

RSS FEEDS

Swapnil Shah(Team Leader)

Kaustav Das

Niket Bagwe

Chandra Pal Singh

Team:23

Department of Computer Science Engineering
IIT Bombay

November 11, 2010

- 1 Introduction
- 2 Problem Statment
- 3 Design
- 4 Screen shots
- 5 Implementation Techniques
- 6 Test cases and Results
- 7 Work Distribution
- 8 References

What is RSS?

- RSS (Really Simple Syndication) is one of the web feed formats used to publish frequently updated works such as blog entries, news headlines, audio, and video in a standardized format.

What is RSS?

- RSS (Really Simple Syndication) is one of the web feed formats used to publish frequently updated works such as blog entries, news headlines, audio, and video in a standardized format.
- They benefit readers who want to subscribe to timely updates from their favored websites or to aggregate feeds from many sites into one place.

What are advantages RSS FEED?

- RSS is becoming popular day by day as it is a free and easy way to promote a site and its content without the need to advertise or create complicated content sharing partnerships.

What are advantages RSS FEED?

- RSS is becoming popular day by day as it is a free and easy way to promote a site and its content without the need to advertise or create complicated content sharing partnerships.
- A website willing to publish their content using RSS, creates one RSS feed and keeps it on an web server.

What are advantages RSS FEED?

- RSS is becoming popular day by day as it is a free and easy way to promote a site and its content without the need to advertise or create complicated content sharing partnerships.
- A website willing to publish their content using RSS, creates one RSS feed and keeps it on a web server.
- The big benefit to RSS is that individuals opt-in to content of interest, totally controlling the flow of information that they receive. User once subscribed can get updates whenever new information is available in that website. Users need not be bothered polling that website periodically to check updates.

What are disadvantages of RSS?

- Though it is a very advantageous technology main drawback is RSS sends extra information besides the intended one to a subscriber. The motivation is why subscriber bother about information not relevant to him.

What are disadvantages of RSS?

- Though it is a very advantageous technology main drawback is RSS sends extra information besides the intended one to a subscriber. The motivation is why subscriber bother about information not relevant to him.
- Main part of our project is the filtering of news to a specific query(i.e., sports,politics,music etc).

Problem Statement

- In this project we are going to modify that part i.e., filtering news based on subscriber's requirement and notify them through email.

Problem Statement

- In this project we are going to modify that part i.e., filtering news based on subscriber's requirement and notify them through email.
- Then we have seperated the news for each user and send those news to their mailbox.

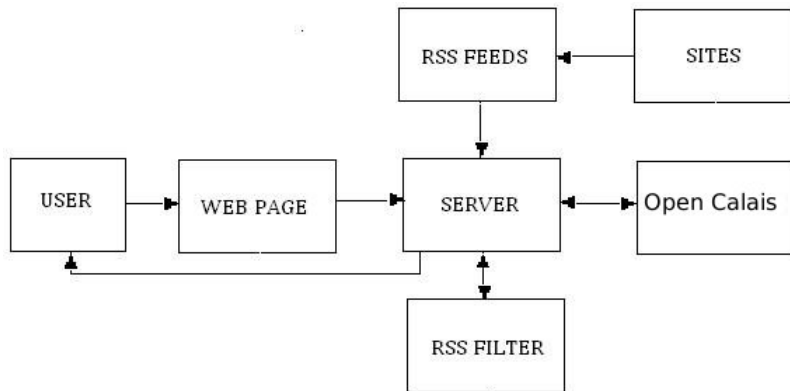
Motivation

- The motivation is why subscriber bother about information not relevant to him.

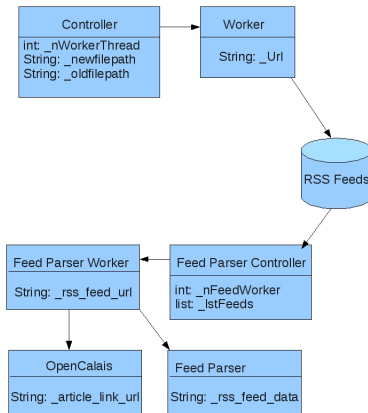
Motivation

- The motivation is why subscriber bother about information not relevant to him.
- Our project mainly deals with that part.Extracting the required information as per subscriber's query and send those through email.

