainsi que sur l'Atlas des flux RSS, là aussi pour sa base de flux en français, ainsi que pour la possibilité qu'il offre d'importer l'ensemble des flux d'une catégorie afin de les ajouter rapidement à son agrégateur. Une bonne surprise donc.

### Les annuaires des agrégateurs de flux RSS

Il existe une autre catégorie d'annuaire à prendre en compte. Il s'agit de ceux mis à disposition par les agrégateurs de flux RSS en ligne. Ils ne sont généralement accessibles qu'après que l'on ait créé un compte gratuit, mais bénéficient de la puissance de ce qui enregistrent leurs utilisateurs. L'autre intérêt ici est que l'on accède ici directement à des flux RSS et non sur des blogs qu'il faut ensuite explorer pour trouver le flux.

Bien sûr, tous les agrégateurs n'en proposent pas (ou parfois uniquement en version Premium comme Feedly), mais nous en avons tout de même repéré quatre exploitables à partir d'un compte gratuit :

Nom	Requêtage	Remarques	Intérêt pour le sourcing
-----	-----------	-----------	--------------------------

### Feeder

<https://feeder.co/>

Navigation par catégories.  
Moteur de recherche peu pertinent. Utiliser Google :  
site:https://feeder.co/discover/\*  
"mot-clé"

Les flux détectés sont protégés afin d'être enregistrés directement dans Feeder. Les récupérer pour les ajouter à un autre agrégateur peut-être chronophage.

### Inoreader

<https://www.inoreader.com/>

Recherche par mots-clés (pas d'annuaire).  
Choisir l'onglet « Flux »

Le moteur de recherche proposé est puissant et permet l'usage des booléens, notamment des guillemets. Son corpus de résultats est conséquent.

### Feedspot

<https://www.feedspot.com/>

Recherche par mots-clés ou par catégories.

Cliquer sur le bouton rouge « Add new site » pour accéder aux catégories (à gauche) ou au moteur de recherche par mots-clés (guillemets non utiles)

### Feedreader

<https://feedreader.com/>

Ni moteur de recherche ni annuaire...  
Utiliser Google:

site:https://feedreader.com/observe/\*  
intitle:"mot-clé" -inurl:index.php -  
inurl:article -inurl:thread -inurl:view

□□

Dans cette catégorie, Inoreader et Feedspot nous semblent être les deux services à privilégier.

## Les moteurs de recherche

Reste enfin les moteurs de recherche généralistes que sont Google et Bing. Commençons par ce dernier qui, une fois n'est pas coutume, dépasse ici les possibilités offertes par le leader du marché. En effet, Bing propose depuis longtemps deux opérateurs permettant de rechercher spécifiquement des flux RSS ou des pages qui en disposent.

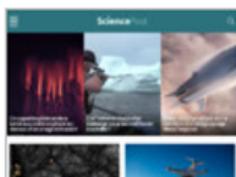
Le premier est « feed: », qu'on utilise ainsi dans une requête :

*feed:mot-clé*

Par exemple :

**feed:"intelligence artificielle"** pour trouver des flux relatifs à ce sujet

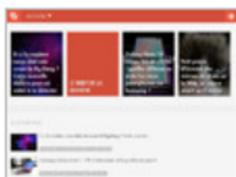
Il est important de comprendre qu'ici Bing ne recherche pas les items (articles, billets, posts,...) issus de flux RSS et dans lesquels votre mot-clé est cité mais bien des flux encapsulés dans des pages web comportant ce mot-clé dans leur titre (donc des sources potentielles). Ce qui leur donne toutes les chances d'être pertinents :



## Intelligence artificielle – Sciencepost

<https://sciencepost.fr/theme/intelligence-artificielle/feed>

Une intelligence artificielle juge Peut importe leur nature, les gouvernements ne sont généralement pas là où nous avons plus de chance [...] More. Une IA pourrait-elle être un bon juge ? L'Estonie pense que oui ! La société Nvidia a pensé aux personnes qui ne savent pas dessiner ou qui cherchent à gagner du temps. En effet, sa nouvelle intelligence artificielle est étonnamment ...



## Intelligence artificielle – Numerama

<https://www.numerama.com/tag/intelligence-artificielle/feed>

Intelligence artificielle – Numerama L'Agence européenne pour les droits fondamentaux a rendu un rapport qui traite de plusieurs problématiques éthiques liées à l'IA.



