# Social Media News Tracking System

Project proposal


By

**Mr. Pichet Potha 542115041**

**Mr. Sittipong Borripan 542115066**

Department of Software Engineering

College of Arts, Media and Technology

Chiang Mai University


Project Advisor

_____

**Dr. Chartchai Doungsa-ard**

# Document History

| Document name | Version | Status | Date | Viewable | Reviewer | Responsible |
|---|---|---|---|---|---|---|
| **Documents** | | | | | | |
| **Social Media News Tracking System-ProjectProposal_v.0.1.docx** | 0.1<br>• Introduction and background<br>• Business review<br>• Business tools and software review<br>• Technology review<br>• Development tools review<br>• System architecture<br>• Deliverables and limit<br>• Future work<br>• Reference | Draft | June 1,2015 | PP, SB, CD | PP, SB | PP, SB |
| **Social Media News Tracking System-ProjectProposal_v.0.2.docx** | 0.2<br>• Business Tools and Software review<br>• Quality standard<br>• Motivation<br>• Aim<br>• Objective<br>• Schedule and Milestones<br>• Reference | Draft | June 7,2015 | PP, SB, CD | PP, SB | PP, SB |
| **Social Media News Tracking System-ProjectProposal_v.0.3.docx** | 0.3<br>• Introduction and background<br>• Business review<br>• Technology review<br>• Schedule and Milestone<br>• Reference | Draft | June 8,2015 | PP, SB, CD | PP, SB | PP, SB |
| **Social Media News Tracking System-ProjectProposal_v.0.4.docx** | 0.4<br>• Introduction and background<br>• Technology review<br>• Deliverables and limit | Draft | June 8,2015 | PP, SB, CD | PP, SB | PP, SB |

# Table of Contents

Chapter One | Introduction and background……...…….…….……...……1

Chapter Two | Literature Review…...……..…...2

2.1 Business review……...…….2

2.2 Business tools and software review……...……..2

2.3 Technology review……...……...6

2.4 Development tools review…….……....10

Chapter Three | Quality standard……..……..……..11

3.1 Software project management in very small entities (VSE)…….…….11

Chapter Four | Project plan……...…….……...12

4.1 Motivation…….……12

4.2 Aim…….……...12

4.3 Objective…….……..12

4.4 System architecture…….……..13

4.5 Deliverables and limit…….…….……..14

4.5.1 Deliverables……..…….14

4.5.2 Limit……..……15

4.6 Future work…….…..15

4.7 Schedule and Milestones……...15

4.7.1 Schedule plan……...….15

Chapter Five | Reference……...……21

# Chapter One | Introduction and background

News is the things that people used to update the events that occur around the world whether the entertainment, science and tech, politics, or sports news. Now a day people can access to the news by using many media such as television, radio, newspaper, magazine, and especially social media. Because of the communication device that can access the internet has a low price, so everyone can use social media anytime from anywhere.

Social media can help people to update the news from anywhere around the world easily, but most of the news that we can see on the social media is a local or regional news. The news always comes from the journalist, but a lot of social media news are come from the netizen. The netizen usually shares the news, status, post, message or the content on the website or any social network. They also usually post the own event that occurs with themselves or their initial people on the social network or website. They may want to warn other people, to request the helping, to request the opinion or the comment, or just want to share the news with their friend or the netizen.

From the behavior of a netizen that usually post and share on the social network, something just maybe a rumor. Some rumors can effect to the people, place, and organization in both of good way and bad way. As there are a lot of people use the social network and their habit that usually post and share their status, posts, or article, they may not sure that it is true or wrong story. In addition, there are a lot of netizens also read and comment on their own opinion that may make the rumor is going to the better way, or the worse way. These will make the rumor spread over a wide area. For example in the previous time, there were a lot of bad rumors about Chiang Mai University on the social network. The CMU Officer comes so late to make the rumor clear, so the rumor was spread over a wide are on the social network. This event makes CMU reputation get worse.

Social Media News Tracking System is the system which is proposed for searching the status, posts, an article on the website "Pantip [9], Facebook [10], and Twitter [11]" that be the popular social network in Thailand. The system will find the content that match with the word that already set or enter by the user, and gathering those contents from all of the posts in these three websites and show them to the user. The system can tags that word for alert the user to know when there is the new posts or content that relate to the tags occurred. It will make the user knows about this so fast. The user can go to answer or do something to make the rumor clear and save the reputation before it was damaged.

In conclusion, this application can help the people who relate to the rumor, have a chance to explain or make the rumor clear before the other people will understand in the wrong way. After the people who get the rumor, understand in the right way about the rumor, they will not spread the rumor and also help to explain to the other people to understand to the right way. Understanding the rumor in the right way will help to decrease the conflicts between the people who have the different opinion.

# Chapter Two | Introduction and background

## 2.1 Business review

### 2.1.1 Overview

Social media network is increasing especially in Thailand [1] such as Facebook, Twitter, Pantip, etc. because anybody wants to communicate with each other to exchange attitude, activity, events or business. Social media allows people to posting varies of text, videos for the other can access it.

Currently, social media become a part of everyday life for millions of people around the world. A higher percentage of Thailand people use social media more than traditional communication like a letter. Because of it can save time and money just a few step on social media service.

Social media provides many advantages but on the other hand it can quickly disseminate unexpected news on social media, and it will be the cause of some occurred tragedy.

Nowadays, there are many system tracking for social media to monitor or control unexpected things.

## 2.2 Business tools and software review

There are many tools or software that have a capability to tracking, analyzing with a social media news. So we choose the most efficient and interesting software system to show at following.

### 2.2.1 NewsTracker

The NewsTracker [2] is an online social network focusing on news, allows the user to create their account for sharing information (videos, text), or event for any of others user. NewsTracker include of four features following.