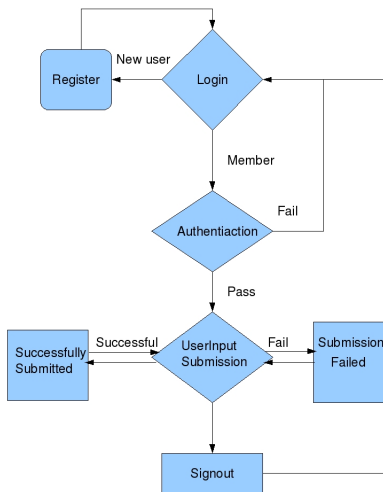
Block Diagram



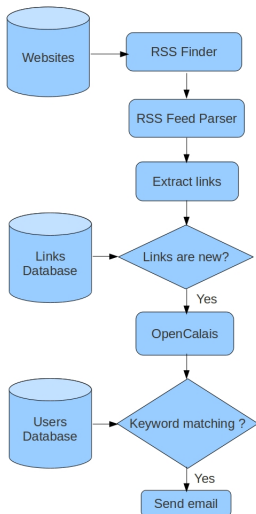
Class Diagram



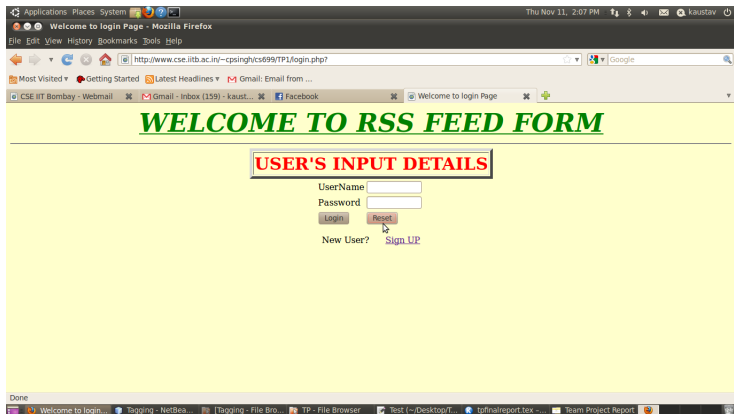
PHP Block Diagram



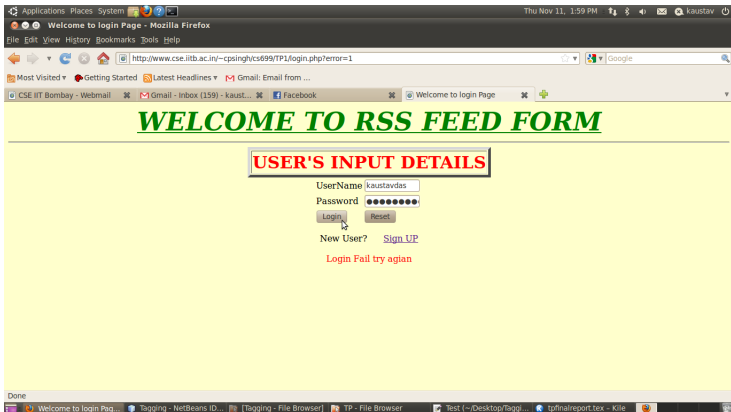
RSS Block Diagram



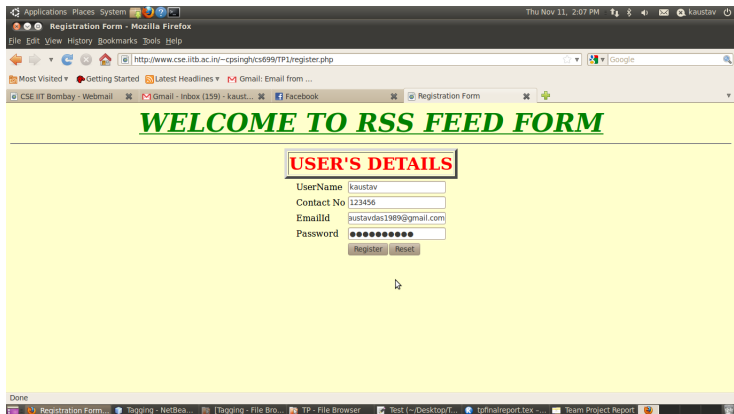
Login Page



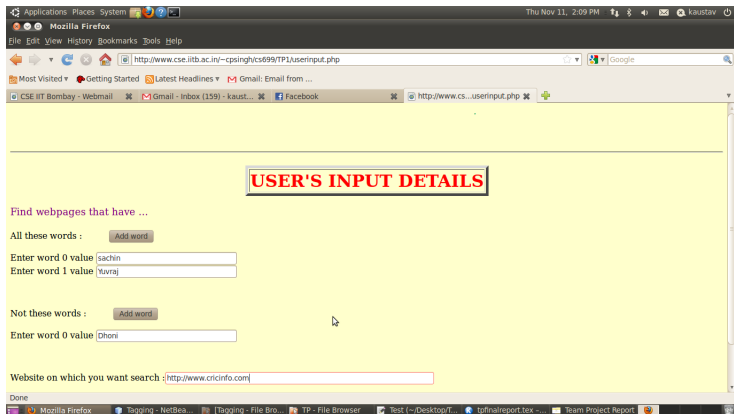
Invalid Login



Registration Page



RSS Feed Form



RSS ScreenShot

```

Parsing RSS XML File http://www.awaretek.com/python/index.xml
sending mail swapnilshah@cse.iitb.ac.in http://www.python.org/news/index.html#Mon06Sep20108300200
Reading old links database from Links/5links
Keyword matching...
Retrieved RSS Feeds from http://www.python.org
sending mail swapnilshah@cse.iitb.ac.in http://www.python.org/news/index.html#Tue24Aug20100945-040
Parsing RSS XML File http://www.python.org/dev/peps/peps.rss
Reading old links database from Links/5links
Keyword matching...
Downloading http://www.python.org
Finding RSS Feeds from http://www.python.org
sending mail swapnilshah@cse.iitb.ac.in http://www.python.org/news/index.html#Wed16Aug20101929-040
Parsing RSS XML File http://www.python.org/community/jobs/jobs.rss
Reading old links database from Links/5links
Keyword matching...
sending mail swapnilshah@cse.iitb.ac.in http://www.python.org/news/index.html#Wed04Aug20101809-040
Parsing RSS XML File http://www.reddit.com/r/Python/.rss
Reading old links database from Links/5links
Keyword matching...
sending mail swapnilshah@cse.iitb.ac.in http://www.python.org/news/index.html#Sun01Aug20108300200
sending mail swapnilshah@cse.iitb.ac.in http://www.reddit.com/r/Python/comments/e3yga/getting_star
