

IMPROVING ACCESSIBILITY OF NIGERIAN NEWSPAPER WEBSITES THROUGH RSS FEEDS

Ayo C. K. (Ph. D)*

and

Adeyeye O. M. (B. Sc)**

*Dr. Charles K. Ayo holds B.Sc, M.Sc and a Ph.D in Computer Science. His research interests include: Mobile Computing, Interest Programming, etc. He is member of the Nigerian Computer Society. He is currently the head of Computer and Information Science Department of Covenant University, Ota, Ogun State, Nigeria. E-mail: ckanyo@yahoo.com

**O. M. Adeyeye (B.Sc.) Postgraduate (M.Sc) Student, Department of Electrical and Information Engineering, Covenant University, Ota, Ogun State, Nigeria. E-mail: micadeyeye@yahoo.com

ABSTRACT

*Really Simply Syndication or Rich Site Summary (RSS) now becomes common way of summarizing content of an information website. This is based on Resource Description Format (RDF), which is an extensible markup language(XML) standard for analyzing information. Many an organization such as CNN, YAHOO generates summary of the content of her website in this xml format called *rss feed*, which can be read by news aggregators such as *rss reader* and *feed reader* or processed by some programming languages. News aggregators make it possible to get updated information from various websites simultaneously without visiting the websites.*

*Today in Nigeria, there are over twenty national dailies. While few are online, yet it is mandatory to visit these websites individually to get latest information. To improve on the usability of these sites, it becomes necessary to generate *rss feeds* for those who prefer getting information from these websites using news aggregators.*

The paper shows how to generate and display an xml document called rss feed using one of the common Nigerian Newspapers. In addition, it looks into benefits of providing rss feeds by information service websites.

KEYWORDS: XML-RPC, News Aggregator, XMI, RSS, RDF, Web Service

1.0 INTRODUCTION

RSS previously stood for Rich Site Summary and was developed by Netscape. At present, it is referred to as Resource Description Framework (RDF) Site Summary. RSS is an XMI document format intended to describe, summarize, and distribute the contents of

a website as a 'channel' [1]. Emphatically, RSS is a format for syndicating news and the content of news like sites. The name 'RSS' is an umbrella term for a format that spans several different versions of at least two different (but parallel) formats [2]. The versions are summarized below.

Version	Owner	Pros	Status	Recommendation
0.90	Netscape		Obsoleted by 1.0	Don't use
0.91	UserLand	Drop dead simple	Officially obsoleted by 2.0, but still quite popular	Use for basic syndication. Easy migration path to 2.0 if you need more flexibility
0.92, 0.93, 0.94	UserLand	Allows richer metadata than 0.91	Obsoleted by 2.0	Use 2.0 instead
1.0	RSS DEV Working Group	RDF based, extensibility via modules, not controlled by a single vendor	Stable core, active module development	Use for RDF based applications or if you need advanced RDF specific modules
2.0	UserLand	Extensibility via modules, easy migration path from 0.9x branch	Stable core, active module development	Use for general purpose, metadata rich syndication

TABLE 1: RSS Versions and Recommendations

Table 1.0 shows the various RSS versions, owners, pros, status and recommendations. As an xml document, to display RSS document requires the availability of XML/XSL processor or parser. This is achieved by a using latest browsers or RSS-aware programs called news aggregators such as RSS reader, FeedReader and many more. Also, XML-RPC can be used to display the RSS document or feed, a process known as **Web service**. The "RPC" in XML-RPC stands for Remote Procedure Call. A client system "calls" a "procedure" (or method) on a remote server. The concept of a Remote Procedure Call is general to software development, and one which we come across in many places. XML-RPC is an XML standard for data exchange to be used in a Remote Procedure Call [3].

While HTML is a markup language with a fixed set of tags that allows users to specify the appearance of a document, XML allows the user to define and

create new tags in order to provide a document structure appropriate to the task at hand [3]. Hence, XML is referred to as a meta language that is, a describing markup language [4]. Fig 1.0 shows HTML and XML document systems. For HTML, the same document carries information relating to appearance, content, and visual structure while XML has different means employed for defining its structure, appearance and content. Usually, Document Type Definition (DTD) or Schema file describes the structure of a document; the appearance is specified by an eXtensible style sheet (XSL), and the content is provided in the XML document [3]. Such content often times is text information. RSS file as an XML document simply provides content (summary of a website) while the structure and appearance are governed by browser or news aggregator.

