

Beste de savoir

Ne loupez plus une publication : utilisez
les flux RSS !

12 août 2019

Table des matières

1.	Qu'est-ce que la syndication de contenu RSS/Atom ?	2
1.1.	La syndication de contenu, ou souscription à du contenu	2
1.2.	RSS vs Atom	4
1.3.	Concrètement, comment ça marche ?	4
2.	Quelques exemples de lecteurs RSS	5
2.1.	Clients lourds	5
2.2.	Clients web « dans les nuages »	6
2.3.	Clients web à auto-héberger	6
3.	Comment récupérer les adresses des flux ?	7
3.1.	Le site rend ses flux RSS visibles	7
3.2.	Le site propose des flux RSS, mais les planque	7
3.3.	Le site ne propose pas de flux RSS	8
4.	Comment proposer de tels flux sur mon site ?	8
4.1.	Générer les flux	8
4.2.	Présenter les flux à l'utilisateur	9

Vous aussi vous râlez parce que Facebook ne vous affiche plus qu'une fraction¹ des publications de vos contacts ? Vous pestez contre Youtube qui a décidé de ne plus vous envoyer les notifications de vos vidéastes préférés alors que vous êtes abonnés à leurs chaînes ?

Plus généralement, vous détestez louper des informations intéressantes parce qu'un algorithme abscons a décidé qu'il ne devait pas vous les montrer ?



La solution existe pour ne plus rien louper : c'est la **syndication de contenu** à l'aide de flux RSS ou Atom, et c'est ce dont on va parler ici !

Le plus ironique ? C'est que cette solution aux mauvais comportements des plateformes existe depuis 1999, bien avant les outils problématiques eux-mêmes...

1. le compte Facebook compte présentement 933 abonnés, mais le partage du [tuto sur le Jazz de nohar](#) n'a été vu que par 61 personnes en 4 jours...

1. Qu'est-ce que la syndication de contenu RSS/Atom ?

1. Qu'est-ce que la syndication de contenu RSS/Atom ?

1.1. La syndication de contenu, ou souscription à du contenu

1.1.1. Se tenir au courant des nouveautés ?

Imaginons : vous êtes intéressé par un *producteur de contenu* quelconque sur Internet (un site, une chaîne Youtube, une institution...) dont vous désirez suivre les nouveautés. Vous avez en gros trois solutions pour ce faire :

1. **Vous vérifiez les nouveautés à la main.** C'est long, pénible et sujet à erreurs – ce ne peut être géré que pour le site que vous avez en page d'accueil et une poignée d'autres.
2. **Vous utilisez les réseaux sociaux** ou d'autres outils intégrés qui vous permettent, via un compte, de souscrire aux nouveautés publiées. Ça fonctionne bien, et même très bien puisque c'est aujourd'hui la méthode principale utilisée. Elle souffre néanmoins de ces défauts :
 - soit vous devez avoir un compte par producteur de contenu, et ça devient à peine plus pratique à gérer que la méthode 1 ;
 - soit vous utilisez un réseau social quelconque. Les producteurs y annoncent leurs nouveautés et vous les voyez... *quand les réseaux ont décidé que vous auriez le droit de les voir.*
3. **Vous vérifiez les nouveautés automatiquement** et elles vous sont présentées à l'endroit de votre choix. D'accord, mais comment ? Eh bien grâce à des mécanismes appelés *syndication de contenu* ou *souscription à du contenu* dont on va justement parler ici.

1.1.2. Un peu d'histoire...

Au commencement était Internet, et il était informe et laid². Et comme l'être humain est un animal grégaire^[réf. nécessaire], il a commencé à créer des communautés et à s'organiser sur Internet comme dans la réalité³ : en réseaux. Comme c'était des réseaux, qu'ils avaient une fonction avant tout sociale, et qu'on est quand même des grosses feignasses quand il faut inventer des noms, on a appelé ça des *réseaux sociaux*.

En parallèle, pour éviter de devoir suivre plein de sites chaque jour juste pour savoir s'il y a eu des nouveautés, on a inventé la syndication de contenu – mais j'y reviendrai.