Retrieved RSS Feeds from http://www.python.org
sending mail swapnilshah@cse.iitb.ac.in http://www.python.org/news/index.html#Thu29Jul2010830-0700
sending mail swapnilshah@cse.iitb.ac.in http://www.reddit.com/r/Python/comments/e43b5/python_progr
sending mail swapnilshah@cse.iitb.ac.in http://www.python.org/dev/peps/pep-3333
Downloading http://www.python.org
Finding RSS Feeds from http://www.python.org
sending mail swapnilshah@cse.iitb.ac.in http://www.python.org/news/index.html#Wed14Jul20108300200
sending mail swapnilshah@cse.iitb.ac.in http://www.reddit.com/r/Python/comments/e2hi0/hey_reddit_

```

Working

- RSS finder downloads RSS files from the websites and passes it to the RSS feed Parser.

Working

- RSS finder downloads RSS files from the websites and passes it to the RSS feed Parser.
- RSS feed parser extracts links from the RSS file.

Working

- RSS finder downloads RSS files from the websites and passes it to the RSS feed Parser.
- RSS feed parser extracts links from the RSS file.

Working

- RSS finder downloads RSS files from the websites and passes it to the RSS feed Parser.
- RSS feed parser extracts links from the RSS file.
- The new links are then passed to the open Calais to get the semantic information of the article.

Working

- RSS finder downloads RSS files from the websites and passes it to the RSS feed Parser.
- RSS feed parser extracts links from the RSS file.
- The new links are then passed to the open Calais to get the semantic information of the article.
- This semantic information is then matched with the keywords given by the user.