- **Track** feature allows users to enter the keyword to search, and NewsTracker will organize all relates news in Newsline for users can view.
- **Write** feature, allows users to write their post with the insertion of images and videos.
- **Customization** feature, allows the user to subscribe to tags of their choice and they can choose to see only the news containing their tags, written by the authors that following, or recommended by their friends.
- **Me** feature, allows users to owns a personal page in the 'Me' page. And users can upload their profile image, select a cover, and add the personal description.

A NewsTracker is illustrated in Figure 1, 2.

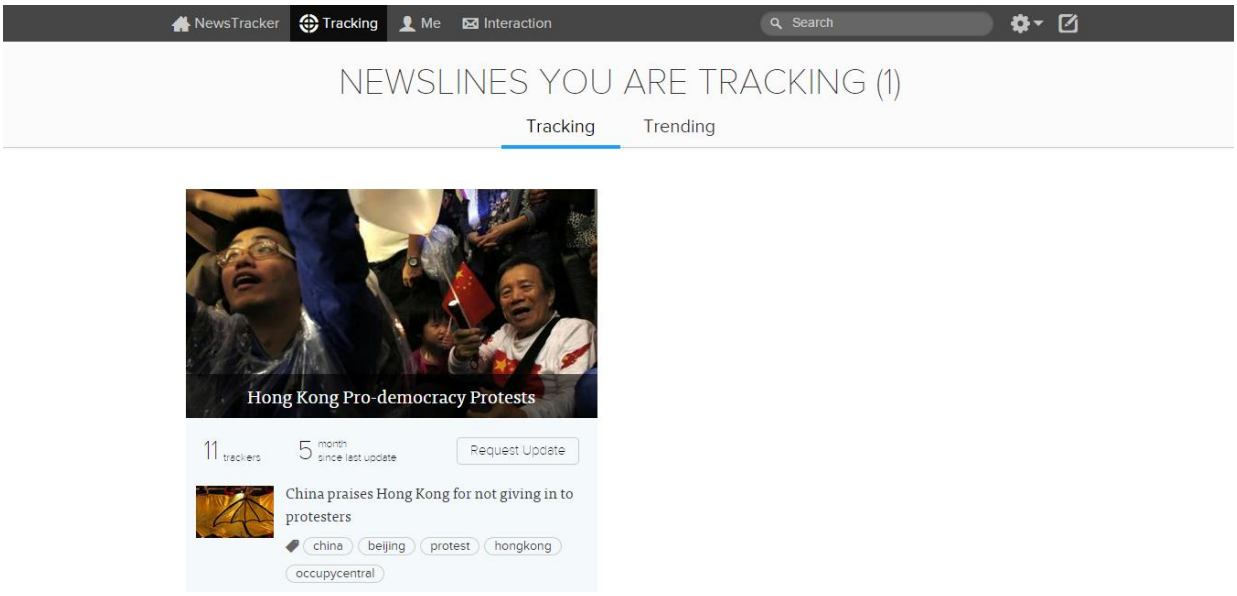



Figure 1: Newsline interface

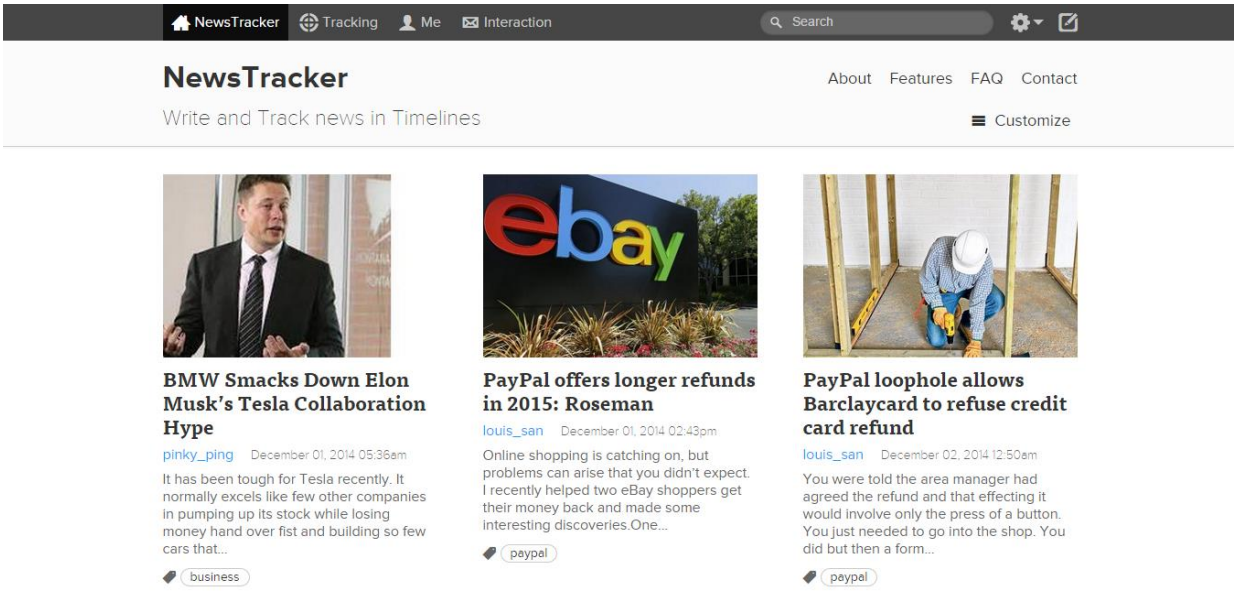



Figure 2: NewsTracker homepage

Pros

- NewsTracker has capability to tracking news
- NewsTracker has a good looking interface
- NewsTracker has responsive web design

Cons

- NewsTracker can track only own site news

### 2.2.2 Twazzup

Twazzup [3] is a social media monitoring tool for Twitter users only. Allows users enter a keyword, to get a lot of unique results like top tweets, photos and most tweeted about topics, subject, a real-time string of updates regarding the subject under consideration, list of the top Twitter users to have posted about that particular area, top 10 keywords in relation that users chosen. A Twazzup is illustrated in Figure 3, 4.