Avec le temps, les réseaux sociaux sont devenus de plus en plus importants et généralistes. Jusqu'aux mastodontes difficilement contournables que sont aujourd'hui Facebook, Twitter, et Youtube en ce qui concerne les vidéos.

Cette puissance de frappe et une bonne communication a fait dire aux créateurs de contenu que ces réseaux sociaux étaient **la** solution pour publier leurs nouveautés. Pensez donc !

- Une grande partie de la communauté qui est intéressée par vos contenus est déjà présente sur ces réseaux.

2. Genre, *vraiment* laid. Avec des GIF animés en fond d'écran.

3. Ce qui est sur Internet est-il moins réel que ce qui est hors Internet ? Vous avez deux heures.

1. Qu'est-ce que la syndication de contenu RSS/Atom ?

- Ces réseaux, via leurs fonctions sociales (partages, suggestions...) sont un excellent moyen de toucher un public nouveau.

Le paradis pour les créateurs !

Or, maintenir un réseau social d'envergure, ça coute cher. *Très* cher. Genre, des quantités colossalement invraisemblables d'argent. Pour vous donner une petite idée, Twitter a été fondé le 21 mars 2006 et n'a engrangé ses premiers bénéfices qu'au [quatrième trimestre 2017](#) [↗](#), soit plus de dix ans plus tard – n'importe quelle entreprise traditionnelle aurait coulé longtemps avant⁴.

Comment gagner autant d'argent ? Comme presque toujours sur Internet, avec de la publicité. Mais voilà que les marchés publicitaires ne sont pas au beau fixe⁵. Qui donc aurait de l'argent ? Les créateurs de contenu !

Bon, en réalité, beaucoup n'ont *pas vraiment* d'argent, mais ce sont néanmoins eux qui sont prêts à en mettre. L'astuce est simple : les réseaux sont devenus tellement importants pour leur visibilité qu'ils peuvent difficilement s'en passer. Il suffit donc de monnayer cette visibilité, et le tour est joué !

Et ça, les responsables des réseaux sociaux l'ont bien compris, même [s'il semble qu'ils commencent à demander trop cher](#) [↗](#). En fait, c'est [The Oatmeal qui résume encore le mieux la situation](#) [↗](#).

[Je m'étais dit que je ferais court, c'est raté. Tant pis].

1.1.3. Automatisation de la récupération de contenu ?

Or donc, l'idéal serait d'avoir un système automatique qui va vérifier la présence de nouveau contenu pour nous et nous la présente sous un format agréable.

On pourrait imaginer un système à base de liste de diffusion de courriels, ou une *newsletter*. Ça existe, mais c'est assez peu pratique. À la fois pour le producteur (il faut envoyer tous ces courriers, en évitant de se faire bloquer pour spam, à gérer les formatages, les clients qui n'affichent pas les images...) et pour l'utilisateur (les boîtes mail sont vite saturées, et un nouveau contenu n'a sans doute pas la même priorité qu'un courriel *normal*).

Non, le mieux serait d'avoir un robot qui va vérifier automatiquement les sites qui nous intéressent et nous alerte en cas de nouveau contenu. Et pour bien faire, il faudrait un standard qui évite d'avoir à interpréter chaque site pour deviner s'il y a eu des nouveautés ou non.

i

Eh bien, **la syndication de contenu, c'est exactement ça**. C'est :

1. Des robots qui vont vérifier la présence de nouveau contenu sur une liste arbitraire et paramétrable de sites.
2. Une interface pour voir et gérer ces contenus (lus, non lus, pas intéressé, ajouter,

4. D'aucuns pensent « bulle ». Les géants ne sont pas immortels – pensez à MySpace par exemple. Ou à la bulle Internet de l'an 2000.

5. Ça mériterait une série d'articles, mais on pourrait disserter longtemps sur le comportement des publicitaires eux-mêmes, qui a *fortement incité* le grand public à s'équiper de bloqueurs, au cours d'une folle course au pistage toujours pas finie...

1. Qu'est-ce que la syndication de contenu RSS/Atom ?

i

modifier ou supprimer des producteurs de contenus...)

