

RSS et syndication

Nouvelle technologie de veille et de diffusion de l'information

Michel Roland – Urfist PacaC Nice, le 23 février 2007



RSS et syndication

- Qu'est-ce que RSS ?
- Comment ça marche ?
- Quels contenus?
- Utiliser des flux RSS
 - Intégrer les fils RSS dans mon workflow
 - Importer des fils RSS dans mon contenu publié
- Plus si affinités:
 - Les formats
 - RSS & documentation
 - Agrégateurs
 - Produire des fils



Qu'est-ce que RSS?

- RSS:
 - Really Simple Syndication (RSS 2.0)
 - Rich Site Summary (RSS 0.91, RSS 1.0)
 - RDF Site Summary (RSS 0.9 and 1.0)
- RSS est un format de syndication de contenu Web, codé sous forme XML (Wikipedia).



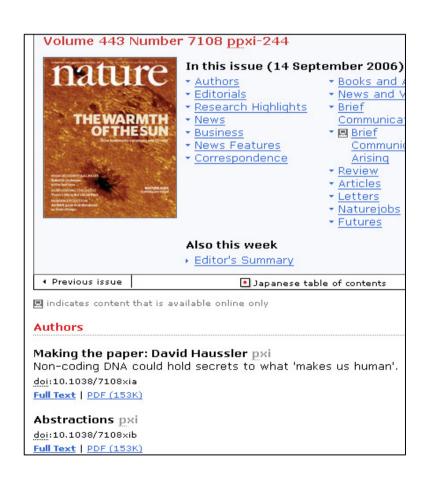


La syndication de contenu

- la syndication dans la presse (ex. <u>articles</u>, <u>comics</u>)
- plusieurs formats de fils (ou flux) de syndication: rss et ses diverses versions, Atom...
- la plupart des outils intègrent les différents formats



Exemple: la revue Nature



```
</rdf:Seq>
    </items>
  </channel>
- < image rdf: about="http://www.nature.com/includes/rj_globnavimages/na
    <title>Nature</title>
  - <ml>
      http://www.nature.com/includes/rj_globnavimages/nature_logo.gif
    </ur>
    link>http://www.nature.com/nature/</link>
  </image>
- <item rdf:about="http://dx.doi.org/10.1038/7108xia">
    <title>Making the paper: David Haussler</title>
    http://dx.doi.org/10.1038/7108xia
  - <description>
      Non-coding DNA could hold secrets to what 'makes us human'.
    </description>
    <dc:title>Making the paper: David Haussler</dc:title>
    <dc:identifier>doi:10.1038/7108xia</dc:identifier>
    <dc:source>Nature 443, xi (2006) </dc:source>
    <dc:date>2006-09-13</dc:date>
    prism:publicationName>Nature/prism:publicationName>
    <prism:publicationDate>2006-09-13</prism:publicationDate>
    <prism:number>7108</prism:number>
    <prism:section>Authors</prism:section>
```



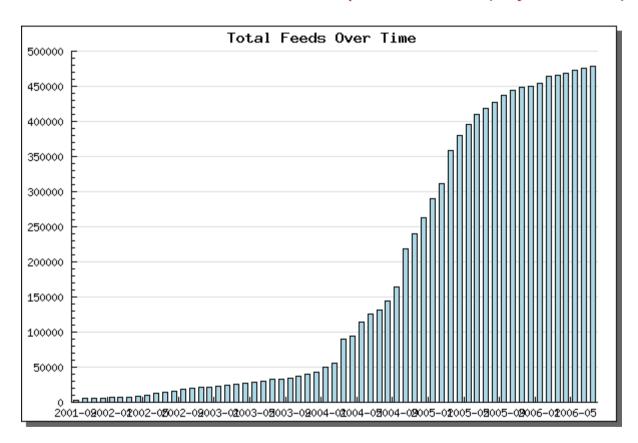
Pourquoi RSS?