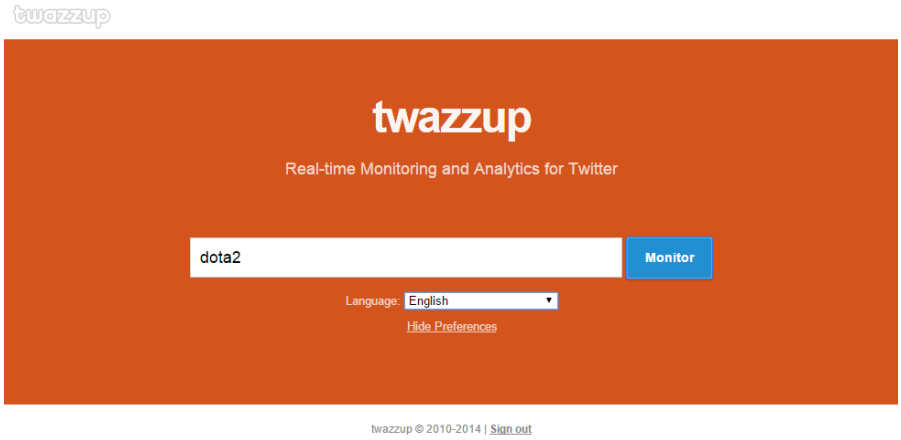



Figure 3: Twazzup homepage

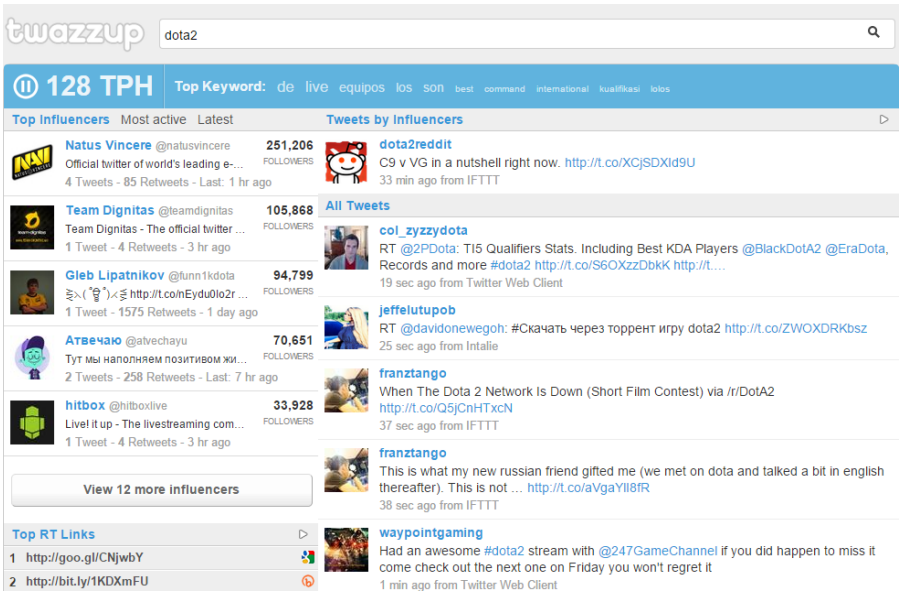



Figure 4: Twazzup search result

Pros

- Twazzup has real-time update
- Twazzup provide specific searching
- Twazzup return analysis result of text search

Cons

- Twazzup cannot track news
- Twazzup only work on twitter

## 2.3 Technology review

### 2.3.1 Spring Framework

Spring Framework [4] is an application framework and inversion of control container for the Java platform. The framework's core features can be used by any Java application, but there are extensions for building web applications on top of the Java EE platform. Spring Framework includes several modules that provide a range of services such as data access, MVC and testing.

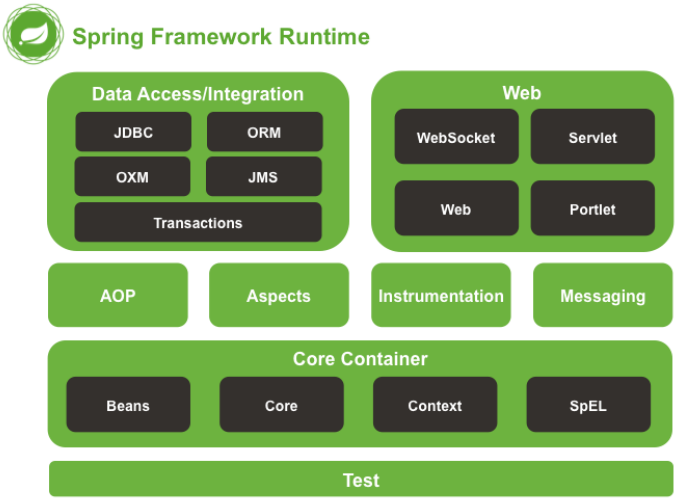



Figure 5: Spring framework architecture

In figure 5 show that Spring Framework provides about 20 modules which can be used based on an application requirement.

**Core Container**

The Core Container consists of the core, Beans, Context, and Expression Language modules whose detail is as follows [5]:

- The Core module provides the fundamental parts of the framework, including the IoC and Dependency Injection features.
- The Bean module provides BeanFactory which is a sophisticated implementation of the factory pattern.
- The Context module builds on the solid base provided by the Core and Beans modules and it is a medium to access any objects defined and configured. The ApplicationContext interface is the focal point of the Context module.
- The Expression Language module provides a powerful expression language for querying and manipulating an object graph at runtime.

**Data Access/Integration**

The Data Access/Integration layer consists of the JDBC, ORM, OXM, JMS and Transaction modules whose detail is as follows [5]:

- The JDBC module provides a JDBC-abstraction layer that removes the need to do tedious JDBC related coding.
- The ORM module provides integration layers for popular object-relational mapping APIs, including JPA, JDO, Hibernate, and iBatis.
- The OXM module provides an abstraction layer that supports Object/XML mapping implementations for JAXB, Castor, XMLBeans, JiBX and XStream.
- The Java Messaging Service JMS module contains features for producing and consuming messages.
- The Transaction module supports programmatic and declarative transaction management for classes that implement special interfaces and for all your POJOs.