3. Un standard pour que tout le monde se parle sans se prendre le chou.

Et ça fonctionne bien. En fait, ça fonctionne tellement bien que c'est le mécanisme principal de tout lecteur de podcast, ce qui semble être un bon exemple de robustesse du système, ne trouvez-vous pas ?

1.2. RSS vs Atom

Depuis le début, je parle de RSS et d'Atom. Vous avez peut-être entendu parler de « flux RSS ».

Ce sont deux standards de syndication. RSS est le premier, mais posait certains problèmes de droits et de cohérence. Atom a été une réponse normalisée [dont j'ai d'ailleurs parlé ici ↗](#) (et ce tuto n'est probablement plus à jour, au moins dans ses liens), mais qui n'a pas été adoptée partout.

En vrai, vous vous fichez totalement du format utilisé techniquement, les deux offrent les mêmes fonctionnalités – pour celles qui sont réellement utilisées. Si les deux vous sont proposés (comme sur Zeste de Savoir, mais c'est plutôt une exception) et que votre lecteur les gère, utilisez plutôt Atom : comme il est mieux défini il devrait y avoir moins de surprises à l'affichage.

Bon. On sait maintenant ce qu'est la syndication de contenu, pourquoi elle est là et qu'il existe un protocole technique pour l'utiliser et dont on ne reparlera plus. Et si on s'attaquait aux points 1 et 2 de ma définition, à savoir : **comment on se sert de ce truc ?**

1.3. Concrètement, comment ça marche ?

Imaginons que je veuille suivre les actualités d'un site, disons au hasard, les nouveaux articles de Zeste de Savoir.

1. Je trouve l'adresse du flux RSS qui m'intéresse : [https://zestedesavoir.com/articles/flux/rss/ ↗](https://zestedesavoir.com/articles/flux/rss/) et le rentre dans mon client (cf la partie suivante).
2. Un nouvel article, au hasard « Ne loupez plus une publication : utilisez les flux RSS ! » est publié sur Zeste de Savoir. Cet article est ajouté au flux par le site.
3. Au prochain lancement de mon lecteur RSS (ou périodiquement, selon la technologie utilisée), l'outil va trouver la version mise à jour du flux. Il me montre ce nouvel article.
4. Ici, deux possibilités :
 - Soit c'est un flux qui contient les articles complets, il suffit de lire dans le lecteur – mais ce n'est pas le cas des RSS de Zeste de Savoir.
 - Soit le flux permet à l'outil d'afficher un court résumé, et s'il m'intéresse il me donne aussi un lien pour lire le contenu sur le site.

2. Quelques exemples de lecteurs RSS

À la grande époque des flux RSS⁶ la technologie était assez visible : les sites avaient souvent un logo qui permettait d'accéder au site, ou à défaut les navigateurs affichaient un logo directement dans la barre d'adresse⁷. Et ces navigateurs permettaient de lire et gérer ces flux, ce qui était bien suffisant quand on en avait pas trop.

Mais le monde change, et les usages avec.

Les réseaux sociaux ont pris la main, les sites Internet n'affichent plus leurs flux RSS, et les navigateurs ne semblent plus capables de les gérer – au moins de manière *simple*.

Il va falloir trouver un lecteur, et avec les usages actuels ça couvre trois cas très différents.

2.1. Clients lourds

Un « client lourd », c'est tout simplement un logiciel qui s'installe sur votre ordinateur. C'est pratique parce que ça permet d'avoir toutes les données en local et disponibles hors-connexion.

Alors quel client choisir ? Eh bien, n'importe quel bon logiciel de gestion des **courriels** gère aussi la lecture de RSS ! C'est par exemple le cas avec :

- [Outlook](#) ↗
- [Thunderbird](#) ↗

Si vous n'utilisez pas de logiciel dédié pour lire vos courriels, vous pouvez aussi utiliser un plugin pour votre navigateur ; il en existe plein, et je vous laisse présenter vos préférés dans les commentaires. [Opera](#) ↗ a rétabli la fonctionnalité, et [Firefox](#) ↗ intègre toujours un support minimal via les marque-pages dynamiques. Dans les deux cas le raccourci direct dans la barre d'URL est maintenant masqué par défaut.

