

AberLink - Creating a link between University and Discord accounts

CS39440 Major Project Report

Author: Joel Adams (joa38@aber.ac.uk)

Supervisor: Dr. Neal Snooke (nns@aber.ac.uk)

31st March 2021

Version: 1.0 (Draft)

This report was submitted as partial fulfilment of a BSc degree in Computer Science
and Artificial Intelligence (GG4R)

Department of Computer Science
Aberystwyth University
Aberystwyth
Ceredigion
SY23 3DB
Wales, U.K.

Declaration of originality

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 Registry (AR) 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: Joel Luca Adams

Date:

Consent to share this work

By including my name below, I hereby agree to this project's report and technical work being made available to other students and academic staff of the Aberystwyth Computer Science Department.

Name: Joel Luca Adams

Date:

Acknowledgements

Mr. Dan Monaghan - for giving me the original inspiration for this project and for helping me along the way with understanding key concepts.

Dr. Neal Snooke - for collaborating with me on multiple Discord bots and improve the unis online Discord presence. His bots were main motivation behind this project and I'd like to thank him for allowing me to re-create them for this project.

Dr. Alun Jones - for his continued support throughout this project and his resourcefulness in providing instructions on how to setup and use uni services for this project.

Mr. Leslie Johns - for having setup the API endpoint for student attendance that has been crucial in getting this project off the ground.

My family and friends - for providing support and helping me get through university. A special thank you to everyone in the Comp Sci After Dark Discord server for being a great group of people and friends.

Dr. Angharad Shaw, Mr. Chris Loftus and the Computer Science Department - for helping to populate the Comp Sci Community Hub which is what lead to the unis online Discord presence, without which wouldn't have lead to the creation of Neal Snooke's bots. A special thank you to Chris Loftus for acknowledging my work on DemoHelper and turning it into a job that inspired me to work on more Discord bots and the eventual creation of AberLink.

Abstract

AberLink contains 2 components; a website which is used to link up Discord [1] accounts to students' accounts and a Discord bot called AberLink that is used for verifying students' Discord accounts. This bot is also responsible for marking students' attendance during practicals that have the flag of "Discord" on their student record.

This system was developed on the Linux distribution Debian 10 (Buster) [2]. It uses the open source HTTP server Apache2.0 [3] to web-host along with the Python framework Django [4] for the website. The website is locked behind the Open-ID Connect [5] login system that only authenticates users with Aber accounts. Once logged in a user can link multiple Discord accounts to their Aber account.

Information is saved to a back-end PostgreSQL [6] database using primary keys to link the uni and Discord accounts together. The Discord bot (AberLink) using the Python library Discord.py [7] then verifies Discord accounts by checking if that account is present in the database and then grant users access to the Discord server. AberLink also has an API endpoint for marking attendance on Student Record and uses the database to perform a reverse search on Discord users to get their uni information and send that to the endpoint.

The AberLink Discord bot is based on two previous bots called Aber Verify Bot [8] and I am here! [9], both created by Neal Snooke. These bots verified students' Discord accounts and helped to mark attendance during practicals respectively. AberLink greatly modifies the original implementation of these bots to use a proper back-end database and an API to update attendance.

Contents

1	Background & Objectives	1
1.1	Background	1
1.1.1	Research	1
1.2	Analysis	2
1.2.1	Objectives	2
1.2.2	Possible Security Issues	3
1.3	Process	4
1.4	Functional Requirements	5
2	Design	7
2.1	Overall Architecture	7
2.2	Website Architecture and Design	8
2.3	Discord Bot Architecture and Design	10
2.3.1	Database Interaction	11
2.3.2	Complicated Behaviours	12
2.4	Database design	13
2.5	User Interface	14
2.5.1	Website	14
2.5.2	Discord bot	16
2.6	Version Control System and IDE	17
3	Implementation	18
3.1	Code Implementation & Third-Party libraries	18
3.1.1	Website Building	18
3.1.2	Discord Bot	18
3.2	Unforeseen Issues	19
3.3	Review Against Planned Requirements	19
4	Testing	21
4.1	Sample Data	21
4.2	Automated Testing	21
4.2.1	Unit Tests	21
4.2.2	Stress Testing	23
4.3	User Interface Testing	24
4.4	User Testing	24
4.5	Integration Testing	24
5	Evaluation	25
	Annotated Bibliography	26
	Appendices	29
A	Third-Party Code and Libraries	31
B	Ethics Submission	33

C	Code Examples	36
3.1	Random Number Generator	36

List of Figures

2.1	Architecture diagram of the overall system	7
2.2	Architecture diagram of website	8
2.3	Website flowchart for authentication of accounts and display of user data	9
2.4	Architecture diagram of Discord bot	10
2.5	Tree diagram showing Discord file structure	10
2.6	AberLinkDiscord UML diagram	11
2.7	Sequence diagram for Discord bot to database	12
2.8	Entity relationship diagram for database	13
2.9	Website mock-up for 0 users	15
2.10	Website mock-up for 1 user	15
2.11	Website mock-up for 2 users	16
2.12	Discord Embed example for getting users information	16
2.13	Discord Embed example for source command	17
2.14	Example of GitLab repository commits	17
3.1	Updated Entity relationship diagram for database	19
4.1	Django URL render test	22
4.2	Django database render test	22

List of Tables

2.1	Aberystwyth user table example	14
2.2	Discord user table example	14

Chapter 1

Background & Objectives

1.1 Background

This project required a substantial amount of discussion with IS and CS-support due to the sensitivity of this project and the data that it interacts with/stores. As meetings can sometimes take weeks to organise and further weeks to decide what can be used I began work on the background and spike work as soon as possible. By the end of the first week I had already had completed most of the spike work required to form the Project Outline and send it out to IS and CS-support. This blog post also has some details about the first week of research here <https://cs39440blog.wordpress.com/2021/02/01/week-1-25-01/>.

Note: The dev folder of the technical submission also contains spike work that may not be listed below and I encourage that you have a look there too when reviewing this section.

1.1.1 Research

Web Hosting & Containers - This was definitely one of the sections that I had the least experience working with but thanks to help from CS-support they guided me through the process and setup a Debian 10 (Buster) [2] container. Debian is definitely the best OS for this project as it is flexible and has many useful libraries for my project. This came preloaded with Apache2 [3] which is used to host websites so was the logical choice for me to work with. NGINX [10] was also considered for this project but was scrapped in favour of Apache2 due to the assistance available from CS-support and online documentation.

Coding Languages - Over the summer I worked on the DemoHelper [11] project and learnt how to create my own Discord bots in Python. This spawned many Discord bot projects that can be seen in the comp sci server. From these projects I gained a solid foundation of Python and is the reason behind me choosing this as my primary coding language this project. It also had a knock on effect for choosing the database and web framework that worked well or used Python. I also used HTML, CSS and some Javascript to develop the website pages.

Databases - Early on I decided that a relational database was best suited for the data as

I am only storing simple user information. The data would be split up into two tables; one for Aberystwyth user information and one for storing Discord account information. These would then be linked using a primary key in the Aberystwyth user table and then a foreign key in the Discord user table. The system would also be designed to allow users to have multiple Discord accounts so the database used a one-to-many relationship.

There are a few databases out there that support these features but PostgreSQL (PSQL) [6] was used as I am already familiar with using it in a second year project and it has native support for the web framework that I will discuss later.

Website Frameworks - For the website framework I wanted to use something that would scale easily, have lots of documentation and preferably be written in Python. There are lots of frameworks that fit this category such as Flask and web2py but I settled on Django. I chose this because it came with integrated features such as custom user models, pre-made admin pages and authentication support. It also came with built in features to write data to and read from PSQL [6] so that was a big bonus too.

1.2 Analysis

1.2.1 Objectives

After completing the spike work and prior research the next step was breaking the project down into sections and deliverables. Below is a list of the items that were used as milestones for the project completed in descending order.

- **Research & Discussion with IS/CS-support.**
 - Discuss how to access uni data and how to login users who are only on the uni network using OpenID Connect [5].
 - Build the attendance API so that it is secure and only marks students for current practical and not previous practicals.
 - Setup PSQL [6] and make it secure from outside attacks.
- **Version control, documentation and setup.**
 - Create a container on the uni network to remain secure and locked under VPN access.
 - Documentation and version control using git on the university's GitLab upon the departments request so that it can be easily redeployed as a complete service later on.
 - A blog (usually bi-weekly) to document the process and progress of this project.
- **Creation of Python back-end for website and the database.**
 - Build the website using the Python framework Django [4].
 - Establish a PSQL [6] database for data storage.

- **Re-writing 'AberVerify' and 'I am here' into single Discord bot.**
 - Recreate the two mentioned bots in Python instead of JavaScript.
 - Make the bot use a Discord users information to lookup their Aber ID in the database using relational keys.
 - Make the bot update attendance using the uni provided API endpoint.
- **Interface for lecturers and students on website.**
 - Create webpages for users to add Discord accounts, view the module servers they are in and the one's that they aren't.
 - Create Admin pages only visible to staff to view all the connected students, remove them and configure their roles in servers for which they have administrative privileges.
- **Resource links and further webpages.**
 - Create a Discord bot function to get information such as the Aber account that is linked to the Discord account.
 - Discord bot function to display other useful discord bots that can be added to the server. e.g. DemoHelper [11]
 - Create webpages to display information such as the 'privacy policy', 'ethics form', 'blog' and 'about this project'.
- **Further potential work**
 - Integrate DemoHelper Discord bot into AberLink
 - Add multiple language support

1.2.2 Possible Security Issues

Throughout this project I have attempted to minimise the amount of security risks that can occur. The main identified weakpoints are as follows:

- **Direct Database Access** - This is the first point of attack that could be used to get access to Discord and Aber account details. This has been secured by using a PostgreSQL [6] database that is located in Departments servers so it is already behind a very secure and up to date firewall. To access the database you also need to have a registered Aber account and be using the campus WiFi or the uni's VPN so that adds another layer of security. If the user however has access to both of these then brute forcing the PSQL login system is difficult as it has many security levels and incoming connections are monitored by the university.
- **Unauthorized Admin Access** - This is the possibility that the user will try and brute force access to the admin page of the website. Before they attempt this they would need to get past the website's OpenID Connect [5] authentication system that requires an Aberystwyth account to authenticate. The website also has another layer

of security as it requires the user to be connected to the campus' internet or be logged in on the VPN. If they get past both methods there is no easy way to spoof the system to gain access to the administrator panel as the backend uses cookies to keep the user authenticated.

- **Accessing Database Credentials through Back-end Code** - All database credentials have been hidden in files that are not stored in the git repository and are loaded from JSON and .env files. This is good practice and creates a simple way for maintainers to setup and change variables such as database passwords easily instead of going through source code and manually editing it.
- **Accessing Data through AberLink Discord Bot** - The Discord bot doesn't contain any sensitive data, it merely queries the database using generic queries that are changed depending on the input variable. It can however be used to gain an understanding of the database model that is used.
- **Worst Case Scenario** - If the user somehow gains unauthorized access to the data they will only be collecting a list of emails and linked Discord accounts. No password data is ever stored in the database because OpenID Connect [5] is used to authenticate Aber accounts and Discord accounts are linked using OAuth2 [12]. This means that no password data is ever exchanged from the authentication systems to the website or database.

1.3 Process

For this project I found that Extreme Programming (XP) would best fit my project as early on I discovered that I worked best by breaking the project down into deliverables and then into components of work. You can see from the above section 1.2.1 the list of objectives (Functional Requirements) that I was working towards. Below are a list of the processes that I have followed or worked around in XP:

- **Project iterations** - These usually last around one week, however sometimes last longer due to unforeseen issues.
- **Pair Programming** - Due to the nature of this project being independent pair programming was not a viable approach. To compensate for this every morning after code was written I would go back through and review the code adding comments or refactoring the code.
- **TDD vs BDD vs FDD** - My style of coding usually revolves around writing code to pass some specific goal that has been set followed by user testing and finally unit testing (when applicable). Feature Driven Development (FDD) definitely fits my style best for this project.
- **When do iterations run?** - These run usually from Monday to Friday as I try to keep my weekends free to work on other projects and think over the next project iteration.

- **Where are requirements recorded?** - The project requirements can be found in the above section 1.2.1 and in the next section of this document 1.4. There is also a board of issues on the GitLab page that can be used to review the timestamps for when work was completed.

1.4 Functional Requirements

- **FR1** - Authenticate Aber User accounts and return useful Metadata
 - **FR1.1** - Only authenticate Aber accounts and reject all other accounts.
 - **FR1.2** - Return Aber user information along with the usertype.
 - **FR1.3** - Make users with "staff" usertype have administrative privileges.
- **FR2** - Authenticate Discord users accounts and link to Aber account
 - **FR2.1** - Once authenticated with Discord metadata is returned about account.
 - **FR2.2** - Discord account gets saved and linked to the Aber account using the primary key.
 - **FR2.3** - Multiple accounts can be linked to the same Aber account.
- **FR3** - Home webpage
 - **FR3.1** - Webpage displays Aber username and full name.
 - **FR3.2** - Webpage displays linked Discord accounts.
 - **FR3.3** - Discord ID's of linked accounts are sent to Discord API to get profile picture and username.
- **FR4** - Admin webpage
 - **FR4.1** - Webpage only allows accounts which have the "is_admin" permission.
 - **FR4.2** - Webpage displays a list of Aber user accounts.
 - **FR4.3** - Webpage displays a list of Discord user accounts that are linked to Aber accounts.
 - **FR4.4** - Webpage doesn't allow Admins to change any of the accounts data apart from their admin access.
 - **FR4.5** - Webpage allows admins to delete a user or Discord account entry.
- **FR5** - Discord bot verification
 - **FR5.1** - Bot can configure a server for verification.
 - **FR5.2** - Bot can verify students when command is called or when the user joins and is already authenticated.
 - **FR5.3** - Bot has a message for Alumni verification.
 - **FR5.4** - After user is verified their Aber email gets appended to their username. e.g JoelinRome#4893 becomes JoelinRome [joa38].

