

RSS - Really Simple Syndication

Table des matières

<u>RSS - Really Simple Syndication</u>	1
<u>Qu'est-ce qu'un RSS ?</u>	2
<u>Comment accéder aux flux RSS ?</u>	2
<u>Le standard RSS</u>	2
<u>Principales balises RSS</u>	3
<u>Les balises descriptives obligatoires du canal</u>	4
<u>Balises optionnelles pour le canal</u>	4
<u>Ajouter une image</u>	4
<u>Les articles du canal</u>	5
<u>Plus quelques indications optionnelles pour l'article</u>	5
<u>Comment indiquer la présence d'un flux RSS au navigateur</u>	5
<u>Rendre le flux accessible</u>	5
<u>Mieux : afficher l'existence d'un flux dans l'url du site</u>	5
<u>Flux RSS 2, exemple de structure validée</u>	6
<u><category>Musique, poesie, voiture,littérature</category></u>	6
<u>Valider</u>	6
<u>Exercice de synthèse</u>	7
<u>La méthode de sortie HTML</u>	8
<u>Creez un fichier en php qui converti le flux en fichier html</u>	9
<u>Afficher des flux RSS 2.0 avec PHP</u>	10
<u>Pour les motivés !!</u>	10




Qu'est-ce qu'un RSS ?

C'est un format d'échange de données, défini en XML version 1.0. Vous pouvez délivrer des données en les intégrant dans ce format, et l'on peut recueillir des données de sources multiples dans ce format. On désigne les sources par fil ou flux RSS.

Le format « RSS » (traduisez « **Really Simple Syndication** ») permet ainsi de décrire de façon synthétique le contenu d'un site web, dans un fichier au format XML, afin de permettre son exploitation par des tiers. Le fichier RSS, appelé également flux RSS, canal RSS ou fil RSS, contenant les informations à diffuser, est maintenu à jour afin de constamment contenir les dernières informations à publier.

Comment accéder aux flux RSS ?

Vous verrez en bas de votre navigateur ou dans la barre d'adresse cet icône .

Cliquez sur cet icône puis sélectionnez le flux de votre choix. Celui-ci sera ajouté à vos favoris.

Les navigateurs récents peuvent directement lire des fichiers RSS, mais on peut aussi utiliser un logiciel spécialisé que l'on appelle **lecteur RSS** ou **agrégateur**

Le standard RSS représente un moyen simple d'être tenu informé des nouveaux contenus d'un site web, sans avoir à le consulter.

Le standard RSS

Un flux RSS est un simple fichier texte contenant le titre de l'information, une courte description et un lien vers une page décrivant plus en détail l'information.

On utilise un format XML, permettant de décrire de façon fine le contenu du canal et de chacun des items qui le constituent. Il existe plusieurs versions du standard RSS :

1. RSS 0.91 (Rich Site Summary) représente une évolution mineure des spécifications de format RSS originel mis au point par la société Netscape ;
2. RSS 1.0 (RDF Site Summary) est une spécification alternative, issue du standard RDF ;
3. **RSS 2.0 (Really Simple Syndication)** est une amélioration du standard RSS 0.91, mise au point par Dave Winer.

Mais on parle aussi souvent de RSS pour désigner également le format Atom.

[http://fr.wikipedia.org/wiki/RSS_\(format\)](http://fr.wikipedia.org/wiki/RSS_(format))

L'appellation Atom se réfère à deux standards liés.

- Le Format de Syndication Atom est un format de document basé sur XML conçu pour la syndication de contenu périodique, tel que les weblogs ou les sites d'actualités,

tandis que le

- Protocole de Publication Atom (APP) est un protocole simple basé sur HTTP pour la création et la mise à jour de ressources Web.

Au moment où l'idée d'Atom est esquissée, en juin 2003, la principale manière de « syndiquer »



périodiquement du contenu est d'utiliser la famille des formats RSS.