On trouve aussi des logiciels dédiés, si vous êtes adeptes du paradigme « *une tâche, un outil* » :

- [Newsboat](#) ↗ , un lecteur de flux en mode console.
- [Vienna](#) ↗ , [Leaf](#) ↗ , [ReadKit](#) ↗ ou [Reeder](#) ↗ pour MacOS et iOS uniquement.

Cette section concerne aussi les **applications mobiles** (un smartphone n'étant rien d'autre qu'un ordinateur qui tient dans une poche). Là aussi vous avez des zillions d'applications de lecture de flux RSS, selon votre fournisseur.

6. On va se mettre d'accord tout de suite : je désigne par ce terme n'importe quel flux de syndication que vous pouvez utiliser avec le lecteur de votre choix, quelle que soit sa technologie réelle. Parce que c'est comme ça qu'on les appelle dans le langage courant.

7. Chose possible grâce à une méta-donnée présente dans la page HTML qui indique l'existence et l'emplacement du ou des flux.

2. Quelques exemples de lecteurs RSS

2.2. Clients web « dans les nuages »

C'est beau tous ces clients, mais pas toujours très pratique : les flux sont centralisés sur *un* ordinateur, sans accès depuis d'autres. On peut installer la même liste de flux sur plusieurs ordinateurs, mais sans synchronisation (ai-je déjà vu cet article ? l'ai-je déjà lu ? etc).

Ce qui serait pratique, c'est un site web, accessible de n'importe où, qui centraliserait tous mes flux. Et c'est exactement ce que proposent les clients RSS *dans le cloud* !

Le Grand Ancien très utilisé du domaine était [Google Reader](#) [↗](#), mais il a été assassiné en 2013, officiellement parce que son usage baissait à cause de la montée en puissance des réseaux sociaux⁸.

Aujourd'hui on trouve encore [beaucoup de lecteurs RSS dans les nuages](#) [↗](#), dont entre autres :

- [Framanews](#) [↗](#), qui est une instance de Tiny Tiny RSS fournie par [Framasoft](#) [↗](#)⁹
- [Feedly](#) [↗](#) qui est sans doute l'un des plus connus.
- [Netvibes](#) [↗](#), un Grand Ancien lui aussi, mais qui a survécu (apparemment en se diversifiant : il permet maintenant d'agréger des contenus en provenance de flux RSS mais aussi de réseaux sociaux).
- ... et tous ceux que vous nous présenterez en commentaires !

2.3. Clients web à auto-héberger

D'accord, on peut lire ses flux RSS à l'aide d'outils en mode SaaS. Mais ce n'est pas sans poser quelques soucis :

- Que se passe-t-il si l'outil ferme ?
- Et si l'outil change de modèle économique ?
- Où se retrouvent mes données ? Qui les possède ?
- Qu'est-ce qui est fait avec mes données ?

Si ces questions vous importent **et que vous avez les compétences pour héberger vous-même un site Internet dynamique**, il y a des projets pour ça aussi ! Par exemple :

- [Tiny Tiny RSS](#) [↗](#), l'un des plus connus semble-t-il – et utilisé par Framasoft pour Framanews, cf au-dessus.
- [FreshRSS](#) [↗](#).
- [Leed](#) [↗](#), utilisé par votre renard préféré.
- *[Insérez ici votre projet préféré tel que vous le décririez en commentaire]*

Attention lors de vos recherches : beaucoup de projets de lecteurs RSS en ligne sont apparus avec la disparition de Google Reader, et seulement une petite partie ont survécu jusqu'à aujourd'hui. Le sujet est moins trivial qu'il n'y paraît¹⁰ et la baisse d'intérêt due à la montée des réseaux sociaux n'a pas aidé.

8. On était encore à l'époque où Google tentait d'imposer Google Plus par à peu près tous les moyens. Coïncidence ou causalité ? Difficile à dire, sans avoir les véritables chiffres de fréquentation de tout ce petit monde.