## intelligence artificielle - disciplines.ac-toulouse.fr

<https://disciplines.ac-toulouse.fr/arts-plastiques/tags/intelligence-artificielle/feed?...>

L'intelligence artificielle, une artiste à part entière ? Un détail du portrait d'Edmond de Belamy. Collectif Obvious. Amélie Boutinot, Université de Strasbourg et Helene Delacour, Université de Lorraine. Le 25 octobre 2018, Christie's, la célèbre maison de vente aux enchères, a mis en vente le portrait d'Edmond de Belamy « peint » par un programme d'intelligence ...

*Résultats de Bing avec l'opérateur "feed:"*

Quoique moins directement utile, la syntaxe « hasfeed: » peut également être utilisé en complément. A la différence du premier, cet opérateur n'ouvre pas directement des flux RSS dans votre navigateur, mais des pages web qui en contiennent (« has feed ») et où votre mot-clé est présent. On peut donc penser qu'il s'agit de flux en lien avec cette même thématique mais il faudra bien entendu s'en assurer avant de s'y abonner. On l'utilise ainsi :

**hasfeed:"intelligence artificielle"** pour trouver des pages web où l'expression "intelligence artificielle" est citée et qui comportent un flux RSS.



✓ **Intelligence artificielle** : OpenAI parvient à créer des ...

<https://www.20minutes.fr/high-tech/2948159-20210108-intelligence...> ▼

Il y a 2 jours · L'intelligence artificielle progresse bien. Voici dall.E une IA de openAI qui génère des images à partir d'un texte descriptif. Franchement les chaises avocats on est clairement dans la ...



✓ Cette **intelligence artificielle** différencie les théories ...

<https://www.tomsguide.fr/cette-intelligence-artificielle-differencie...> ▼

07/01/2021 · Des chercheurs de l'université de Californie à Los Angeles ont développé une intelligence artificielle capable de différencier les théories du complot et les véritables conspirations qui ...



✓ **Intelligence artificielle** : un outil détermine la ...

<https://www.tomsguide.fr/intelligence-artificielle-un-outil-determine-la-categorie...> ▼

30/12/2020 · L'intelligence artificielle est alors capable de déterminer si une scène va présenter de la violence ou autre. En effet, le script d'un film ne contient pas que le dialogue des personnages ...

---

*Résultats de Bing avec l'opérateur "hasfeed:"*

Comme on peut le voir dans les résultats, l'expression « intelligence artificielle » est bien présente dans le texte et/ou le titre des pages. Par ailleurs, les vérifications que nous avons effectuées montrent que des flux RSS sont bien présents dans chacune d'elles.

Si Google n'offre pas d'opérateurs de ce type, on peut toutefois utiliser le classique « filetype: » (ou « ext: »).

Ainsi on peut lancer la requête suivante :

**filetype:rss "intelligence artificielle"**

ou encore :

**filetype:atom "intelligence artificielle"**

Les résultats sont nettement moins pertinents que ceux fournis par Bing et nous suggérons d'utiliser plutôt ces requêtes dans Google Alerts, afin de compléter ses sources au fil de l'eau avec les nouveautés détectées par les robots d'indexation de Google.

## Conclusion : faire les choses dans le bon ordre pour des résultats plus efficaces

Le déploiement d'une stratégie de sourcing de flux RSS rapide et efficace nous semble donc devoir être effectué dans le bon ordre et avec les outils suivants :

1. Requête dans Bing avec l'opérateur « feed: » ;
2. Interrogation des annuaires de flux d'Inreader et de Feedspot ;
3. Interrogation des annuaires de flux RSS FeedDigest et Atlas des flux RSS.

On pourra ensuite compléter ces sources au fil de l'eau *via* des alertes Google utilisant la syntaxe « filetype:rss » ainsi qu'avec des flux de requêtes Bing (cf. [cet article pour apprendre à les créer](#)).

Comme on le voit, les flux RSS restent une valeur sûre de la veille et il serait dommage de ne pas continuer à les exploiter.



**Christophe Deschamps, Consultant-formateur : veille stratégique, intelligence économique, social KM, e-réputation, mindmapping, IST (<http://www.outilsfroids.net/>)**