**Web**

The Web layer consists of the Web, Web-Servlet, Web-Struts, and Web-Portlet modules whose detail is as follows [5]:

- The Web module provides basic web-oriented integration features such as multipart file-upload functionality and the initialization of the IoC container using servlet listeners and a web-oriented application context.
- The Web-Servlet module contains Spring's model-view-controller (MVC) implementation for web applications.
- The Web-Struts module contains the support classes for integrating a classic Struts web tier within a Spring application.
- The Web-Portlet module provides the MVC implementation to be used in a portlet environment and mirrors the functionality of Web-Servlet module.

**Miscellaneous**

There are few other important modules like AOP, Aspects, Instrumentation, Web and Test modules whose detail is as follows [5]:

- The AOP module provides aspect-oriented programming implementation allowing you to define method-interceptors and pointcuts to cleanly decouple code that implements functionality that should be separated.
- The Aspects module provides integration with AspectJ which is again a powerful and mature aspect oriented programming (AOP) framework.
- The Instrumentation module provides class instrumentation support and class loader implementations to be used in certain application servers.
- The Test module supports the testing of Spring components with JUnit or TestNG frameworks.

Pros

- Spring has many web supporting
- Spring reduce redundancy of programming
- Spring is good designed

### 2.3.2 Responsive web design (Bootstrap)

Overview

Responsive web design is the practice of building a website suitable to work on every device and every screen size, no matter how large or small, mobile or desktop. Responsive web design is focused around providing an intuitive and gratifying experience for everyone. Desktop computer and cell phone users alike all benefit from responsive websites [6].

Figure 6: Responsive web design in each device

In figure 6 show the flexible of layout in each device, but it is the same content. For good looking and benefit of any user device.

Pros

- Use single code base, for any platform
- Reduce cost of developing
- Looking good design

### 2.3.3 JSOUP

Jsoup [7] is a Java library for working with real-world HTML. It provides a very convenient API for extracting and manipulating data, using the best of DOM, CSS, and jquery-like methods. jsoup is designed to deal with all varieties of HTML found in the wild; from pristine and validating, to invalid tag-soup; jsoup will create a sensible parse tree.

Pros

- Jsoup is simplified URL fetching to the extreme.
- Jsoup is facilitating the use of cookies
- Jsoup allows the of use CSS

## 2.4 Development tools review

2.4.1 NetBeans IDE [6]

Overview

NetBeans IDE [8] is an open-source integrated development environment. NetBeans IDE supports development of all Java application types. Also Users can choose to download NetBeans IDE bundles tailored to specific development needs. Users can download and install all other features at a later date directly through the NetBeans IDE. Such as NetBeans IDE Bundle for Web and Java EE, including the JavaServerPage (JSP), Spring, Hibernate. It includes Apache Tomcat. Below is an illustrated of NetBeans IDE.

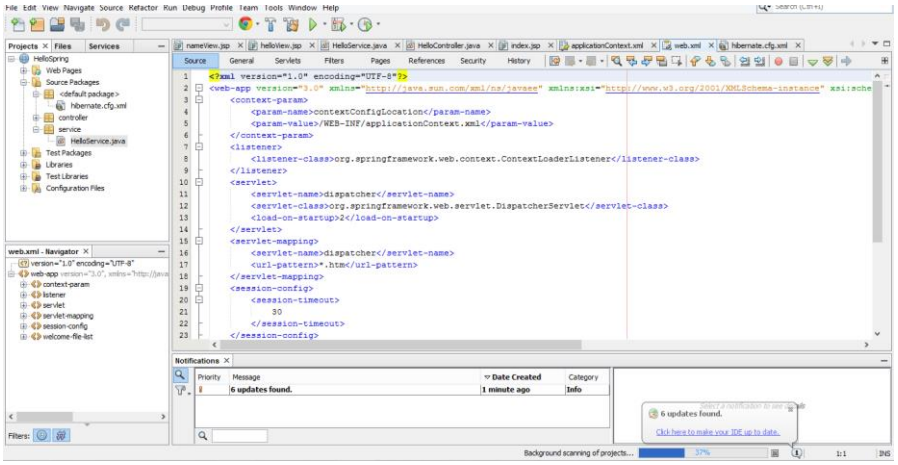


Figure 7: NetBeans interface

Alternative tools

- Eclipse
- Intelji

The selection of this tool

- Convenience to setup spring framework
- Open source

# Chapter Three: Quality Standard

## 3.1 ISO29110 for Very Small Entity (VSE)

ISO29110 is a guide applies to a Very Small Entity (VSE), enterprise, organization, department or project up to 25 people, dedicated to software development. The Guide provides Project Management and Software Implementation processes which integrate practices based on the selection of ISO/IEC 12207-Systems and Software Engineering —Software Life Cycle Processes and ISO/IEC 15289 Software Engineering – Software Life Cycle Process – guidelines for the content of software life cycle process information products (documentation) standards elements.

### 3.1.1 Project Management process

The purpose of the Project Management process is to establish and carry out in a systematic way the tasks of the software implementation project, which allows complying with the project's objectives in the expected quality, time and cost.

**Selected process**

3.1.1.1 Project Planning Process

3.1.1.2 Project Plan Execution Process

3.1.1.3 Project Assessment and Control Process

3.1.1.4 Project Closer Process

3.1.2 Software Implementation process

The purpose of the Software Implementation process is the systematic performance of the analysis, design, construction, integration and tests activities for new or modified software products according to the specified requirements.

