Daily Project Log

Software Engineering Assignment 2

Tomás Kelly

21207079



Software Eng (COMP30830)

UCD School of Computer Science

02/02/2022

Personal Notes

Learning journal

# Sprint 0

## 2022/02/07 Mon:

First informal meeting today. Chose a Scrum process management tool (goodday.work) suggested by Will. Spend an hour or two getting to grips with how it works. Decide to use it exclusively for Scrum process and to keep some other files (like this log) on Google drive. Real work begins tomorrow, creating realistic tasks, assigning them, getting stuck in.

## 2022/02/08 Tue:

First lab today. Went well. Team appears to mesh well. We got a couple of RDS database instances configured. We got API credentials for weather, google maps and JCDecaux – and agreed where to store them. Getting everyone sorted on Github took a while – there were visibility issues that neither we nor the TA’s could resolve (but the work around we arrived at is good enough for the project). The project folder structure is starting to take shape and we’ve agreed to start stand-ups from Thursday. All in all a productive start, given the limited timeframe. The goals this week are to have everyone up and running, see *some* code committed for the scraper, kick off the scrum process, get the database started and start filling out an SRS/our product backlog Ambitious?

## 2022/02/10 Thu:

Second lab today. Went well. Focused on planning pretty much exclusively, although did finally start some code (data loader scheduling). Daily scrums up and running. Some discussion on github for sharing, and working out where we will store everything (which docs go where?, folder structures, minutia…)

## 2022/02/11 Fri:

Scrum standups still a bit stilted. We decided to walk the board for input rather than round-robin. Hopefully as we get more used to it it will be easy to stick to the format and avoid segways into issue discussion. Very productive meeting on project direction. First wireframe for where we’re going (very happy with that). More minor issues on the sprint backlog. Looking like by the end of Sprint 0 we’ll have a much clearer direction of where we’re going. The question is… will we manage to get some actual dev work in too…

## 2022/02/12 Sat:

Document the Branching strategy we agreed yesterday. Try to implement it too. Leave main as main, create a dev branch, create a feature branch for the data loader and commit to it. Finish out the feature so I can then merge that branch back to dev. Screenshots, documentation etc. so we can share the process on Monday.

## 2022/02/13 Sun:

Finish the GIT documentation. A few hours ago I was still new to Git. Now - not so much. I tested everything I was doing, screenshot it all and put it all together in the Branching Strategy document from yesterday. All in all - I’m happy with the progress. I just wish I knew this stuff before doing this project, so I could get some coding done.

## 2022/02/14 Mon:

Desperately want to get some coding done before tomorrow. Not much time on Mondays but get some late python coding in. Get a Data Loader/Scraper started. It runs, connects to the API, collects JSON, and tries to connect to the database but crashes. Has to be good enough – have a start point for tomorrow.

## 2022/02/15 Tue:

No coding in the two-hour practical – we’re still finding our feet as a group. We go through the process of GIT branching together. We go through the process of replicating environments together. Good progress here – everyone seems to be on the same page. Resolve the issue with my database connections when I get home. Clean up our shared credentials spreadsheet. Add all my AWS keys etc. – the others will need those to access my instance. Commit everything. Communicate it all to our team. There was a bit of tension today as we’re getting pulled in twenty directions at once – but as our comfort levels grow with the team environment (where we share docs, where we commit code, how we sync environments, our project management tool) then things start to look more stable.

## 2022/02/16 Wed:

Stand-up today (later in the day) was a bust. Clearly tension remains from yesterday. We talked it through which was great. We agreed some changes to the process – and I reckon we’ll do that again (maybe at the weekend). It has been a frustrating first sprint as we’ve been dealing more with environment and coordination issues than coding. Get an hour or two in in the evening – review the first pass of the db’s, make some changes, code up the DB storage part of the data loader. Great to finally see some data in the db. But not quite complete – need to fix some bugs, add support for the weather db and get it scheduled on our EC2.