Working

- RSS finder downloads RSS files from the websites and passes it to the RSS feed Parser.
- RSS feed parser extracts links from the RSS file.
- The new links are then passed to the open Calais to get the semantic information of the article.
- This semantic information is then matched with the keywords given by the user.
- If the match succeeds then the link of the article is sent as mail to the user.

PHP Design

- In our design there is one php page for login which provides facility for new users to sign up. Once new user is registers his/her information is stored in database and becomes member of the system.

PHP Design

- In our design there is one php page for login which provides facility for new users to sign up. Once new user is registers his/her information is stored in database and becomes member of the system.
- Existing members can avail facility of RSS FEED. After successful authentication user can put his query. Query is in the terms of set of wanted keywords and set of unwanted keywords.

PHP Design

- In our design there is one php page for login which provides facility for new users to sign up. Once new user registers his/her information is stored in database and becomes member of the system.
- Existing members can avail facility of RSS FEED. After successful authentication user can put his query. Query is in the terms of set of wanted keywords and set of unwanted keywords.
- User also gives intended website url from where he requires information. After filling up all those user submits his/her query.

Open Calais

- The OpenCalais Web Service automatically creates rich semantic metadata for the content. It gives semantic aspects of content. Calais analyzes document and finds the entities within it.

Open Calais

- The OpenCalais Web Service automatically creates rich semantic metadata for the content. It gives semantic aspects of content. Calais analyzes document and finds the entities within it.
- Those entities have been filtered with respect to given key-words and page has been sent to user mail-box if it satisfies query criteria.

Web Module

- We did many tests and found positive results as follows.

Web Module

- We did many tests and found positive results as follows.
- When user give wrong password or username it gave error and don't let user enter without proper authentication.

Web Module

- We did many tests and found positive results as follows.
- When user give wrong password or username it gave error and don't let user enter without proper authentication.
- When user succesfully login then user don't have need to again login.

Web Module

- We did many tests and found positive results as follows.
- When user give wrong password or username it gave error and don't let user enter without proper authentication.
- When user succesfully login then user don't have need to again login.
- When user succesfully logout from his account then he can't login without reauthentication.

Web Module

- We did many tests and found positive results as follows.
- When user give wrong password or username it gave error and don't let user enter without proper authentication.
- When user succesfully login then user don't have need to again login.
- When user succesfully logout from his account then he can't login without reauthentication.
- After login user give details about for what he want to register at perticular website. Then his details are succesfully stored in proper file if not then it gave proper error.

Web Module

- We did many tests and found positive results as follows.
- When user give wrong password or username it gave error and don't let user enter without proper authentication.
- When user succesfully login then user don't have need to again login.
- When user succesfully logout from his account then he can't login without reauthentication.
- After login user give details about for what he want to register at perticular website. Then his details are succesfully stored in proper file if not then it gave proper error.
- After succesfully entered details user will get update from the registered website periodically by email.

RSS feed processing

- The application will not process duplicate links i.e. timestamp of links are being checked before processing.

RSS feed processing

- The application will not process duplicate links i.e. timestamp of links are being checked before processing.
- Open Calius module was tested for different sets of keywords. If relevance of wanted keywords is greater than 0.4 and of unwanted keywords is less than 0.15 then only the link is sent to the user.

Work Distibution

- Swapnil Shah did work mostly on Project Design, RSS FEED PRO-CESSOR, RSS FEED PARSER and OPEN CALAIS Module.
- Niket Bagwe did work mostly on RSS FEED PROCESSOR and some work on OPEN CALAIS Module.
- Chandra Pal Singh I did work mostly on PHP and some work on OPEN CALAIS module.
- Kaustav Das did work mostly on PHP and OPEN CALAIS module.

References

- RSS wikipedia <http://en.wikipedia.org/wiki/RSS>

References

- RSS wikipedia <http://en.wikipedia.org/wiki/RSS>
- RSS http://www.codeproject.com/KB/books/rss_samplechapter.aspx

References

- RSS wikipedia <http://en.wikipedia.org/wiki/RSS>
- RSS http://www.codeproject.com/KB/books/rss_samplechapter.aspx
- OpenCalais <http://www.opencalais.com>