Visualising plants and metadata

Final Report for CS39440 Major Project

Author: Siôn Griffiths (sig2@aber.ac.uk)
Supervisor: Dr. Hannah Dee (hmd1@aber.ac.uk)

4th March 2016
Version: 1.0 (Draft)

This report was submitted as partial fulfilment of a BEng degree in
Software Engineering (G600)

Department of Computer Science
Aberystwyth University
Aberystwyth
Ceredigion
SY23 3DB
Wales, UK

Declaration of originality

In signing below, I confirm that:

- This submission is my own work, except where clearly indicated.
- I understand that there are severe penalties for Unacceptable Academic Practice, which can lead to loss of marks or even the withholding of a degree.
- I have read the regulations on Unacceptable Academic Practice from the University's Academic Quality and Records Office (AQRO) and the relevant sections of the current Student Handbook of the Department of Computer Science.
- In submitting this work I understand and agree to abide by the University's regulations governing these issues.

Name .....

Date .....

Consent to share this work

In signing below, I hereby agree to this dissertation being made available to other students and academic staff of the Aberystwyth Computer Science Department.

Name .....

Date .....

Acknowledgements

I am grateful to...

I’d like to thank...

Abstract

Visualising plants and metadata is a project delivering a web-based system which enables the convenient exploration of plant images and associated metadata captured as part of experiments carried out at the National Plant Phenomics Centre(NPPC).

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Background & Objectives

This section should discuss your preparation for the project, including background reading, your analysis of the problem and the process or method you have followed to help structure your work. It is likely that you will reuse part of your outline project specification, but at this point in the project you should have more to talk about.

Note:

- All of the sections and text in this example are for illustration purposes. The main Chapters are a good starting point, but the content and actual sections that you include are likely to be different.
- Look at the document on the Structure of the Final Report for additional guidance.

1.1 Background

What was your background preparation for the project? What similar systems did you assess? What was your motivation and interest in this project?

1.2 Analysis

Taking into account the problem and what you learned from the background work, what was your analysis of the problem? How did your analysis help to decompose the problem into the main tasks that you would undertake? Were there alternative approaches? Why did you choose one approach compared to the alternatives?

There should be a clear statement of the objectives of the work, which you will evaluate at the end of the work.

In most cases, the agreed objectives or requirements will be the result of a compromise between what would ideally have been produced and what was felt to be possible in the time available. A discussion of the process of arriving at the final list is usually appropriate.

1.3 Process

Plan driven approaches traditionally associated with software development projects usually expect that all system requirements are understood and collected prior to any further work on design or implementation. A number of factors made such an approach unsuitable for this project, chiefly a lack of domain knowledge made up-front requirement gathering difficult and the requirements themselves were likely to be poorly defined and subject to change. With these considerations in mind it was decided that an agile approach would be best.

A SCRUM-inspired approach was adopted for the project methodology, featuring time-boxed iterations in the form of sprints with regular releases of the software. Work would be tracked in the form of user-stories, the planning and organisation of work would revolve around a defined functionality goal for each sprint and release.

Chapter 2

Design

You should concentrate on the more important aspects of the design. It is essential that an overview is presented before going into detail. As well as describing the design adopted it must also explain what other designs were considered and why they were rejected.

The design should describe what you expected to do, and might also explain areas that you had to revise after some investigation.

Typically, for an object-oriented design, the discussion will focus on the choice of objects and classes and the allocation of methods to classes. The use made of reusable components should be described and their source referenced. Particularly important decisions concerning data structures usually affect the architecture of a system and so should be described here.

How much material you include on detailed design and implementation will depend very much on the nature of the project. It should not be padded out. Think about the significant aspects of your system. For example, describe the design of the user interface if it is a critical aspect of your system, or provide detail about methods and data structures that are not trivial. Do not spend time on long lists of trivial items and repetitive descriptions. If in doubt about what is appropriate, speak to your supervisor.

You should also identify any support tools that you used. You should discuss your choice of implementation tools - programming language, compilers, database management system, program development environment, etc.

Some example sub-sections may be as follows, but the specific sections are for you to define.