## 2022/02/17 Thu:

Last practical before the end of the sprint. In a bizarre way – today was a reset. We had our stand-up before the session and that gave us a handful of small tasks to clear (changes to the initial database design, replicated on two servers, additional items for the board, notes on our env, etc..). I handed the data-loader over to Jörg, Will got stuck into our PM tool (we all agree, we are not using it correctly) and I kicked off the installer process. Think the team left on a high. Had a quick look at DB access from home – seems like making the db private and then accessing it via the EC2 instance (using our SSH keys) is a way to go.

## 2022/02/18 Fri:

Debugging the ‘setup.py’ process that installs out application using pip install from github. I thought this would be trivial but takes me quite a while to get running. An not sure (to be honest) if it will work for the second package (our data loader and main application are in two different packages).

## 2022/02/19 Sat:

Saturday – have to focus on other modules today. Pick away at the scheduler through the day as I have little ideas on how to progress it. In theory – I believe it’s more or less complete now. But of course… it’s never been run! Perhaps tomorrow…

## 2022/02/20 Sun:

Big late-night Sunday push to get a clean data loader over the line. Get all the weather data loaded (from the db tables up). Resolve all the outstanding issues with the scheduler (the difference between thinking it’s done and running it!). Merge everything, tag a version and install. Was late – but I’m excited by having the first sprint more or less complete.

# 

# Sprint 1

## 2022/02/21 Mon:

Introduce better logging (as in… actually generate a log!...) for the running data loader and install before the scrum daily

## 2022/02/22 Tue:

Discuss remote access issues to Amazon RDS with Jörg

## 2022/02/23 Wed:

Team met up today in the hub to have yet another debate on the wireframes. For a single page - it’s been subject to a lot of discussion

## 2022/02/24 Thu:

So…. the decision to use Setup.py to make life easier… hasn’t made my life much easier. I was struggling to get Setup.py to install an endpoint for a script I had. Turned out… after much testing… I was just not familiar enough with the tool. Felt like the lab today was wasted - I spent far too long trying to resolve the issue with the script. Once it was installed (to manage the start/stop and schedule actions of our data collection) I quickly hit the issue that the command had to be issued using sudo… but our virtual environment is not enabled for the root user… so kept getting command not found.

## 2022/02/25 Fri:

Issue with our virt. environment not available to root solved using something as simple as “sudo -E env "PATH=$PATH"”. Probably spent too long today patting myself on the back. Did go through

## 2022/02/26 Sat, 2022/02/27 Sun:

<not much this weekend. other modules demanding some time… did make up a fun logo for the website when it goes live (some copyright liberties were taken! Lol…>

## 2022/02/28 Mon:

I started installing Apache on my EC2 instance today. I expected it to be be easy - but I’m getting an error when I try to launch my Python project using mod\_wsgi

## 2022/03/01 Tue:

I think I’ve resolved most of the WSGI configuration issues - I read extensively about mod\_wsgi and what files go where, etc. Create logical from the Apache /var/www/html/ to our project directory. Remain convinced I have 99% of everything in the right place - but still blocked as I can’t get rid of this python 3.8 error (when everything I have is running Python 3.9.

## 2022/02/02 Wed:

Take the nuclear option today. Completely remove all my software (apache, python, conda, dudeWMB) from my EC2 instance. Reinstall everything from scratch and… still exactly the same error! Tearing my hair out a bit here. I haven’t been blocked like this since we started the project. Can see an error. Just can’t rationalise it. Have been using all my google-fu to no avail on this one.

## 2022/03/03 Thu:

Finally! Apache sits up and lives on the EC2 instance. Issues with mod\_wsgi being compiled under a different python version resolved. Was able to demo a ‘Hello World’ while on a call to the team. Very happy. Just need to plug some content in now… Meet Mandy in the morning to discuss where we are. Explain progress has taken a hit this week due to having deadlines in five modules over the next week. Quiet practical… very few people in.

## <exams and assignment deadlines put everything on pause…>

## 2022/03/14:

Pick up the reins again. Document the Apache installation procedure so the team can duplicate the install on the backup EC2 instances. The plan has always been to have at least two running in case of failure. At the moment – two are data collecting. Now that we have a running server, start work on the fun (web) part of the project. Create a branch for ‘basic project layout’ and start to work on that. Talk to the team at a scrum standup and we roughly divide the implementation of this first basic product version into layout/maps/getting-data-from-database.