Principales balises RSS

rss. Le conteneur global.

channel. Un canal. Il contient plusieurs balises descriptives, et une série de balises "item", les informations.

```
<?xml version="1.0" ?>
```

```
<rss version="2.0">
```

```
<channel>
```

```
</channel>
```

```
</rss>
```

Liste de tous les éléments pouvant se trouver dans la balise <channel>

Balise	Description
title	Titre du channel
link	URL du site contenant le channel
description	Description du channel
language	Langue du channel
copyright	Info sur le copyright du channel
managingEditor	Mail de la personne responsable du contenu
webMaster	Mail du webmaster
pubDate	Date de publication
lastBuildDate	Date de la dernière publication
category	Catégorie à laquelle le channel appartient
generator	Programme utilisé pour générer le channel
docs	Lien vers la documentation du format utilisé dans le fichier RSS
cloud	Permet à un programme de s'enregistrer pour être notifié des modifications de ce channel
ttl	Time to live, avant le prochain rafraîchissement
image	Image affichée avec le channel
rating	note PICS
textInput	Ajouter une zone de saisie de texte
skipHours	Heures que les agrégateurs peuvent ignorer
skipDays	Jours que les agrégateurs peuvent ignorer

- **Items :** Un channel peut contenir n'importe quel nombre d'items. Chaque item représente un



article, une news et peut posséder notamment un titre, une description, un lien vers l'article en question et une date de publication.

Les balises descriptives obligatoires du canal

- **title.** titre du canal, devrait contenir son nom.
- **link.** Adresse (URL) du site web pour ce canal.
- **description.** Une phrase pour décrire le canal.
- plus au moins une balise **item** pour le contenu.

```
<rss version="2.0">
<channel>
  <title>mon flux</title>
  <link>http://www.moi.be</link>
  <description> mes élucubrations</description>
  <item>
  </item>
</channel>
</rss>
```

Balises optionnelles pour le canal

- language La langue dans laquelle le texte est écrit.
- docs Adresse web de la documentation sur le format,
- webMaster E-mail du responsable.
- pubDate Date de publication.
- etc. Il y a d'autres options à voir dans la spécification.

Ajouter une image

C'est facultatif. Créer une petite image (88x31 par exemple) au format usuel (gif, jpg, png) et placez la dans le même répertoire que ce fichier RSS.

La balise image est un élément de la balise channel

- url est l'adresse de l'image elle-même.
- link est l'adresse de la page affichée quand on clique sur l'image. <channel>

```
<channel>
  <title> </title>
  <link> </link>
  <image>
    <url>http://www.moi.be/monImage.gif</url>
    <link>http://www.moi.be/index.php</link>
  </image>
</channel>
```



Les articles du canal

Chaque balise **item** correspondant à une information contient les balises suivantes:

- **title**. Titre de l'article ou autres genre d'information.
- **link**. L'URL de la page.
- **description**. Résumé, "synopsis" de l'article.

```
<item>
  <title>nouvelle neuve</title>
  <link>http://Derniere.html</link>
  <description>quoi de neuf </description>
</item>
```

Plus quelques indications *optionnelles* pour l'article

- **pubDate**. Date de publication.(format américain)
- **guid**. Une chaîne de caractère qui est un identifiant unique de l'item.
- **category**. Une catégorie à laquelle peut appartenir l'article.
- etc.

Comment indiquer la présence d'un flux RSS au navigateur

Rendre le flux accessible

Pour cela on peut placer un bouton RSS sur la page d'accueil. En cliquant sur le bouton on doit afficher le fichier précédemment créé.  ou 

```
<a href="http://www.xul.fr/rss.xml" >  </a>
```

Si l'image ne s'affiche pas, utiliser une URL complète comme <http://www.xul.fr/rss.gif>.