Intérêts majeurs de rss :

- pouvoir consulter les mises à jour sans avoir à visiter chaque site
- Pouvoir consulter plusieurs mises à jour en un seul endroit → « guichet unique »
- Symétriquement proposer son contenu sous une forme manipulable par l'usager

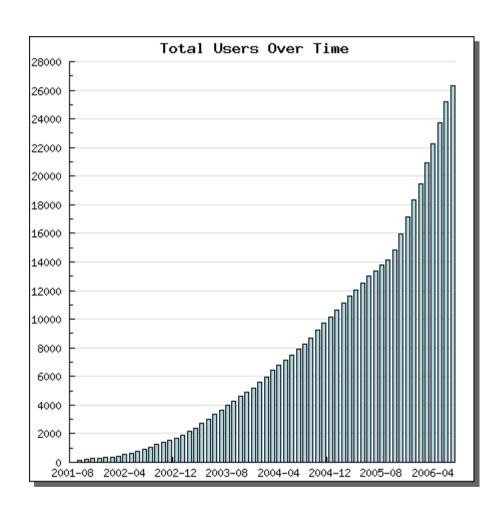


Augmentation des fils RSS proposés depuis septembre 2001 (-> juillet 2006)



Source: http://www.syndic8.com





Augmentation des utilisateurs des fils RSS depuis août 2001 (-> juillet 2006)

Source:

http://www.syndic8.com



Quels types de contenus ?

 De l'actualité : tous les sites de presse écrite et audiovisuelle, les agences de presse proposent plusieurs fils rss: Le Monde, TF1,

Liste des fils RSS du Monde.fr

Pour récupérer facilement les adresses des flux rss, vous pouvez les recopier depuis les boites texte ci-dessous.



http://www.lemonde.fr/rss/sequence/0,2-3208,1-0,0,xml

International

http://www.lemonde.fr/rss/sequence/0,2-321

> AFP - RSS



http://www.lemonde.fr/rss/sequence/0,2-321

A propos du flux RSS d'informations AFP

Les dernières informations de l'AFP sont désormais disponibles gratuitement et dans la mesure d'une utilisation individuelle et non commerciale, via le flux RSS d'informations AFP.

Quelle est la signification et l'utilité de la technologie RSS?

RSS est un format XML de syndication Internet de plus en

_E FIGARO

Grâce aux fils RSS du Figaro, vous pouvez désormais recevoir les derniers titres du Figaro, du Figaro Economie, du Figaro Magazine, de Madame Figaro et du Figaro Etudiant directement sur votre Pc ou Mac. Parmi la liste ci-dessous, choisissez le ou les fils qui correspondent à vos centres d'intérêts, puis cliquez sur l'icône du service que vous utilisez habituellement. Nous vous suggérons de personaliser votre page Yahoo!, Google ou Netvibes. Vous pouvez également evodiquer les fils dans un "agrégateur BSS"

Nos flux RSS d'informations

RSS AFP - L'actualité (Français)

RSS AFP - Wire News (English)

RSS AFP - La Actualidad (Español)

RSS AFP - Topthemen (Deutsch)

RSS AFP - Wire News (Arabic)

RSS AFP - Actualidade (Portugues)



Quels types de contenus ?

- Des informations pratiques :
 - Météo (<u>Weather channel</u>, <u>National hurricane</u>
 <u>center</u>...), Trafic routier (<u>www.traffic.com</u>...)
 - Offres d'emploi (<u>Keljob</u>, <u>Michael Page</u>,
 <u>Directemploi</u>, ...)
 - Loisirs: bandes-annonces, programmes tv/radio (BBC ...), sites touristiques...
- Des contenus documentaires



Les contenus documentaires

- Bases de données
 - Publiques : <u>Pubmed</u>, certains catalogues de bibliothèques, des bases de signets (<u>Intute</u>, <u>signets du</u> <u>CERIMES</u>, ...)
 - Payantes : Business Source Premier, Compendex, Inspec, Emerald, Proquest...





