

Online Real Estate Feed Reader

By

Ahmad Ridzuan Bin Ibrahim

Dissertation Submitted in partial fulfillment of

The requirements for the
Bachelor of Technology (Hons)
(Business Information Systems)

January 2009

Universiti Teknologi PETRONAS

Bandar Seri Iskandar

31750 Tronoh

Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

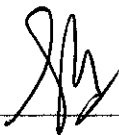
Online Real Estate Feed Reader

By

Ahmad Ridzuan Bin Ibrahim

A project dissertation submitted to the
Business Information System Programme
Universiti Teknologi PETRONAS
In partial fulfillment of the requirement for the
Bachelor of Technology (Hons)
(BUSINESS INFORMATION SYSTEMS)

Approved:



(MS. Shakirah Bt. Mohd. Taib)

UNIVERSITI TEKNOLOGI PETRONAS

TRONOH, PERAK

January 2009

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.



AHMAD RIDZUAN BIN IBRAHIM

ABSTRACT

This report basically discusses the preliminary research done and basic understanding of the proposed topic, which is “Online Real Estate Feed Reader”. This online real estate feed reader is an idea in order to give services to people who lives in the developed area where people will face the difficulty to find land or houses. This also gives people an easy alternative way to find the agent for real estate from using the traditional way like newspaper or directly to agent. User can just click search button to display the result and also can just subscribe RSS in order to get time to time updated. This Online Real Estate feed reader website with the objective to help the user to search exactly the real estate website available from internet and not with the anonymous result which sometimes no related at all to the real estate web. User also can get information about what they are searching in term of description, contact person or picture from this online real estate feed reader, this will make sure user will get right information from the selected real estate website. The website also contains with the RSS feeds where it can be subscribe from the top of the page, this RSS will be update frequently automatically will show the outline of the information and also the picture if there are available. For the introduction, the scope for this online real estate feed reader will be focusing to search or the of real estate in KL area only, where this website only lists the real estate from KL area and stores all the information into the database. The methodology use in this project is prototyping methodology, where it consist of several phase which are planning, analysis, design and implementation phase, which the planning, analysis and design are perform repeatedly until the system is completed .The analysis also has been perform under result and discussion session, most of the user give a positive feedback for the system and for the conclusion author hope this project will be success and achieve it scope and objectives like planned and user can get benefit by using this system.

ACKNOWLEDGEMENT

First of all I would like to say Alhamdulillah, as my gratitude to Allah for blessing me to finish this project successfully and also give me a good health during the process of finishing this project.

Here, I also would like to give my appreciation and thankfulness to my supervisor, Ms Shakirah bt Mohd Taib for the help and guidance in helping from the beginning of this project by giving the idea for the project and also good advice for me in order to improve my project to make it more relevant for the final output. Thank you very much.

Last and not least I would like to give my greatest gratitude to my family especially my father and mother Mr Ibrahim and Ms Azizah to give me full support in finishing my project also not forget to my friends, Izhar and Azwan for lend their hand for helping me to finish this project, also deepest appreciation to Nur Nabihah by inspire me to finish this project and also all my friends, Khairul, Izzuan and Syahid who helped me directly and indirectly with this project.

Thank you.

TABLE OF CONTENTS

CERTIFICATION OF APPROVAL	i
CERTIFICATION OF ORIGINALITY	ii
ABSTRACT	iii
ACKNOWLEDGEMENT	iv
CHAPTER 1:	INTRODUCTION	1
	1.1	Background of Study	1
	1.2	Problem Statements	2
	1.3	Objectives	3
	1.4	Scope of study	3
CHAPTER 2:	LITERATURE REVIEW	4
	2.1	Introduction	4
	2.2	How to sell house in slow real estate market	4
	2.3	Webs Feed	6
	2.4	Text Mining	11
CHAPTER 3:	METHODOLOGY / PROJECT WORK	19
	3.1	Prototyping Methodology	19
	3.2	Software /Tools	21
CHAPTER 4:	RESULT AND DISCUSSION	27

	4.1	Introduction	.	.	.	27
	4.2	Analysis	.	.	.	27
	4.3	Result	.	.	.	30
CHAPTER 5:		CONCLUSION	.	.	.	42
	5.1	Conclusion	.	.	.	42
	5.2	Limitation of the project	.	.	.	43
	5.3	Recommendation	.	.	.	44
REFERENCES			.	.	.	45

LIST OF FIGURES

Figure 2.1: www.realestate.co.nz	5
Figure 2.2: www.realestate.co.nz real estate web traffic statistic	6
Figure 2.3: common web feed icon	6
Figure 2.4: example of web feed on website	7
Figure 2.5: example of web feed on website	8
Figure 2.6: concept idea using RSS	9
Figure 2.7: example of how the RSS will be structured in the Real Estate feed	10
Figure 2.8: process chart for conducting text mining	11
Figure 2.9: text mining life cycle	14
Figure 2.10: web text mining process taken into account	15
Figure 2.11: web crawler architecture	17
Figure 2.12: rules and annotators scheme	17
Figure 3.1: prototyping methodology	19
Figure 3.2: screenshot of RSS builder	22
Figure 3.3: screenshot of FTP site manager and view XML in RSS builder	23
Figure 3.4: Joomla! Logo	24
Figure 3.5: Joomla! Web screenshot	24
Figure 3.6: WordPress logo	25

Figure 3.7: WordPress web screenshot.....	26
Figure 4.1: Online real Estate website using joomla!	31
Figure 4.2: RSS icon at the page.....	32
Figure 4.3: RSS feed page.....	32
Figure 4.4: screenshot of sample RSS code	33
Figure 4.5: WordPress Real Estate Web screenshot	34
Figure 4.6: WordPress Real Estate Feed.....	35
Figure 4.7: Home page of Online Real Estate Feed Reader	36
Figure 4.8: Admin Login Page	37
Figure 4.9: Admin Home Page	37
Figure 4.10: add new feed	38
Figure 4.11: remove feed	39
Figure 4.12: search database.....	40
Figure 4.13: Table feed	41
Figure 4.14: Table items	41

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND OF STUDY

Everyday in this new most of the transaction of buying and selling is already goes online as an example that we can see now is Lelong.com, mudah.com and many more. All this kind of website provide user with the ability to search and get the information for what kind of product that they interest about and also can deal the price with the seller.

On this situation the author try to come out with the new idea how to make the buying and selling or renting house become online, as the result Real Estate Feed Reader try to be create by author where it can do the same like lelong.com or mudah.com where it can search from the various real estate website and find what exactly the user want find from that website.

This site will be able to search and get all the data that needed in the renting or buying and selling transaction, display the information in graphical when it available. User also can get the contact number of the seller base on the search result return

1.2 PROBLEM STATEMENT

The main problem that the author had been identified is as we see now days, our development in technology has grown very rapidly, and this will result in increasing the quantity of the buildings in developing are. Most of the potential land now has been used in order to support out technology growth. As a result demand of housing area increase for this are and people are searching for it via various type of medium. One of the most popular today is by using the internet, and for the example is Google site where the user can find various of the result by just click the search button, but the big problem using these techniques is there are more redundancy when the user search for real estate agent especially when we search by using the keywords. This is become worst when we have to find it using the old traditional way like newspaper and contact directly to the agent which are more costly and lack of efficiency.

1.3 OBJECTIVE

The objectives of this project are to:

- To find information that related to Online Real Estate.
- To study the concept of RSS Feed
- To integrate RSS Feed Reader with Online real estate website
- To build a system prototype for the system.

1.4 SCOPE OF STUDY

The primary focus of the project is to create the search function that will be able to return the result of the real estate website on the internet in the KL area. It also can find and store the real estate information into the database and show the update using the Online Real Estate Feed Reader.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In this literature review, Author try to focus on how to create the feed reader which capable to collect all the RSS headline or updates from the real estate website on the web, where it can combine all the RSS together in the website and combine all the RSS in one page. This literature review also includes the text mining, where user can get the information or Real Estate location that located in KL area only.

2.2 How to sell house in slow real estate market

Real estate market runs in cycle because it also has the peak and plunge situation, and this affecting the bought and sold between the buyers and sellers. The up and down of the market in real estate can effect the seller, mostly when the market price goes down where it's hard to sell their house. As we know some of the buyers are very too picky in buying the house, we must ready to fulfill what the criteria than buyers want in order to attract them to have a business with us. (Derick Van Ness, 2008)

2.2.1 These are the ways for a seller to sell a house more quickly when this situation happened

1. Be the lowest price in order to understand the “buyer mentality”, where most of the buyers only want the cheapest house.

2. Use a professional, where the good or bad real estate agent can effect selling, where if we can look on the MLS (multiple listing service), where we can find the selling of house in 6 past month period. This is the one of the best way to sell house quickly in slow real estate market.
3. Sell on a 'lease to own' terms, where most way to attract the buyers in slow real estate market by using our creativity, because this will make buyers have limited option.
4. Exposure yourself to sell quickly, this we can say that we must attract buyer to see the house by putting more advertisement in order to exposure to people.

