2014

Project Plan

DOCUMENT CLASSIFICATION SYSTEM

BY ACUMEN

Contents

Background

Living in the world of big data, data collection seems never be a problem for people anymore. Whatever we search, there are always plenty of data on the internet. However, the existence of astonishing amount of data creates additional difficulties on pinpointing the information we desire among the sea of data. Aiming to settle this problem, our group decides to develop a program which takes articles in plain text as input, and makes accurate predictions about the classification of this article. By attaching classification information to articles, people will be able to locate their desired articles with less trouble.

Project Description

Firstly, as a student majored in Computer Science, our program focuses on classification of technology news. Moreover, to improve the correctness of clustering result, we will implement various data mining techniques during development. Not only can user classify article, they can also establish their own clustering dictionary and categories with the help of our program.

Goals and Objectives

Nowadays, huge amount of information come out everyday. People are busy and may not be able to get every piece of information, so we need a program that can make it easy for us to locate our desired posts, and can provide us with concise summarisations for news posts. This may improve the efficiency of getting information in this information world.

- To make it easy for users to locate their desired post
- To provide users concise summarizations for news posts
- To practice software development methodology

Scope

Scope Definition

The project will introduce the following technology

- Key words dictionary generation
- Information Gain Calculation
- Automatic sorting

Stakeholders

Stakeholder Register				
Name	Project Role	Internal/External	Position	Contact Information
Mingyang WANG	Project Manager	Internal	Manager	mingywang3-c@my.cityu.edu.hk
Jie PU	Assistant Project Manager	Internal	Manager	jiepu2-c@my.cityu.edu.hk
Zhou FANG	Configuration Manager	Internal	Manager	zfang6-c@my.cityu.edu.hk
Xikang FENG	Team Member	Internal	Developer	xikanfeng2@gmail.com
Yiji WANG	Team Member	Internal	Developer	yijiwang3-c@my.cityu.edu.hk
Renjie ZHU	Team Member	Internal	Developer	renjiezhu2-c@my.cityu.edu.hk
Dr. W.K.Chan Ricky	Project Superviser	External	President	wkchan@cityu.edu.hk

Methodology

Software Development Methodology

In this project, we chose test-driven development as our development methodology. It is the method that divide the project into small test cases, which drive the software to pass all the test cases. This method can lead to simple designs in this project. Object-oriented programming paradigm is applied in this project.

Development tools

We use several development tools in this project. The main development software is Eclipse, because we use Java as the language of our software.

Project Team Organization

In terms of project management, the team organization for our group is the following.

Team Member	Student ID	Nickname	Position	Responsibility
WANG Mingyang	52640166	Young	Project Manager	Plan, coordinate, motivate and oversee the project
PU Jie	52640234	Paul	Assistant PM	Make decisions whenever PM is offline
FANG Zhou	52639689	Alex	Configuration Manager	Configuration Methodology, version control
FENG Xikang	52639689	Eric		
WANG Yiji	52639040	Jason		
ZHU Renjie	52639014	Leo		

Among each cycle, the members of our team rotate to take charge of certain functionality requirements in our project. And managers also take extra responsibilities for supervising and coordinating the progresses.

Cycle 1

Team Member	Student ID	Nickname	Responsibility
WANG Mingyang	52640166	Young	- algorithm design - coding
PU Jie	52640234	Paul	- documentation preparation
FANG Zhou	52639689	Alex	- algorithm design - coding
FENG Xikang	52639689	Eric	- algorithm design - coding
WANG Yiji	52639040	Jason	- testing/debugging
ZHU Renjie	52639014	Leo	- documentation prepration

Cycle 2

Team Member	Student ID	Nickname	Responsibility
WANG Mingyang	52640166	Young	- documentation - testing/debugging
PU Jie	52640234	Paul	- coding

FANG Zhou	52639689	Alex	- documentation - algorithm improvement
FENG Xikang	52639689	Eric	- documentation - algorithm improvement
WANG Yiji	52639040	Jason	- coding
ZHU Renjie	52639014	Leo	- coding

Cycle 3

Team Member	Student ID	Nickname	Responsibility
WANG Mingyang	52640166	Young	- documentation - testing/debugging
PU Jie	52640234	Paul	- documentation
FANG Zhou	52639689	Alex	- coding
FENG Xikang	52639689	Eric	- testing/debugging
WANG Yiji	52639040	Jason	- documentation
ZHU Renjie	52639014	Leo	- documentation

Configuration Management

Configuration management monitors any change to the project including program codes, documentation, data and development tools. The management is responsible by the configuration manager, who ensures that new functionalities or documentation are merged and archived, conflicts between pull requests are handled, and every member gets the most updated version of the project's products.