**Selected process**

3.1.2.1 Software Implementation Initiation Process

3.1.2.2 Software Requirements Analysis Process

3.1.2.3 Software Architectural Design Process

3.1.2.4 Software Construction Process

3.1.2.5 Software Integration and Test Process

3.1.2.6 Software Delivery Process

# Chapter Four| Project Plan

## 4.1 Motivation

The rumor always occurs in everyday life. Although it occurred by deliberately or not, it will have the effect on the people who were mentioned in each rumor. For the example, in the recently time there is the bad rumor about Chiang Mai University (CMU). The netizen post, share, comment, and talk about this rumor. Some of them show the negative opinion because they believe that this bad rumor is true. However, there are some of the netizens who don't agree or believe in this rumor try to protect the reputation by showing their positive opinion about this rumor. Although the CMU officer try to fix this problem, they so late because the rumor already spread over a wide on the social network.

Follow the issue above, Social Media News Tracking System can help user can search and check the rumor on the threes popular social network in Thailand including "Pantip, Facebook, and Twitter". It will help them can found and fix the problem before the rumor were spread over a wide area.

## 4.2 Aim

The aim of this project is to develop a responsive web application that support any browser device. Social Media News Tracking System using Jsoup to parse HTML, file, or string to match a content in popular social media in Thailand such as Facebook, Twitter and Pantip[6]. Also provide a convenience tacking function to the user to receive news update.

## 4.3 Objective

- To help user can found a bad rumor that already occur in the social network.
- To make user can found a bad rumor in the social network faster when the news one occur, the user can clear it before it spread over a wide area.

## 4.4 System Architecture

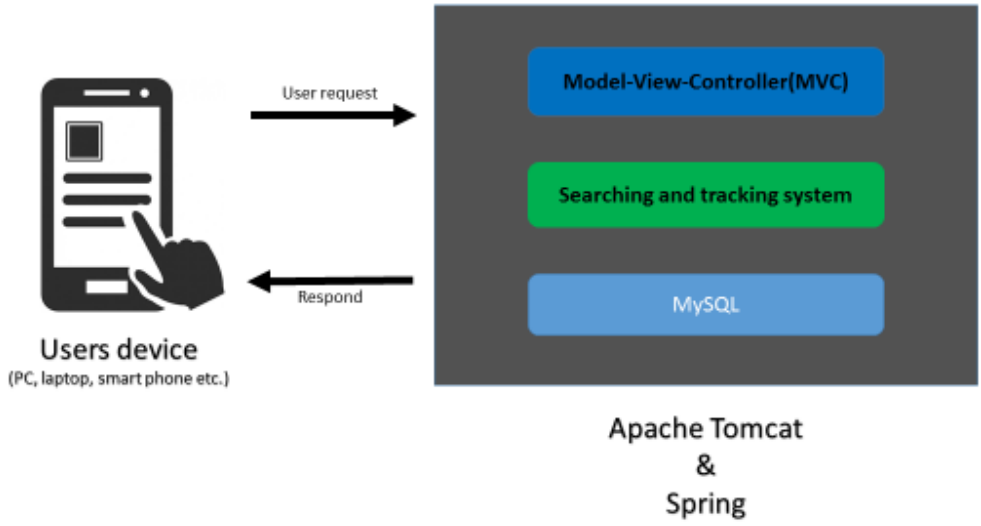



Figure 8: System architecture.

Figure 8 show the system architecture of Social Media News Tracking System. Firstly, the system receives an input keyword from user via user's device. Then it send data to the server and system on server will recommend an appropriate solutions for user using "searching". After that, system returns a searching keyword result to user and system allows user to tracking by let user register to system. Finally the system will save user account and tracking information into database.

## 4.5 Deliverables and Limits

### 4.5.1 Deliverables

4.5.1.1 Architecture Overview

**Feature 1: User Management system**

There are two types of user in this application that can use this application with the different function. The first one is User, this type of user just can only use searching System to see the result in each time that he searched. The second is Member, this type of user can use both of Search and tracking system for receive the update the result after the previous time that he was searching for each tag.

**Feature 2: Searching system**

This system using to search the result by crawl on "Pantip, Facebook, and Twitter" and analyzes the media that posted by the users. The system will parser to the HTML, CSS, or class of each website to match the tag, elements, or the content that related to the keyword that tagged or inputted by the user. Then, it will get the content of the media is related to what the users' interested.

**Feature 3: Tracking system**

This system will tag the keyword that Member interest by the member must select to tag that keyword and the system will continue searching the result to find the post that has been created after the previous searching.

**Feature 4: Real-time alert system**

This system will alert the Member when there are the new posts about the tags that Member interest has been created. If there are the posts has been created after the previous searching. The system will alert and update the result for the Member.

4.5.1.2 Documents

- Proposal
- Project Plan
- Software Requirement Specification
- Software Design Document
- Testing Document
    - Test Plan
        - Unit Test Plan
        - System Test Plan
    - Test Report
        - Unit Test report
        - System Test report
- Traceability Record

### 4.5.2 Limit

- User require the device that have the internet browser and can connect to the internet.
- User can see the result from website “Pantip, Facebook, and Twitter” only.
- User require the account of “Facebook” and “Twitter” to access to the content in this two websites
- User can get only the result that relate with own account.

## 4.6 Future work

- The system will have interface the interface that more easy to use.
- The system can crawl on more social media application.
- The system can show the result with the location that the event occurred.

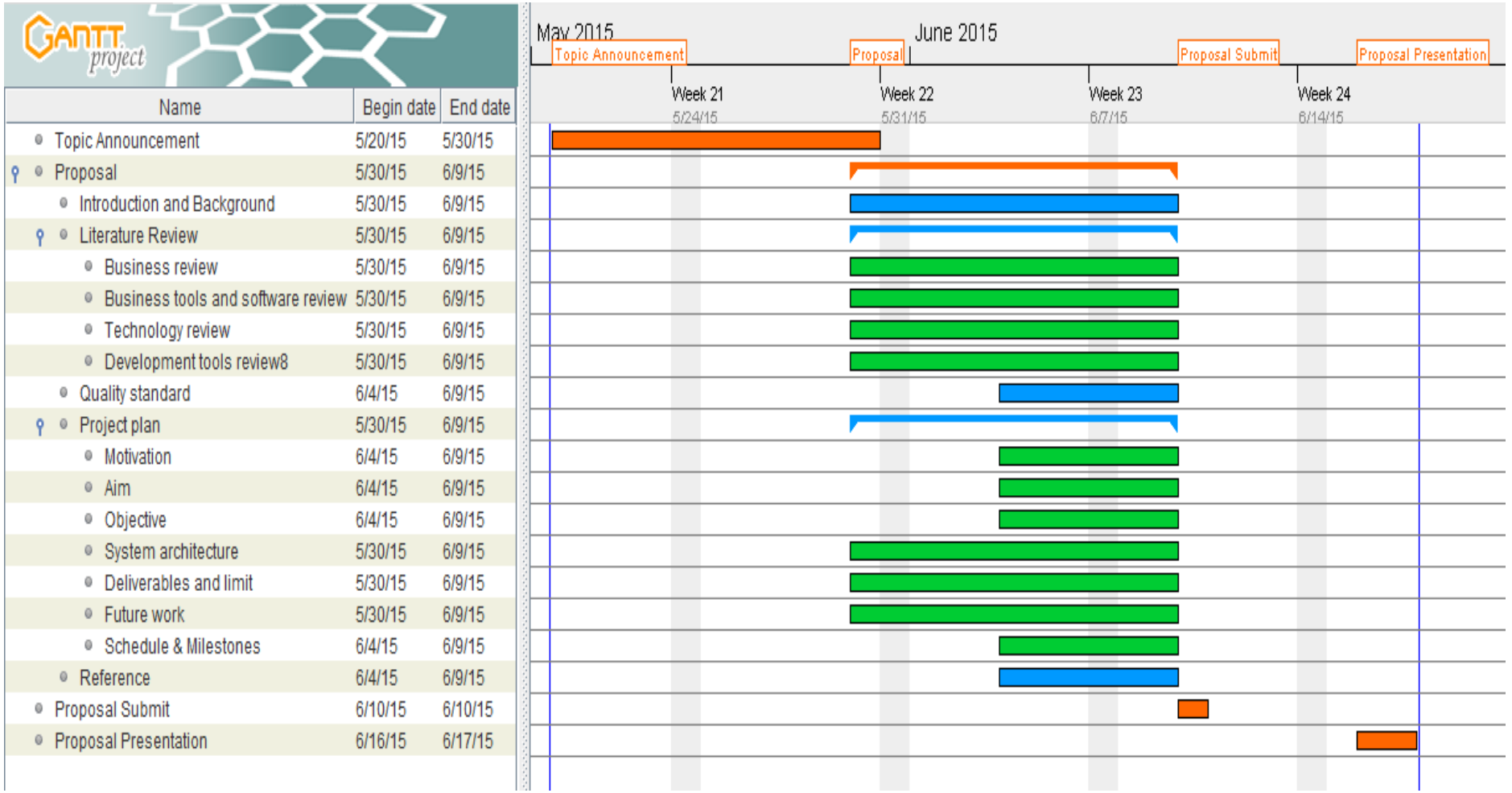
## 4.7 Schedule and Milestones

The schedule and milestones of Social Media News Tracking System that will show the plan and all process during start until end the project.

### 4.7.1 Schedule Plan

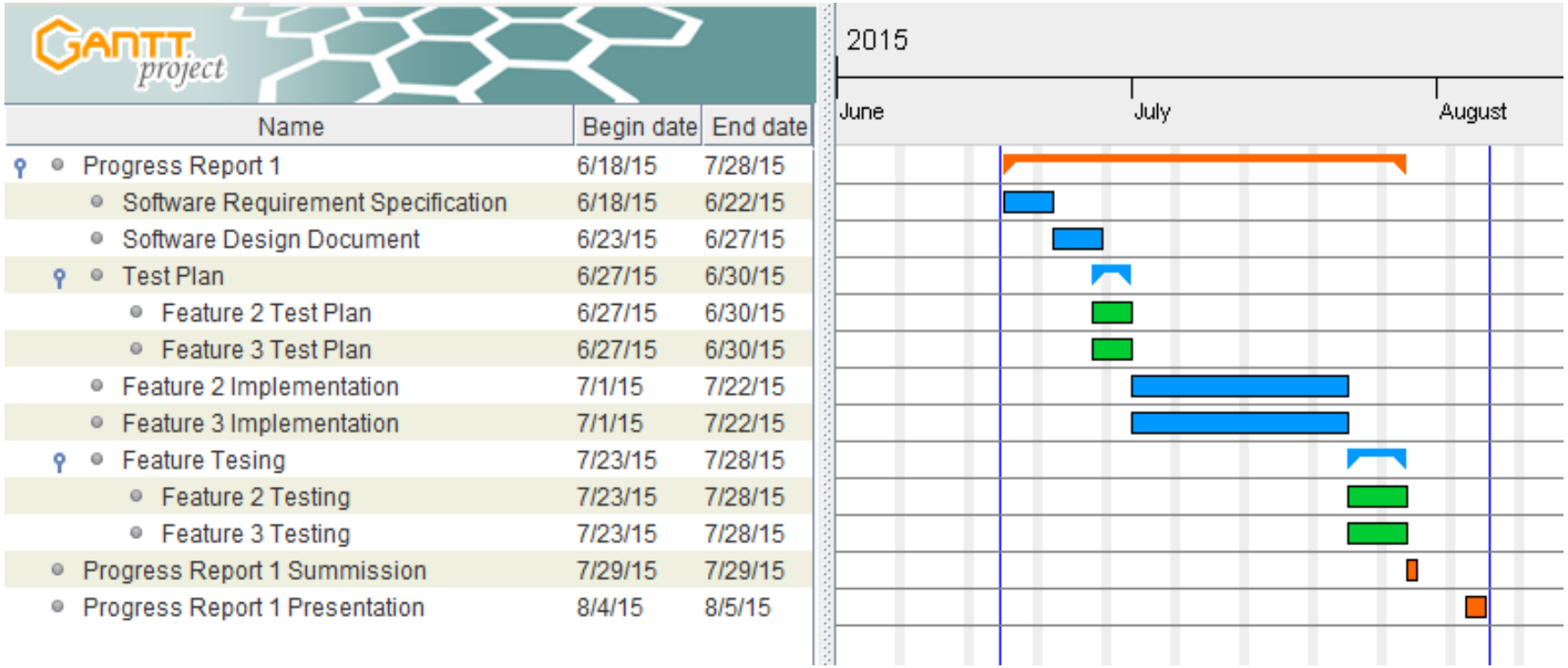
| Milestone | Task | Milestone Criteria | Planned date |
|---|---|---|---|
| 1 | Proposal | • Topic defined<br>• Create Proposal<br>• Proposal reviewed<br>• Proposal submitted<br>• Proposal presentation | June |
| 2 | Progress Report I | • Software requirement specification<br>• Feature 2 , Feature 3<br>• Software design document<br>• Test Plan<br>• Feature implemented<br>• Feature test report<br>• Traceability record progress I<br>• Progress report I submitted<br>• Progress report I presentation | June - August |