## 2022/03/15:

More work on the basic nature of the flask application. Getting stylesheet references correct (templating, url\_for), getting includes correct (literally, flask includes), loading configuration for host, ip, port, maps keys etc.. Sort out db connection using flask-sqlalchemy. Define models for the tables we have so far. Implement a JSON load of basic station data. DudeWMB starting to take some shape!

## 2022/03/16:

<no DudeWMB today… take the day to focus on Java…>

## 2022/03/17 Thu:

St. Patricks Day! :-)

## 2022/03/18 Fri:

Team met up today in UCD to take a step back and look at the project documentation etc.. We did merge a feature branch (for the maps) while we were there but mostly focused on looking at what documentation we had and if we were fulfilling the project brief. Was quite a productive session - we organised the documentation we had and agreed what gaps had to be filled etc..

## 2022/03/19 Sat:

Integrate the commit on the maps yesterday into the main (dev) branch. Not a traditional merge - more pushing all our CSS into a single resource, all our JS likewise. Take the ‘shows one station’ version of the map and get it to load and display all the stations. Basic information now displayed per station. Also take a quick call from Jörg on merging his work into dev - all good there. The basic layout is not complete.

## 2022/03/21 Mon:

First day back in term. All I did today was update the release file, merge dev into main, and then do an install so we’d have something to show the product owner in the morning. Feels great to have a site finally up if I’m honest. Even though it really is a ‘minimum deliverable’ at the moment.

## 2022/03/22 Tue:

<sick today - like, really sick. barely stood - didn’t do an DudeWMB today>

## 2022/03/23 Wed:

Did a quick catch up with the guys on what happened yesterday. Still ropey.

## 2022/03/24 Thu:

<another day – I’ve missed two labs now – pretty down about that>. The team have started looking at the ML model we need to integrate. It’s looking like I’m going to be mostly in a supporting role for this sprint. Jorg is leading up the effort to resample our data to hourly (we desperately need this done to drive both the ML model and the intended display of data on the front page).

## 2022/03/25 Fri:

Spend an hour just looking at some of our project documentation – turns out our documentation is out of date. We’ve been reasonable at keeping the scrum process going I think. But as the pressure has built with deadlines for the other modules our own scrum documentation has gotten lighter. Plan to spend some time at the end of this sprint rectifying that. Meet the lads online… what a day… one for my personal-personal logs…

## 2022/03/28 Mon:

I drew up and committed a couple of colorised icons for the bikes on the map. They’re useless now – but as soon as we get access to our new dataset it should be trivial to display them. I added an endpoint for the occupancy data. After missing a couple of days last week it seems I’m picking up the small jobs so the team can focus on the two big tasks taken.

## 2022/03/29 Tue:

At the lab today I went through a design of the data-resampler with Jorg to agree an “in English, in principle” approach to the process he’s working in. It’s split into two tasks really. One is the running process I want to schedule. The other is the job to run this process for our historical data. Design wise it seems straightforward enough. The devil is… of course… always in the detail! Also go through the process with Will of pickling the model (it’s almost finished) and then installing the model to our project (just dropping it in – that’s neat) and accessing it using a Flask endpoint.

## 2022/03/30 Wed:

Jörg was working away on the resampler and is feeling the pinch. Collaborate with him for an hour this evening to get past the SQL hump and to get the process running. Begin adding the resampler to the scheduler. But… it’s not finished yet so I can only guess what it will look like and add a framework. Will return to this to nail it once the python process is finished.

## 2022/03/31 Thu:

More time spent with Jörg today trying to get the prototype data resampler working. Most of the delay seems to be big chunks of time blocked. Do my best to keep removing impediments so the process doesn’t stall.

## 2022/04/01 Fri:

And again - just support today. A total of two and a half hours on zoom. Tiring stuff. Frustrating. Would rather be coding myself.

## 2022/04/02 Sat:

<no DWMB today – big push to clear DS&A Comp20230 Project>

## 2022/04/03 Sun:

<no DWMB today – big push to clear DS&A Comp20230 Project>

## 2022/04/04 Mon:

Sort out the google maps API key this evening. Can’t tell you how happy I am to see the site without the developer lines/warnings on it. Also do a release and install the data re-sampler to my own instance. Test it a couple of times (just some teething issues related to the environment). But it sits up and runs so we should have a nice clean data source now for charts etc..