Any changes to the code or documentation should be reported to the configuration manager through Github Pull Request with detailed comments about modifications or improvements. Every member could fork the whole project from the group version, make modification locally, and push the new version to the group version until certain functionalities are implemented. Any changes out of the project scope should be reported to the configuration manager.

Items	**Release 1**	**Release 2**	**Release 3**
Project Plan	Ver. 1	Ver. 2	Ver. 3
OOA/D	Ver. 1	Ver. 2	Ver. 3
Configuration Manual	-	Ver. 1	Ver. 2
Test Plan	Ver. 1	Ver. 2	Ver. 2
User Manual	-	-	Ver. 1
Bug Report	Ver. 1	Ver. 2	Ver. 3
Program	Release 1	Release 2	Release 3
Source Code	Ver. 1	Ver. 3	Ver. 5
JavaDoc	Ver.1	Ver. 3	Ver. 5
Test Code	Ver. 1	Ver. 2	Ver. 2.1
Java SDK	SE 6	SE 6	SE 6

Platform	Win 7/8, Linux	Win 7/8, Linux	Win 7/8, Linux
Development Environment	Win 7/8, Linux	Win 7/8, Linux	Win 7/8, Linux
JUnit Framework	Ver. 4.0	Ver. 4.0	Ver 4.0
Eclipse IDE	Ver. 3.6.1	Ver. 3.6.1	Ver. 3.6.1
Version Control Tool	Git/Github	Git/Github	Git/Github

