

RSS 1.0

[Eric van der Vlist](#)

XML expert (?) and RDF newbie.

vdv@dyomedea.com

RSS 1.0 and its taxonomy module.

Bringing Metadata back into RSS

23 October 2000

[Eric van der Vlist](#) - RSS 1.0 and its taxonomy module. - 23 October

2000 - slide 1/13

© Dyomedea 2000. Reproduction interdite.

Plan

RSS 1.0	1
Plan	2
RDF Rich Site Summary	3
Working Group members	4
The RSS 1.0 proposal	5
Other design goals	6
Example	7
First conclusions	8
From DC to taxonomy.	9
Extracting semantic from XMLNews-Story	10
rdfDB	11
Closing the loop.	12
Preliminary conclusions	13

RDF | Rich Site Summary

RDF = "RDF|Rich Site Summary"

0.9 introduced in 99 by Netscape as a RDF vocabulary.

Rapidly followed by 0.91, with more features and no RDF.

Its usage has been derived to headlines syndication (Userland, Moreover, O'Reilly, ...).

Seemed to reach a dead end mid-2000.

No real focus after the many additions of 0.91

Many requests for improvement without structures/criterias to advance them.

No way to add metadata.

[Eric van der Vlist](#) - RSS 1.0 and its taxonomy module. - 23 October

2000 - slide 3/13

© Dyomedeia 2000. Reproduction interdite.

Working Group members

RSS 1.0 Working Group Members:

Gabe Beged-Dov
Dan Brickley
Rael Dornfest
Ian Davis
Leigh Dodds
Jonathan Eisenzopf
David Galbraith
R.V. Guha
Ken MacLeod
Eric Miller
Aaron Swartz
Eric van der Vlist

Working Group Chair:

Rael Dornfest

[Eric van der Vlist](#) - RSS 1.0 and its taxonomy module. - 23 October
2000 - slide 4/13
© Dyomedeia 2000. Reproduction interdite.

The RSS 1.0 proposal

<http://purl.org/rss/1.0/> (CR)

Definition of a refocused core.


Core functionalities equivalent to RSS 0.9

Definition of namespaces based modules.

To let people add their functionalities without impacting the core.

Reintroduction of RDF.

To enable addition of metadata.

 Modules are also defined to cover the functionalities of RSS 0.91 and to add Dublin Core elements.

[Eric van der Vlist](#) - RSS 1.0 and its taxonomy module. - 23 October

2000 - slide 5/13

© Dyomedea 2000. Reproduction interdite.

Other design goals

The vocabulary should be as simple as possible for simple XML tools.


Simple to explain and simple to parse means that one RDF syntax is chosen and the other are forbidden.

Also true for URIs that are compared as strings and not allowed to be relative.

Attributes are used only for `rdf:about` and `rdf:resource`.

The data model should be as similar as possible for both XML and RDF applications.

Literals are not used in the core nor in the basic modules.

 Both design goals are often contradictory (ex of item ordering).

[Eric van der Vlist](#) - RSS 1.0 and its taxonomy module. - 23 October 2000 - slide 6/13

© Dyomedeia 2000. Reproduction interdite.

Example

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  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://meerkat.oreillynet.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 (mailto:rael@oreilly.com)</dc:creator>
    <dc:rights>Copyright © 2000 O'Reilly & Associates, Inc.</dc:rights>
    <dc:date>2000-01-01T12:00:00</dc:date>
    <sy:updatePeriod>hourly</sy:updatePeriod>
    <sy:updateFrequency>2</sy:updateFrequency>
    <sy:updateBase>2000-01-01T12:00:00</sy:updateBase>

    <image rdf:resource="http://meerkat.oreillynet.com/icons/meerkat-powered.jpg" />

    <items>
      <rdf:Seq>
        <rdf:li resource="http://c.moreover.com/click/here.pl?r123" />
      </rdf:Seq>
    </items>

    <textinput rdf:resource="http://meerkat.oreillynet.com" />
  </channel>

  <image rdf:about="http://meerkat.oreillynet.com/icons/meerkat-powered.jpg">
    <title>Meerkat Powered!</title>
    <url>http://meerkat.oreillynet.com/icons/meerkat-powered.jpg</url>
    <link>http://meerkat.oreillynet.com</link>
  </image>