2.0 DESIGN METHODOLOGY

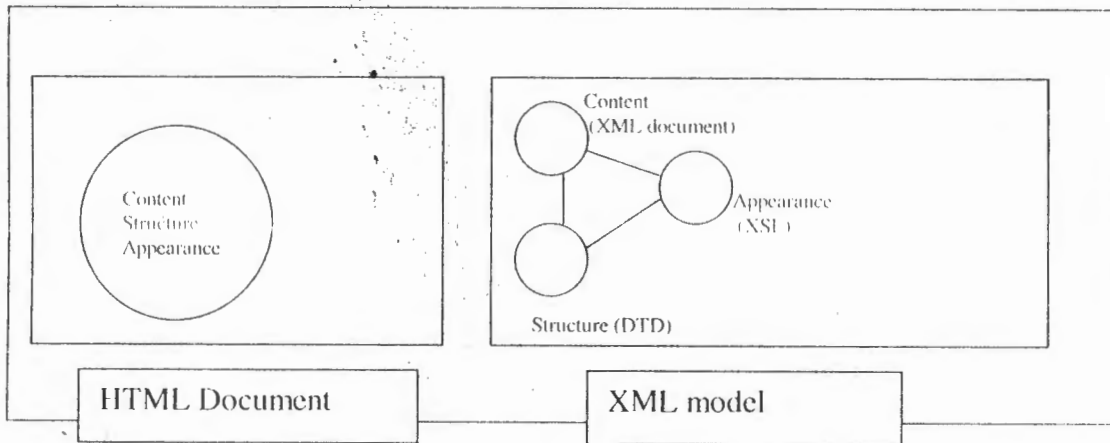


Fig 1 HTML and XML document systems

From table 1, there are different versions of RSS. The commonly used is RSS 1.0 developed by RSS-DEV Working Group. Like HTML, RSS follows a chronological order when generating or writing the codes. Each version of RSS

has its approach though similar to one another. Shown below are the diagrammatic differences of writing RSS 0.91, 1.0 and 2.0.