2.2.2 Example of online real estate web page

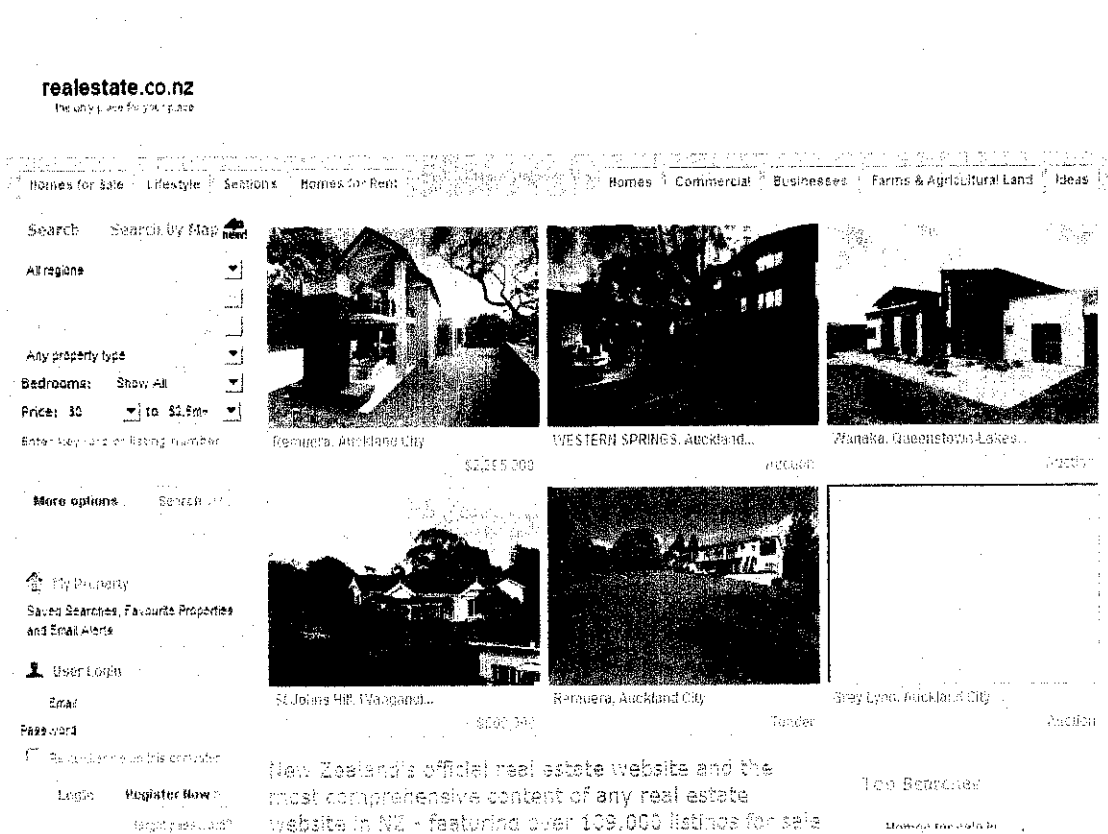


Figure 2.1- www.realestate.co.nz

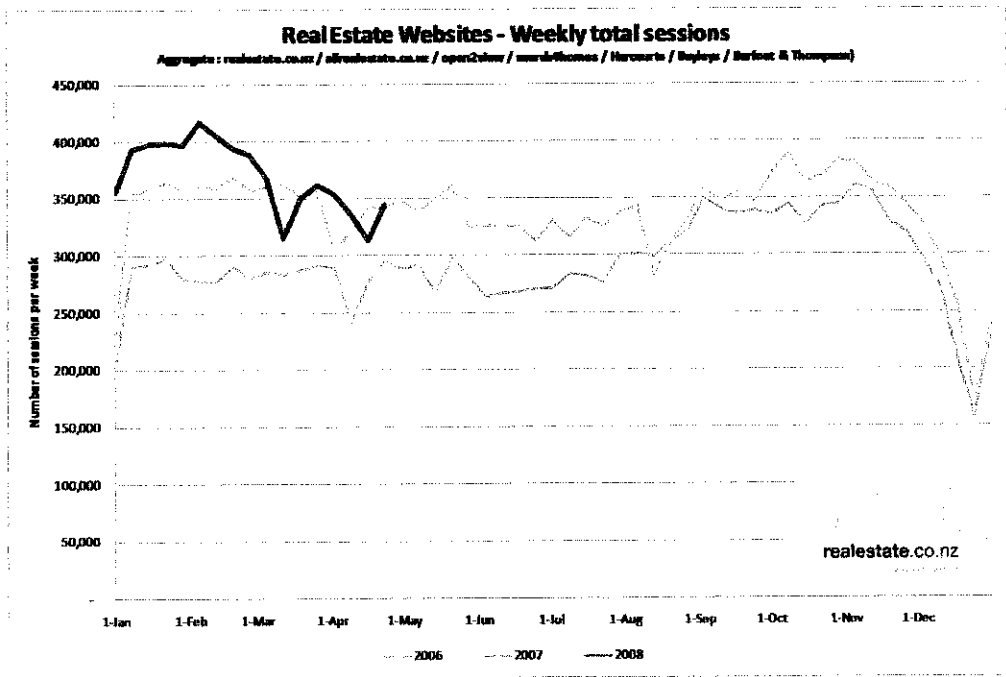


Figure 2.2- www.realestate.co.nz real estate web traffic statistic

2.3 Web Feed

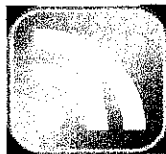


Figure 2.3– common web feed icon

Web feed is a tool used to update web frequently such as entry, headlines in standard format. Web feed also channel, are capable to summarize the content of the text. By using this, people will be able to get update news from the site frequently. Most of the web feed today can be read by browser itself, where most of the browser today already

installed with the web feed reader as a build in function of the browser and some of the web feed reader is desktop based, where it can read it using widgets.

One of the most famous web feed is Really Simple Syndication (RSS). RSS come with the xml format that makes it capable to publish and viewed by many kinds of program. User can just subscribe RSS by just click the RSS icon at the navigation bar. RSS reader will check subscribes feeds regularly it there are new content available on the web. RSS can be referring to the following format Really Simple Syndication (RSS 2.0), RDF Site Summary (RSS 1.0 and RSS 0.90) and Rich Site Summary (RSS 0.91). (Mark Pilgrim, 18 December 2002)

2.3.1 Example web feed on the website

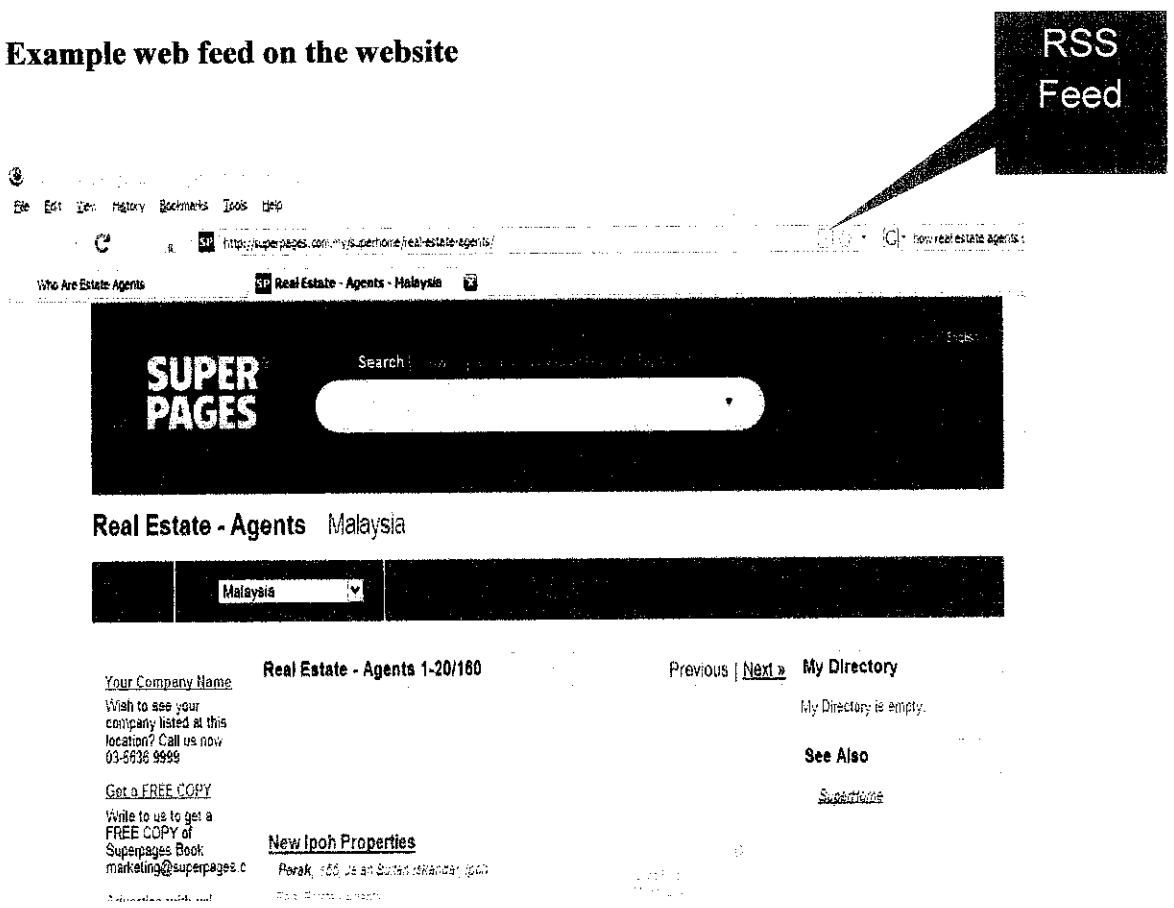


Figure 2.4— Example of web feed on Website

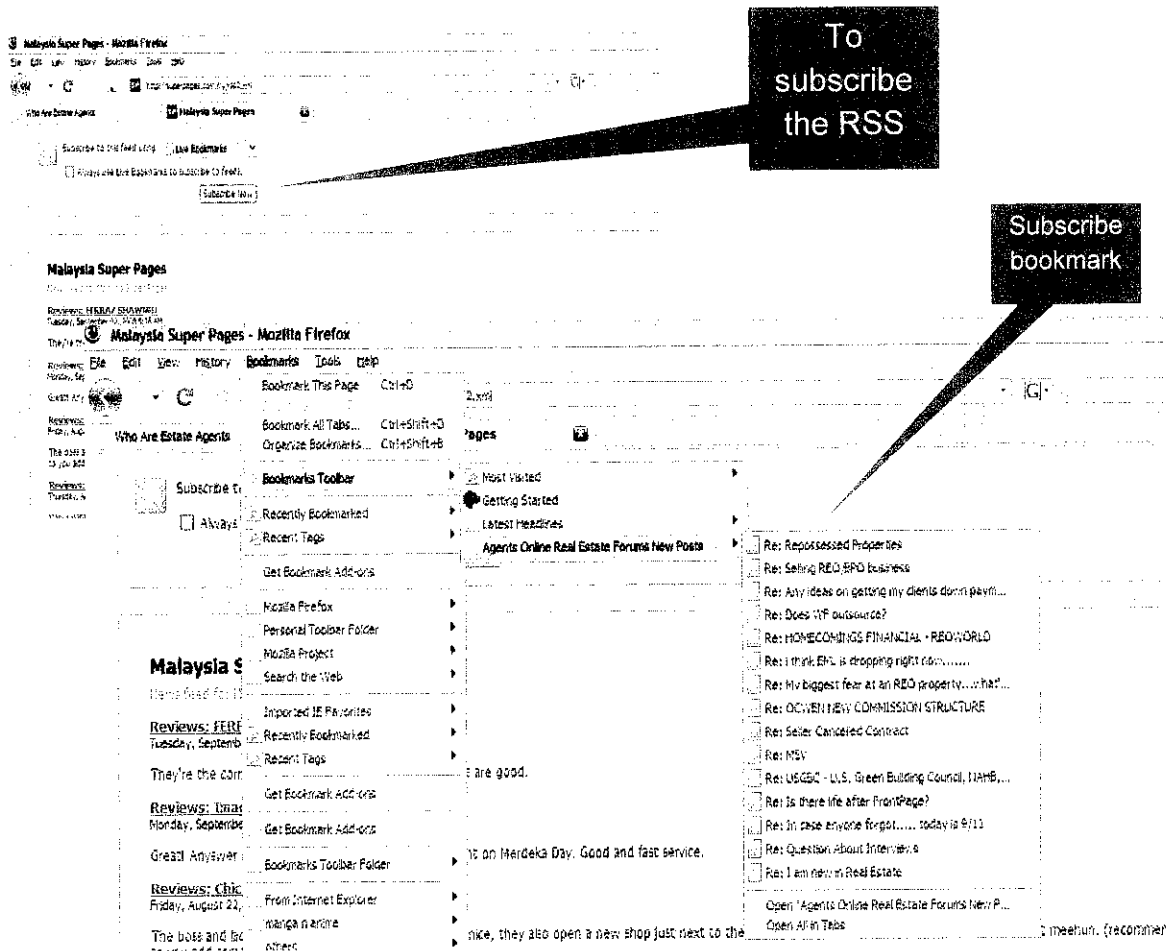


Figure 2.5– Example of web feed on Website

2.3.2 Concept idea of RSS

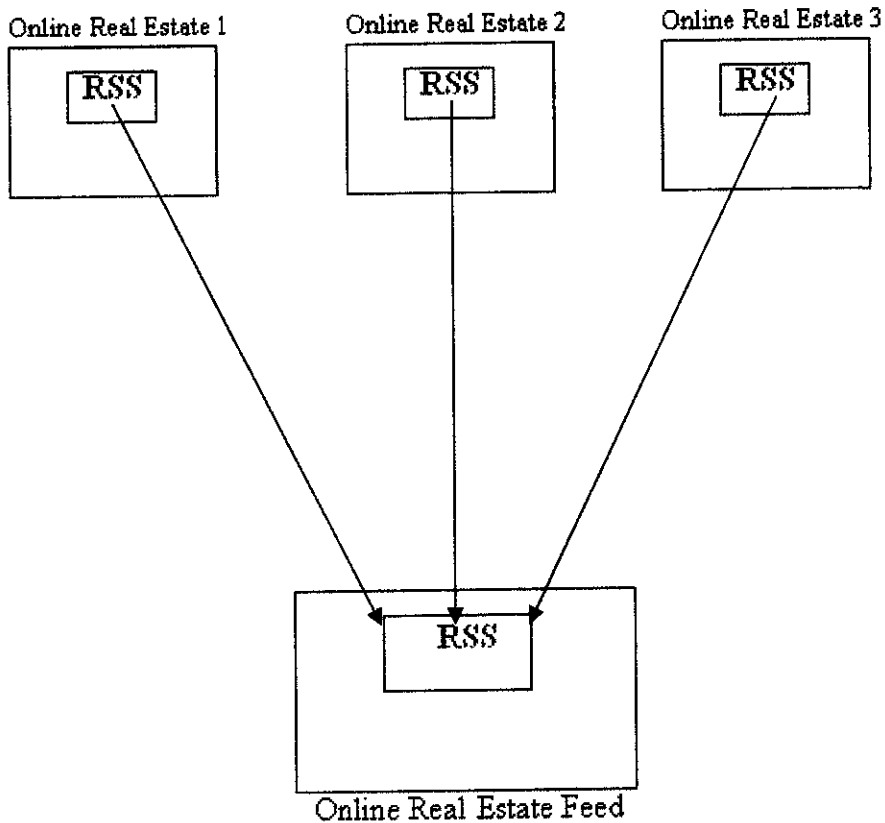


Figure 2.6– Concept idea using RSS

The figure 2.6 shows how the RSS will function in this project. The Online Real Estate Feed Reader will monitor all the feeds from another Online Real Estate websites, where all the information that contains in the RSS from another Real Estate Website will be mining and will be list into the structured way in order to make user more understand the content that list in the RSS after subscribed. Figure 2.7 is an example how the RSS will be listed from the Online Real Estate Feed Reader.