Les contenus documentaires

- Revues scientifiques
 - American Chemical Society
 - American Institute of Physics
 - American Physical Society
 - Blackwell
 - BioMed Central
 - BMJ Journals
 - Cambridge Journals
 - Cell Press

- EDP Sciences
- Institute of Physics
- <u>Lippincott Williams &</u>
 Wilkins
- Nature Publishing Group
- Oxford University Press
- Springer
- ..
- Surtout sciences, mais aussi sciences sociales
 - Droit



Les contenus documentaires

- Sites web
 - Information officielle
 - Vie-publique.fr, Premier ministre, ...
 - Education nationale, Science.gouv.fr, AMUE, ...
 - <u>Edufeeds</u>: les fils des universités et des grandes écoles US
 - <u>FirstGov rss directory</u>: annuaire des fils rss du gouvernement US
 - Blogues professionnels
 - d'enseignants-chercheurs (<u>en droit</u>, <u>en économie</u>, <u>en sciences</u>...)
 - de <u>bibliothèques</u>
 - comme support pédagogique





Our Catalogue | Ask Us

*

Electronic Journals Beginning with A



A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

There are 41 journals beginning with A with an RSS feed

Journals 1 to 41 of 41

- 1. Abacus
- 2. Accounts of chemical research
- 3. Acta crystallographica. Section A, Foundations of crystallography
- Acta crystallographica. Section B, Structural science
- 5. Acta crystallographica. Section C, Crystal structure communications
- 6. Acta crystallographica. Section D, Biological crystallography
- 7. Acta numerica
- 8. Acta zoologica
- 9. African affairs
- 10. Age and ageing

Wondering what the icon means?

http://library.usask.ca:9003/ejournals/rss_feeds.php



Comment ça marche?

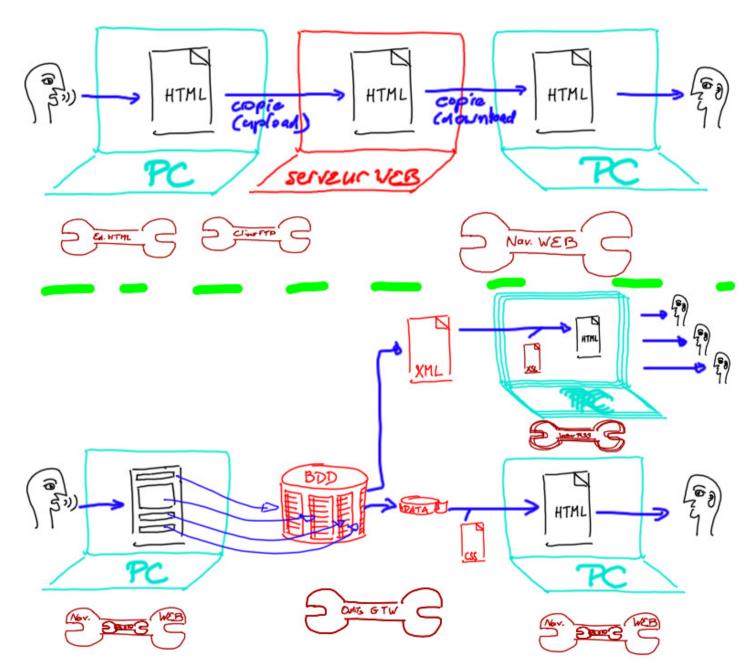
- Choix d'un agrégateur pour lire les contenus
 - Local ou en ligne
- Repérage des fils/flux rss
 - Sur des annuaires, des moteurs de recherche
- Abonnement aux fils rss
 - Ajout des fils par un simple copier-coller