  <item rdf:about="http://c.moreover.com/click/here.pl?r123">
    <title>XML: A Disruptive Technology</title>
    <link>http://c.moreover.com/click/here.pl?r123</link>
    <dc:description>
      XML is placing increasingly heavy loads on the existing technical
      infrastructure of the Internet.
    </dc:description>
  </item>
</rdf:RDF>
```

```
</dc: description>
<dc: publisher>The O'Reilly Network</dc: publisher>
<dc: creator>Simon St. Laurent (mailto:si.monstl@si.monstl.com)</dc: creator>
<dc: rights>Copyright © 2000 O'Reilly & Associates, Inc.</dc: rights>
<dc: subject>XML</dc: subject>
<co: name>XML.com</co: name>
<co: market>NASDAQ</co: market>
<co: symbol>XML</co: symbol>
</item>

<textinput rdf: about="http://meerkat.oreillynet.com">
  <title>Search Meerkat</title>
  <description>Search Meerkat's RSS Database...</description>
  <name>s</name>
  <link>http://meerkat.oreillynet.com/</link>
  <ti:function>search</ti:function>
  <ti:inputType>regex</ti:inputType>
</textinput>

</rdf: RDF>
```

[Eric van der Vlist](#) - RSS 1.0 and its taxonomy module. - 23 October 2000 - slide

7/13

© Dyomedeia 2000. Reproduction interdite.

First conclusions

Strong pushback on RDF.


Sometimes seems irrational and linked to the very RDF "brand".

The data model (infoset) is what is important.

Discussions on the syntax are counter-productive before a data model is established.

XSLT is placing us on a moving ground.

Since it's so easy to "generate semantics" from a known vocabulary, why should one bother to write RDF tags by hand?

 Should we define a canonical format and several derived formats with their transformations ?

[Eric van der Vlist](#) - RSS 1.0 and its taxonomy module. - 23 October

2000 - slide 8/13

© Dyomedeia 2000. Reproduction interdite.

From DC to taxonomy.

The (WD) taxonomy module will allow associating metadata to DC elements.

The syntax isn't fixed yet.

The current one is subject to discussion:

```
<dc: subject>  
  <rdf: Description resource="http://xml fr. org/i ndex/obj ect. ti tle/DOM/">  
    <rdfs: Label >DOM</rdfs: Label >  
  </rdf: Descri ption>  
</dc: subj ect>
```

But already allows some experiments.

[Eric van der Vlist](#) - RSS 1.0 and its taxonomy module. - 23 October 2000 -
slide 9/13

© Dyomedea 2000. Reproduction interdite.

Extracting semantic from XMLNews-Story

XMLNews-Story:

```
<person>Paul Tchi stopol ski </person> a publi é  
<a href="http://www.paul t. com/prod/Hi awatha"><obj ect. ti tle>Hi awatha</obj ect. ti tle></a>  
0.1, un <q>Serveur Web pour ceux qui parlent couramment  
<obj ect. ti tle>XSL</obj ect. ti tle></q>  
écrit en <obj ect. ti tle>Java</obj ect. ti tle> et distri bué sous licence  
<obj ect. ti tle>open source</obj ect. ti tle>.
```

Can easily generate (through XSLT):