Milestones

Project charter signed by Sep. 12th, 2014

Project plan approved by Sep. 19th, 2014

1st Iteration by Oct. 16th, 2014

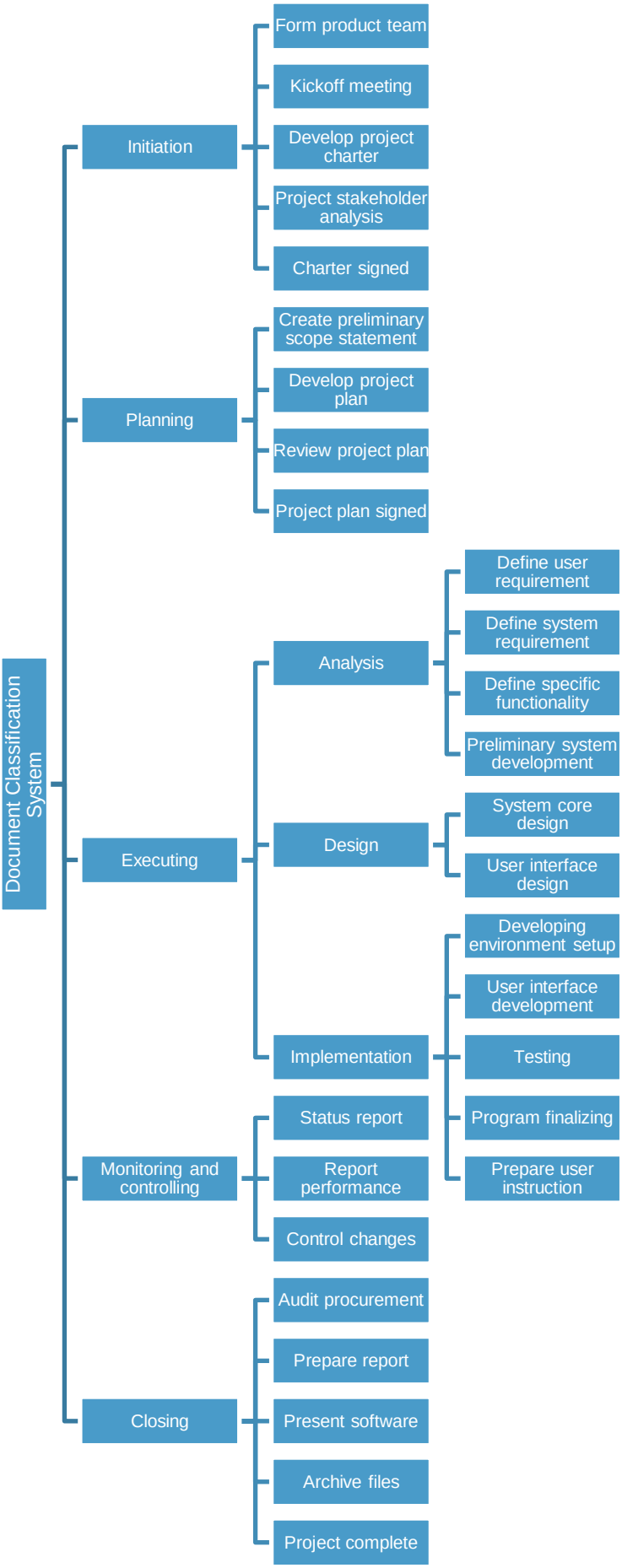
2nd Iteration by Nov. 12th, 2014

3rd Iteration by Dec. 3rd, 2014

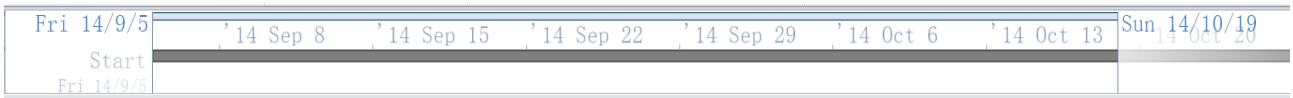
Project Completed by Dec. 4th, 2014

Work Breakdown Structure

1. Project
    1.1. Initiation
        1.1.1. Form product team
        1.1.2. Kickoff meeting
        1.1.3. Develop project charter
        1.1.4. Project stakeholder analysis
        1.1.5. Charter signed
    1.2. Planning
        1.1.1. Create preliminary scope statement
        1.1.2. Develop project plan
        1.1.3. Review project plan
        1.1.4. Project plan signed
    1.3. Executing
        1.1.1. Analysis
            1.1.1.1. Define user requirement
            1.1.1.2. Define system requirement
            1.1.1.3. Define specific functionality
            1.1.1.4. Preliminary system development
        1.1.2. Design
            1.1.2.1. System core design
            1.1.2.2. User interface design
        1.1.3. Implementation
            1.1.3.1. Developing environment setup
            1.1.3.2. User interface development
            1.1.3.3. Testing
            1.1.3.4. Program finalizing
            1.1.3.5. Prepare user instruction
    1.4. Monitoring and controlling
        1.1.1. Status report
        1.1.2. Report performance
        1.1.3. Control changes
    1.5. Closing
        1.1.1. Audit procurement
        1.1.2. Prepare report
        1.1.3. Present software
        1.1.4. Archive files
        1.1.5. Project complete