9. Framasoft, l'association qui arrive tellement bien à *dégoogliser Internet* [↗](#) qu'ils *dégooglisent* un service que Google ne propose plus, et ça c'est la classe !

10. Notamment parce que beaucoup de flux ne respectent pas les standards RSS ou Atom : un lecteur incapable d'improviser correctement face à des flux foireux est un lecteur inutilisable.

3. Comment récupérer les adresses des flux ?

3. Comment récupérer les adresses des flux ?

Parce que c'est mignon d'avoir un beau lecteur RSS tout neuf, mais encore faut-il le remplir. Et là, trois possibilités s'offrent à nous.

3.1. Le site rend ses flux RSS visibles

Il « suffit » de trouver le bon lien sur le site. Selon le site, ça peut être plus ou moins compliqué, mais généralement on y arrive.

Bref, pas grand-chose de passionnant ici.

3.2. Le site propose des flux RSS, mais les planque

Si les boutons RSS ont pratiquement disparu des interfaces web au profit de boutons de type « réseaux sociaux », beaucoup de sites proposent encore cette fonctionnalité sans y donner directement accès¹¹.

Ici, le plus simple est encore de regarder le code source de notre page, ce qui se fait généralement avec `ctrl-u` (vue de la source) ou `ctrl-alt-i` (console de debug, plus puissante mais plus complexe). Il y a deux possibilités, décrites ci-dessous. Dans tous les cas, n'hésitez pas à contacter les responsables du site pour leur demander de remettre en place leur flux : beaucoup de fonctionnalités ne sont maintenues que si elles sont utilisées !

3.2.1. Le flux est correctement intégré

Si le flux est correctement intégré, on trouve une balise¹² `<link rel="alternate" type="application/rss+xml" href="http://www.lemonde.fr/rss/une.xml" title="Le Monde.fr : A la une" type="application/rss+xml">` avec un attribut `type="application/rss+xml"` ou un attribut `type="application/atom+xml"`. Cette balise contient le lien qui nous intéresse, à l'intérieur de l'attribut `href`.

Exemple : si on cherche dans les sources de la page <http://www.lemonde.fr> ↗, on trouve ceci :

```
1 <link rel="alternate" href="http://www.lemonde.fr/rss/une.xml"
  title="Le Monde.fr : A la une" type="application/rss+xml">
```

Le flux des actualités à la une est donc : <http://www.lemonde.fr/rss/une.xml> ↗

11. D'aucuns estiment que la présence de flux RSS concerne encore environ 40 % des sites, malgré la baisse constatée depuis des années. Chiffre à prendre avec une pelleuse, hein !

12. En fait une balise par flux, donc plusieurs balises s'il y a plusieurs flux. Mais cette explication est déjà bien assez compliquée, et les sites qui planquent leurs flux prennent rarement la peine d'en gérer plusieurs.

4. Comment proposer de tels flux sur mon site ?

3.2.2. Le flux est intégré bizarrement

Parfois, le site fournit un flux RSS mais sans le mettre proprement dans la balise dédiée. On peut toujours lancer une recherche dans le code source de la page, par exemple sur `rss`, et espérer tomber sur quelque chose.

Par exemple, si on fait cette manœuvre sur [une page de chaîne YouTube](#) , on trouve ceci :

```
1 "rssUrl": "https://www.youtube.com/feeds/videos.xml?channel_id=UCaNLbnghwlsGF-
```

Ce qui est bien ce que l'on recherche (quoique [ce ne semble pas être la meilleure manière d'ajouter une telle chaîne à son lecteur RSS](#)).

Digression de l'auteur : ce qui serait pratique, c'est un outil à qui on passe une URL et qui renvoie directement la liste des flux RSS liés. Je n'en ai pas vu après une recherche rapide sur le net, mais il est minuit passé. Si quelqu'un en connaît ou en développe un, qu'il fasse signe !

3.3. Le site ne propose pas de flux RSS

Là, c'est le drame.

La première chose à faire est de contacter le webmaster du site et de lui demander de rajouter la fonctionnalité. Qui sait, s'il y a suffisamment de demandes, peut-être que ça finira par arriver.