2.1 Overall Architecture

MVC - for ease of testing, scalability, separation of concerns, maintainability through familiarity(people know mvc and what to expect), maturity of supporting technologies

3-tier based approach to data layer / service / presentation stuff that may not entirely fit with the mvc pattern

2.2 Framework and Programming Language

The sheer range of MVC frameworks available to developers is incredible and the decision of which to use is potentially difficult. It was not within scope to review all the available choices

2.3 Tools and third-party services

2.3.1 Intellij

2.3.2 Git and Github

2.3.3 Jira

2.3.4 Codeship

2.3.5 Maven

2.4 Some detailed design

2.4.1 Even more detail

2.5 User Interface

2.6 Other relevant sections

Chapter 3

Implementation

The implementation should look at any issues you encountered as you tried to implement your design. During the work, you might have found that elements of your design were unnecessary or overly complex; perhaps third party libraries were available that simplified some of the functions that you intended to implement. If things were easier in some areas, then how did you adapt your project to take account of your findings?

It is more likely that things were more complex than you first thought. In particular, were there any problems or difficulties that you found during implementation that you had to address? Did such problems simply delay you or were they more significant?

You can conclude this section by reviewing the end of the implementation stage against the planned requirements.

3.1 Stuff

Model plant domain DB building Show plants in page, Data reading and routing via annotated csv Ajax submission of forms -¿ Graphing

Chapter 4

Testing

4.1 Overall Approach to Testing

4.2 Automated Testing

4.2.1 Unit Tests

4.2.2 User Interface Testing

4.2.3 Stress and Performance Testing

Performance and stress testing was carried out through the use of Apache Jmeter [1], an open source Java application built to measure site and application performance under controlled loads. Jmeter enables the simulation of a number of concurrent users accessing a given site, these simulated agents follow a defined sequence of actions as specified in the test script.

For the purposes of this project Jmeter was used to assess whether pages in the site would load within defined time limits and whether implementation decisions have an effect performance. In general the goal was to have pages served 300ms with a hard limit of 1000ms or one second, although this does not include image load times. A target of 300ms is well under the 1 second limit for keeping a users flow of though as defined by Nielsen [7]. Running the tests regularly could also help highlight issues that may not be uncovered under other forms of testing such as intermittent problems that could result in request errors that would be difficult to reproduce otherwise. Unless otherwise stated the tests run with ten concurrent simulated users and the tests are repeated thirty times in order to smooth out any outliers in the data and try to pick up intermittent issues if present.

General results as output by Jmeter are included in figure 4.1 for an experiment which has been initialised with data. The initialisation is an important distinction because the amount of experiment data significantly affects the initial page response time for the Graphs page, other pages are affected somewhat but to a much lesser degree. Figure 4.2 displays the results of running the same test without the data having been added to the experiment and it's clear to see the effect on the load time for the Graphs page.

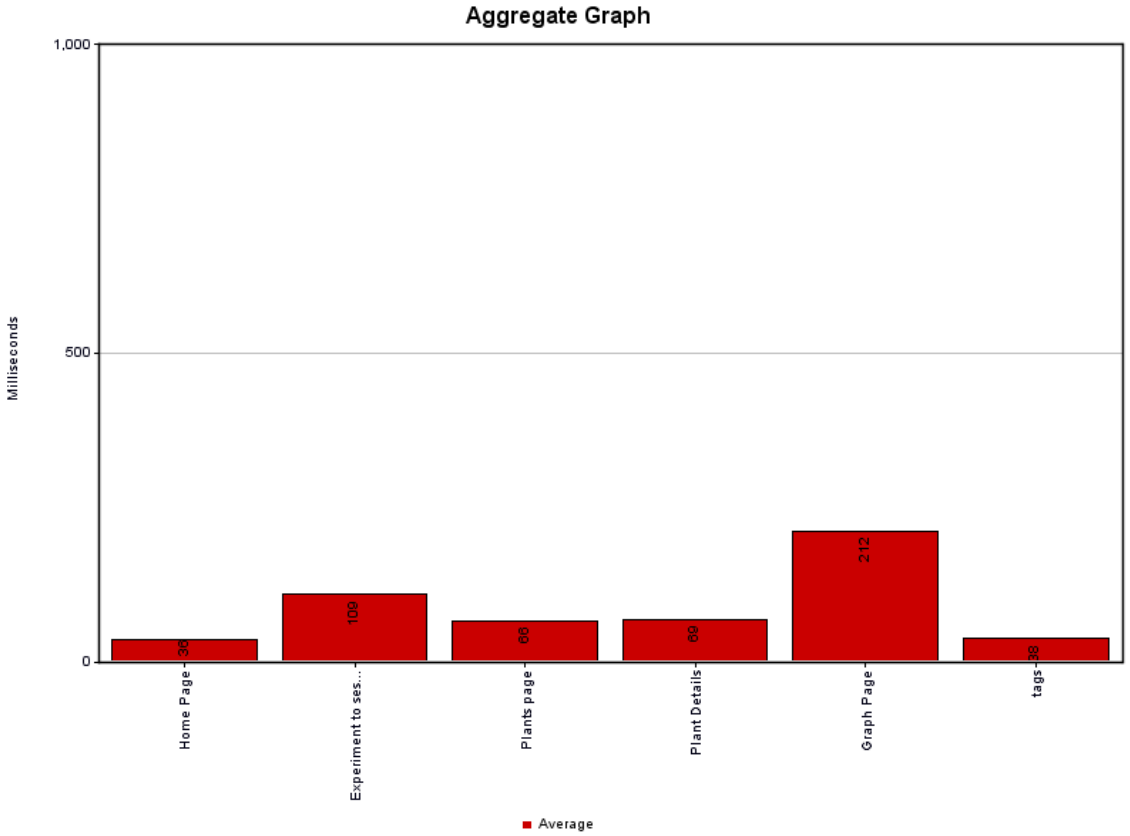

Figure 4.1: Visulisation of Jmeter test result of a fully initialised experiment

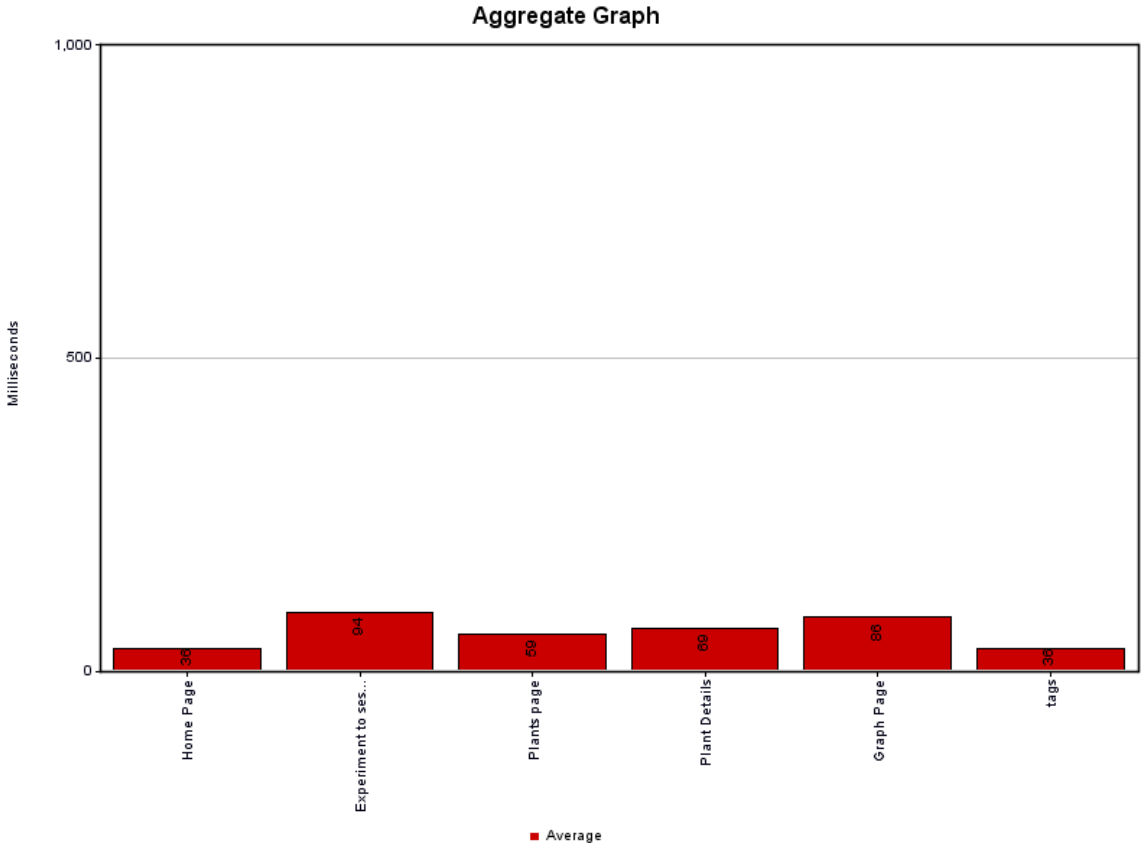

Figure 4.2: Visulisation of Jmeter test result of a partially initialised experiment

One particular area of the system which benefited from this form of testing was the choice of default pagination options on the plants and plant details pages.

4.3 Manual Testing

For areas of the system where automated testing was impractical or insufficient to verify results, a manual approach was taken and test tables used to verify functionality is as expected. Much of the functionality on the Admin page relies on an active network connection to the NPPC data repository

4.3.1 Admin Page Test Table

Test	Input	Expected Output	Pass
Attempt to access admin area without login	Go to /admin without login	Redirected to administrator login page	✓
Attempt to access admin area with correct login	Go to /admin with login	Admin is page is displayed	✓
Attempt admin login with incorrect credentials	Submit admin login form with incorrect credentials	Error displayed to user.	✓
Admin log out	Click logout button from admin page	Redirect to home page and authorisation cleared from session	✓
Initialise Experiment	Click initialise button for uninitialised experiment	Experiment begins initialising - plants are created	✓
Update experiment	Click Update button on initialised experiment	Experiment begins update, plants are updated or created	✓
Import data with valid csv	Click Init Data button on initialised experiment	Data is imported from csv	✓
Import data with invalid csv	Click Init Data button on initialised experiment	Invalid csv data is ignored	✓
Delete data	Click delete data on experiment	Data is deleted from the experiment	✓
Delete plants	Click delete plants button on experiment	Plant data and images are deleted	✓

Table 4.1: Test Table for Admin page functionality

4.3.2 Graph Page Test Table

4.4 Integration Testing

4.5 User Testing

Chapter 5

Evaluation

Examiners expect to find in your dissertation a section addressing such questions as:

- Were the requirements correctly identified?
- Were the design decisions correct?
- Could a more suitable set of tools have been chosen?
- How well did the software meet the needs of those who were expecting to use it?
- How well were any other project aims achieved?
- If you were starting again, what would you do differently?

Such material is regarded as an important part of the dissertation; it should demonstrate that you are capable not only of carrying out a piece of work but also of thinking critically about how you did it and how you might have done it better. This is seen as an important part of an honours degree.

There will be good things and room for improvement with any project. As you write this section, identify and discuss the parts of the work that went well and also consider ways in which the work could be improved.

Review the discussion on the Evaluation section from the lectures. A recording is available on Blackboard.

Appendices

Appendix A

Third-Party Code and Libraries

If you have made use of any third party code or software libraries, i.e. any code that you have not designed and written yourself, then you must include this appendix.

As has been said in lectures, it is acceptable and likely that you will make use of third-party code and software libraries. The key requirement is that we understand what is your original work and what work is based on that of other people.

Therefore, you need to clearly state what you have used and where the original material can be found. Also, if you have made any changes to the original versions, you must explain what you have changed.