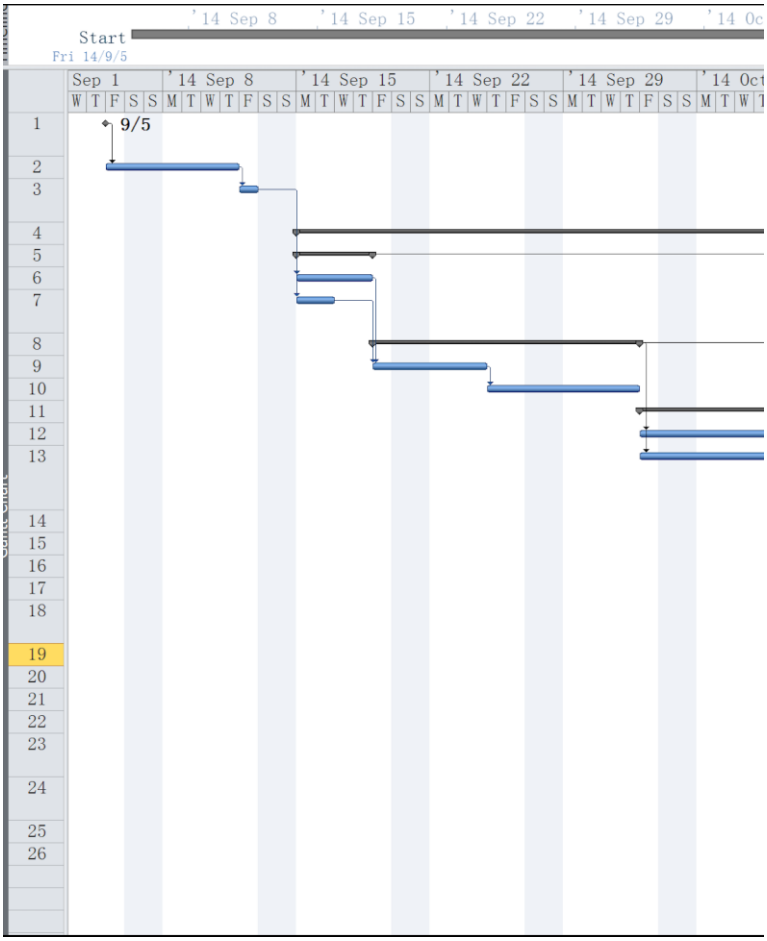

Document Classification System
Initiation
Form product team
Kickoff meeting
Develop project charter
Project stakeholder analysis
Charter signed
Planning
Create preliminary scope statement
Develop project plan
Review project plan
Project plan signed
Executing
Analysis
Define user requirement
Define system requirement
Define specific functionality
Preliminary system development
Design
System core design
User interface design
Implementation
Developing environment setup
User interface development
Testing
Program finalizing
Prepare user instruction
Monitoring and controlling
Status report
Report performance
Control changes
Closing
Audit procurement
Prepare report
Present software
Archive files
Project complete

