School of Information and Communication Technology

Griffith University

7820ICT – Industry Affiliates Program

Sniffing For Services

Project Report

**Wenjin Li s5120196**

Date of Submission (Include Year and Semester)

**CSIRO Land & Water**

**Industry Supervisor: Matthew Stension, Nicholas Car**

**Academic Advisor: Andrew Lewis**



*This document provides a template for the Project Report to be submitted as Assessment for IAP. The content is based upon Australian Standard AS 4071 (IEEE Std 1058.1): Standard for Software Project Management Plans.*

*This template is intended to be a guide for developing the project report. Items that are intended to stay in as part of your document are in bold; italic text is used for explanatory information that should be removed when the template is used.*

*The content of Section 1, Project Plan, forms one assessment item. Sections 2, Post-Project Review, and 3, Project Portfolio (including the Executive Summary, and the appendices containing the project deliverables), should be added in turn and the whole document re-submitted when the latter assessment items become due.*

**Executive Summary**

*Provide a one (1) page summary of your project, including an overview of your industry partner, what their business entails, the context of your project, the project objectives and major requirements, the proposed solution, any major issues that arose and a brief summary of the project outcomes.*

# 1 Project plan

# 1.1 Introduction (what to do)

## 1.1.1 Project Overview

The Commonwealth Scientific and Industrial Research Organisation (CSIRO) is an independent Australian Federal government agency responsible for scientific research. As a result, CSIRO publishes a large amount of scientific data both internally and externally. Some of data it publishes is via web services, both standard and custom. Despite sophisticated server management tools, it is very hard to know what services are installed across all CSIRO publication systems. It’s even harder to know, centrally, what the published data is about.

## 1.1.2 Project Objectives

The main goal of the project is to help CSIRO identify and manage the data publication system. The outcome of the project should help users like research scientists in CSIRO use the service conveniently to retrieve data they want without programming knowledge or experience. A [NASA web service](https://dataserver.nccs.nasa.gov/thredds/catalog.html), based on Thredds data server, is a good example for the project to start with to make query, which contains a catalogue for different datasets.

## 1.1.3 Proposed Solution Design and Requirements

This project is to write code to interrogate CSIRO’s internal and external web presences and hunt for particular service technologies. This involve polling for open ports known to support specific applications and testing endpoints with a range of requests standard services are known to answer. Once discovered and identified, it is required to design a data structure like a database that is used to make queries of the descriptions of data exposed by the services and map them to standard dictionaries of data description terms.

## 1.1.4 Project Deliverables

|  |  |  |
| --- | --- | --- |
| Components | Description | Tools |
| A network sniffing bot (Web Crawler) | The bot has ability to slot in multiple tests. It also can identify specific service types to extract description terms of the services and to compare extracted description terms to dictionaries. | Mainly use NMAP and netCDF4 library in python |
| A database | The database is well-designed and has ability to response the queries and request rapidly. | MongoDB - NoSQL |

## 1.1.5 Work Packages and Dependencies

*A work breakdown structure must be documented that includes all the tasks necessary to produce the project deliverables. Task dependencies must be shown clearly. For software development projects, the WBS must be based on an identified software development lifecycle model. By preference the WBS should be organised as a WBS tree diagram.*

*What to do*

## 1.1.6 Resource Requirements and Allocation

*The procedure for estimating resource requirements must be documented. A top-level estimate of size for the project as a whole should be generated and used as a basis for defining estimates for effort and duration. Estimates for all the activities and tasks listed in the WBS must be documented.*

*How long*

## 1.1.7 Schedule

## *The Schedule should show, in graphical or textual form, the planned starting and finishing dates of each task listed. The critical path showing the minimal set of tasks for successful project completion should be clearly shown. There must be provision for the schedule to be updated following changes to the Plan. Any dependencies on contributions by other personnel should be indicated. In addition, the schedule should include the delivery of milestone reports at the completion of major project phases.*

## 1.1.8 Reference Materials

*A complete list of all documents and other sources of information referenced in the Project Plan. It is sufficient to demonstrate knowledge of industry best practice as it relates to the project. This section may identify other planning documents regarded as important; references to these plans should clearly identify the content required. All documents referenced should be cited using an appropriate bibliographic style.*

## 1.1.9 Definitions and Acronyms

*A list of all terms and acronyms necessary to understand the Project Plan.*

NMAP (Network Mapper): It is a security scanner, used to discover hosts and services on a computer network. It has a python library which helps in using nmap ports scanner.

MongoDB: It stores data in JSON-like documents, which makes the database very flexible and scale. One of the most popular NoSQL databases is MongoDB.

Flask: It is a microframework for Python.

# 1.2 Technical Processes (how to)

*This section of the Plan describes what tools, methods and techniques will be used in the Project, and how their use will be documented. It should include the following elements:*

## 1.2.1 Methods, Tools and Techniques

Method: a brief story to show how I use the following tools

Communication tools:

1. Email:
2. Face to face meeting

Development tools:

1. Pycharm: It is a major IDE to do the project
2. Sublime: I will use it when I need to do a test on some snippets of code for the project.
3. Bitbuckets: I will commit and push all the files in the project folder when major changes have made.
4. NMAP: This tools is used to scan open ports and services of each host in CSIRO network across Australia.

Technical:

1. Python
2. NMAP commands
3. Linux or Unix commands

## 1.2.2 Quality Management

*An outline of the specific quality objectives for the project, and of the quality management strategy to be applied. This should be brief, and may refer to a Quality Plan that will be developed separately; it should identify the processes relevant to quality management that will be implemented, and how these relate to the overall process model and life cycle for the project.*

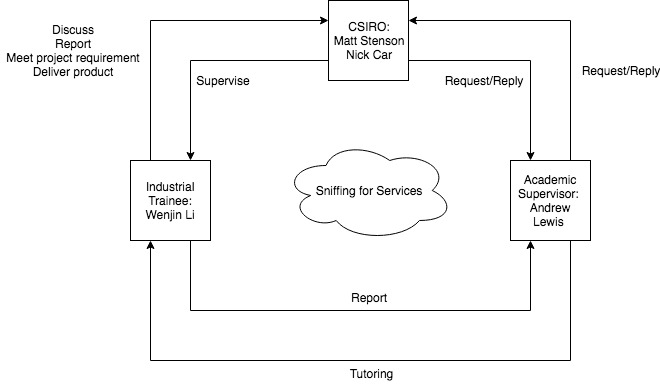
**Testing – how good quality in (survey – feedback, how meet the needs)**

**What I wna to know**

**What is it going to do for me**

# 1.3 Managerial Processes (how to manage)

## 1.3.1 Project Organisational Structure



Daily stand-up will be held to make a three-minute pitch to report what I have done and any potential issues and any impediments in my way to my supervisors Matt and Nick. I will be also email Matt if I need help and any clarification in regard to the project.

## 1.3.2 Project Success Criteria

|  |  |  |
| --- | --- | --- |
| Project Objectives | Success Criteria | Prioritisation |
| A network sniffing bot (Web Crawler) | Detect all open ports of all active hosts for a particular IP address within a range of addresses | Mandatory |
| Retrieve type of services of each host e.g. THREDDS data server | Mandatory |
| Identify service types by making query to the hosts to check if they are the data server  e.g. 192.168.1.1/thredds | Mandatory |
| Sort out what kind of servers and dataset it hosts e.g. OpenDAP, WMC and WCS etc. | Mandatory |
| Figure out the metadata of the dataset | Mandatory |
| A database | A well-designed database to store the testing result | Mandatory |
| Short response time | Desirable |

## 1.3.3 Risk Management

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ID | Threat/Vulnerability | Measurement | Likelihood | Consequences | Management Techniques |
| 1 | Someone tries to copy code from my computer | Require password to log in and put computer in sleep modes | Possible | Moderate | Avoidance |
| 2 | Accidental fire or flood | Remote backup to Bitbucket | Possible | Major | Risk mitigation |
| 3 | Leave e.g. sick leave | Manage time properly and adjust the time schedule to ensure the project can be done on time | Possible | Moderate | Risk mitigation |
| 4 | Unplanned activities e.g. fire drill | Manage time properly and adjust the time schedule to ensure the project can be done on time | Possible | Moderate | Risk mitigation |
| 5 | Hardware problem e.g. hard drive is dead | Backup to Bitbucket when make any major change of the project | Possible | Major | Risk mitigation |

## 1.3.4 Change Control

*This should be an outline of the strategy for change control that will be implemented, and how this relates to the overall process model and life cycle for the project. Particular attention should be given to managing project scope.*

# 2 POST-PROJECT REVIEW

# 2.1 Success Criteria

*Review the criteria documented in Section 1.3.2. Evaluate to what extent each success criterion was met.*

# 2.2 EstimateReporting

*Review the estimates documented in the Project Plan. Evaluate to what extent each was met.*

## 2.2.1 Analysis of planned vs actual schedule

*Outline the planned schedule documented in Section 1.1.7. Describe any major variance in the actual schedule, analyse the causes, and describe the actions you took in response.*

## 2.2.2 Analysis of planned vs actual size

*Outline the planned size or scope of the project documented in the Project Plan, particularly in Sections 1.1.4 and 1.1.5. Describe any major changes in the actual project scope, analyse the causes, and describe your response (for example, using change management procedures documented in Section 1.3.4, and the outcome.)*

## 2.2.3 Analysis of planned vs actual effort

*Outline the planned effort documented in Section 1.1.6 (your time dedicated to the project.) Describe any major variance in the actual effort (which may have been caused, for example, in response to a change in project scope), analyse the causes, and describe your response.*

# 2.3 Review of Techniques

*A review of any techniques* ***that you may have used****, including, for example:*

• *Requirements Elicitation*

• *Analysis*

• *Design*

• *Implementation*

• *Testing*

• *Quality Assurance*

• *Change Management*

*(Note: this is not a prescriptive, or complete list. Review the techniques* ***you*** *used in* ***your*** *project, and refer to Section 1.2.1 of your report.)*

# 2.4 Quality Achievements

*Provide an analysis of achievement against your quality goals, as documented in Section 1.2.2.*

# 2.5 Lessons Learned

*What did you learn?*

*What would you do the same again and why?*

*What would you do differently, how and why?*

*Try to use an established framework for reflective writing to clearly demonstrate what and how you have developed professional competencies through undertaking the project. Try not to limit skills acquired to technical skills only, but consider transferrable skills also.*

# 3 Appendices: Project Deliverables

*Provide a list and brief description of the project deliverables, which may include source code, user manuals, technical documentation, test and milestone reports, screenshots, etc. Anything that you delivered to your industry partner should be included. All deliverables should be attached in appendices, or an accompanying compressed archive, as appropriate. The material in the appendices will be assessed.*

# 3.1 Appendix A

*Content of the first deliverable listed, etc.*