## 2022/04/05 Tue:

Spot our shiny new resampler not working! And even worse… was my fault! The endpoint I had added was ‘resample’ - but the scheduler was calling ‘resampler’. Fix immediately - luckily thanks to a nice design the resampler recovers automatically, so not data lost. Also renamed a chunk of stuff now that we’re scheduling more than one process. Just wanted the naming to be a bit more coherent.

## 2022/04/06 Wed:

Because… you know… we should start something new with days to go… all the data we had loaded so far used SQL (SQLAlchemy) in just raw/paling text format. But we did decide to use ORM for the main web application. So - figure out how to use SQL Alchemy ORM today. I find the SQL Alchemy documentation really light. And because it works in a couple of different ways finding working examples online can be challenging. But anyway, by the end of the day we’re querying data, filtering it, etc.. Really happy with this - it’s at least some python I reckon I’ll use in the future.

## 2022/04/07 Thu:

Another big day (am spotting a pattern). Do a flask endpoint to call the prediction model (this will not be workable… model is just too big - will address tomorrow). This endpoint is 100% need though so the work on the front end can progress. Also do an endpoint to load and return the occupancy data - not a small amount of work in itself. Added some code (really placeholders) to kick off the slider implementation. Can only hope it gets turned around.

## 2022/04/08 Fri:

Our initial data model was 3GB. The ‘revised one’ was 300mb. It is simply not feasible to load these models 110 times to run predictions live. Grab Will on a zoom today and rework the “single monolithic model” into 111 pickles for all the individual stations. Also reduce the forest sizes. We end up with a model that’s still quite an acceptable R2

**2022/04/09 Sat:**

Today was a huge, all day, where do I start push… On the small side I fixed a bug with the weather icon. On the big side I finished out the work on the the gviz api for the charts on the front. But the biggest work today was the gut punch in the morning… ever since the pip install of sklearn our EC2 has been flakey. I’ve already reinstalled apache. But today… for some reason… our app won’t load at all. Out of frustration, shut down the EC2 instance and create a new one from scratch. Thank god I documented the process. I managed a clean install based entirely on those docs (missed only one step) and the server sat up and ran. But still… dudeWMB won’t load. I can see the issue (credentials aren’t loading). Turns out using relative paths is a bad idea. Introduce a formal check of the module load path before attempting to load the credentials. BOOM! Back in the game. To be doing this with one week to go was just mental. But anyway… it’s done now. I can breathe again… I had visions of not even having a website to hand in…

## 2022/04/10 Sun:

Cool - there was an issue with our maps where the markers weren’t focusing on click correctly. I wasn’t involved in that feature initially so hadn’t had a chance to look at it. But finally got to look at it today. Definitely part of the ‘closing things out vibe’. Was just a scope issue, but makes a big difference to the functionality of the UI (basic as it is). Done, tested, committed… nice.

## 2022/04/11 Mon:

<no dude WMB today - mad java day!>

## 2022/04/12 Tue:

Into the hospital early today… so back to UCD by 07:30. Get the hotfix version of DudeWMB installed live. Delighted to see the site back on the air again. One or two small JS bugs dealt with straight away. Also make banking code human readable. We can deploy those later in the week, time permitting. Have to focus on the report today.

## 2022/04/12 Wed:

Spent today putting together the architecture section of the project report. Space constraints quite tight - think the report will fill itself quite quickly.

## 2022/04/12 Thu:

Update the front end as it wasn’t switching the display to the predicted data, after it was pulled in using the slider. That’s the last of the big functionality. We now have (past) occupancy, (current and future) status displayed on the map. Reasonable text prompts and navigation tools. All that’s left is a few css tweaks and some tidying up. Plan to meet J in the morning to clear those.

## 2022/04/12 Fri:

Meet Jörg in UCD for the final push. We sort out any remaining display issues for the front end - and add a cool little optimisation so the page won’t reload data unnecessarily. Would rather not be making changes at this late stage. But delighted we got it over the line in a serviceable if basic state.

# Personal review:

# Retrospective of the project: