University of Surrey
Department of Computing

Faculty of Engineering & Physical Sciences

COM3001 Professional Project

Project Idea: Pinboard website

Student: Chara Katiri | 6166668 | ck00113@surrey.ac.uk

Supervisor: Dr. Steve Wesemeyer

Examiner: Dr. Anna Lisa Ferrara

Academic Year 2014-15

Overview

- Present situation (the problem)
- Project Idea
- Project objective & goal statement
- Project Benefits, ROI Statement
- Project Stakeholders
- Background knowledge
- Technical investigation
- Planning
- References
- Discussion/Answering questions

Present situation (the problem)

Nowadays, electronic mail is the predominant digital communication platform.

85% of human population is connected online and communicate through email despite the number of social channels available. (Reuter's survey [1])

100 billions

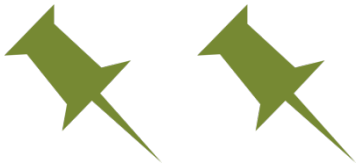
emails send and received daily [2]

97.4 billions

are **spam** emails [3]

The question is: What leads to the creation of so many spam emails **within the university environment?**

Project Idea



The creation of a **Pinboard** solution

(in the form of web application)

as a method to

minimize the number of spam emails send and received daily

within the University of Surrey .

Project objective & goal statement

The **objective** of this dynamic Web application is to allow users to interact with the Pinboard database and **sell/buy second hand items.**

The solution has the potential to significantly reduce the number of emails send between students regarding **second-hand books and room swaps.**

Additional functionalities can enrich the search user experience by offering a number of browsing categories like unwanted tickets for events, bicycles, electronics etc.

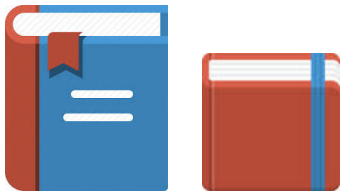
Project Benefits, ROI Statement

Introducing a **new communication channel** that will allow the students to:

- sell second-hand books (and other equipment),
- search for housemates/swap rooms on campus.

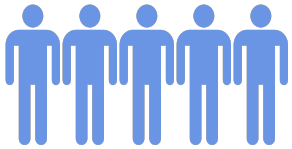
This will help:

- reduce network traffic,
- minimize the creation of spam/junk emails regarding second-hand books, room swaps etc.
- replace of the physical Pin-boards located around the campus, reduce paper copies and promote environmental friendly behavior.



The Pinboard solution will enhance the quality of student experience with the use of multimedia and make students requests/hunt for houses and books easier.

Project Stakeholders



- University of Surrey students:
  - The main users of the web application:
  - **Students** interested to **buy** products
  - **Students** interested to **sell** products

- University of Surrey IT and Service Desk:
  - The support team.

- Chara Katiri:
  - Computer Science Student, the developer.

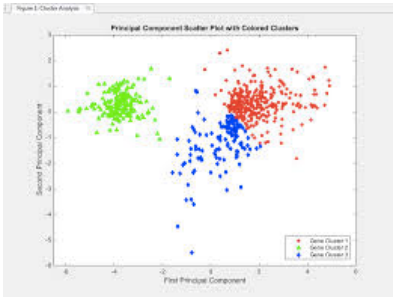
Background Knowledge

Alternative solutions used by others to solve the problem:

- Recognizing Spam domains by extracting features from Spam emails using Data Mining.[4]
  - Discovering patterns in data sets.

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

- Identifying spam e-mail messages using an intelligent algorithm. [5]
  - Use of Bayes rule and keyword patterns.



- Text and image based spam email classification using KNN, Naïve Bayes and Reverse DBSCAN algorithm. [6]
  - KNN, Naïve Bayes and Reverse DBSCAN algorithm.

Technical Investigation

- Application type:
  - Dynamic Web application.
  - Scalable for use on a variety of devices and screen sizes.

- Availability:
  - Hosted on the University of Surrey intranet.

- Security:
  - Identification: the user claims to be a university student.
  - Authentication: the user provides his username and password details to declare that his identity is valid.
  - Authorization: the system interacts with the AD group to authorize and grand access to the user.