Mieux : afficher l'existence d'un flux dans l'url du site

Dans la partie <head> de la page, ajouter :

```
<link rel="alternate" type="application/rss+xml" href="flux.xml" title="tester rss">
```

à adapter



Flux RSS 2, exemple de structure validée

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<rss version="2.0">
  <channel>
    <title>nom de votre site</title>
    <link>http://www.mon-site/</link>
    <description>Mon site est le plus beau des sites</description>
    <language>fr-fr</language>
    <managingEditor>email de l'éditeur</managingEditor>
    <webMaster>mailwebmaster@monHebergeur.org (nom webmaster)</webMaster>

    <category>Musique, poesie, voiture, littérature</category>

  <item>
    <title>itre de la news</title>
    <link>http://www.lien de la news/index.php</link>
    <guid isPermaLink="true">http://www.lien de la news</guid>
    <description>je suis bien content de decrire cette
      news</description>
  </item>

  <item>
    <title>itre de la news2</title>
    <link>http://www.lien de la news2/index.php</link>
    <guid isPermaLink="true">http://www.lien de la news2</guid>
    <description>je suis bien content de decrire cette
      news2</description>
  </item>

  </channel>
</rss>
```

vous pouvez mettre à jour en ajoutant un item même structure, mettez en premier la nouvelle news pour faciliter la lecture.

Valider

Aller sur
<http://feedvalidator.org/>

FEED Validator

FOR ATOM AND RSS AND KML

RSS Profile Support Added ([more](#))

[Home](#) · [About](#) · [News](#) · [Docs](#) · [Terms](#)

Copyright ©2002-7 [Sam Ruby](#), [Mark Pilgrim](#), [Joseph Wallon](#), and [Phil Ringnalda](#)



Exercice de synthèse

Voir doc <http://xmlfr.org/w3c/TR/xslt/>

Considérons un site de 3 pages:page1.html, page2.html et page3/html

1) créer un fichier RSS

```
nouvelle 3
description: troisième essai d'un flux
Voir la page : http://page3.html
publiée le 1/03/2008
nouvelle 2
description: deuxième essai d'un flux
Voir la page : http://page2.html
publiée le 10/02/2008
nouvelle 1
description: premier essai d'un flux
Voir la page : http://page1.html
publiée le 1/02/2008
```

Rappels:structure

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<rss version="2.0">
  <channel>
    <title>exercice de synthèse</title>
    <link>http://www.monsite.com</link>
    <description>Exercice pour les webmasters</description>
    <language>français</language>
    <copyright>Plomteux Michel</copyright>
    <pubDate>Wed, 20 Feb 2008 09:07:42 GMT</pubDate>
    <lastBuildDate>Thu, 21 Feb 2008 10:03:03 GMT</lastBuildDate>
    <item>
      <title>nouvelle 3</title>
      <link>http://page3.html</link>
      <guid isPermaLink="false">code unique 3</guid>
      <description>troisième essai d'un flux</description>
      <pubDate>Sat, 1 Mar 2008 10:00:00 GMT</pubDate>
    </item>
    <item>
      <title>...</title>
      ...
    </item>
  </channel>
</rss>
```

Ou UTF-8

les informations sur le flux lui-même

- * <title> : Le titre de votre flux .
 - * <link> : Le lien vers votre site web.
 - * <description> : La description de votre flux
- + quelques autres informations sur la langue et la date de publication (entre autre)

Chaque nouvelle
<title> titre
<description>
<lien>
...

Le format de la date de publication est un **format américain**, n'oubliez pas de traduire.

Pour les jours, Mon, Tue, Wed, Thu, Fri, Sat, Sun.

Pour les mois, Jan, Feb, Mar, Apr, May, Jun, Jul, Aug, Sep, Oct, Nov, Dec.

Deux règles sont à respecter cependant :

- Il ne faut pas précéder le numéro du jour d'un zéro (Ne faites pas "05 Aug" mais "5 Aug").
- Il faut précéder les heures d'un zéro (Ne faites pas "7:25:00 GMT" mais "07:25:00 GMT").



2) Ajoutez une ligne à vos pages 1,2,3 pour que **l'indicateur de flux (logo du feed RSS dans la barre d'adresse si vous êtes sous Firefox)** apparaisse dans les navigateurs récents. (code entre les balises head)

3) Créez un fichier xsl qui permet d'afficher le flux correctement pour les navigateurs qui ne possèdent pas d'agrégateur de flux.