As an example, you might include a definition such as:

Apache POI library The project has been used to read and write Microsoft Excel files (XLS) as part of the interaction with the clients existing system for processing data. Version 3.10-FINAL was used. The library is open source and it is available from the Apache Software Foundation [**?**]. The library is released using the Apache License [**?**]. This library was used without modification.

Appendix B

Ethics Submission

This appendix includes a copy of the ethics submission for the project. After you have completed your Ethics submission, you will receive a PDF with a summary of the comments. That document should be embedded in this report, either as images, an embedded PDF or as copied text. The content should also include the Ethics Application Number that you receive.

Appendix C

Code Examples

3.1 Random Number Generator

The Bayes Durham Shuffle ensures that the psuedo random numbers used in the simulation are further shuffled, ensuring minimal correlation between subsequent random outputs [**?**].

Annotated Bibliography

[1] "Apache JMeter - Apache JMeter." [Online]. Available: http://jmeter.apache.org/

An open-source Java based performance and load testing tool originally designed for web applications.

[2] R. Boyle, F. Corke, and C. Howarth, "Image-based estimation of oat panicle development using local texture patterns," *Functional Plant Biology*, vol. 42, no. 5, p. 433, 2015. [Online]. Available: http://www.publish.csiro.au/?paper=FP14056

Paper detailing a technique used to detect oat panicles via computer vision techniques. Development of panicles can be directly corrolated with certain growth stage (around GS55) in oats

[3] "Build software better, together," GitHub, Inc. [Online]. Available: https://github.com

An online Git repository hosting service

[4] D. Kendal, C. E. Hauser, G. E. Garrard, S. Jellinek, K. M. Giljohann, and J. L. Moore, "Quantifying Plant Colour and Colour Difference as Perceived by Humans Using Digital Images," *PLoS ONE*, vol. 8, no. 8, p. e72296, Aug. 2013. [Online]. Available: http://dx.doi.org/10.1371/journal.pone.0072296

Paper detailing how a humans perception of colour in images of plants can affect judgements made about these images.

[5] M. N. Merzlyak, A. A. Gitelson, O. B. Chivkunova, and V. Y. Rakitin, "Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening," *Physiologia Plantarum*, vol. 106, no. 1, pp. 135–141, May 1999. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1034/j.1399-3054.1999.106119.x/abstract

Paper detailing senescence detection in plant images by analysis of colour

[6] "National Plant Phenomics Centre," National Plant Phenomics Centre. [Online]. Available: http://www.plant-phenomics.ac.uk/en

National Plant Phenomics Centre

[7] J. Nielsen, "Response times: the three important limits," 1994.

Article discussing tollerable wait times for web page loads

[8] J. R. Quinlan, "Induction of Decision Trees," *Mach Learn*, vol. 1, no. 1, pp. 81–106, Mar. 1986. [Online]. Available: http://link.springer.com/article/10.1023/A%3A1022643204877

Paper detailing the ID3 decision tree algorithm

[9] J. M. Tanner, R. H. Whitehouse, W. A. Marshall, M. J. R. Healty, and H. Goldstein, "Assessment of Skeleton Maturity and Maturity and Prediction of Adult Height (TW2 Method)," 1975. [Online]. Available: http://core.tdar.org/document/125299

Paper detailing the atlas approach used in this instance to predict adult height in human's from skeletal features in children

[10] "Travis CI User Documentation," Travis CI GmbH. [Online]. Available: https://docs.travis-ci.com/

A continuous integration tool which can hook into other online resources such as GitHub

[11] G. W. Williams, "Comparing the Joint Agreement of Several Raters with Another Rater," *Biometrics*, vol. 32, no. 3, pp. 619–627, 1976. [Online]. Available: http://www.jstor.org/stable/2529750

A paper describing the comparison of expert opinion with that of other experts or a group of experts

[12] J. C. Zadoks, T. T. Chang, C. F. Konzak, and others, "A decimal code for the growth stages of cereals," *Weed res*, vol. 14, no. 6, pp. 415–421, 1974. [Online]. Available: http://old.ibpdev.net/sites/default/files/zadoks_scale_1974.pdf

Paper detailing the decimal growth stages of cereal plants