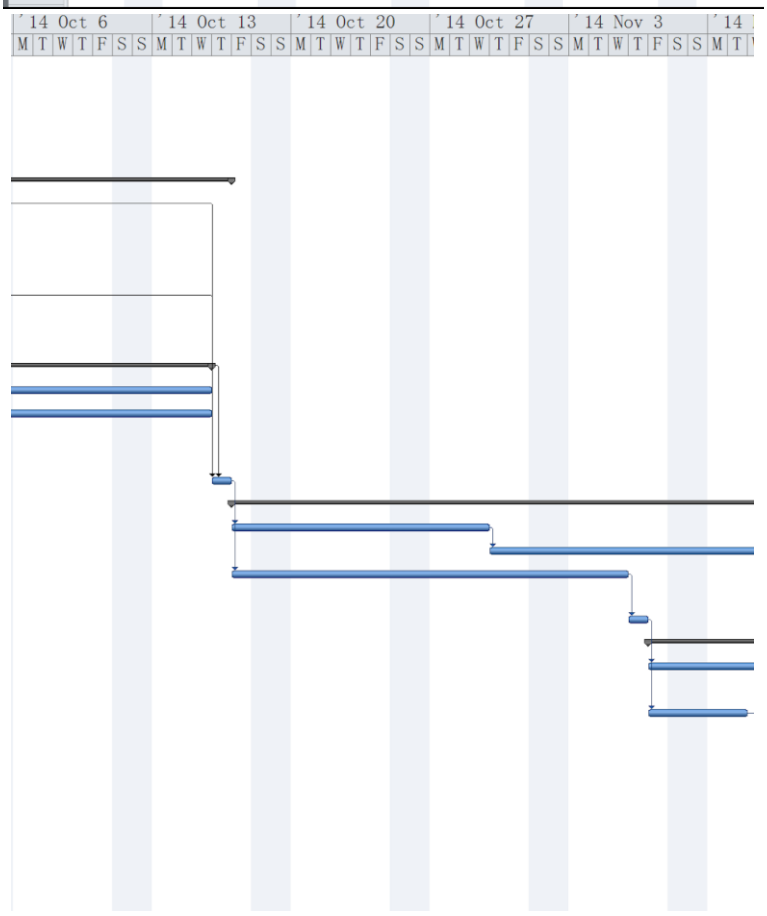
Project Schedule

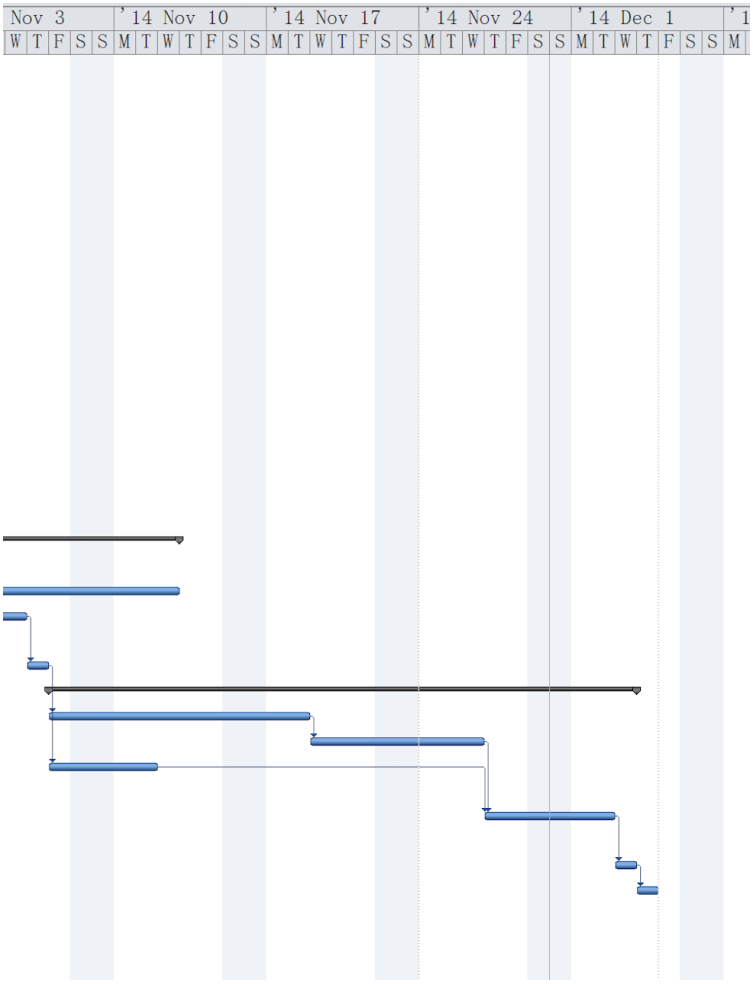


	Task Mode	Task Name	Duration	Start	Finish	Predecessors	Resou
1		Project Kickoff Meeting	0 days	Fri 14/9/5	Fri 14/9/5		
2		Requirement Gathering	5 days	Fri 14/9/5	Thu 14/9/11	1	
3		Project Charter Signoff	1 day	Fri 14/9/12	Fri 14/9/12	2	
4		**Project Iteration 1**	**24 days**	**Mon 14/9/15**	**Thu 14/10/16**		
5		**Project Plan**	**4 days**	**Mon 14/9/15**	**Thu 14/9/18**		
6		WBS Setup	4 days	Mon 14/9/15	Thu 14/9/18	3	
7		Gantt Chart Preparation	2 days	Mon 14/9/15	Tue 14/9/16	3	
8		**Design**	**10 days**	**Fri 14/9/19**	**Thu 14/10/2**		
9		System Design	4 days	Fri 14/9/19	Wed 14/9/24	6,7	
10		Detailed Design	6 days	Thu 14/9/25	Thu 14/10/2	9	
11		**Implementation**	**9 days**	**Fri 14/10/3**	**Wed 14/10/15**		
12		Implement Dictionar	9 days	Fri 14/10/3	Wed 14/10/15	8	
13		Implement Information Gain Calculator	9 days	Fri 14/10/3	Wed 14/10/15	8	
14		Release 1	1 day	Thu 14/10/16	Thu 14/10/16	5,8,11	
15		**Project Iteration 2**	**19 days**	**Fri 14/10/17**	**Wed 14/11/12**		
16		Design Refractoring	9 days	Fri 14/10/17	Wed 14/10/29	14	
17		Code Refractoring	10 days	Thu 14/10/30	Wed 14/11/12	16	
18		Preparing documentation	14 days	Fri 14/10/17	Wed 14/11/5	14	
19		Release 2	1 day	Thu 14/11/6	Thu 14/11/6	18	
20		**Project Iteration 3**	**19 days**	**Fri 14/11/7**	**Wed 14/12/3**		
21		Design Refractoring	8 days	Fri 14/11/7	Tue 14/11/18	19	
22		Code Refractoring	6 days	Wed 14/11/19	Wed 14/11/26	21	
23		Preparing User Manual	3 days	Fri 14/11/7	Tue 14/11/11	19	
24		Documentation Finalizing	4 days	Thu 14/11/27	Tue 14/12/2	22,23	
25		Release 3	1 day	Wed 14/12/3	Wed 14/12/3	24	
26		Project Finish	1 day	Thu 14/12/4	Thu 14/12/4	25	

Gantt Chart


Start
Fri 14/9/5
Sep 1
'14 Sep 8
'14 Sep 15
'14 Sep 22
'14 Sep 29
9/5


'14 Oct 6
'14 Oct 13
'14 Oct 20
'14 Oct 27
'14 Nov 3


Nov 3
'14 Nov 10
'14 Nov 17
'14 Nov 24
'14 Dec 1
W T F S S M T W T F S S M T W T F S S M T W T F S S M T W T F S S M