Technical Investigation cont.

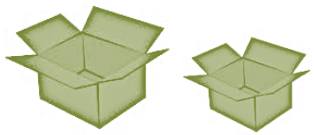
- **Technologies**:
  - Client and Server side development using Java and Spring MVC framework.

- **Architecture:**
  - The architecture of the developed application will be separated in several layers to support the needs of the application:
    - Presentation Layer: JSP views and presentation of data (HTML5, CSS).
    - Security layer: authentication and authorization (JavaScript).
    - Business logic: services.
    - Database layer: Apache2, Tomcat7 and a MySQL.

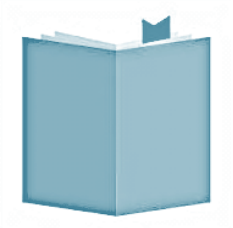
Technical Investigation cont.

Key issues for the project:

- **Product Basket**
  - Keep the user's basket secure and ensure that is not visible to other users.



- **Update the products list:**
  - Make the new items visible:
    - update the list of products to include the last uploaded items.
  - Expire the sold items:
    - update the list of products to exclude the last purchased/sold items.
- **The nature of the products:**
  - Each product is unique (quantity 1).

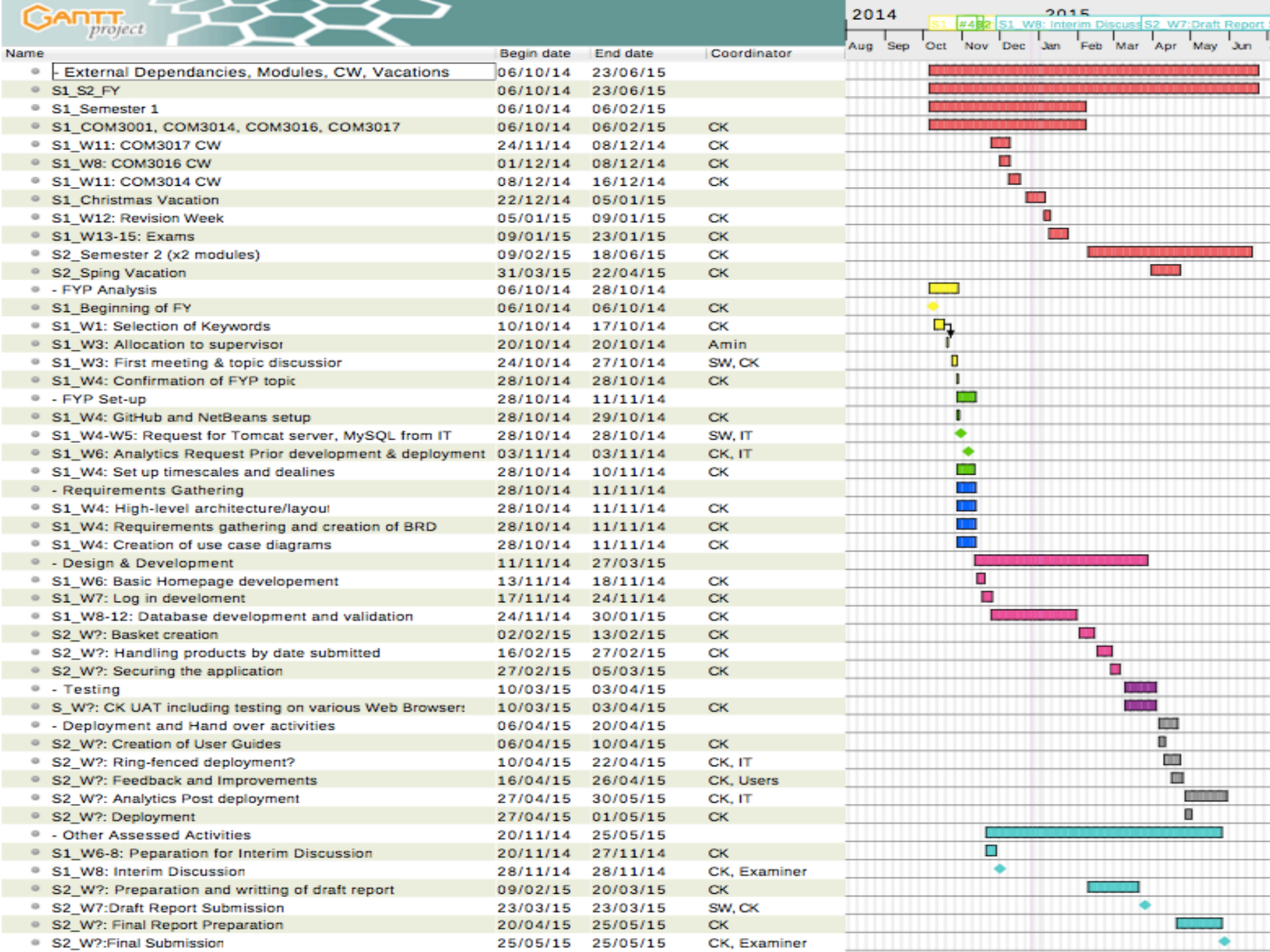


- **Use of Spring MVC framework:**
  - Is a learning curve.

- Host the website on a web server. In this case, if the solution is not accepted for deployment it can still be tested by volunteers.

Planning

- Major steps divided in achievable chunks, progress so far following the **Agile Development** lifecycle.
- Dependencies between tasks.
  - Planning
  - Requirements Analysis
  - Design
  - Development
  - Testing
  - Implementation
  - Maintenance
- Evaluation of goals.



Name	Begin date	End date	Coordinator
- External Dependancies, Modules, CW, Vacations	06/10/14	23/06/15	
S1_S2_FY	06/10/14	23/06/15	
S1_Semester 1	06/10/14	06/02/15	
S1_COM3001, COM3014, COM3016, COM3017	06/10/14	06/02/15	CK
S1_W11: COM3017 CW	24/11/14	08/12/14	CK
S1_W8: COM3016 CW	01/12/14	08/12/14	CK
S1_W11: COM3014 CW	08/12/14	16/12/14	CK
S1_Christmas Vacation	22/12/14	05/01/15	
S1_W12: Revision Week	05/01/15	09/01/15	CK
S1_W13-15: Exams	09/01/15	23/01/15	CK
S2_Semester 2 (x2 modules)	09/02/15	18/06/15	CK