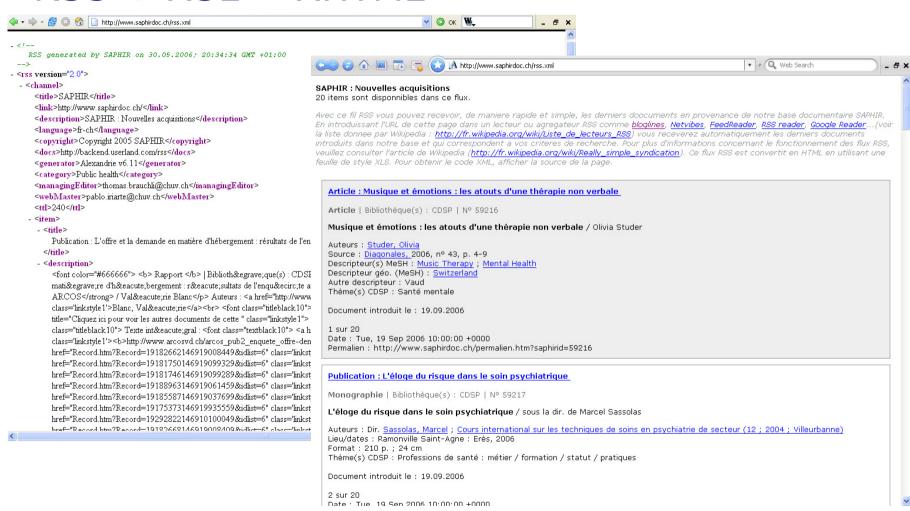


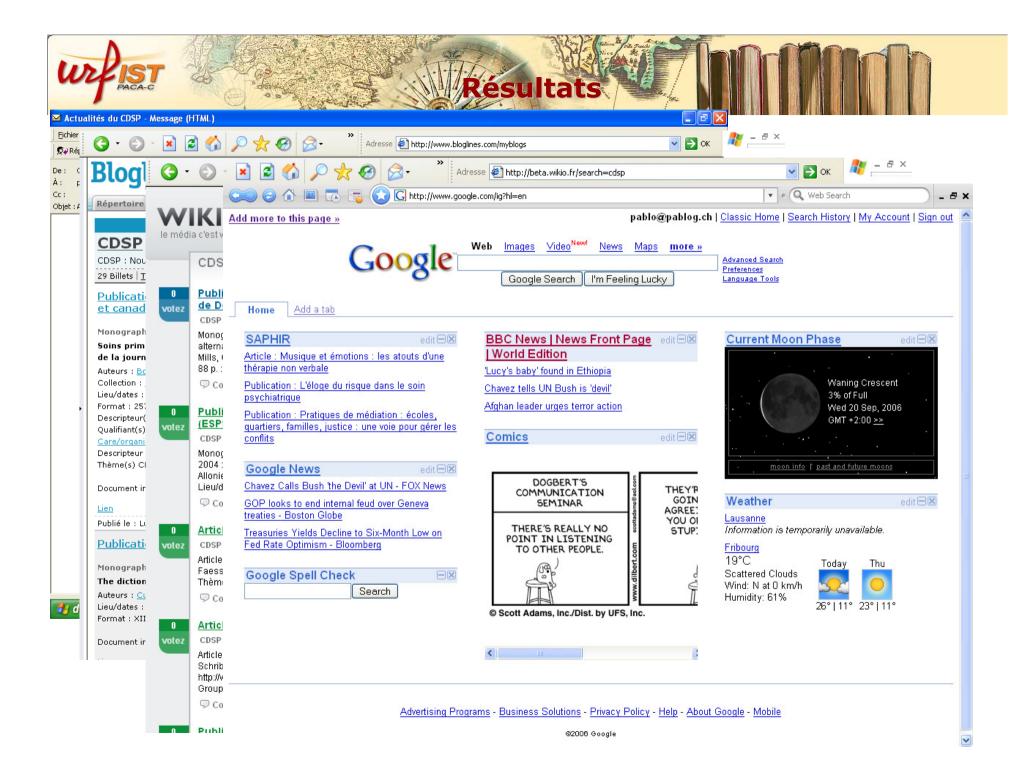






RSS + XSL = XHTML







RSS et blogosphère

- Le fil rss est un outil qui permet de surveiller les mise à jour des pages « dynamiques »
- La plupart des plateformes de blogues intègrent la production de fils rss
- Les blogues sont ainsi plus faciles à surveiller que d'autres types de sites
- La technologie rss est un des facteurs expliquant la constitution du sous-ensemble du web appelé la « blogosphère »



Les formats de syndication



RSS: Really Simple Syndication / RDF Site Summary

Format de syndication crée en 1997. La structure RSS est une enveloppe XML qui normalise la structure des données d'un article ou plus largement d'un objet web.

