### DH RSE Scenario 1

The Annals of Amberwald - Guidance

# Objectives

The aim of this set of exercises is to acquaint groups of potential Research Software Engineers with a workflow and documentation process which encourages them to think about defining requirements and identifying dependencies within sets of requirements.

The main outputs should be group discussions, with there being no definitive right answers but rather an understanding of the complex nature of competing priorities.

### Task 1

The imaginary RSE team, Digital Benchpress, have received an e-mail enquiry asking for guidance and costs for a potential funding bid for a project.

It’s fairly dense , whilst also managing to be light on details. However, there are aspects that should prompt discussion about the viability of this project and some further questions for the PI.

#### Allow 5 minutes in a breakout session to discuss and then return to the main session using the list below as a prompt for sharing points. Hopefully, the groups will have other suggestions to add to this mix

Some possible take-aways from the e-mail might include:

* Questions about the format of the manuscript to be studied
  + Is it accessible?
  + Is it digitised?
  + Does it need to be OCR’d?
* The term Digital Methods is used vaguely
  + What is meant by digital methods?
  + Does this vagueness imply tokenism for the digital research aspects?
* The deadline of the project is close
  + Is it a hard deadline?
  + Can an adequate feasibility study be carried out in time?
* The aspirations of the project are very boldly stated
  + How certain are we that this research has merit and will have impact?
* The Lab is under external pressure to meet targets
  + Does this project meet the criteria for Digital Benchpress to move forward?
* There is a lot of information missing that would help inform a group decision
  + What is the budget?
  + Is the timeline flexible?
  + What are the actual requirements of the digital aspect?
  + Are there similar projects that we can look to build on?

For the purposes of this exercise, we will assume that the team presses on with the SDLC process, despite any reservations we uncover.

### Task 2

More information is provided to the groups about the outputs expected, some are digital, and some are more traditional, but they give a better overall sense of what is needed.

#### This exercise should take a little longer, so allow 10 minutes of discussion, and ask that a team scribe record the requirements identified in the table format shown in the document.

The prioritised requirements identified should be informed by the research goals and the user journeys, i.e., can a particular research goal be achieved without delivering a technical aspect? What is essential, and what is “nice to have”?

An example for discussion is provided:

|  |  |  |
| --- | --- | --- |
| **Priority** | **Requirement Definition** |  |
| M | Content Management System for publication of project articles and blog posts |  |
| M | Method for recording palaeographic features in the manuscript |  |
| M | Method for recording codicological features of the manuscript |  |
| M | Method for identifying entities such as places and people within the text |  |
| M | Method for encoding events and themes within the text |  |
| M | Interface for searching through the recorded entities and palaeographical features and presenting summary results for analysis |  |
| S | Front-end presentation of the full text of the manuscript |  |
| S | Parallel viewer of text and original images of the manuscript |  |
| S | Front-end Map view presenting results of searches spatially |  |
| S | Embed links to external resources |  |
| S | Provide method for easy citation of data gathered during the project |  |
| C | Enable workflow for public contribution |  |
| C | Include digitised objects from the local museum where relevant to search results |  |
| C | *Employ Natural Language Processing approaches on the text to identify topics and automatically identify Persons and Places.* |  |
| C | *Enable semantic search* |  |

### Task 3

This task is intended to get the participants to think about pulling threads on the requirements and identifying dependencies. It can’t be expected that a group will make very accurate estimates of the time needed, but they can think about the roles involved. It should be clear that apparently small changes to deliverables have potentially far reaching chains of consequences.

### Task 4

From their estimates, the participants should be able to calculate a cost for the Product Quote. This is a simple exercise, but demonstrates a process. Regardless of the total, Professor Neilman will insist that the budget can’t meet the product quote.

He has his own perspectives on where money can be saved, but the team should now debate whether the Professor’s position is reasonable, and also look at reprioritising the requirements.

## Notes

This exercise was carried out with groups of 4 and 5 in breakout rooms, each with access to a Miro board to focus the activities; a template can be found here.

<https://miro.com/app/board/o9J_l4tx5nQ=/>

It is helpful to allow the teams 5 minutes or so to acquaint themselves with Miro. To keep sessions focussed, the Timer function can be used which provides a countdown.