- **FR6** - Discord bot attendance
 - **FR6.1** - Bot can get a users information from database using a reverse lookup using Discord ID and get Aber and linked Discord accounts.
 - **FR6.2** - Bot can mark students attendance in practicals by doing a reverse lookup of their aber username and sending it to an API endpoint. Then send the user a DM with the status of their attendance e.g. Attended module CS10000 at 11:00 on 13/04/2021.
- **FR7** - Discord bot utilities
 - **FR7.1** - Bot has a config file to choose wether a users name gets modified to include their Aber email at the end.
 - **FR7.2** - Bot will display the current server configurations.
 - **FR7.3** - Bot can clear all messages from a channel.
 - **FR7.4** - Bot displays a source command showing where source code can be found.
 - **FR7.5** - Bot displays a list of other useful bots that should be added to the server e.g. DemoHelper [11].
 - **FR7.6** - Bot can be pinged to get latency and check wether databse connection is working correctly.
- **FR8** - Discord bot database connections
 - **FR8.1** - Connects to database and returns information on connection.
 - **FR8.2** - Can reconnect to database in case of database disconnect during up-time.
 - **FR8.3** - Python function to get a Discord user's information using their Discord ID.
 - **FR8.4** - Python function to get a Aber user's information using their aber email.
 - **FR8.5** - Python function to get database information e.g. connection status, latency and polling.

Chapter 2

Design

2.1 Overall Architecture

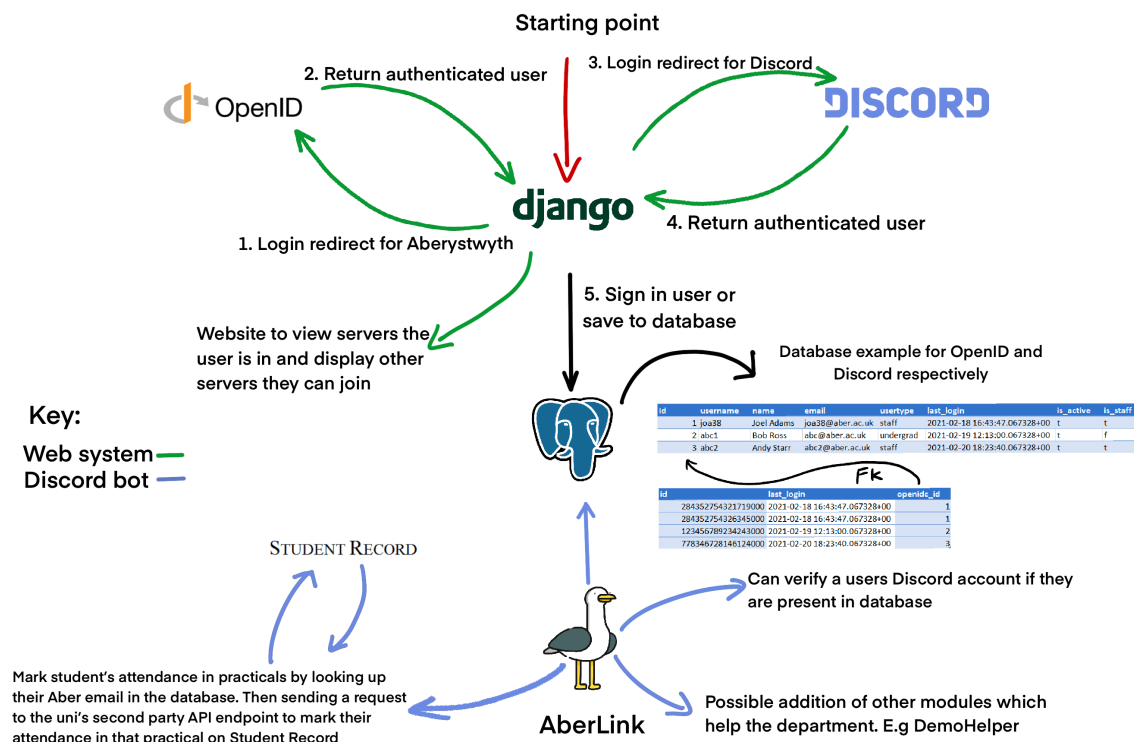


Figure 2.1: Architecture diagram of the overall system

As seen above here is a brief overview of the main components used in this project which I go into more detail later on in this document. Below is a list of the main folders and the files contained within them. The folder structure is based largely off of the second year group project folder structure.

- **config** - Describes how to setup the project; a bash script for installing the relevant

dependencies, a README.md for manual configuration of the project and example configuration files for Apache2 [3], authentication [13], Discord bot and Django [4].

- **dev** - Contains the spike work completed throughout this project and I strongly encourage that you review it when reading through this document.
- **docs** - Contains the documents created for this project; including this document, the project outline, the ethics for and the usability form.
- **src** - Contains the source code required to run this project. There are two folders for easier maintainability; a folder for the website and a folder for the Discord bot.
 - **AberLinkAuthentication** - A folder containing all the website resources and the Django [4] code. It also contains a virtual environment [14] folder that is not stored in git because dependencies update all the time.
 - **AberLinkDiscord** - A folder containing the Discord bot.

2.2 Website Architecture and Design

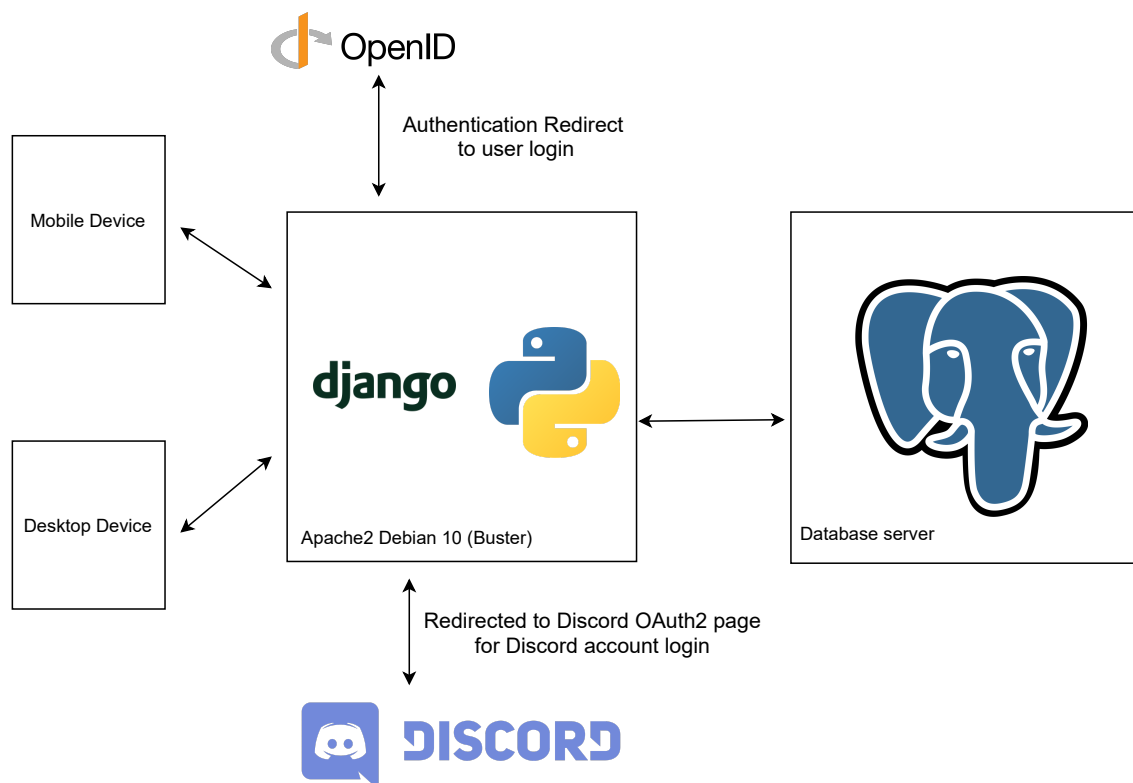


Figure 2.2: Architecture diagram of website

The website is built using the Python framework Django [4]. This framework has been used to generate a template of the code required and only requires minimal tweaking to setup. When you visit the website you then get redirected to the OpenID Connect [5] authentication framework which helps to prevent unwanted access by users who aren't permitted to view the website. Once authenticated the users data is then saved to the database and they can then add Discord accounts by using the button displayed on the screen. This then redirects them to Discord's OAuth2 authentication page and once they've logged in with a discord account then it redirects them back to the my website and saves that information to the database as well. I have included a flow chart below to explain the process.

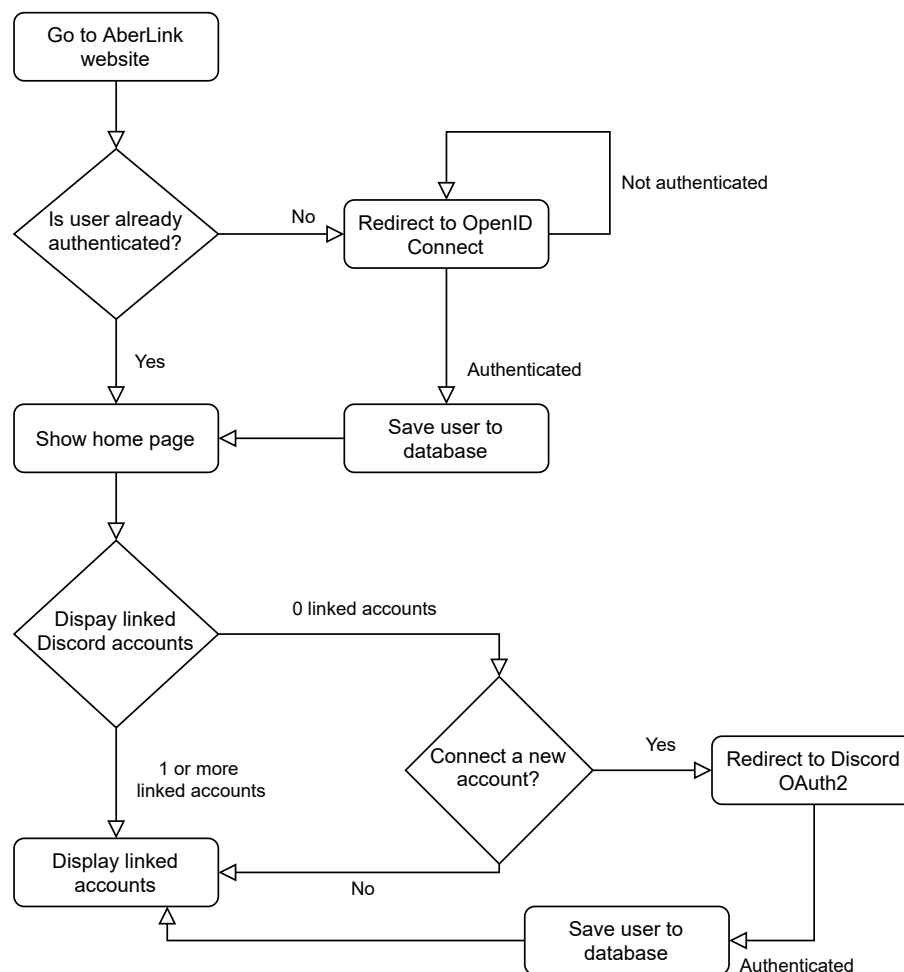


Figure 2.3: Website flowchart for authentication of accounts and display of user data

The website can be found here <https://mmp-joa38.dcs.aber.ac.uk/> and is only accessible over VPN.

2.3 Discord Bot Architecture and Design

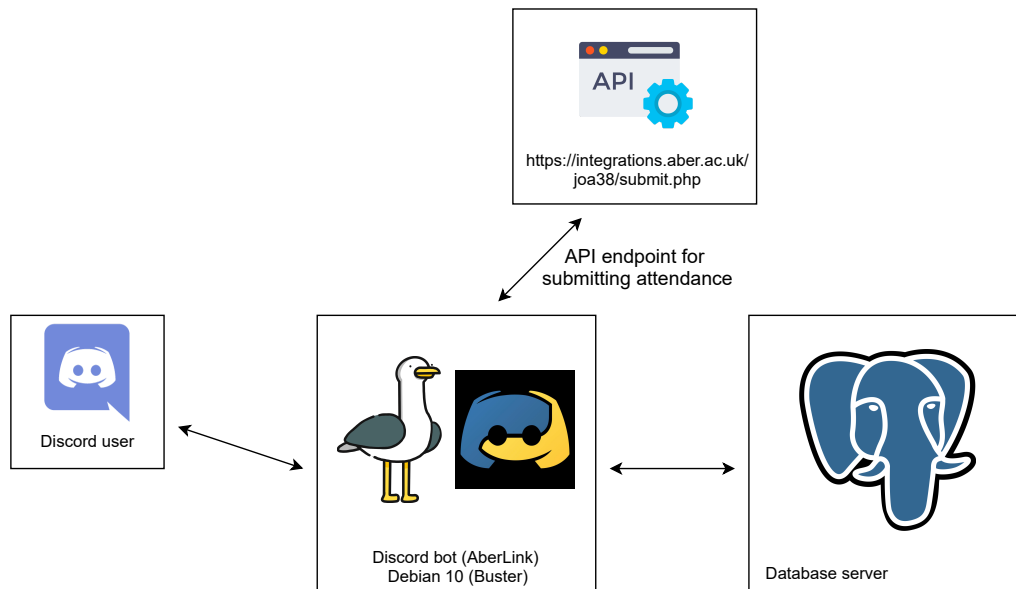


Figure 2.4: Architecture diagram of Discord bot

The Discord bot (AberLink) is usable through the Discord application available on mobile, desktop and their website. AberLink has many commands which can be easily found by typing the 'help' command in Discord on a server where the bot is present or via a direct message to the bot. The basic premise however is that it can access the database to retrieve information about a user but it never saves any information to the database however. The file structure is displayed below.

```

AberLinkDiscord
├── AberLink.py
├── Pipfile
├── .env
├── cogs
│   ├── __init__.py
│   ├── db.py
│   ├── here.py
│   ├── utilities.py
│   └── verify.py

```

Figure 2.5: Tree diagram showing Discord file structure

As seen above in the file structure the main file that is used to run the program called AberLink.py, initialising the bot instance and loading all the files from the cogs folder. Discord.py [7] uses a smart system where code is not written into the main file (in this case AberLink.py) but instead broken down into components or 'cogs' as they're called here. The .env file is used to store sensitive variables such as the Discord token required to run the bot and the data used to connect to the database server. This file is also not uploaded to git so the data is never exposed to the wider web and is only stored locally. The Pipfile is used to store the dependencies required to setup the bot such as discord.py [7], requests [15], psycpg2 [16] and the version of Python required to run the bot. The Pipfile can be initialised using the Python virtual environment pipenv [14].

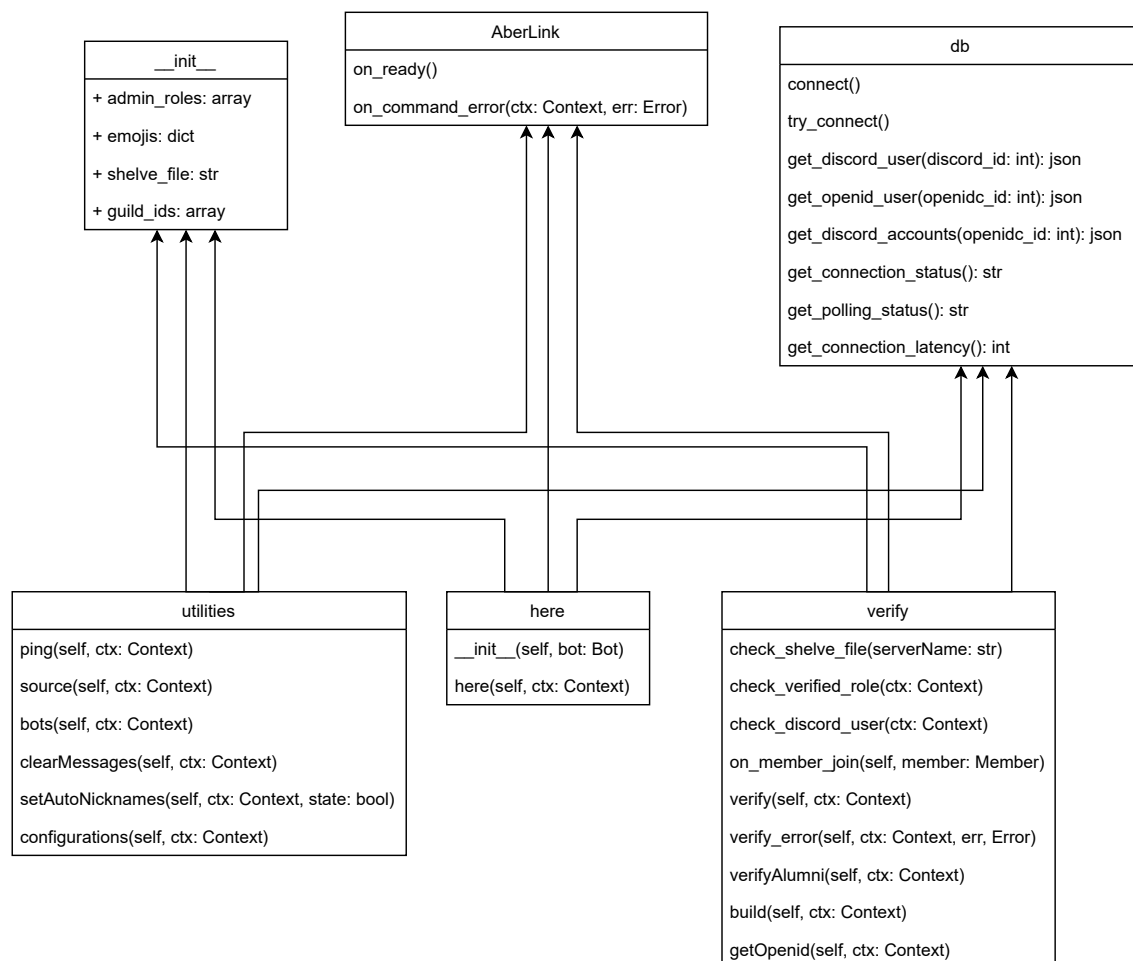


Figure 2.6: AberLinkDiscord UML diagram

2.3.1 Database Interaction

Below is a sequence diagram explaining how AberLink connects to the database and sends requests back and forth. It uses the Python library psycpg2 [16] to connect and send SQL requests back and forth; the details of which can be found in the db.py file. Once connected to the database if an error occurs while making a request then the code

executes a function (`try_connection()`) to reconnect to the database. Please see Appendix A for more details on `psycpg2`.