Photo	Price	Address City Zip	Sq. Ft	Bed / Bath	Property Type	Listed
	\$500	2 king street Los Angeles , CA 90009	0	9 / 3	Single Family Home	02/04/2009
	\$44,500	Rupley Dallas , TX 75218	0	3 / 2	Single Family Home	02/04/2009
	\$398,500	12795 Bengal Ave Rosemount , MN 55068	2780	3 / 2	Single Family Home	02/04/2009
	\$499,000	741 Middletrail Ct. Mount Airy , MD 21771	0	6 / 3	Single Family Home	02/04/2009
	\$510	638 Andover Rd Newtown Square , PA 19073	3000	3 / 4	Single Family Home	02/04/2009

Figure 2.7– Example of how the RSS will be structured in the Real Estate Feed

After the feed has been subscribed the RSS will be display like Figure 2.7 above. All the information will be display in the structured way to the user such as picture, Price, Address, type and other more detail if it available for the type of property.

2.4 Text Mining

Text mining, sometimes alternately referred to as text data mining roughly equivalent to text analytics. It also can be identify as a process to deriving high quality information into text and it will be derive again into pattern. In text mining, it usually involves the process of structuring text into the database to be evaluation and interpretation of the output. (Ronen Feldman and James Sanger, the Text Mining Handbook)

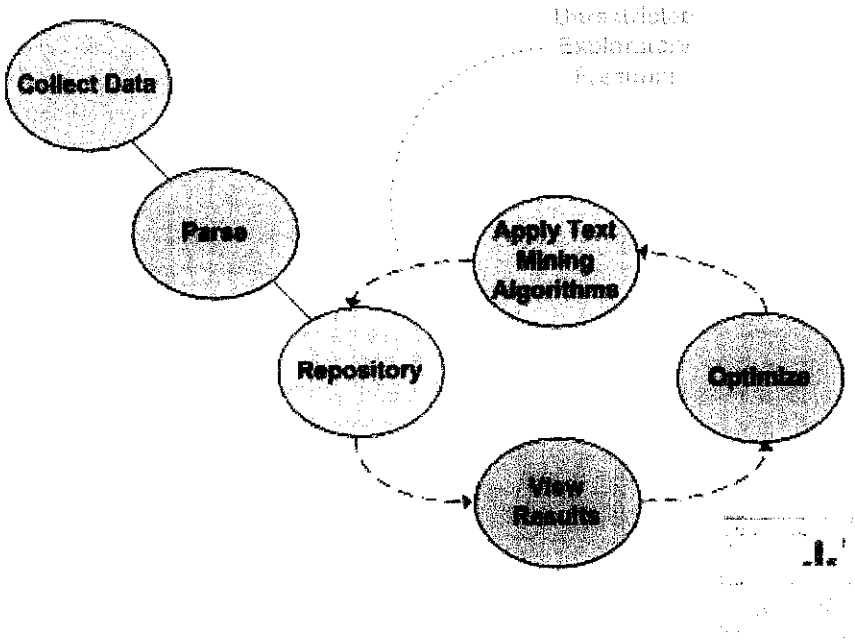


Figure 2.8– Process Chart for Conducting Text Mining

2.4.1 A Web Text Mining Flexible Architecture

Web text mining can be describe as an application to mining data techniques to discover the pattern or the web. According to the analysis web mining have three different types which web usage mining, web content mining and web structure mining. We can get content from the web page by using the web content mining where the data will be process to get the useful information from the internet. In the web page there will have a lot of resource including text, audio, and video. Content of the web also can be classified as a text which also can be called web text mining, in web text mining the interesting and non-trivial pattern or knowledge data will be process and extract from the web and “The technologies that are used in web text mining are HLP, natural language processing and IR information retrieval” (W. Zhong, X. Tang, 2008)

We need the specific tools to discover the text and analyze pattern of text, there are lot of tool that relate d to the text mining that we can find today, this tools must be able to transform the pattern into a new knowledge structured. This structure must be more manageable and importantly, are able to provide higher level abstractions that facilitate the solution of text related tasks. (M. Castellano, G. Mastronardi, A. Aprile, and G. Tarricone, 2008)

2.4.2 Web Text Mining Lifecycle

Text Refining is a phase that transforms input raw text in structured information. Text Refining input are not-structured data as texts or semi-structured data as HTML pages coming from Web. These data are submitted to the opportune process composed by recovery of documents, cleaning of the draw out documents and extraction of the useful information to the Text Mining phase. Text refining output is stored in database, XML file or any other structured form that is called *Intermediate Form*. (A.H. Tan, 1999)

Document clustering is the assignment of entity multivariate to few categories (classes, groups) not previously defines. The criterion is to gather among them similar entity. It is necessary to establish the variable of classification, the measure of the proximities among the entities, the number of groups in which to leave again the entities and finally to pass to an integration phase with other methods of analysis.

Document categorization is the phase where the objects must be attributing to one or more classes or categories already notes in start. In text reference, this technique of mining is said Text Categorization (TC). Text Categorization is often developed through algorithms of Machine Learning. (B. Mobasher, R. Cooley, and J. Srivastava, 1999)

In the *Pattern Extraction* some patterns are identified following the analysis of associations and of tendencies. The meaningful correlations among frequent whole data are found and this process can be applied to textual data if these data are been transformed previously in a structured form. The following techniques are carried out:

- Predictive Text Mining which is used for identifying the tendencies in collected documents in a time period.
- Association Analysis which identifies the relationships among the attributes, like if the presence of a pattern implicates the presence of another pattern in documents.

Finally, *presented techniques* are used for introduce the patterns and for visualize the results.

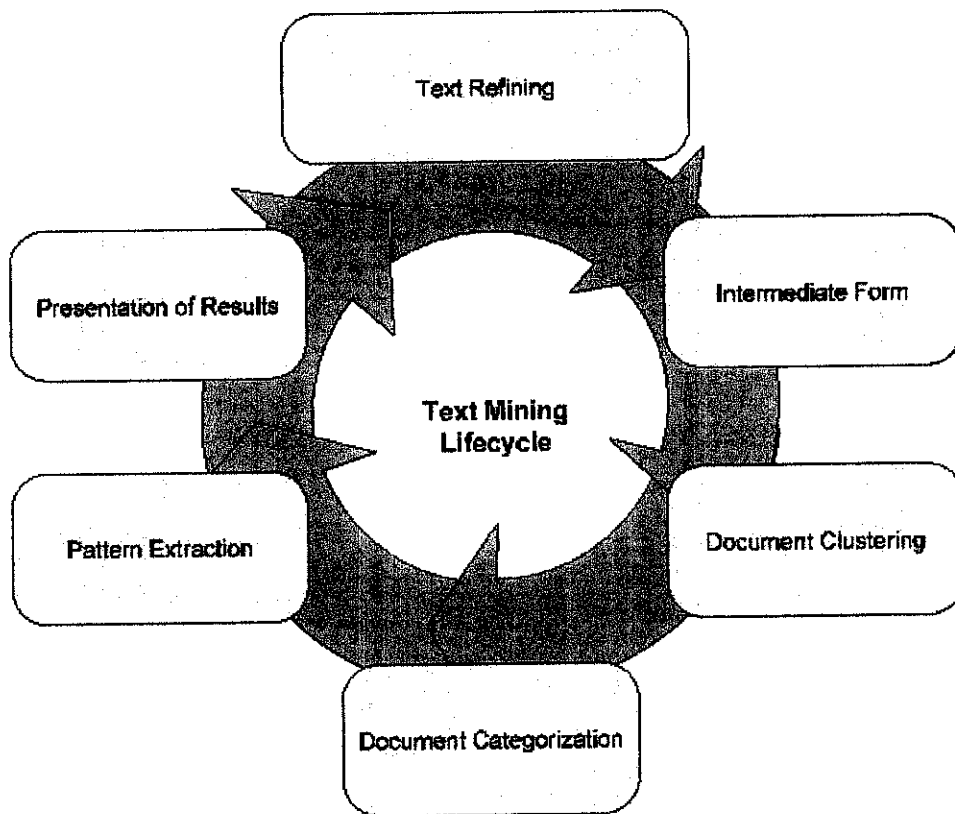


Figure 2.9– Text Mining Lifecycle

2.4.3 Web Text Mining Process

All the information that we get from the internet is the one of the role in the Knowledge Discovery process. To examine web contents and to extract useful information to a purpose through techniques of Mining is scientifically called Web Text Mining. (R. Kimball and R. Merz, 2000)

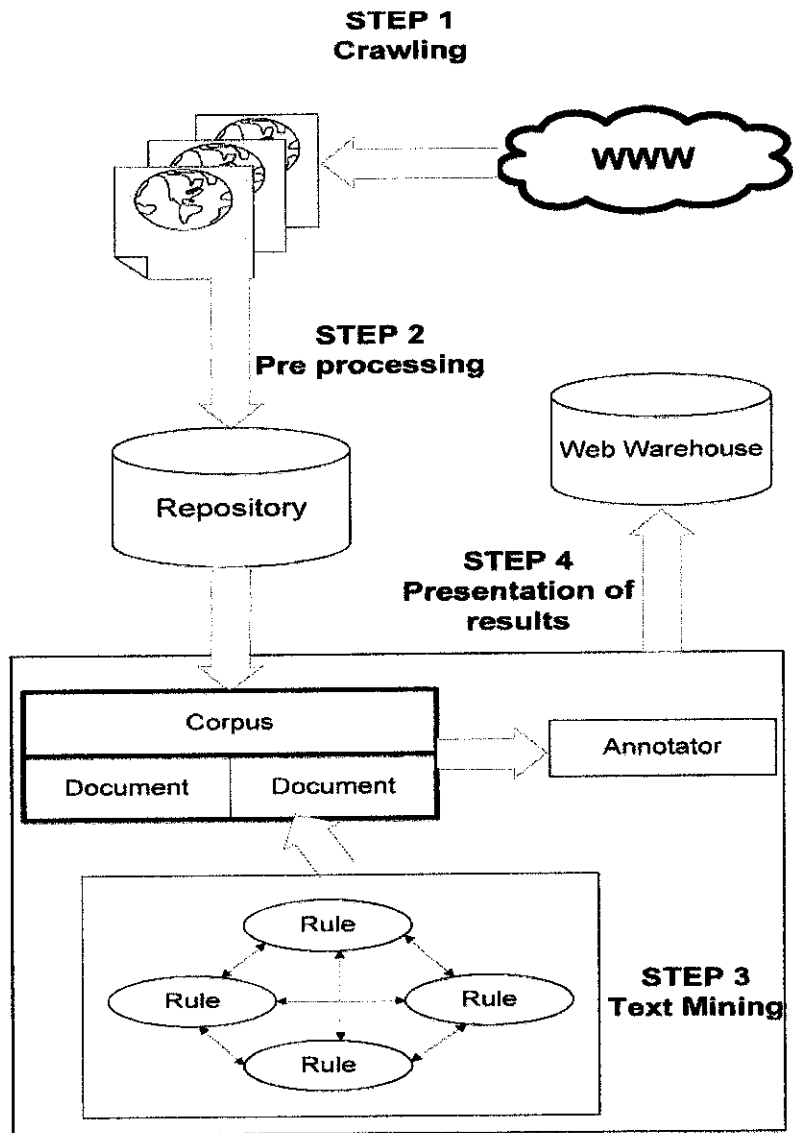


Figure 2.10– Web Text Mining Process taken into account.