Il existe différentes versions : RSS 0.91 RSS 1.0 et RSS 2.0

Atom ATOM XML FEED

Atom est le successeur du format RSS, plus complet et basé sur RDF.

Atom a été <u>standardisé par l'Internet Engineering Task Force</u>, (IETF est une organisation de standardisation qui aide à définir l'Internet, et qui se soucie avant tout de permettre l'écriture et la lecture de contenus dans un format unique.)





Naissance d'un nouveau moyen de diffusion

1992 : CERN, What's new?

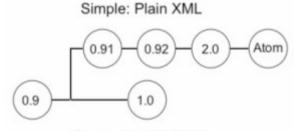
1997 : Dave Winer, syndication XML

1999 : Netscape, RDF Site Summary

2002 : RSS 2.0 « Winer »

- 2003 : « blog bang » et extension de RSS aux médias en ligne. Format Atom
- 2005 : web 2.0 et implémentation de RSS dans les BD (A&I) archives, moteurs...

(Pablo Iriarte - BiUM/CDSP - CHUV Lausanne)



Extensible: RDF/XML



RSS 0.91. Cest le premier pour Netscape Viner pour Netscape

```
<rss version="0.91">
- <channel>
    <title>WriteTheWeb</title>
    link>http://writetheweb.com</link>
    <description>News for web users that write back</description>
    <language>en-us</language>
    <copyright>Copyright 2000, WriteTheWeb team.</copyright>
    <managingEditor>editor@writetheweb.com</managingEditor>
    <webMaster>webmaster@writetheweb.com</webMaster>
  - <image>
       <title>WriteTheWeb</title>
       <url><url>http://writetheweb.com/images/mynetscape88.gif</url></url>
       link>http://writetheweb.com</link>
       <width>88</width>
       <height>31</height>
       <description>News for web users that write back</description>
    </image>
  <item>
       <title>Giving the world a pluggable Gnutella</title>
       http://writetheweb.com/read.php?item=24</link>
    - <description>
         WorldOS is a framework on which to build programs that work like Fr
       </description>
    </item>

    ≤item≥.

       <title>Syndication discussions hot up</title>
       http://writetheweb.com/read.php?item=23</link>
    - <description>
         After a period of dormancy, the Syndication mailing list has become act
       </description>
    </item>
  <item>
    < <title>
         Personal web server integrates file sharing and messaging
       </title>
       http://writetheweb.com/read.php?item=22</link>
    - <description>
         The Magi Project is an innovative project to create a combined person;
```



</rdf:RDF>

RSS 1.0: c'est un format écrit par un consortium ouvert de développeurs pour faire évoluer RSS 0.91 en intégrant diverses normes mondiales du consortium du Web (W3C), y compris RDF.

RSS 1.0 ajoute par exemple l'utilisation des espaces de nommage XML. Ces derniers vous permettent de combiner différentes définitions XML dans un même document.

<?xml version="1.0" encoding="utf-8"?> <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdfsvntaxes#" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:sy="http://purl.org/rss/1.0/modules/syndication/" xmlns:co="http://purl.org/rss/1.0/modules/company/" xmlns:ti="http://purl.org/rss/1.0/modules/textinput/" xmlns="http://purl.org/rss/1.0/" > <channel rdf:about="http://xxx.com/? fl=rss1.0"> <title>Meerkat</title> <link>http://meerkat.oreillynet.com</link> <description>Meerkat: An Open Wire Service</description> <dc:publisher>The O'Reilly Network</dc:publisher> <dc:creator>Rael Dornfest > <dc:rights>Copyright © 2000.</dc:rights> <image rdf:resource="http://xxx.ipg" /> <items> <rdf:Sea> <rdf:li resource="http://xxxxxxxx" /> </rdf:Seq> </items>