```
<i tem rdf: about="http://xml fr. org/actual ites/tech/001120-0001. xml ">  
<ti tle>Un nouveau serveur Web Java.</ti tle>  
<li nk>http://xml fr. org/actual ites/tech/001120-0001. xml </li nk>  
<dc: descri ption>Paul Tchi stopol ski a publi é Hi awatha 0.1, un Serveur Web pour  
ceux qui parlent couramment XSL écrit en Java et distri bué sous licence  
open source.</dc: descri ption>  
<dc: creator>Eric van der Vli st, Dyomede a (vdv@dyomede a. com).</dc: creator>  
<dc: date>2000-11-20</dc: date>  
<dc: subj ect>  
<rdf: Descri ption resource="http://xml fr. org/i ndex/obj ect. ti tle/Java/">  
<rd fs: labe l>Java</rd fs: labe l>  
</rdf: Descri ption>  
</dc: subj ect>  
<dc: subj ect>  
<rdf: Descri ption resource="http://xml fr. org/i ndex/person/Paul _Tchi stopol ski i /">  
<rd fs: labe l>Paul Tchi stopol ski i </rd fs: labe l>  
</rdf: Descri ption>  
</dc: subj ect>  
<dc: subj ect>  
<rdf: Descri ption resource="http://xml fr. org/i ndex/obj ect. ti tle/XSL/">  
<rd fs: labe l>XSL</rd fs: labe l>  
</rdf: Descri ption>  
</dc: subj ect>  
<dc: subj ect>  
<rdf: Descri ption resource="http://xml fr. org/i ndex/obj ect. ti tle/open_source/">  
<rd fs: labe l>open source</rd fs: labe l>  
</rdf: Descri ption>  
</dc: subj ect>
```

rdfDB

These RSS feeds are loaded into a rdfDB database. Related key words are proposed for each article:



The screenshot shows a website header with navigation links: PARTENAIRES, INFOS, WAP, DOCUMENTATION. Below the header, there is a section titled "Mots clés." (Keywords) with a list of related terms: Apache, Blackdown, C, Dyomedeia, Eric van der Vlist, Hiawatha, and HTML. To the left of the keywords, there is a partial view of another section titled "Serveur écrit en" (Server written in).

With link to a dynamic web page corresponding to the URI of our taxonomy system:



The screenshot shows a website header with navigation links: ACTUALITÉ, DISCUSSIONS, EMPLOIS, LIENS, PARTENAIRES. Below the header, there is a section titled "Apache" with the text "Nombre de page(s) : 5". Below this, there is a list of three articles:

- [Un nouveau serveur Web Java.](#) (2000-11-20)
Paul Tchistopolskii a publié Hiawatha 0.1, un Serveur Web pour ceux qui parlent couramment XSL écrit en Java et distribué sous licence open source.
- [RDF vous guide sur XMLfr](#) (2000-09-19)
XMLfr vous propose un nouvel index thématique, basé sur l'utilisation de rdfDB, une base de données RDF.
- [Module XSI T Apache.](#) (2000-07-21)

Closing the loop.

The query is sent and the results are processed within a XSLT transformation:

```
<xsl:template match="rdfDB">
  <xsl:param name="uri" select="$request-tree/request/variables/variable[name='uri']/value"/>
  <xsl:param name="request" select="concat(
    'select ?label ?item ?title ?date ?description from rss where
    (resource ?des http://xmlfr.org', $uri, ')
    , (http://www.w3.org/2000/01/rdf-schema##label ?des ?label)
    , (http://purl.org/dc/elements/1.1/#subject ?item ?des)
    , (http://purl.org/dc/elements/1.1/#date ?item ?date)
    , (http://purl.org/dc/elements/1.1/#description ?item ?description)
    , (http://purl.org/rss/1.0/#title ?item ?title)')"/>

  <xsl:variable name="rtf"><xsl:copy-of select="rdfDB:pool edRequest($request)"/></xsl:variable>
  <xsl:variable name="ns" select="xt:node-set($rtf)"/>
  <h3 class="titre-fenetre"><xsl:value-of select="$ns/result-set/row[1]/col [name='label']/value"/></h3>
  <p>Nombre de page(s) : <xsl:value-of select="count($ns/result-set/row)"/></p>
  <ul><xsl:apply-templates select="$ns/result-set/row" mode="rdfDB">
    <xsl:sort select="col [name='date']/value" order="descending"/>
  </xsl:apply-templates></ul>

</xsl:template>
```

See <http://4xt.org/downloads/JrdfDB/>

[Eric van der Vlist](#) - RSS 1.0 and its taxonomy module. - 23 October 2000 - slide 12/13

© Dyomedeia 2000. Reproduction interdite.

Preliminary conclusions

RDF can be easy!

A specific RDF query syntax is really useful.

It's missing sorting and aggregation feature.

How can I evaluate the "distance" between articles or between keywords?

The tool I have used is very fast but its administration features are limited.