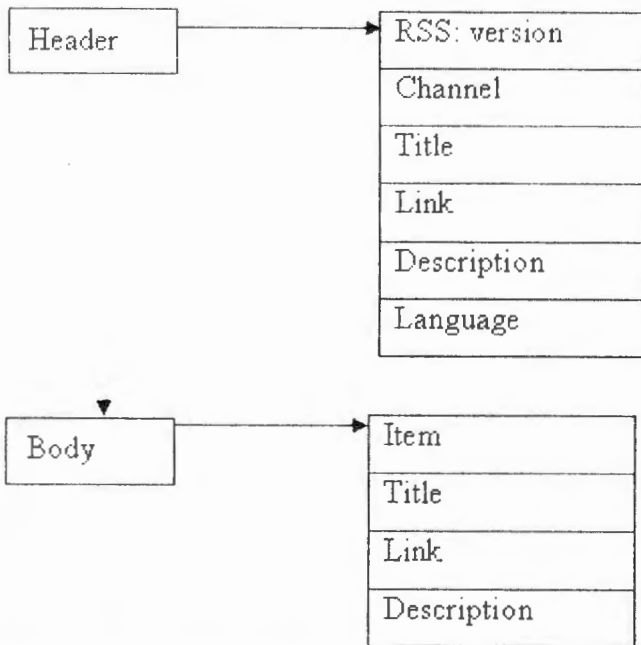


Fig 2 RSS 0.91 Format

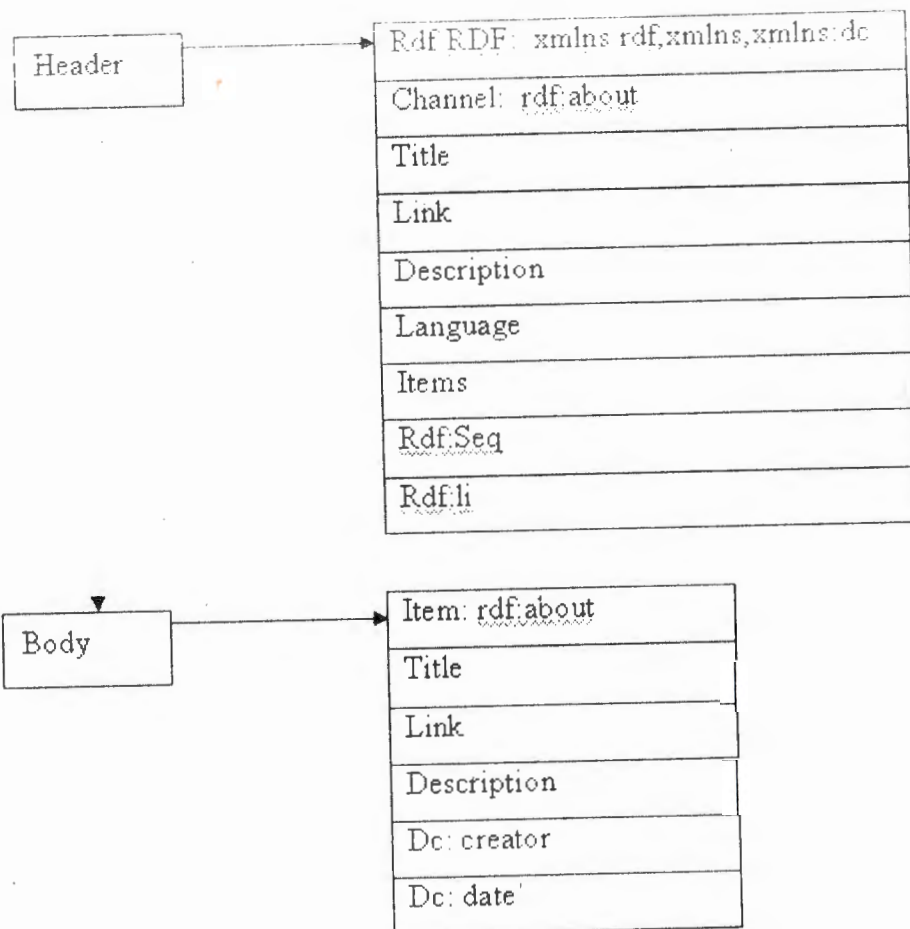


Fig 3 RSS 1.0 Format

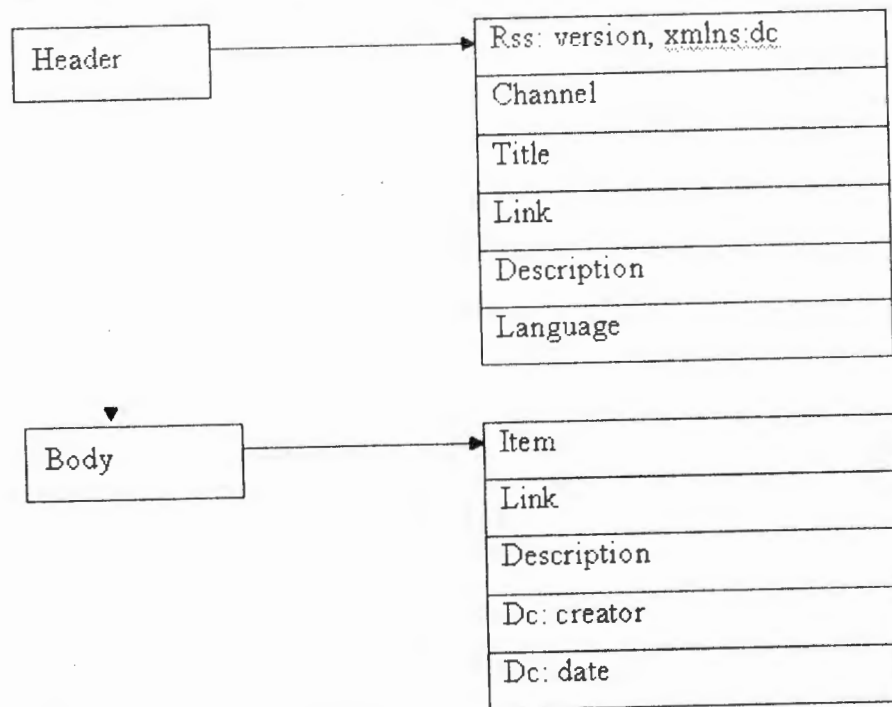


Fig 4 RSS 2.0 Format

CASE STUDY: How to generate RSS 1.0 for Nigerian Guardian Newspaper

Guardian Newspaper has enormous columns which are updated daily. This paper focuses on generating RSS 1.0 for a column named "education". In addition, an excerpt of the index source code is used in this case. There are two widely-used methods for programming

languages to read XML documents. These are event-based Application Programming Interfaces (APIs) and Document Object Model (DOM) APIs methods. Alternatively, a news aggregator can be used to read this XML document called RSS. In this paper, a news aggregator is used to display the produced RSS feed.

3.0 SYSTEM DESIGN

The RSS Feed generated from Education Column of Guardian Newspaper

```
<?xml version="1.0" encoding="ISO-8859-1" ?>

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns="http://purl.org/rss/1.0/"
  xmlns:dc="http://purl.org/dc/elements/1.1/">

  <channel rdf:about="http://localhost/RDFDemoServer">
    <title>Guardian Newspaper Headlines - Education Column.</title>
    <link>http://www.guardiannewsngr.com</link>
    <description>Guardian newspaper is published in Nigeria. This column gives
information relating to education in Nigeria.</description>
    <items>
      <rdf:Seq>
        <rdf:li rdf:resource="http://www.guardiannewsngr.com/education/article01" />
        <rdf:li rdf:resource="http://www.guardiannewsngr.com/education/article02" />
        <rdf:li rdf:resource="http://www.guardiannewsngr.com/education/article03" />
        <rdf:li rdf:resource="http://www.guardiannewsngr.com/education/article04" />
      </rdf:Seq>
    </items>
  </channel>
</rdf:RDF>
```


<rdf:li rdf:resource="http://www.guardiannewsngr.com/education/article05" />

</rdf:Seq>

</items>

</channel>

<item rdf:about="http://www.guardiannewsngr.com/education/article01">

<title>Why NIEPA is relevant to the education sector, by Ajayi</title>

<link>http://www.guardiannewsngr.com/education/article01</link>

<description>Prof. Taiwo Ajayi is the Director of the National Institute for Education Planning... </description>

<dc:date>2005-04-12</dc:date>

</item>

<item rdf:about="http://www.guardiannewsngr.com/education/article02">

<title>NUC approves first University of Education</title>

<link>http://www.guardiannewsngr.com/education/article02</link>

<description>Ogun

State stood practically tall in Abuja last Monday, as the... </description>

<dc:date>2005-04-12</dc:date>

</item>

<item rdf:about="http://www.guardiannewsngr.com/education/article03">

<title>Bida Poly to benefit from exchange programme</title>

<link>http://www.guardiannewsngr.com/education/article03</link>

<description>The Rector Federal Polytechnic Bida, Niger state Engineer...

</description>

<dc:date>2005-04-12</dc:date>

</item>

<item rdf:about="http://www.guardiannewsngr.com/education/article04">

<title>Don faults inability of satellite to discover Bellview airline's crash site</title>

<link>http://www.guardiannewsngr.com/education/article04</link>

```

<description>Head
of the department of Computer Engineering of the Federal... </description>
<dc:date>2005-04-12</dc:date>
</item>
</rdf:RDF>

```

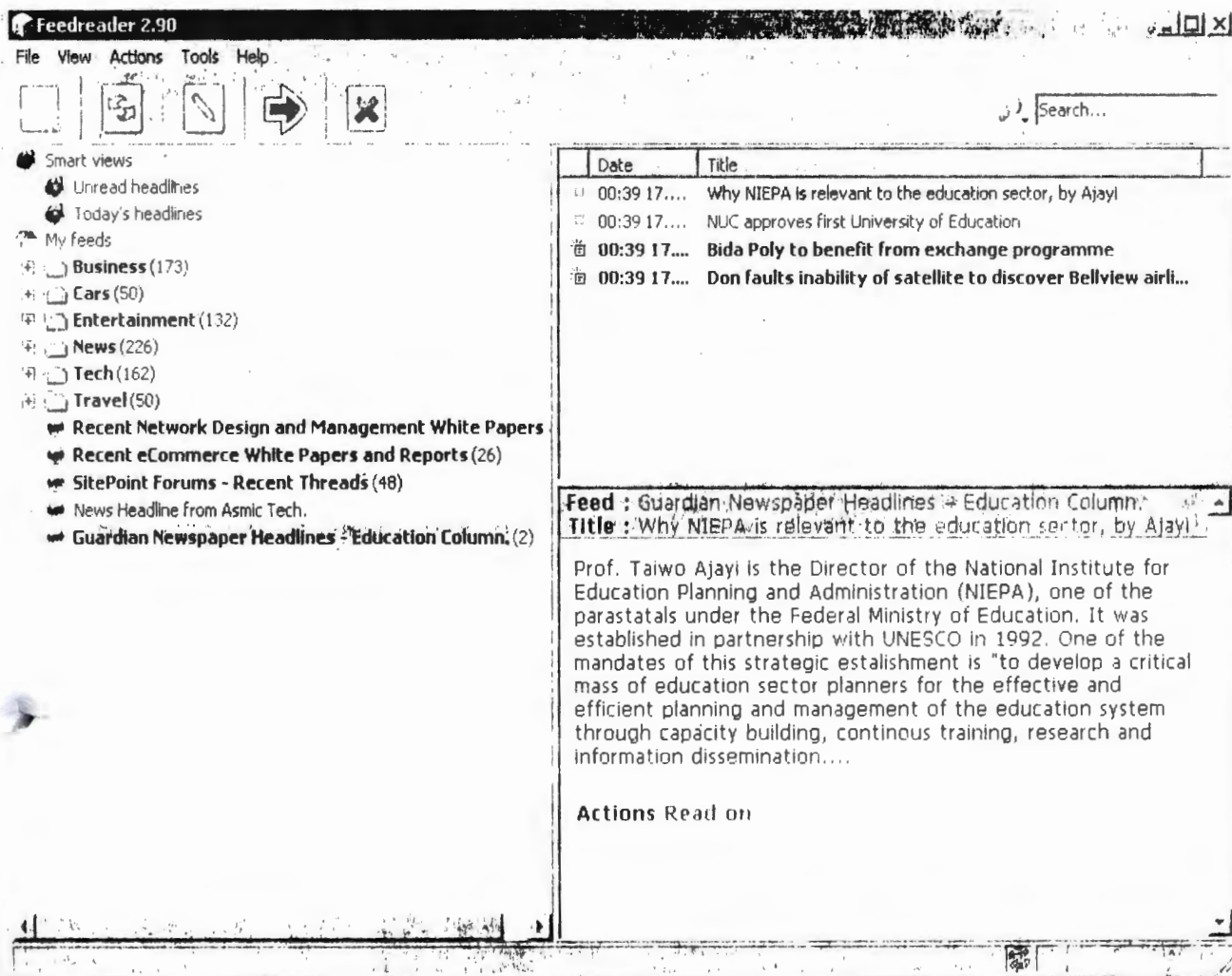


Fig 5 Rss feed displayed in News Aggregator

4.0 DISCUSSION

Generating RSS 1.0 is simple so long the principles guiding the version are followed. Irrespective of the version, the

presentation remains the same. The differences in the versions are based on the tags, attributes and values used. Observing Fig 2, 3 and 4 as modeling RSS, Version

1.0 contains more tags than the version 0.91 and 2.0; the header has more tags and attributes. Interestingly, some tags are optional while some are mandatory and must be typed sequentially. Web development Integrated Development Environments (IDEs) like Macromedia Dreamweaver can be used to generate RSS rather than hand coding it. Often times, such IDEs require using third party tool(s) to generate RSS automatically. Having generated RSS feed, displaying or reading it requires a browser with XML parser or processor. Also, it requires web languages such as ASP.Net, PHP, XSL to mention a few in order to display a formatted information. While all these require visiting the website, one can use a news aggregator such as RSS Reader, FeedReader to pull the information (respective RSS file) without visiting the site. With an interface in Fig 5.0, it becomes interesting reading news or information from various columns or websites if need be without visiting them individually. First, a news aggregator displays titles and summary of an entire channel. On clicking each title, the summary of the title is displayed. Where each item (a title and a summary) refers

to webpage, the news aggregator has a built-in browser that makes it possible to load web pages that are of interest for further reading.

5.0 CONCLUSION

On a global scene, Nigeria is such a country where there is a severe dearth of adequate information. On the internet, there is no information about some topical development issues.

Similarly, the libraries that are supposed to be a repository of information for research and development has fallen short for lack of appropriate tools. Thus with RSS, alternative way of reading information by the populace is now possible and stress free using news aggregators.

The backlog of information about topical issues can be generated in retrospect; and online services can be developed for all on various captions as Education, Vacancies, Health, Security e.t.c. It has the overall objective of encouraging quicker online access particularly for students and jobseekers that may not have the money to buy the newspapers.

In addition, RSS as an xml-based description tool has found use in emerging technologies like podcasting because of ability to render platform independent and updated information.

6.0 REFERENCES

<http://www.sitepoint.com/article/php-xml-parsing-rss-1-0>, (Feb, 2005)

<http://www.xml.com/pub/a/2002/12/18/dive-into-xml.html>, (Feb, 2005)

<http://www.sitepoint.com/print/own-web-service-php-xml-rpc>, (Feb, 2005)

Ellen Pearlman and Eileen Mullin (2004). Programming the web using XML, New York, McGraw-Hill, 1st Edition, pp 1-9

Richard Light (1997), Presenting XML, USA, Sams.net, 1st Edition, pp 2, 202-209

Abhijit Chaudhury and Jean-Pierre Kuilboer (2002). E-business and E-commerce Infrastructure, New York, McGraw-Hill, 1st Edition, pp 261-282