There are four steps involve in data mining process. For the first step, the information that contains the textual content in the web page will go through the recovery step, and it will covers the first step in the web text mining process using automatic retrieval for the documents and ensuring at the same time that the not-relevant ones macaws fetched as few as possible. In the recovery step the information will be crawl from one or more URL. (Cooley, R. et al, 1997)

The crawling is implemented by a Web Crawler, as shown in the Figure 9. Web Crawler's architecture fits the guide's lines of to Focused Crawler (Chakrabarti, S. et al, 1999). The second step is the pre-processing, where it create the repository of information from execution of the first step. In this step it must provide the effect of the control and a cleaning during the first step. Template in this process is identify as keywords that contains of the universe and allow content to be examine on the web page. On the template there are cleaning of the content is affected to eliminate all what to be held "garbage" that so to say tag, banner, information not tightly connect to the content of the same page. Otherwise the page will be discard if found correspondence with one of the template define for the system, all the useful information of the page are memorize in the repository.

The third step is the fundamental step on which the whole Web Mining process founds him. In this step a first phase of Information Extraction and the following phase of Text Mining are effected. It founds on the use of Text Mining tools that receiving in input the data which are contained in the repository of the second step and they are able to extract "intrinsic information" contained in the document or in all documents. (M. Castellano, G. Mastronardi, A. Aprile, and G. Tarricone, 2008)

The discovery information in this phase is momentarily memorized in the Annotators which have the assignment of memorize all useful Information in a Document or in a Corpus. The further analysis on the Annotators, that is to say, elimination of the duplicates, elimination of the errors, etc. are the start of the discovery of the real information called Skill. The Skills are keywords which allows to the system to have a general vision of the semantic content of the Document or the Corpus. (GATE, 2008)

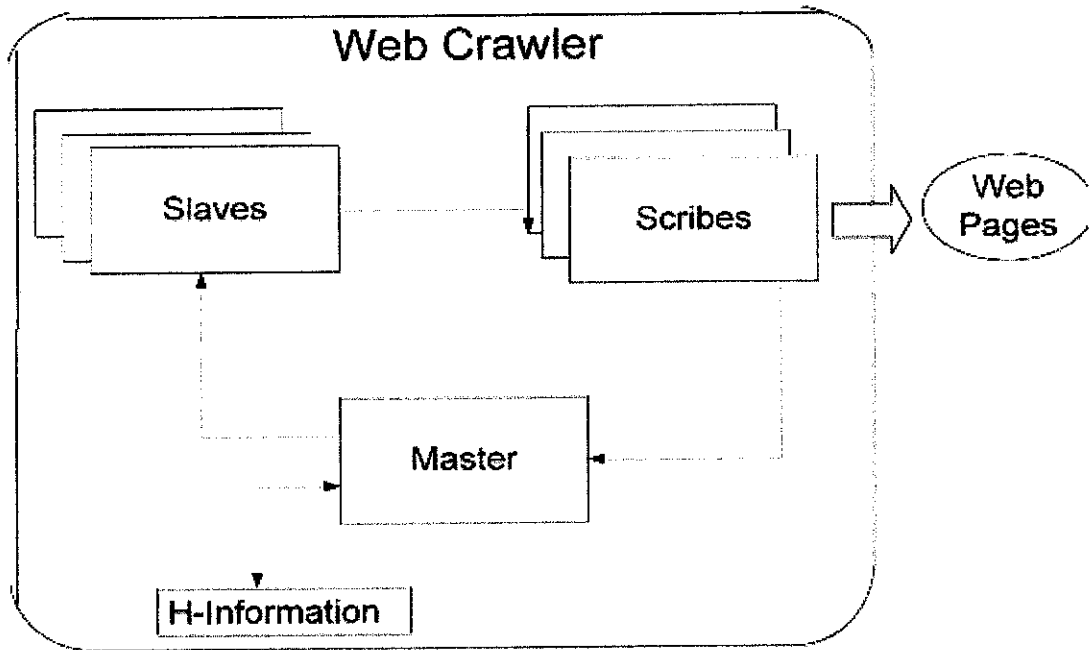


Figure 2.11- Web Crawler Architecture

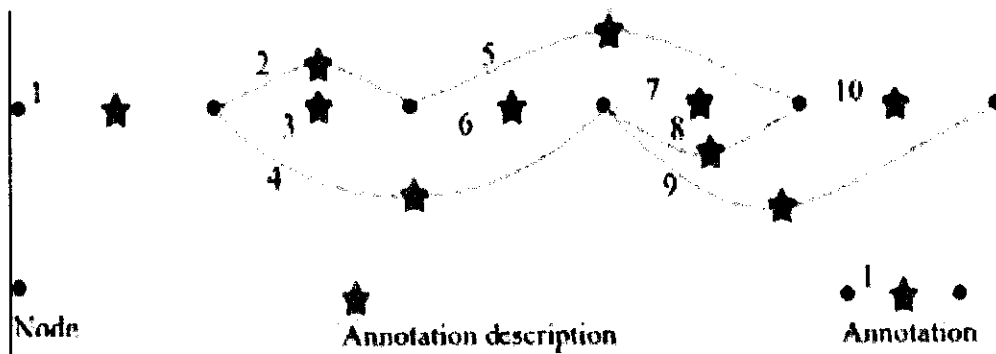


Figure 2.12- Rules and Annotators scheme

The fourth is last step of the process; it foresees the phase of presentation of the obtained results. In our architecture, the information drawn out during the execution of the third step is stored in a second repository. This repository created during this phase

will follow the rules of the Web Warehousing according to the Whoweda schemas. The repository is composed by URLs and nodes. URLs are the reference URLs of the document, that is to say, "referenced from" and "it references who" whereas the nodes are the skills found in the third step To fill this repository is necessary to effect a parsing of the result in XML supplied by software during the third phase for recover the necessary information. Use of Whoweda schemas is fundamental in the implementation of Web Warehouse because our system is directed to the execution of Data Mining on the data produced by the whole Web Text Mining process. (M. Castellano, G. Mastronardi, A. Aprile, and G. Tarricone, 2008)

CHAPTER 3

METHODOLOGY / PROJECT WORK

3.1 Prototyping Methodology

The methodology that has been used in developing this project is prototyping methodology. In this methodology there consist planning, analysis, development and lastly implementation.

The conventional purpose of a prototype is to allow users of the software to evaluate developers' proposals for the design of the eventual product by actually trying them out, rather than having to interpret and evaluate the design based on descriptions. Prototyping has several benefits: The software designer and implementer can obtain feedback from the users early in the project.

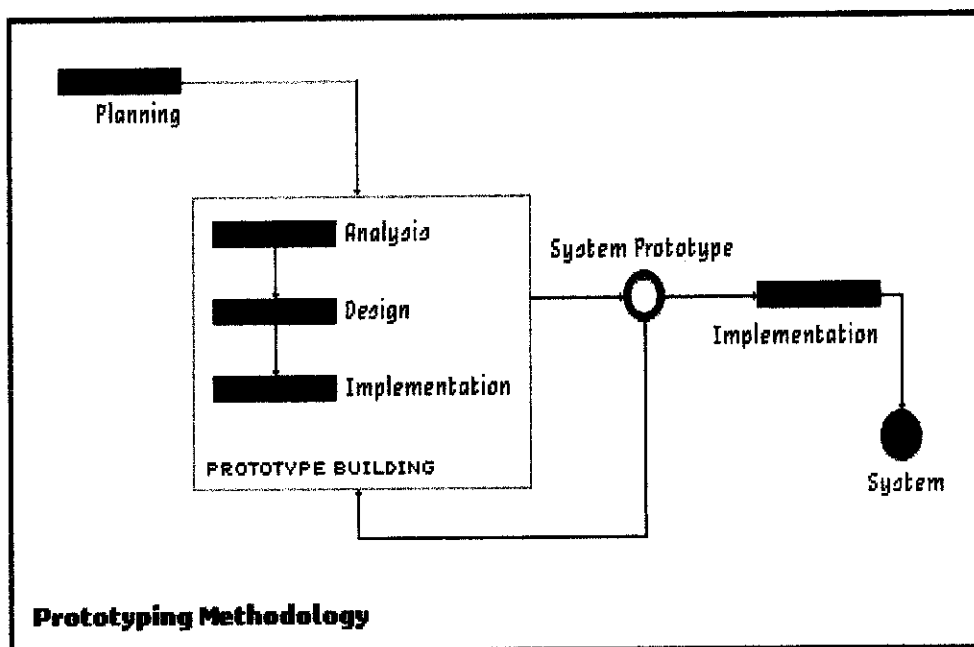


Figure 3.1– Prototyping Methodology

3.1.1 Planning

In the planning phase, view of the project is generated and the goal of the project is determine .In this stage the tools that need to be use in this project is choose.

3.1.2 Analysis

In the analysis phase the requirements gathering technique is use where all the information related to the project is gathered to gather, for example is who is the target user, what is the scope of the project and what is the objective and some additional information about the project . In this phase also need to have research related to the project by reviewing the problem and article related to the project.

3.1.3 Development

In the development stage, the interface of the web is design on order to make it rational to the type of web that author are tried to develop. Database structure also needs to be design in this stage in order to make all the information is easy to be stored and to be retrieving. The additional functions need to be including in this phase such as web feed is and also search function.

3.1.4 Implement

In this phase the prototype of the project is build until it meets the requirement, in this phase all the tools like php, mysql, web feed is used. The output will be tested, if the output is not meet the requirement the process will return back to the analysis phase but if there the web is meet the requirement it will be consider as a new system and already can be used by the user.

3.2 Software / Tools

Tools that will be use in developing this system are:

Apache web server

This software referred as simply Apache, a public-domain open source Web server developed by a loosely-knit group of programmers. The first version of Apache, based on the NCSA httpd Web server, was developed in 1995. Core development of the Apache Web server is performed by a group of about 20 volunteer programmers, called the Apache Group. However, because the source code is freely available, anyone can adapt the server for specific needs, and there is a large public library of Apache add-ons. In many respects, development of Apache is similar to development of the Linux operating system. The original version of Apache was written for UNIX, but there are now versions that run under OS/2, Windows and other platforms. The name is a tribute to the Native American Apache Indian tribe, a tribe well known for its endurance and skill in warfare.

PHP

Self-referentially short for PHP: Hypertext Preprocessor, an open source, server-side, HTML embedded scripting language used to create dynamic Web pages. In an HTML document, PHP script (similar syntax to that of Perl or C) is enclosed within special PHP tags. Because PHP is embedded within tags, the author can jump between HTML and PHP (similar to ASP and Cold Fusion) instead of having to rely on heavy amounts of code to output HTML. And, because PHP is executed on the server, the client cannot view the PHP code. PHP can perform any task that any CGI program can do, but its strength lies in its compatibility with many types of databases. Also, PHP can talk across networks using IMAP, SNMP, NNTP, POP3, or HTTP.PHP.

MySQL

Pronounced “my ess cue el” (each letter separately) and not “my SEE kwil.” MySQL is an open source RDBMS that relies on SQL for processing the data in the database. MySQL provides APIs for the languages C, C++, Eiffel, Java, Perl, PHP and Python. In addition, OLE DB and ODBC providers exist for MySQL data connection in the Microsoft environment. A MySQL .NET Native Provider is also available, which allows native MySQL to .NET access without the need for OLE DB. MySQL is most commonly used for Web applications and for embedded applications and has become a popular alternative to proprietary database systems because of its speed and reliability. MySQL can run on UNIX, Windows and Mac OS. MySQL is developed, supported and marketed by MySQL AB. The database is available for free under the terms of the GNU General Public License (GPL) or for a fee to those who do not wish to be bound by the terms of the GPL.

RSS builder

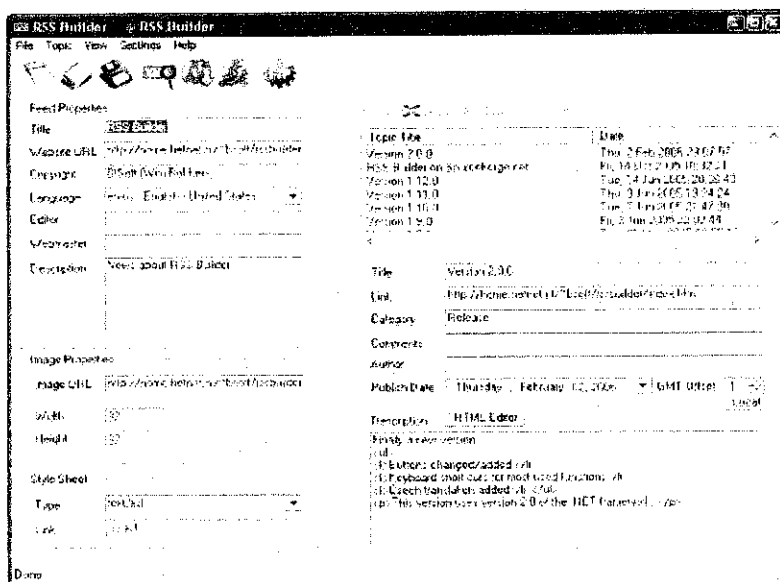


Figure 3.2– Screenshot of RSS builder

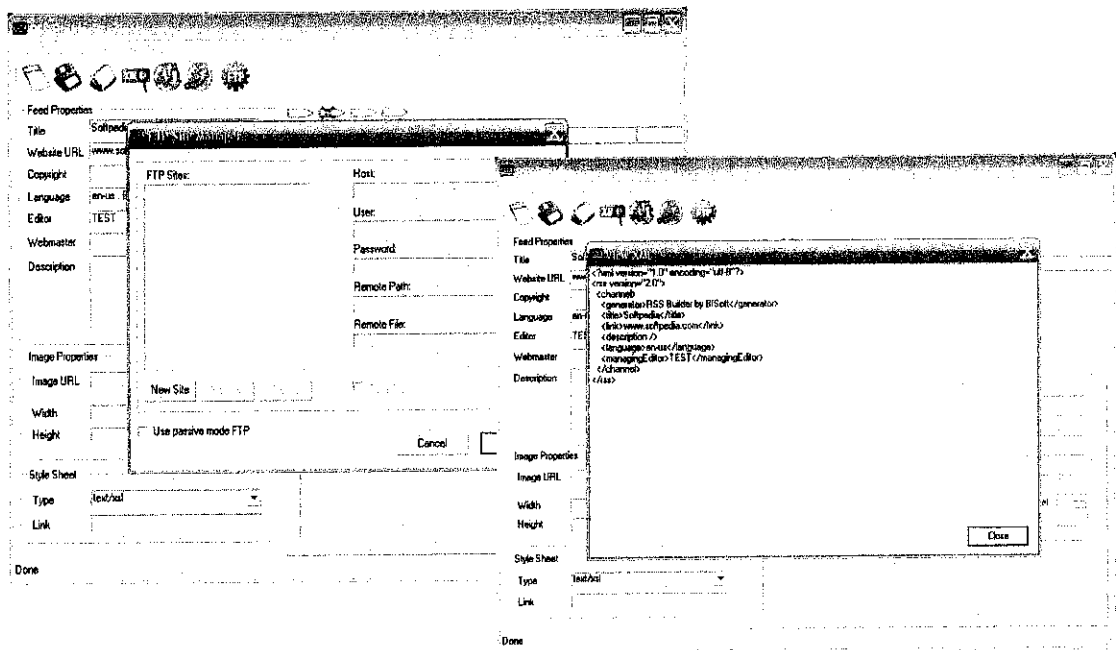


Figure 3.3– Screenshot of FTP site manager and View XML in RSS builder

RSS Builder is an easy to use program to create RSS news feeds and podcasts for the web site. It provides a simple interface which we can add topics, links, content, and enclosures, then publish the RSS (v2.0) feed to the web server, using the built-in FTP client. An RSS news feed is a list of topics which is made available from a web site, using RSS. To read the feed, the user installs a newsreader and enters the URL for the news feed. Podcasting is the distribution of audio (or video) files using RSS for listening on mobile devices and personal computers. The audience can listen (watch) when they want, where they want, and how they want.

Joomla!



Figure 3.4– Joomla! Logo

Joomla! is a free open source content management system for publishing content on the World Wide Web and intranets as well as a Model–view–controller (MVC) Web Application Development framework. The system includes features such as page caching to improve performance, RSS feeds, and printable versions of pages, news flashes, blogs, polls, website searching, and language internationalization. Joomla is licensed under the GPL, and is the result of a fork of Mambo. It is written in the PHP programming language and uses the MySQL database.

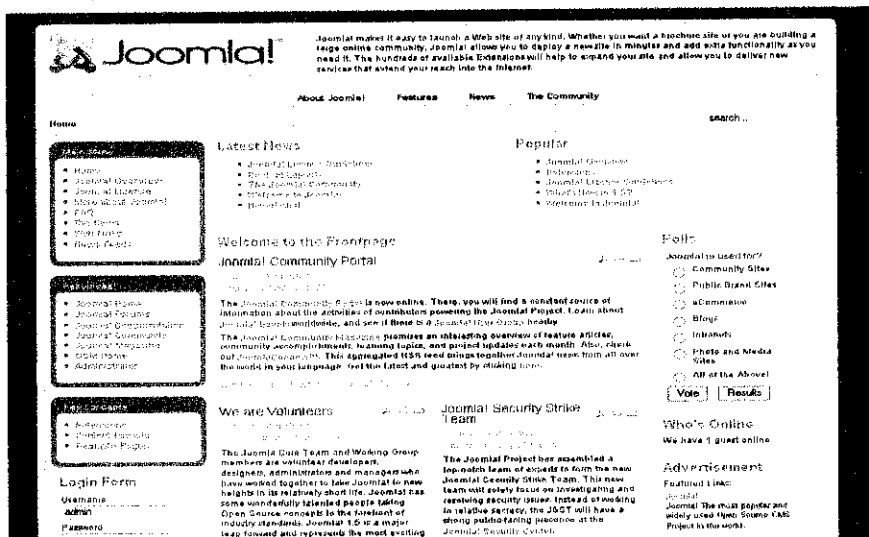


Figure 3.5– Joomla! Web screenshot

The Joomla! Package consists of many different parts, which allow modular extensions and integrations to be made easily. An example of such is extensions called "Plugins". These are background extensions that extend Joomla! with new functionality. The WikiBot, for example, allows the author of Joomla! content to use "Wikitag" in Joomla! articles which will auto-create dynamic hyperlinks to Wikipedia articles when displayed. There are over 3,500 extensions for Joomla! available via the Extensions Directory.

WordPress



Figure 3.6– WordPress logo

From Wikipedia, the free encyclopedia on the web, WordPress is an open source blog publishing application. WordPress is the official successor of b2\cafelog which was developed by Michel Valdrighi. This WordPress has a templating system, which includes widgets that can be rearranged without editing PHP or HTML code, as well as themes that can be installed and switched between. The PHP and HTML code in themes can also be edited for more advanced customizations. WordPress also features integrated link management; a search engine-friendly, clean permalink structure; the ability to assign nested, multiple categories to articles; multiple author capability; and support for tagging of posts and articles. Automatic filters that provide for proper formatting and styling of text in articles (for example, converting regular quotes to

smart quotes) are also included. WordPress also has a rich plug-in architecture which allows users and developers to extend its functionality beyond the features that come as part of the base install. Finally, after find some interesting features in WordPress author had decided used the WordPress to create the prototype of the real estate website instead of Joomla! This WordPress also can be installed on the local machine.

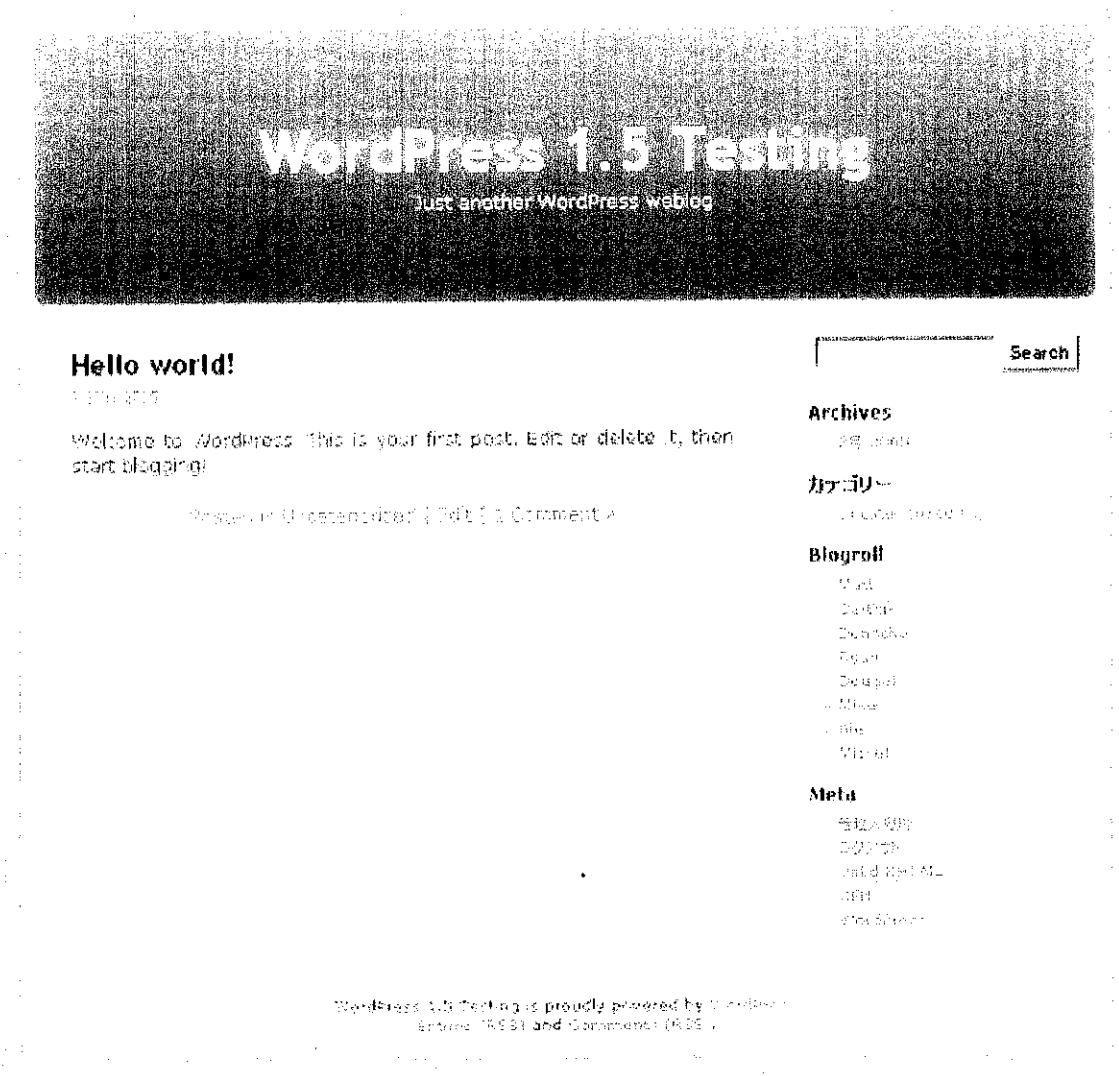


Figure 3.7 - WordPress Web screenshot

CHAPTER 4

RESULT AND DISCUSSION

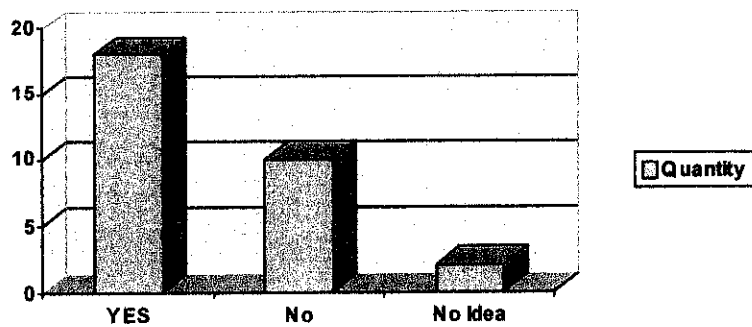
4.1 Introduction

The author of the project has distributed 30 questionnaires to the random people that have a potential with online real estate. In order to get their view regarding the online Real Estate Feed Reader. The questionnaire that contains 5 questions about real estate is distributed. The result of the questionnaire is analyzed and transform into graph in order to get clear view about online real estate.

4.2 Analysis

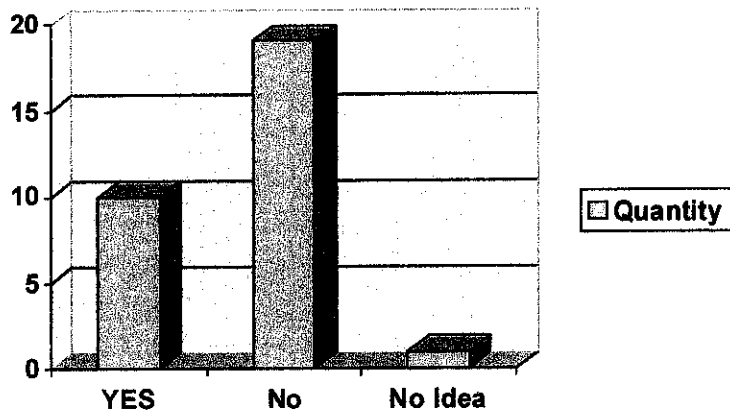
1. Do you know what is Online Real Estate

	YES	No	No Idea
Quantity	18	10	2



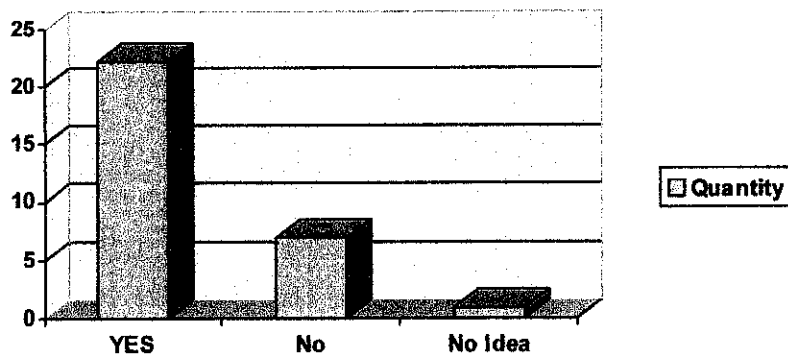
2. Do you browse the Online real estate website

	YES	No	No Idea
Quantity	10	19	1



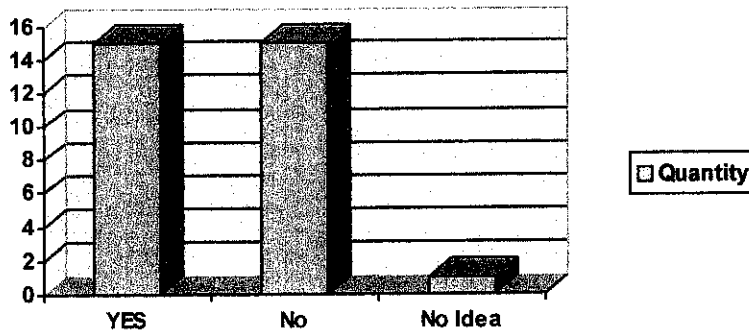
3. Do you know what is RSS feeds

	YES	No	No Idea
Quantity	22	7	1



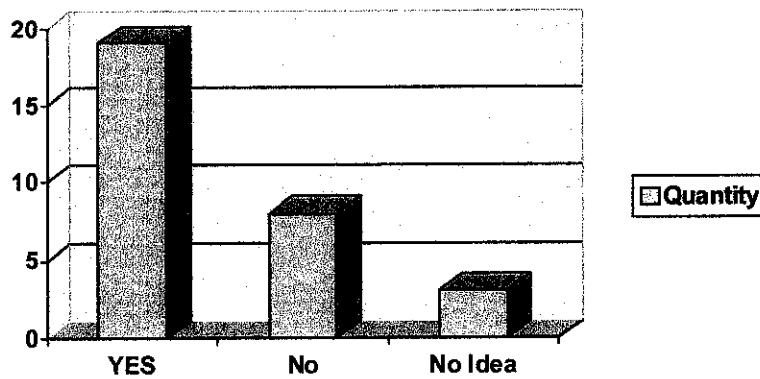
4. Do you use RSS Feed?

	YES	No	No Idea
Quantity	14	15	1



5. Do you agree if real estate feed is combining in one Website?

	YES	No	No Idea
Quantity	19	8	3



The result from the analysis, for the first question author can say that most of the

student know about what is actually online real estate, the term of real estate give them a clue what is website about, but from the second question not many of them browse this kind of websites, maybe they still not familiar because not related to their study and still to early for them to find home or land to rent or buy in their age. The third question is about the RSS feeds, majority of them know about the RSS maybe this familiar in blog or certain website on the internet. But the student who is not using the RSS is greater from the student who uses the RSS by one person and don't know about it. And for the last question most of the student agreed when the author ask their opinion about combining the RSS feed in the one website.

4.3 Result

4.3.1 Joomla! Real Estate website

Author manage to find open source content management system in order to build the web called Joomla! .After the study done on how to use that kind of system, finally author have managed to develop one website with the RSS functional.

This open source system is very helpful for the author to understand how the RSS work and simply can cut the duration for developing the web. But this open content management system is quite hard to understand because it have lot of function to be understand, and the most important is author have to find how to structure the RSS to achieve the objective of the project. However author still come out with a result to be view.

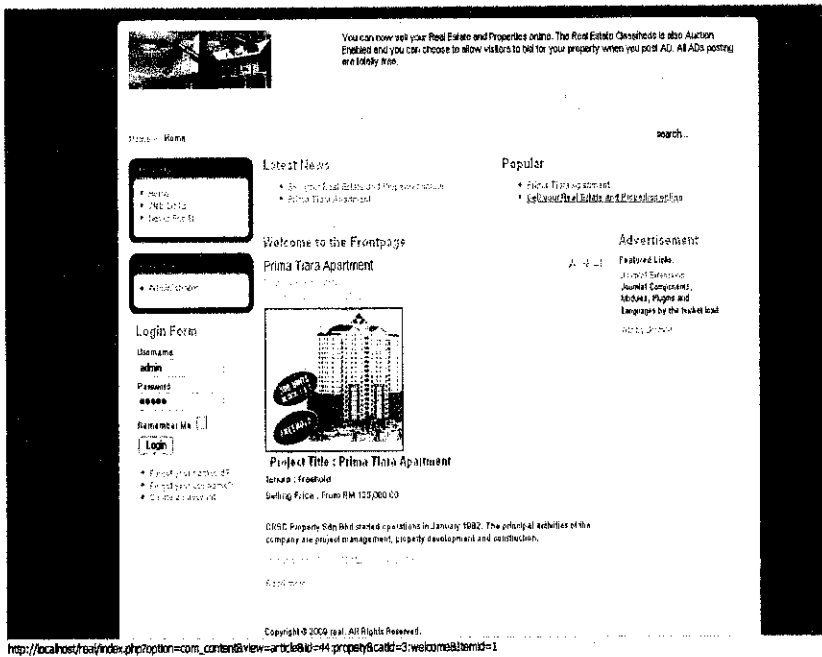


Figure 4.1– Online Real Estate website using Joomla!

For the front page, for the administration, admin can login in order to change or add the content of the webpage. To add the information, admin just can click the administrator link at the left of the web page and will be redirect to the admin login page.

After login, user will be redirect to the admin menu where it already builds to the administrator to change the content or add more function to the web using plug in.

For the RSS, user can subscribe the RSS feed by click at the RSS icon at the top of the navigation bar or at the bottom of the web.

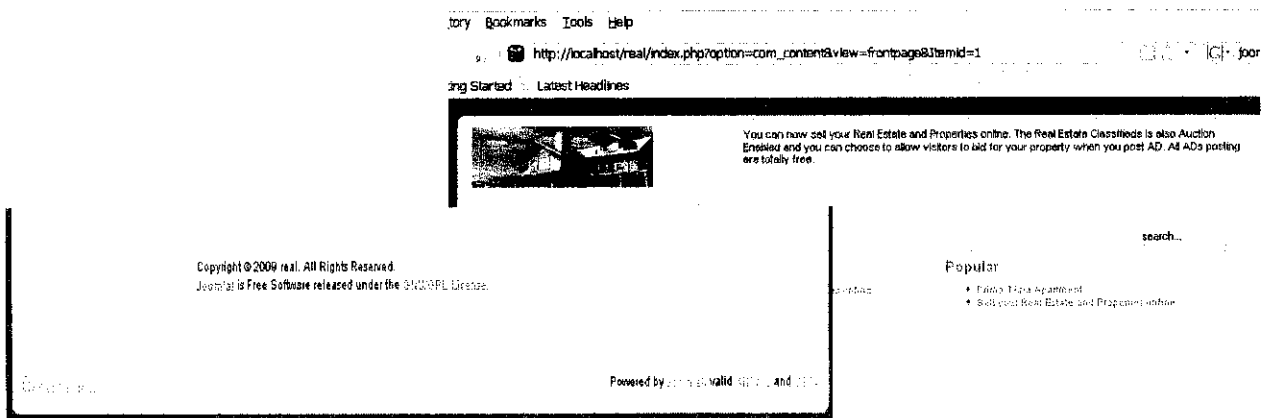


Figure 4.2– RSS icon at the page

After click the icon, the page will be redirect to the XML page, which the feed will be displayed to be subscribed, where the image of the content and detail will be displayed, but this feed is still under construction to meet specific requirement.

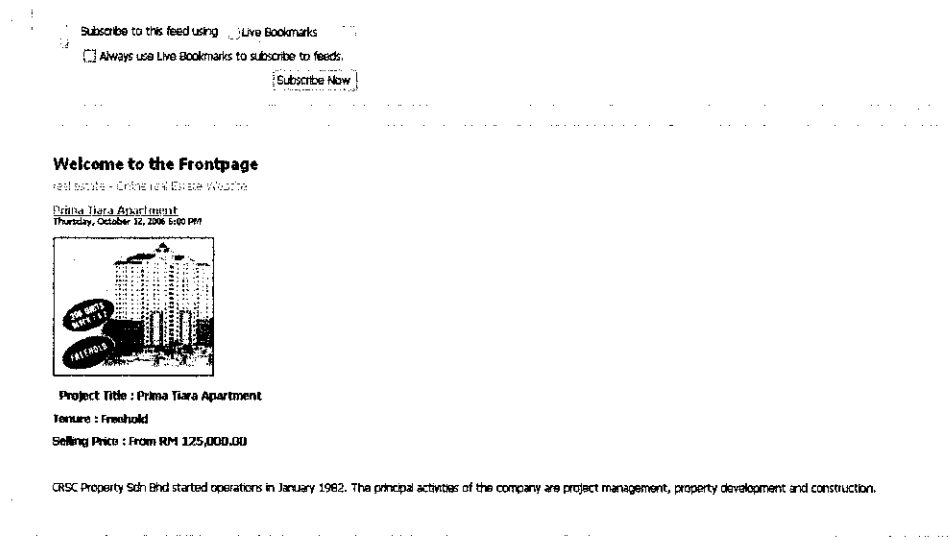


Figure 4.3– RSS feed page

```

<?xml version="1.0" encoding="utf-8"?>
<!-- generator="Joomla! 1.5 - Open Source Content Management" -->
<rss version="2.0">
  <channel>
    <title>Welcome to the Frontpage</title>
    <description>real estate - Online real Estate Website</description>
    <link>http://localhost/real/index.php?option=com_content&view=fr
    <lastBuildDate>Fri, 13 Mar 2009 03:36:52 +0000</lastBuildDate>
    <generator>Joomla! 1.5 - Open Source Content Management</generator>
    <language>en-gb</language>
    <item>
      <title>Prima Tiara Apartment</title>
      <link>http://localhost/real/index.php?option=com_content&view=fr
      <description><![CDATA[<p><!--[if !ms0]> <style> v\:* (behavi
      <category>frontpage</category>
      <pubDate>Thu, 12 Oct 2006 10:00:00 +0000</pubDate>
    </item>
  </channel>
</rss>

```

Figure 4.4— screenshot of Sample RSS code

4.3.2 WordPress Real Estate website

The other development regarding the Online Real Estate Feed Reader is the development of Real Estate Website using the WordPress. Since WordPress has rich of plug-in, author try to do some research about it and the RSS plug-in for the WordPress. This website will be used for the testing purpose with the Online Real Estate Feed Reader. After several day of research and configuration author had managed to build one complete real estate website using WordPress with the RSS feed. Figure below show the home page of the Real Estate website that using the WordPress and the Feed page.

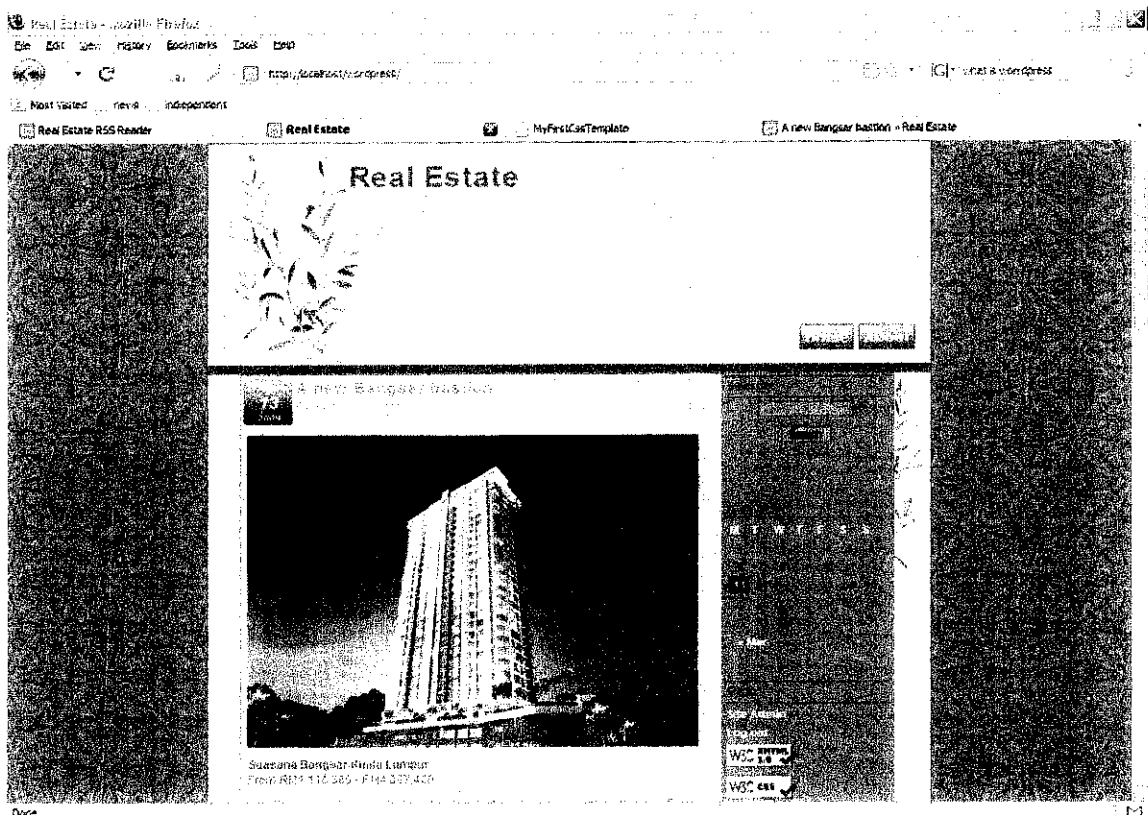
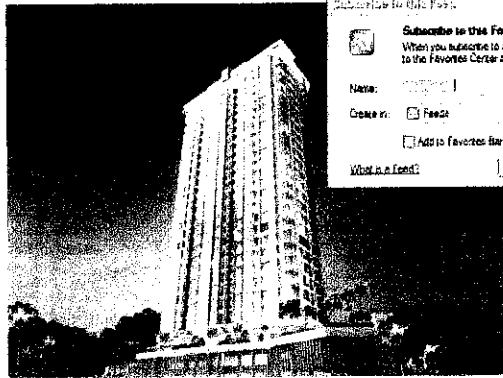


Figure 4.5 – WordPress Real Estate Web screenshot

A new Bangsar bastion



Subscribe to this Feed

Subscribe to this Feed
When you subscribe to a feed, it is automatically added to the Favorites Center and kept up to date.

Name:

Create in: Feed New folder

Add to favorites bar

What is a feed?

Sunsara Bangsar, Kuala Lumpur
From RM1 116 880 - RM4 867 420

An affluent neighbourhood where local and international residents call home, Bangsar is a flawless blend of old and new. Exquisite

Figure 4.6 – WordPress Real Estate Feed

4.3.3 Online Real Estate Feed Reader Overview

i. Home Page

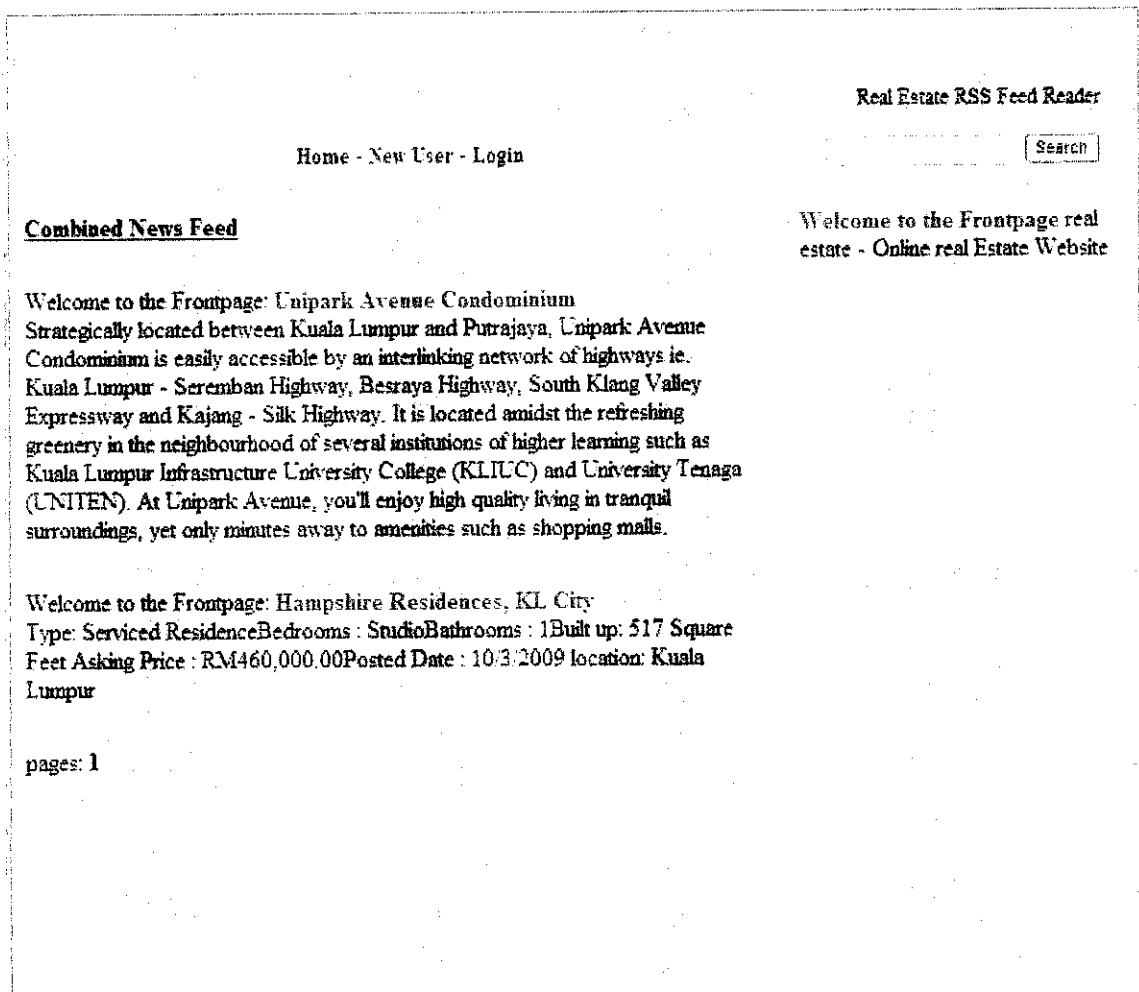


Figure 4.7 – Home page of Online Real Estate Feed Reader

Figure 4.7 shows the home page of the Online Real Estate Feed Reader, Here user can see the outline of the feeds news that taken from the other Real Estate website. From here admin or the web can login to add and remove the feed.

ii. Login Page and Admin Page

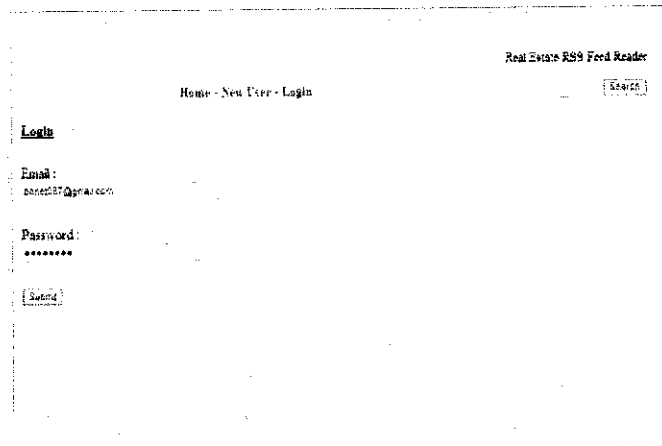


Figure 4.8 – Admin Login Page

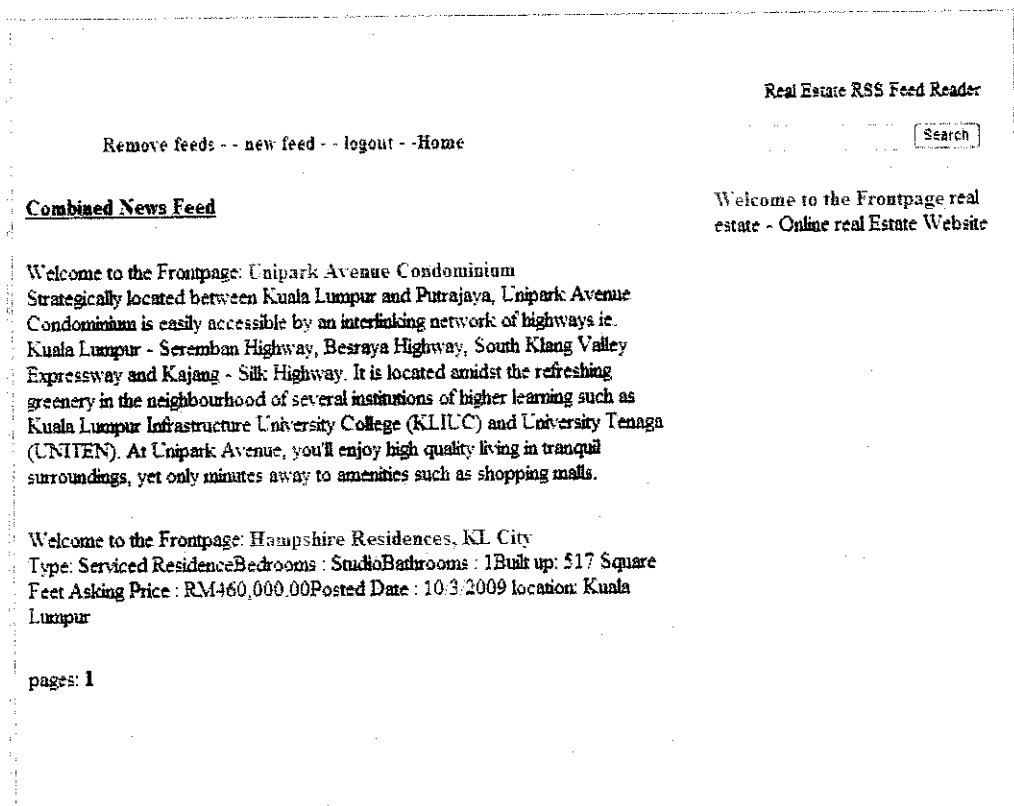


Figure 4.9 –Admin Home Page

Figure 4.8 shows that the admin can login in order to modify the feed and figure 4.9 shows the admin home page where in here feed can be added or remove.

iii. Add feed page

The screenshot shows a web interface titled "Real Estate RSS Feed Reader". At the top right, there is a search bar with a "Search" button. Below the title, there are navigation links: "Remove feeds -- new feed -- logout -- Home". The main section is titled "New Feed" and contains the instruction: "*To subscribe to new feed, enter the URL of the feed into the textbox". Below this instruction is a text input field containing the URL "http://localhost/wordpress/?feed=rss2". At the bottom of the form is a "Submit" button.