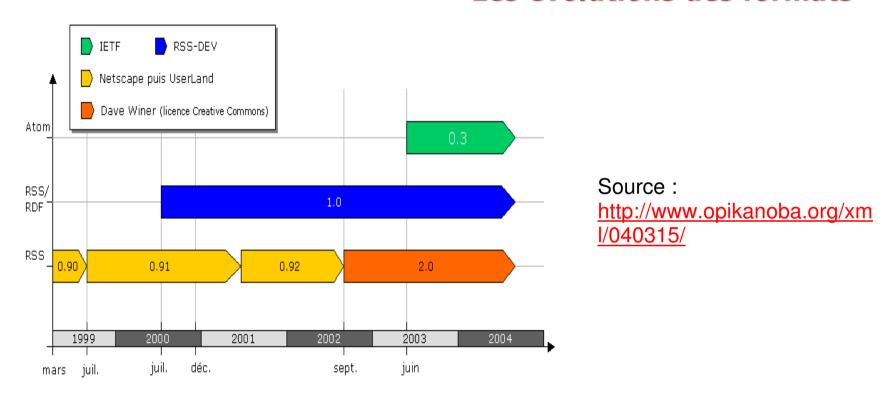


ASS 2.0 ajoute quelques ballises au format ASS 0.97.

```
<rss version="2.0">
< <channel>
    <title>Nouvelles acquisitions de la BU </title>
    http://www-scd.univ-lyon3.fr</link>
  - <description>
      Tout sur les bibliothèques et la documentation à Lyon 3
    </description>
    <copyright>© Scd Lyon 3 2005</copyright>
    <language>fr</language>
  - <image>
       <title>SCD Université Lyon 3</title>
       <ur>url>http://www-scd.univ-lyon3.fr/nouvacq/logo.GIF</url></url>
       http://www-scd.univ-lyon3.fr</link>
    </image>
    <pubdate>Sun, 18 Jun 2006 12:00:08 +0100</pubdate>
  <item>
     <title>
         Karlson Léon. - Vade-mecum du germaniste . - MAN / 438 KAR.
       </title>
     < <li>< link >
         http://multimedia.univ-lyon3.fr/lyon3/recherche/executerRechercheprogress.asp?b.
```



Les évolutions des formats



Pour en savoir plus sur les spécifications et les balises des différents formats : http://www.stervinou.com/projets/rss



OPML

- OPML (Outline Processor Markup Language) est un format XML permettant de regrouper et d'identifier les titres (outlines) d'un texte. Multiples usages, le plus commun d'entre eux étant l'échange de listes de flux RSS
 - Récupérer une sélection faite par une autre personne
 - Migration de ses flux rss d'un outil à l'autre



OPML

```
<opml version="1.1">
  <head>
    <title>FeedDemon Subscriptions</title>
    <dateModified>Wed, 21 Feb 2007 12:15:15 GMT</dateModified>
  </head>
  <body>
    <outline text="Blogs and Blogging">
      <outline text="Salima Naji" title="Salima Naji" type="rss" xml</pre>
      <outline text="InternetActu.net" title="InternetActu.net" type</pre>
      <outline text="le blog du Secours populaire français Comments'</pre>
      <outline text="The Shifted Librarian" title="The Shifted Libra</pre>
      <outline text="Standblog" title="Standblog" type="rss" xmlUrl=</pre>
      <outline text="inFLUX &gt; rechercher sur le net" title="inFI</pre>
      <outline text="Faut le savoir" title="Faut le savoir" type="rs</pre>
      <outline text="LA-BLOG" title="LA-BLOG" type="rss" xmlUrl="htt</pre>
      <outline text="Expressions.be - la 3ème vie... - the third lit</pre>
      <outline text="Journal d'un avocat" title="Journal d'un avocat</pre>
      <outline text="Métaphore, il fait froid dehors !" title="Métax</pre>
```