| | | | |
|---|---|---|---|
| 3 | Progress Report II | • Feature 4<br>• Software design document<br>• Test Planed<br>• Feature implementation<br>• Feature test report<br>• Traceability record progress II<br>• Progress report II submitted<br>• Progress report II presentation | August - September |
| 4 | Progress Report III | • Feature 4<br>• Software design document<br>• Test Planed<br>• Feature implementation<br>• Feature test report<br>• Traceability record progress III<br>• Progress report III submitted<br>• Progress report III presentation | September - November |



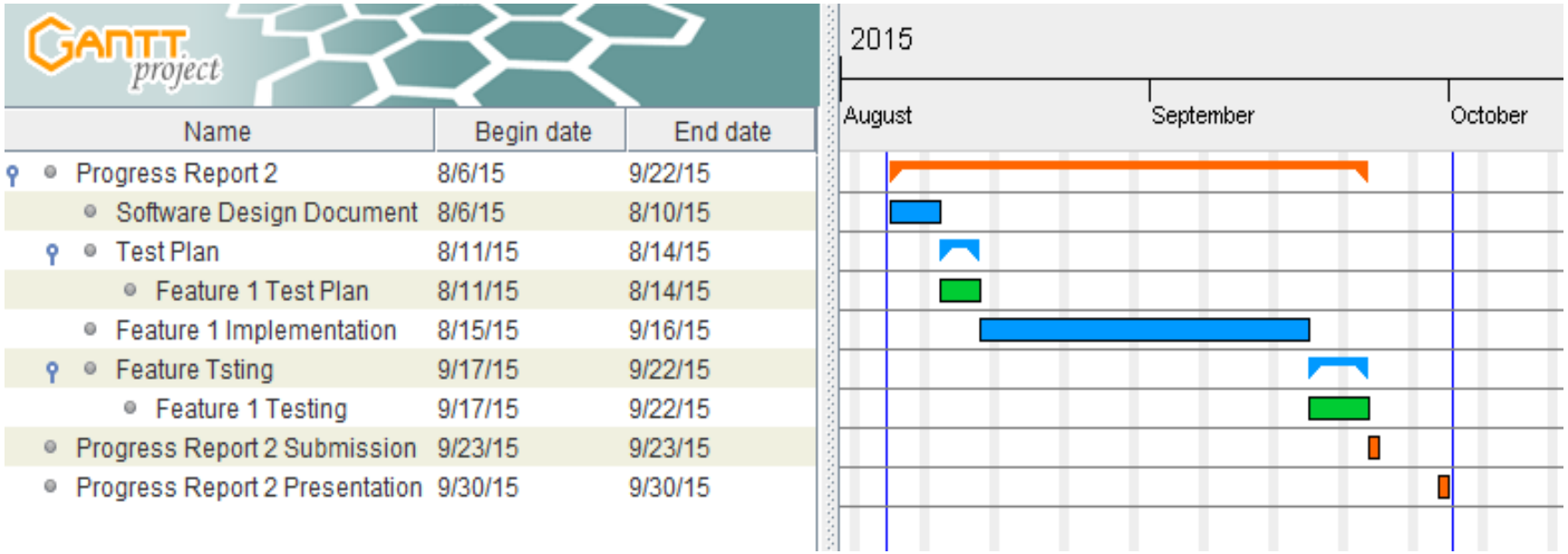

| Name | Begin date | End date |
|---|---|---|
| Topic Announcement | 5/20/15 | 5/30/15 |
| Proposal | 5/30/15 | 6/9/15 |
| Introduction and Background | 5/30/15 | 6/9/15 |
| Literature Review | 5/30/15 | 6/9/15 |
| Business review | 5/30/15 | 6/9/15 |
| Business tools and software review | 5/30/15 | 6/9/15 |
| Technology review | 5/30/15 | 6/9/15 |
| Development tools review8 | 5/30/15 | 6/9/15 |
| Quality standard | 6/4/15 | 6/9/15 |
| Project plan | 5/30/15 | 6/9/15 |
| Motivation | 6/4/15 | 6/9/15 |
| Aim | 6/4/15 | 6/9/15 |
| Objective | 6/4/15 | 6/9/15 |
| System architecture | 5/30/15 | 6/9/15 |
| Deliverables and limit | 5/30/15 | 6/9/15 |
| Future work | 5/30/15 | 6/9/15 |
| Schedule & Milestones | 6/4/15 | 6/9/15 |
| Reference | 6/4/15 | 6/9/15 |
| Proposal Submit | 6/10/15 | 6/10/15 |
| Proposal Presentation | 6/16/15 | 6/17/15 |

## Project Proposal Milestone

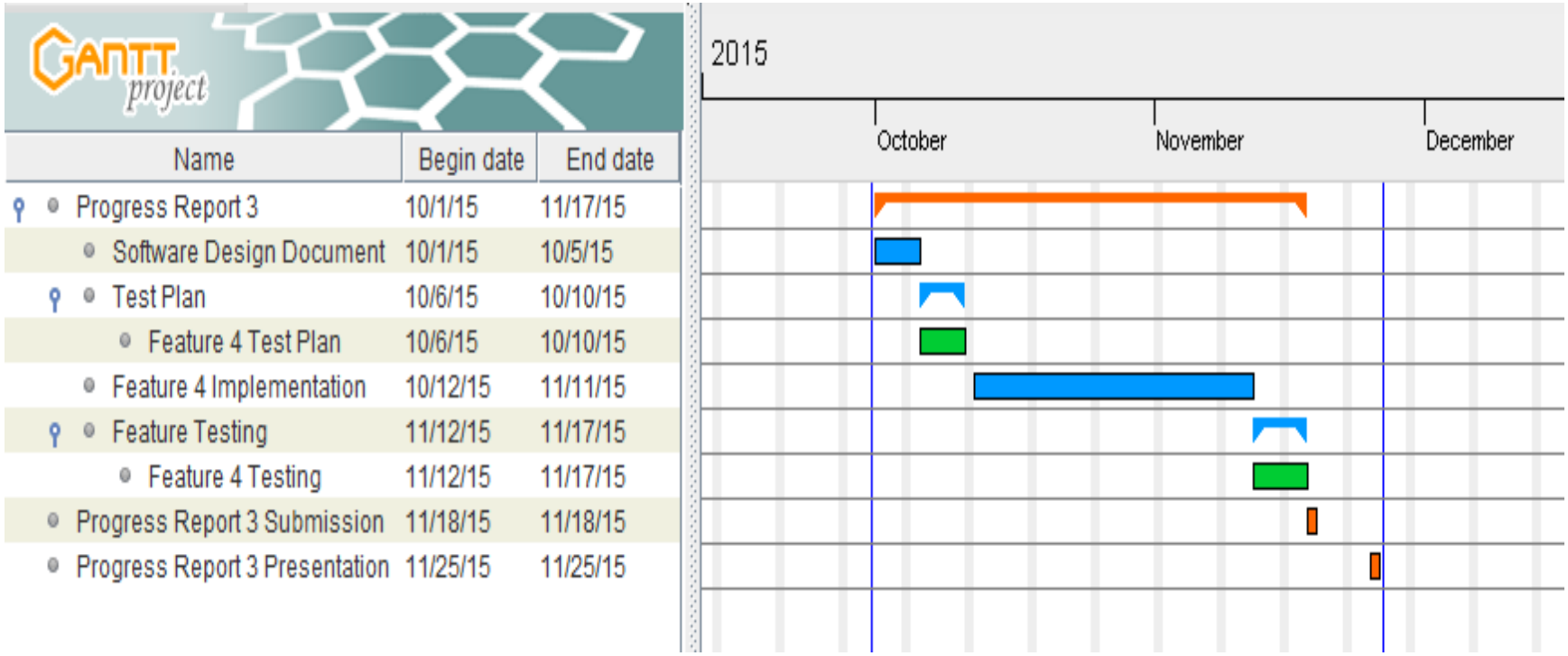



| Name | Begin date | End date |
|---|---|---|
| Progress Report 1 | 6/18/15 | 7/28/15 |
| Software Requirement Specification | 6/18/15 | 6/22/15 |
| Software Design Document | 6/23/15 | 6/27/15 |
| Test Plan | 6/27/15 | 6/30/15 |
| Feature 2 Test Plan | 6/27/15 | 6/30/15 |
| Feature 3 Test Plan | 6/27/15 | 6/30/15 |
| Feature 2 Implementation | 7/1/15 | 7/22/15 |
| Feature 3 Implementation | 7/1/15 | 7/22/15 |
| Feature Tesing | 7/23/15 | 7/28/15 |
| Feature 2 Testing | 7/23/15 | 7/28/15 |
| Feature 3 Testing | 7/23/15 | 7/28/15 |
| Progress Report 1 Summission | 7/29/15 | 7/29/15 |
| Progress Report 1 Presentation | 8/4/15 | 8/5/15 |

**Progress Report 1 Milestone**




**Progress Report 2 Milestone**




**Progress Report 3 Milestone**

# Chapter five | Reference


[1] Jonathan Fairfield, “A complete guide to Internet and social media usage in Thailand” [Online]. Available: http://tech.thaivisa.com/complete-insight-internet-social-media-usage-thailand/3147/ . [Accessed 07 June 2015].

[2] Nicolas, Michelle, Matt, “NewsTracker” [Online]. Available: http://newstracker.me/ . [Accessed 04 June 2015].

[3] Twazzup, “twazzup”, [Online]. Available: http://www.twazzup.com/about . [Accessed 04 June 2015].

[4] Wikipedia, “Spring Framework” [Online]. Available: http://en.wikipedia.org/wiki/Spring_Framework . [Accessed 04 June 2015].

[5] DOCSERVE, “JAVA SPRING HAND WRITTEN NOTES” [Online]. Available: https://docserve.wordpress.com/2013/08/12/java-spring-hand-written-notes/. [Accessed 08 June 2015].

[6] Shay Howe, “Responsive web design” [Online]. Available: http://learn.shayhowe.com/advanced-html-css/responsive-web-design/ . [Accessed 07 June 2015].

[7] Jonathan Hedley, “JSOUP” [Online] . Available: http://jsoup.org/ . [Accessed 07 June 2015].

[8] Wikipedia, “NetBeans” [Online]. Available: http://en.wikipedia.org/wiki/NetBeans . [Accessed 04 June 2015].

[9] Pantip, “Pantip” [Online]. Available: http://pantip.com. [Accessed 09 June 2015].

[10] Facebook, “Facebook” [Online]. Available: https://www.facebook.com. [Accessed 09 June 2015].

[11] Twitter, “Twitter” [Online]. Available: https://twitter.com . [Accessed 09 June 2015