En attendant, il y a des outils pour tenter de pallier ce problème, comme :

- [FetchRSS](#) ;
- [rss-bridge](#) ;
- [Five Filters](#) ;
- [Kill the newsletter](#) ;
- ... toute autre solution présentée en commentaire.

4. Comment proposer de tels flux sur mon site ?

Cette section concerne les *webmasters* : toi qui gères un site Internet et qui es convaincu de la pertinence des flux RSS, comment peux-tu les mettre en place ? Réponse en deux étapes.

4.1. Générer les flux

Si tu utilises un *framework*, la plupart d'entre eux ont des outils intégrés ou des *plugins* pour générer de tels flux. Je ne peux que te conseiller de fouiller la documentation et les dépôts de modules ; voici néanmoins quelques liens en vrac – pas forcément les plus pertinents, n'hésitez pas à en indiquer d'autres en commentaires.

- Wordpress [a une fonctionnalité intégrée](#) .
- On trouve [plusieurs bundle](#) [Symfony](#) .

4. Comment proposer de tels flux sur mon site ?

- Django [a un framework pour ça](#) .
- Ruby on Rails [semble embarquer la fonctionnalité en standard](#) .
- Spring / Spring Boot [s'utilise à priori avec ROME](#) .
- etc.

Au pire, les normes RSS comme Atom décrivent de simples flux XML assez simples à générer avec n'importe quelle bibliothèque de gestion XML.

Il ne reste plus qu'à se poser les bonnes questions :

- Quels flux proposer ? Avec quel contenu pour chacun ?
- Comment est gérée la sécurité si certaines parties du site sont privées ?
- Combien de contenus sont publiés dans chaque flux ? (À voir avec la taille des contenus et la fréquence de publication).
- Comment sont gérées les mises à jour de contenu ?
- Est-ce que les flux sont complets, ou ne servent qu'à renvoyer les utilisateurs sur le site ? (Les deux flux peuvent coexister).

Quelques détails sur le dernier point : les flux complets sont plus confortables pour l'utilisateur, qui peut lire les articles directement dans son lecteur. Ils devraient donc être proposés dès que possible. Néanmoins, il peut être nécessaire de n'utiliser que des flux tronqués :

- Quand la mise en page est complexe et importante (elle peut être très différente dans le lecteur, et impossible à forcer).
- Quand les articles peuvent être très longs. Certains lecteurs RSS *détestent* les articles trop longs.
- Quand le site a besoin de la présence de l'utilisateur sur les pages web (par exemple, pour des raisons de contrôle d'accès ou de publicité).

4.2. Présenter les flux à l'utilisateur

Les flux RSS sont à définir dans l'en-tête HTML avec une balise de ce type (où vous remplacez évidemment le titre et l'URL par vos informations).

```
1 <link rel="alternate" type="application/rss+xml" title="Forum"
  href="/forums/flux/messages/rss/">
```

Il peut y avoir plusieurs balises de ce type sur la page, pour plusieurs flux et/ou plusieurs formats.

À une époque ça suffisait : un bon navigateur détectait la présence de telles balises, et affichait un logo dans la barre d'URL. Mais ce n'est plus le cas.



Aujourd'hui, il faut afficher des liens « explicites » vers les flux RSS.

Zeste de Savoir affiche de tels liens : dans les pages principales des sous-sections, dans le menu de gauche dans le forum.

4. Comment proposer de tels flux sur mon site ?

Et voilà ! Vous avez maintenant un moyen simple et fiable de suivre vos contenus préférés, ou de permettre à vos utilisateurs de suivre vos contenus, et ce **indépendamment des réseaux sociaux**, de leurs algorithmes, de leurs mises à jour voire de leur existence.

N'hésitez pas à implémenter les flux RSS, ou à demander à vos sites préférés de le faire !

Merci à qwerty pour la validation, ainsi qu'à tous les bêta-lecteurs pour leurs retours !

Icône par Mozilla sous triple licence MPL/GPL/LGPL [↗](#).

Liste des abréviations

SaaS Software as a Service – nom générique pour les logiciels utilisables dans le cloud.. 6