Vous pouvez **lier une feuille de style XSLT à votre flux; Pour cela, ajoutez cette ligne (à ajuster) de code juste après la ligne de l'encodage (à la 2ème ligne donc):**

```
<?xml-stylesheet type="text/xsl" href="monStyles.xslt" ?>
```

La méthode de sortie HTML

```
<xsl:stylesheet version="1.0"
                xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<xsl:output method="html"/>
<xsl:template match="/">
  <html>
    <xsl:apply-templates/>
  </html>
</xsl:template>
...
</xsl:stylesheet>
```

Boucle

```
<!--***** boucle ***** -->
  <xsl:for-each select="item">
    ...<xsl:value-of select="title" />
    ...
  </xsl:for-each>
<!--***** fin boucle ***** -->
</xsl:stylesheet>
```

Les template rules sont des balises XSL permettant de définir des opérations à réaliser sur certains éléments du document XML

L'attribut "match" de la balise `<xsl:template>` permet de définir le ou les éléments du document XML sur lesquels s'applique la transformation.

Notons l'utilisation de `xsl:output method="html"`, qui comme son nom l'indique force le type du fichier en sortie.

La restructuration se fait de manière itérative : notre feuille XSL utilise `xsl:value-of` et une boucle `xsl:for-each`

Rem: pour inclure la feuille de style directement dans le fichier xsl, il suffit d'utiliser

```
<![CDATA[...]]>
```

```
<style type="text/css">
```

```
<![CDATA[
h1 {...}
]]>
```

```
</style>
```

Les possibilités déclaratives de XSL permettent le formatage des éléments sources basé sur la position, la valeur des attributs, l'ascendance/descendance

l'instruction `xsl:apply-templates` traite tous les fils du noeud courant, y compris les noeuds textuels



On peut éviter la boucle en utilisant `// //item` Indique tous les descendants d'un noeud

```
<xsl:apply-templates select="//item"/>
```

ou `<xsl:apply-templates select="/rss/channel/item"/>`

qui applique la même règle (ce n'est qu'un exemple) à tous les items et définir cette règle

```
<!-- Item template -->
```

```
<xsl:template match="item">
```

```
  <h2><a href="{link}"><xsl:value-of select="title"/></a></h2>
```

```
  <xsl:value-of select="description" />
```

```
</xsl:template>
```

Creez un fichier en php qui converti le flux en fichier html



Afficher des flux RSS 2.0 avec PHP

<http://www.libordux.org/doc/coding/rss.php>

On peut créer un petit script avec une fonction qui se chargera de scruter le fichier `.xml` pour en extraire les informations que l'on veut. Dans notre cas nous afficherons le titre `title`, le lien vers la page ou l'article `link`, la description `description`, et la date de publication `pubDate`. La fonction se nomme `readrss`.

Vous pouvez recopier le code directement dans une de vos pages ou créer un nouveau fichier, que vous appellerez ensuite grâce à la fonction `include` de PHP. Vous pouvez ajouter un peu de style CSS pour une meilleure présentation:

```
<?php
function readrss($fichier,$objets) {
    if($chaine = @implode("",@file($fichier))) {
        $tmp = preg_split("/<V?\".\"item\".\">/",$chaine);
        for($i=1;$i<sizeof($tmp)-1;$i+=2)
            foreach($objets as $objet) {
                $tmp2 = preg_split("/<V?\".$objet.\">/",$tmp[$i]);
                $resultat[$i-1][] = @$tmp2[1];
            }
        return $resultat;
    }
}
$rss = readrss("rss.xml",array("title","link","description","pubDate"));
foreach($rss as $stab) {
    echo '<div><p><tt>'.$stab[3].</tt><br />
    Post: <strong>'.$stab[0].</strong><br />
    '.$stab[2].<a href="'.$stab[1].'">Suite...</a>
    </p></div>';
}
?>
```

Pour les motivés !!

Voir l'exercice http://www.grappa.univ-lille3.fr/~jousse/enseignement/XML_XSLT/ex_xslt.html