<u>exemple</u>



OPML (outils)

- OMPL Manager
- Bitty.com
- Share Your OPML
- OPMLgenerator





Intégrer les fils rss dans mon workflow

- Intégration avec les outils déjà utilisés
 - Portail personnalisé (type Yahoo! ou Google)
 - Navigateur internet (Firefox, Sage pour Firefox, IE bientôt...)
 - Outil de messagerie
 - Installation locale (Feedreader, Wticker...)
 - Service en ligne (Bloglines, Netvibes, Google IG...)
- RSS > e-mail & e-mail > RSS
 - Rss par Thunderbird
 - RssFwd, RMail
 - Toutes les listes de diffusion du CRU proposent un fil RSS



Agrégateurs

d'après:

 Marie-Hélène Vézina & Sylvie Michaud (Direction des Bibliothèques de l'UdeM)



Comment fonctionne un agrégateur?

- Un fichier XML mis à disposition par le producteur d'information
- À l'aide d'un agrégateur, on s'abonne au fil RSS du producteur
- L'agrégateur fait régulièrement le tour de tous les fils et signale les nouveautés
- Le résultat : une revue de presse personnalisée quoi!



Intégrer les fils rss dans mon site

- Produire des fils rss
 - Augmenter la visibilité de ses contenus
 - Générer du trafic sur son site
- Syndiquer des fils rss
 - Récupérer de l'information produite par d'autres
 - Cibler l'information récupérée



Des fils partout! Hennepin County Library





« Nombreuses applications en bibliothèque, à la fois pour informer le public et lui rendre de nouveaux services au quotidien, et en appui du travail des bibliothécaires, principalement comme outil de veille et d'information. »

ENSSIB. *Utilisation des fils RSS en bibliothèque* http://memsic.ccsd.cnrs.fr/mem 00000249.html



Produire un flux RSS

- Publication intégrée à l'outil de gestion de contenu
 - Paramétrer
 - FeedBurner
- Produire un flux RSS par un script
 - Création et lecture de flux RSS 2.0 en PHP



Syndiquer un flux RSS

- Fonctionnalité intégrée à l'outil de gestion de contenu
- Fonctionnalité intégrée à la publication du flux (ex. Google shared items)
- RSS > JavaScript
- etc.



Un outil multifonctions

<u>xFruits</u>

- Agréger
- -RSS > html
- RSS > mail
- RSS > mobile
- -Mail > RSS
- Etc.



En résumé...

- Avantages de RSS
 - L'information vient à l'utilisateur
 - Sans spam
 - Concentrée sur les contenus mis à jour
- Inconvénients
 - Nécessite l'usage d'un outil spécifique
 - Perte du contexte
 - Addiction



Des questions?

Michel Roland – URFIST PACA-C >> urfist@unice.fr

Pour continuer:

- Sur le site de l'Urfist de Nice: Quelques présentations
- Sur le site de l'Urfist de Paris: Modalités d'appropriation des fils RSS



Webographie

- ENSSIB. *Utilisation des fils RSS en bibliothèque*. Juillet 2005.
 - <http://memsic.ccsd.cnrs.fr/mem_00000249.html>
- Morin, N. RSS: quels usages pour les bibliothèques.
 Octobre 2005. http://morinn.free.fr/articles/>
- Raphael, L. RSS: un fil à suivre pour les documentalistes. Novembre 2004.
 https://dspace.msh-alpes.prd.fr/handle/1801/441>
- URFIST. Modalités d'appropriation des fils RSS, blogs et wikis pour une veille informationnelle active. Novembre 2005.
 - <http://www.ext.upmc.fr/urfist/rss/>
- RSS4Lib: Blogue traitant d'utilisations inovantes de RSS pour les bibliothèques
 - < http://blogs.fletcher.tufts.edu/rss4lib>