Figure 4.10 –Add New Feed

Figure 4.10 is the new feed page, where admin can submit the new RSS link from the real estate website RSS link. This link and its information will be stored in the database.

iv. Remove feed page

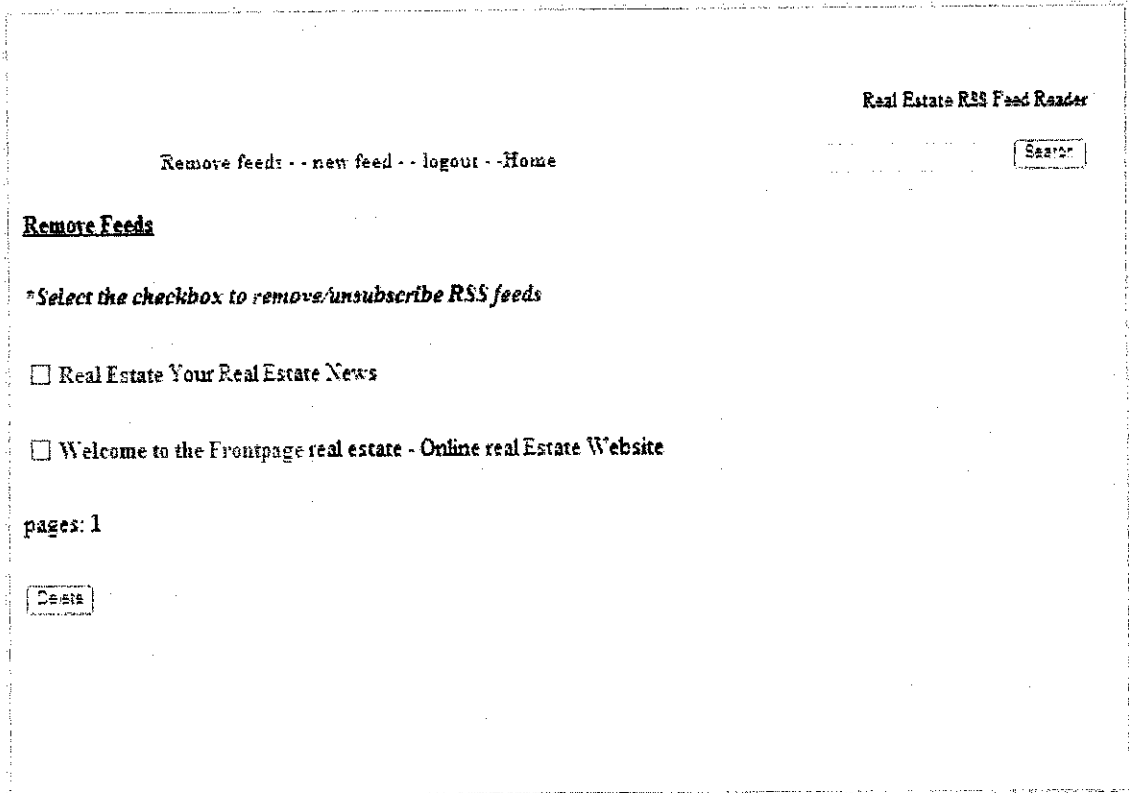


Figure 4.11 –Remove Feed

The figure 4.11 shows the Remove Feeds page, in this page admin need to select the checkbox in order to remove the selected feed.

v. Search The Database

The screenshot displays a web interface for a real estate database. At the top right, there is a search bar with the text "Real Estate RSS Feed Reader" above it. Below the search bar, there are two input fields: one containing "Kuala Lumpur" and another containing "4552". To the left of the search bar, there are navigation links: "Home", "New User", and "Login".

Below the navigation links, there is a section titled "Combined News Feed". This section contains several news items:

- Real Estate: A new Bangsar bacion**
Suzana Bangsar, Kuala Lumpur From RM6,116,980 - RM22,940,420 An affluent neighbourhood where local and international residents call home, Bangsar is a flawless blend of old and new. Existing infrastructures and facilities that give the neighbourhood its charm and personality are complemented by fresh new developments that promise a new wave of exciting energy, creating a multi-racial melting pot [...]
- Real Estate: Prima Tiara Apartment**
Project Title : Prima Tiara Apartament Tenure : Freehold Selling Price : From RM 115,000.00 CRSC Property Sdn Bhd started operations in January 1992. The principal activities of the company are project management, property development and construction. Since its incorporation, CRSC Property has expanded tremendously and is one of the leading [...]
- Welcome to the Frontpage: Unipark Avenue Condominium**
Strategically located between Kuala Lumpur and Putrajaya, Unipark Avenue Condominium is easily accessible by an interlinking network of highways i.e. Kuala Lumpur - Seremban Highway, Banting Highway, South Klang Valley Expressway and Kajang - Silk Highway. It is located amidst the refreshing greenery in the neighbourhood of several institutions of higher learning such as Kuala Lumpur Infrastructure University College (KLKUC) and University Tenaga (UNXTEN). At Unipark Avenue, you'll enjoy high quality living in tranquil surroundings, yet only minutes away to amenities such as shopping malls.
- Welcome to the Frontpage: Hampshire Residences, KL City**
Type: Serviced ResidenceBedrooms : StudioBathrooms : 1Built up: 117 Square Feet Asking Price : RM=60,000.00Posted Date : 10 2 2008 location: Kuala Lumpur

Figure 4.12 --Search Database

This Online Real Estate also can search the Real Estate by location from the database mostly for Kuala Lumpur area. The result will be display to user with link, user can simply click the link to browse the real estate information that they want to see.

vi. Database

port Wizard Export Wizard Filter Wizard Grid View Form View Memo Hex Image Sort Ascending

id	link	title	description	approved
38	http://www.kosmo.com.my/kosmo/rss.asp	Kosmo Online		1
	http://localhost/wordpress/?feed=rss2	Real Estate	Your Real Estate News	1

Figure 4.13 – Table Feeds

Import Wizard Export Wizard Filter Wizard Grid View Form View Memo Hex

items_id	feeds_id	title	description	link	md5	date_in
	39	Zarate agre	Mauro Zarate has	http://rss.	bc20c69d324	2009-05-02 05:44:50
1203	39	United give	Sir Alex Ferguson	http://rss.	3ad4b09b4ee	2009-05-02 05:44:50
1204	39	Wenger bac	Arsenal manager /	http://rss.	b1a96575290	2009-05-02 05:44:50
1205	39	Goalkeeper	Juventus star Gian	http://rss.	fa32bfb5d108	2009-05-02 05:44:50
1206	39	Real coach	Real Madrid coach	http://rss.	c817bacb056c	2009-05-02 05:44:50
1207	39	Mourinho re	Inter Milan coach	http://rss.	76051749141	2009-05-02 05:44:50
1208	39	Veteran Tuğ	Veteran Blackburn	http://rss.	b534b5fcc229	2009-05-02 05:44:50
1209	39	Israeli char	Israeli champions E	http://rss.	4e6817a173a	2009-05-02 05:44:50
1210	39	Hreidarsson	Hermann Hreidars	http://rss.	817bcbe5109	2009-05-02 05:44:50
1211	39	Warnock sig	Crystal Palace mar	http://rss.	56ecc0a6220c	2009-05-02 05:44:50
1212	39	Everton sig	Everton have sign	http://rss.	597d5281463	2009-05-02 05:44:50
1213	39	Cannavaro	Veteran Italian def	http://rss.	4028aa98ff69	2009-05-02 05:44:50

Figure 4.14 – Table Items

Figure 4.13 and 4.14 shows the feed table and the items table that stored in the database about the information of the subscribe feed.

CHAPTER 5

CONCLUSION AND RECOMMEDATION

5.1 Conclusion

As a conclusion, this online real estate feed Reader will improve the existing real estate on the internet whereby it can update the information frequently, so user can get the latest update from all the real estate websites by just viewing this web reader, where all the latest of the update news will be listed in the front page of the website. This means user just can have quick launch to the other real estate website directly through this Online Real Estate Feed Reader, this will automatically save time and more effective than by searching the real estate website using the search engine or using the traditional way where we must find the real estate agent by ourselves. This Real Estate Feed Reader also has the search function where it can help user find the specific location for the real estate in Malaysia especially for the KL area. Finally author hope this online real estate feed can improve the existing real estate website with its unique representation of the Real Estate information.

5.2 Limitations of the project

Below are the limitations faced during the development stage:

- i. The limitation of this project is that the Online Real Estate Feed Reader is depending on the other Real Estate Website to read the Feed, in order to make this Online Real Estate Reader functioning, there are several Real Estate with RSS have to be created and hosted in the local machine, this make the project take long time in development phase before it ready to be tested because there are lots of things need to be learn and check before the system prototyping is ready to be tested.

- ii. There are lack of the knowledge among the user regarding the RSS technologies, because of that some of the people did not understand the concept of the RSS and how it help the user in order to receive the information. This limitation creates a problem in the implementation phase in order to show and make the user understand and use the system.

5.3 Recommendation

The picture of the Real Estate still not available in this Online Real Estate Feed Reader version, because of time constraint, there are more research and configuration need to be done in order to understand the coding and concept used in the system. The search function also needs to enhance in the future development because the current systems just use the basic search function and the result might not accurate under certain condition.

REFERENCES

- A.H. Tan, *Text Mining: The State of the Art and the Challenges*, in PAKDD99 Workshop on Knowledge Discovery from advanced Databases, Beijing, China, April 1999.
- B. Mobasher, R. Cooley, and J. Srivastava: *Creating Adaptive Web Sites through Usage-Based Clustering of URLs* (1999), In Proc. of the 1999 IEEE Knowledge and Data Engineering Exchange Workshop (KDEX'99), November 1999.
- Chakrabarti, S. et al, Focused Crawling, "A New Approach to Topic- Specific Web Resource Discovery", *In Proceeding on the 8th International World Wide Web Conference*, Toronto, Canada, pp. 1623-1640, (1999).
- Cooley, R. et al, "Web Mining: Information and Pattern Discovery on the World Wide Web", *In Proceeding of IEEE International Conference Tools with AI*. Newport Beach, California, USA, pp. 558-567, (1997).
- GATE – General Architecture for Text Engineering, <http://gate.ac.uk/>, retrieved on 27 September 2008

M. Castellano, G. Mastronardi, A. Aprile, and G. Tarricone , “*A Web Text Mining*

Flexible Architecture, Retrieved on 27 September 2008

Ronen Feldman and James Sanger, the Text Mining Handbook, Cambridge University

Press. R.Kimball and R. Merz: “The Data Webhouse Toolkit, Building the Web-Enabled Data Warehouse”, John Wiley & Sons, January 2000.

W. Zhong, X. Tang, “Web Text Mining on XSSC” Institute of System Science,

Academy of Mathematics and System Science, Retrieved on 26 September 2008

Derick Van Ness (2008). “How to sell a house in a slow real estate market – The 5 best ways”, Retrieve on 30 October from

<http://www.hartanah.org/?p=9>

Hari Mailvaganam(2008). “Text Mining Fraud Detection”, Retrieve on 30 October from

http://www.dwreview.com/Data_mining/Effective_Text_Mining.html

Mark Pilgrim (2002).”What is RSS”, Retrieve on 18 December202 from

<http://www.xml.com/pub/a/2002/12/18/dive-into-xml.html>