S2_Sping Vacation	31/03/15	22/04/15	CK
- FYP Analysis	06/10/14	28/10/14	
S1_Beginning of FY	06/10/14	06/10/14	CK
S1_W1: Selection of Keywords	10/10/14	17/10/14	CK
S1_W3: Allocation to supervisor	20/10/14	20/10/14	Amin
S1_W3: First meeting & topic discussior	24/10/14	27/10/14	SW, CK
S1_W4: Confirmation of FYP topic	28/10/14	28/10/14	CK
- FYP Set-up	28/10/14	11/11/14	
S1_W4: GitHub and NetBeans setup	28/10/14	29/10/14	CK
S1_W4-W5: Request for Tomcat server, MySQL from IT	28/10/14	28/10/14	SW, IT
S1_W6: Analytics Request Prior development & deployment	03/11/14	03/11/14	CK, IT
S1_W4: Set up timescales and dealines	28/10/14	10/11/14	CK
- Requirements Gathering	28/10/14	11/11/14	
S1_W4: High-level architecture/layout	28/10/14	11/11/14	CK
S1_W4: Requirements gathering and creation of BRD	28/10/14	11/11/14	CK
S1_W4: Creation of use case diagrams	28/10/14	11/11/14	CK
- Design & Development	11/11/14	27/03/15	
S1_W6: Basic Homepage developement	13/11/14	18/11/14	CK
S1_W7: Log in develoment	17/11/14	24/11/14	CK
S1_W8-12: Database development and validation	24/11/14	30/01/15	CK
S2_W?: Basket creation	02/02/15	13/02/15	CK
S2_W?: Handling products by date submitted	16/02/15	27/02/15	CK
S2_W?: Securing the application	27/02/15	05/03/15	CK
- Testing	10/03/15	03/04/15	
S_W?: CK UAT including testing on various Web Browser:	10/03/15	03/04/15	CK
- Deployment and Hand over activities	06/04/15	20/04/15	
S2_W?: Creation of User Guides	06/04/15	10/04/15	CK
S2_W?: Ring-fenced deployment?	10/04/15	22/04/15	CK, IT
S2_W?: Feedback and Improvements	16/04/15	26/04/15	CK, Users
S2_W?: Analytics Post deployment	27/04/15	30/05/15	CK, IT
S2_W?: Deployment	27/04/15	01/05/15	CK
- Other Assessed Activities	20/11/14	25/05/15	
S1_W6-8: Peparation for Interim Discussion	20/11/14	27/11/14	CK
S1_W8: Interim Discussion	28/11/14	28/11/14	CK, Examiner
S2_W?: Preparation and writting of draft report	09/02/15	20/03/15	CK
S2_W7:Draft Report Submission	23/03/15	23/03/15	SW, CK
S2_W?: Final Report Preparation	20/04/15	25/05/15	CK
S2_W?:Final Submission	25/05/15	25/05/15	CK, Examiner

Planning cont.

Potential risks:

- Time constraints due to 5 other modules that run in parallel with COM3001. All the modules are assessed based on at least one Coursework and an Exam.
- Timescales for set up, development and deployment are aggressive to meet the demands of FYP deadlines and deliveries.
  - **Contingency plan in place**:
    - Timescales were set with the work required by other modules in mind.
    - If for any reason timescales shift then requirements categorized as 'Could' will not be implemented.
- Is possible for the Systems Team not to accept the proposed Pinboad solution. If the solution can't be deployed to the intranet, no data can be collected neither **analytics** can be created to measure the effectiveness of the proposed solution to minimize spam emails.
- UAT testing depends on testers being available when required (CS students, **volunteers**).
- End user expectations unrealistic compared to solution.

References

- [1] Reuters, (2014). *Most of world interconnected through email and social media*. [Online] Available from: http://www.reuters.com/article/2012/03/27/uk-socialmedia-online-poll-idUSLNE82Q02120120327 [Accessed 03/11/2014].
- [2] The Radicati Group Inc., (2013). *Email Statistics Report, 2013-2017*. [Online] Available from: http://www.radicati.com/wp/wp-content/uploads/2013/04/Email-Statistics-Report-2013-2017-Executive-Summary.pdf [Accessed: 03/11/2014]
- [3] Esecurityplanet.com, (2014). *Almost 100 Billion Spam E-mails Sent Daily in Q1 2013 - eSecurity Planet*. [Online] Available at: http://www.esecurityplanet.com/network-security/almost-100-billion-spam-e-mails-sent-daily-in-q1-2013.html [Accessed 3 Nov. 2014].
- [4] Patel, K. (2014). *Recognizing Spam Domains by Extracting Features from Spam Emails using Data Mining*. International Journal of Computer Applications (0975 – 8887), 90(8) [Accessed 23 Nov. 2014].
- [5] Ghaedi, P & Harounabadi, A. (2014). *Identifying spam e-mail messages using an intelligence algorithm*. Decision Science Letters , 3(3), 439-444. [Accessed 23 Nov. 2014].
- [6] Harisinghaney, A., Dixit, A., Gupta, S. and Arora, A. (2014). *Text and image based spam email classification using KNN, Naïve Bayes and Reverse DBSCAN algorithm.* In: Optimization, Reliabilty, and Information Technology (ICROIT), 2014 International Conference on. IEEE, pp.153 - 155. [Accessed 23 Nov. 2014].

Closing

- Questions?
- Remarks
- Comments

Possible Questions

- **Note for CK:** this slide will be removed prior the interim discussion presentation.
- What is the need for this solution?
- Is the structure/logic of the solution appropriate?
- Will the solution be implemented as intended (planning)?
- Will the solution be technically efficient?
- What will the outcome of the solution be? (effectiveness)?
- How will the solution achieve its intended objectives (effectiveness)?
- Will the solution be cost-effective and cost beneficial?