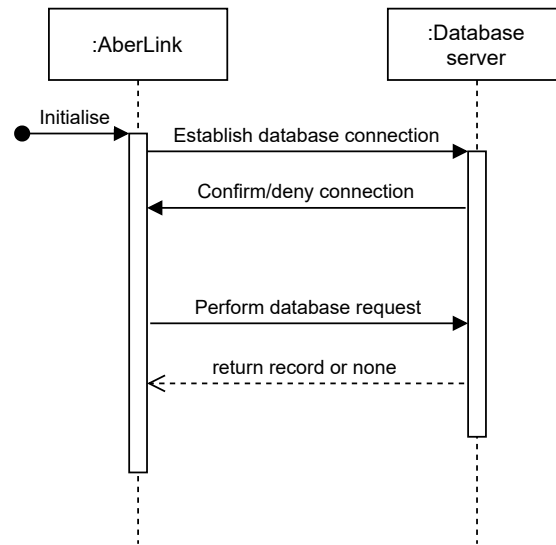


Figure 2.7: Sequence diagram for Discord bot to database

2.3.2 Complicated Behaviours

Most of the Discord bot code is relatively straight forward however there is one function in particular that needs some explanation called `build()` located in `AberLinkDiscord/cogs/verify.py`. This function and Discord command is used to configure a Discord server for verification. Below is a list of steps that the command executes.

1. Begins by simulating that the bot is typing giving the user visual feedback that the command is running as this function usually takes 500ms to 1s.
2. It then attempts to search for the `@everyone` role, `verify` role and `#verify_channel` so that it can update their permissions.
3. The `@everyone` role is then stripped of it's permissions entirely as it is the only role that is given by default to new joining users. By removing all permissions the new user cannot view any channels.
4. It then gets the `verify` role and if it doesn't exist then it is created and given all the default permissions that `@everyone` used to have. This basically makes it so that users that have the verified role can do everything that the `@everyone` role can no longer do.
5. The bot is then given the `verify` role so that they can view the server channels and interact with members that have the new verified role.
6. If the `#verify_channel` doesn't exist then it is created and a message is pinned to the channel asking users to type `!verify` to verify their accounts.

7. The #verify_channel is then configured so that users with the @everyone role can view and type in the channel to get verified.
8. The #verify_channel then removes the ability for the verified role to see the channel and therefore sets up a barrier where @everyone and verified can no longer see each other.
9. Finally once completed the bot stops typing and sends a message containing all the changes which have been made and the time spent making said changes.

2.4 Database design

Below is the Entity Relationship diagram that describes how the database works. The tables were created in Django which then remodels the psql [6] database to fit those requirements.

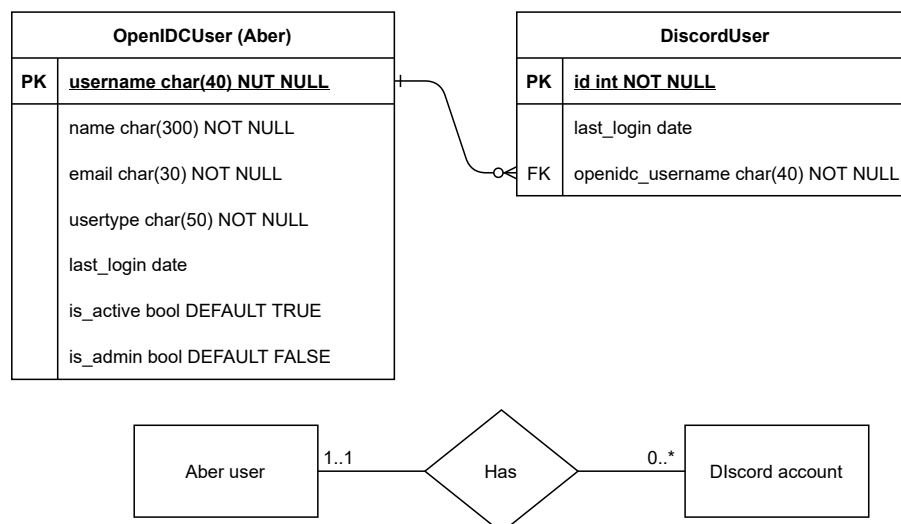


Figure 2.8: Entity relationship diagram for database

The **OpenIDCUser** is the main university account that the user authenticates with and is called so because it uses the OpenID Connect [5] system to authenticate users. The username, name and email are determined from the OpenID Connect response and are all unique. The usertype is also determined by the response and is used to determine if the user will be given administration privileges e.g. if usertype="staff" then admin=True or if usertype="student" then admin=False. Due to Django's custom user models it requires that the authenticated account has a `is_active` column that is kept true unless the account is marked as deactivated and cannot be used on the website.

The **DiscordUser** is simple and contains a unique Discord id that is a snowflake; a technology invented by twitter to keep id's unique. It also contains a `last_login` date to identify when the account was last authenticated with and lastly a `openidc_id` that is a foreign key from the **OpenIDCUser**'s id.

Below are examples of the records that are stored in the database tables.

Note: The column `last_login` in the table 2.1 contains the word `datetime` but should have a `datetime` object as seen in table 2.2. It has been removed so that the table fits nicely in this document.

<u>id</u>	username	name	email	usertype	last_login	is_active	is_admin
1	joa38	Joel Adams	joa38@aber.ac.uk	staff	datetime	t	t
2	jet39	Jenny Thyer	jet39@aber.ac.uk	student	datetime	t	f
3	maw86	Michael Antony West	maw86@aber.ac.uk	student	datetime	t	f

Table 2.1: Aberystwyth user table example

<u>id</u>	last_login	openidc.id*
727834884915331144	2021-02-18 16:43:47.067328+00	1
284352754321719296	2021-02-18 16:43:47.067328+00	1
246998944964542464	2021-02-04 11:14:40.057891+00	2
282248714955784192	2021-02-12 17:35:23.044226+00	3

Table 2.2: Discord user table example

2.5 User Interface

2.5.1 Website

Before beginning work on the website I designed some website mock-ups that I have used as reference when working on the design of the webpages. I created these mockups using the popular design application Figma [17] and they are located below.

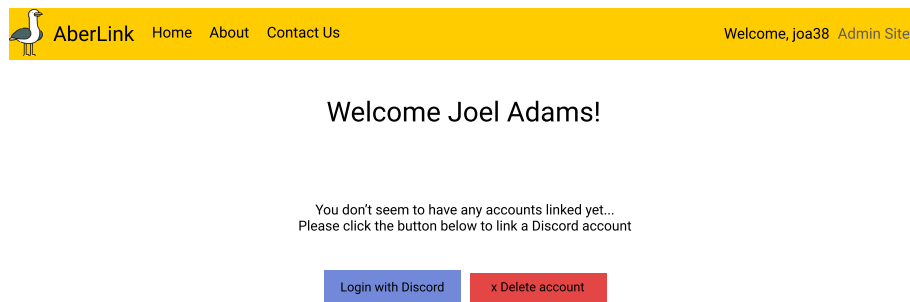


Figure 2.9: Website mock-up for 0 users

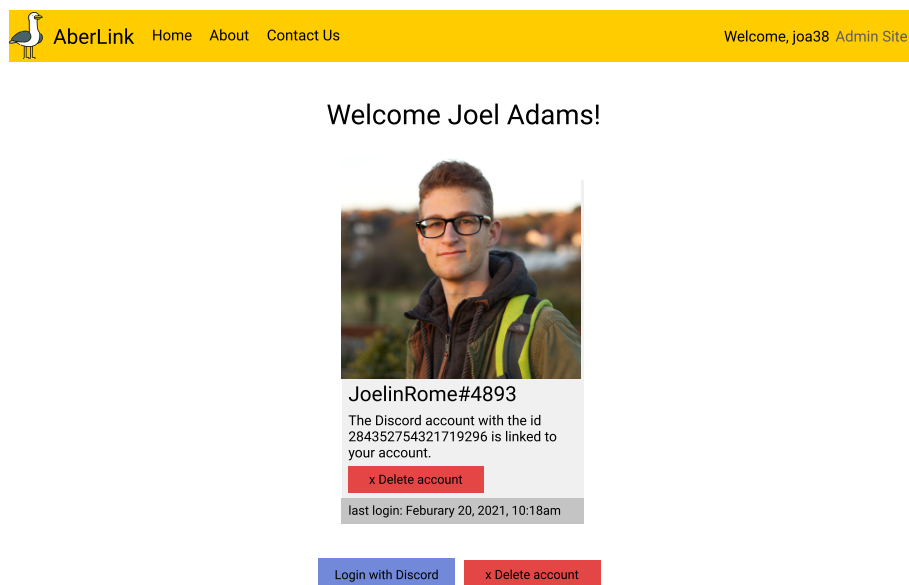


Figure 2.10: Website mock-up for 1 user

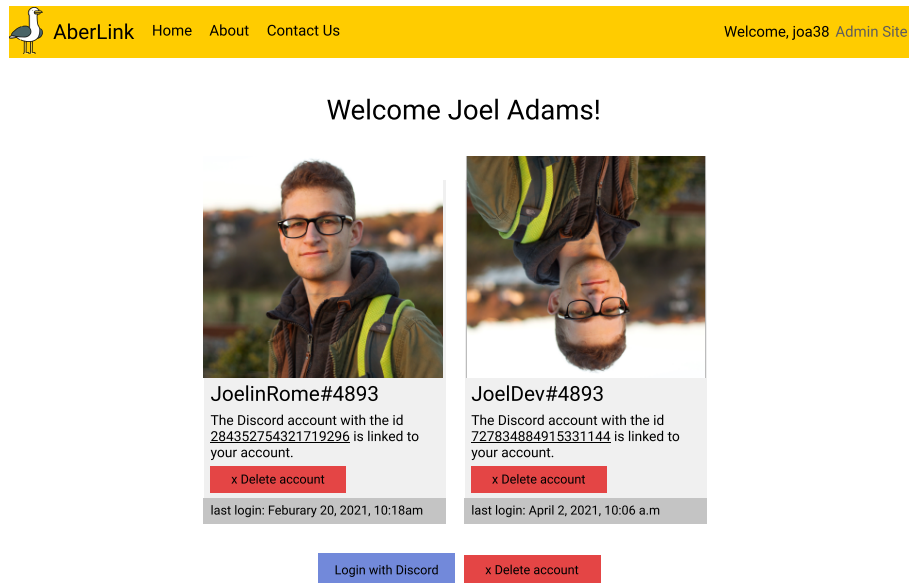


Figure 2.11: Website mock-up for 2 users

2.5.2 Discord bot

Discord provides its own interface for Discord bots and doesn't vary much from the regular interface provided to users. Discord bots however have the additional feature of being able to use Embeds which contain far better formatting than regular messages and are used extensively throughout AberLink. Below are some examples of the output of bot commands that I have created using Embeds.

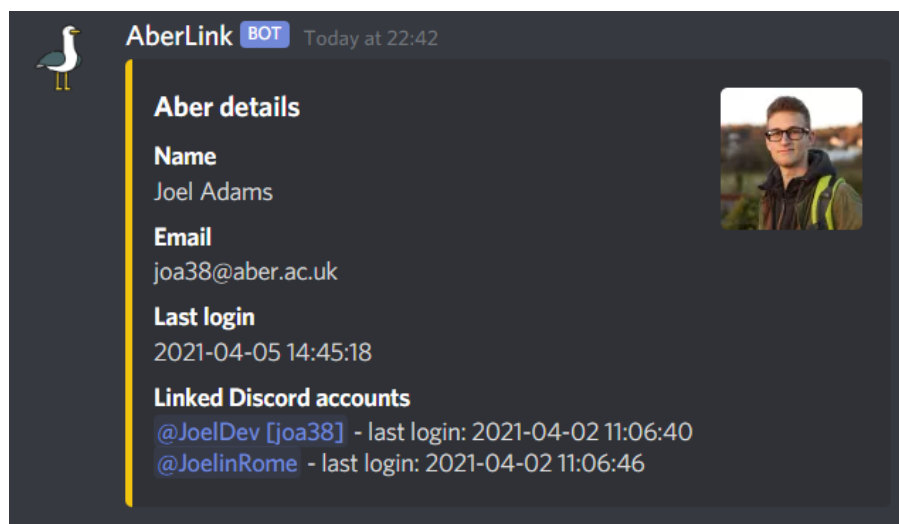


Figure 2.12: Discord Embed example for getting users information

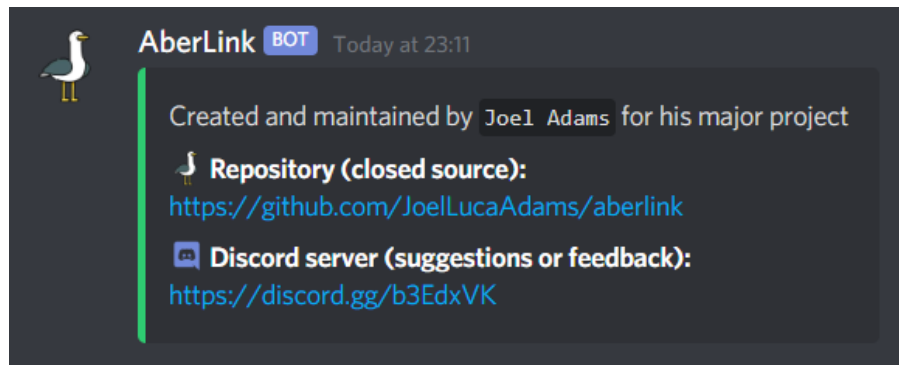


Figure 2.13: Discord Embed example for source command

2.6 Version Control System and IDE

Over the last year of working in Python I have used JetBrains Python IDE PyCharm [18] to code my Python projects but recently switched to Visual Studio Code [19] due to the amount of extensions it supports and the versatility it has. I chose this because it has extensions for querying a psql [6] database, compile Python programs, use virtual environments and write LaTeX documents like this one. It was also chosen as it allows you to remotely connect to a host over SSH which is very useful as I have developed the majority of this program on a Linux (Debian 10 Buster [2]) container. This last feature is not supported in PyCharm hence my choice of Visual Studio Code.

For Version Control system I have used the GitLab instance provided by the university to store my code. I have also added mirroring of the project to another repository called aberlink on my primary GitHub account in case of emergencies. I've included below an example of the commits that are made to the repository.

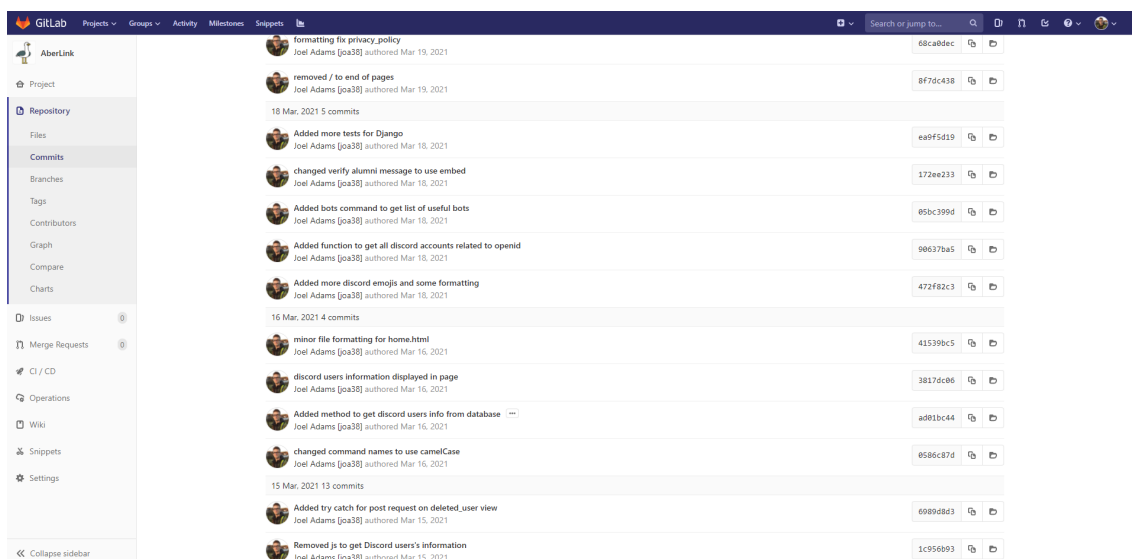


Figure 2.14: Example of GitLab repository commits

Chapter 3

Implementation

3.1 Code Implementation & Third-Party libraries

As seen below the Website and Discord bots have been split up into two separate sections because during implementation I decided to make the systems separate. This helps with code maintainability, readability and allows the administrator who sets up this project to deploy the website and bot services on different containers or networks.

3.1.1 Website Building

To create and run the website I've used a set of different libraries to perform certain tasks. Firstly I've used the Apache2 [3] web hosting framework to host the website along with the library libapache2-mod-auth-openidc to reroute all incoming traffic to OpenID Connect [5] for user authentication. On top of these I have used the linux tool certbot [20] to create a Let's Encrypt certificate for the website to enable HTTPS.

As discussed previously in this document in section 1.1.1 there were many libraries that were considered when deciding on the website framework. Django [4] was the framework of choice and it is open-source and free to use on personal projects (Django Trademark License Agreement). It is also very useful as it generates the majority of the code required to create and run a website so most of the code used will be picked up by the system for UAP. Apart from the general template I created a "Django Application" called login that contains the files and code which is used to run the website and it's respective services.

The website pages also use a third party library called bootstrap [21] that is used to generate responsive mobile-first CSS on the website.

3.1.2 Discord Bot

AberLink uses the Python library Discord.py [7] to make calls to and from the Discord API and to interact with the database it uses the Python library Psycopg2 [16]. Both of these

pieces of software are open-source and free to use in personal projects.

3.2 Unforeseen Issues

An issue that I encountered was with creating the custom user model in Django that would have been used to model the database. The documentation and videos I found online about implementing user models were rather cryptic and difficult to understand, however I eventually found a video explaining how to implement a good custom user model. Once this was completed I realised that Django isn't happy with modelling the primary key of a table using a char so I had to switch to using an int. This turned out to be a good idea as Aberystwyth university tends to recycle old emails so over the next 5 years there could be an issue where the database would try and create a new entry in the database with the same email.

3.3 Review Against Planned Requirements

Most of the planned requirements haven't changed during implementation however as discussed in the section above 3.2 the database model has changed slightly and the updated diagram is included below.

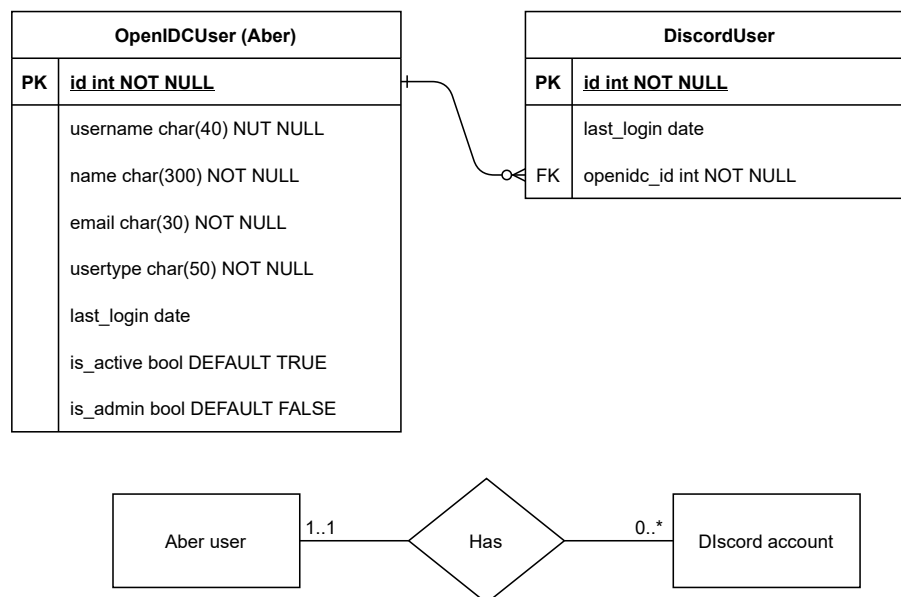


Figure 3.1: Updated Entity relationship diagram for database

The planned requirements also included a section called **Interface for lecturers and students on website** has also only been partially fulfilled. I have completed the second part of this task which was to create admin pages to view the connected staff and students however the first part where users can view what servers they are in is missing. This

was a design choice that I made because I realised that this would require that Discord servers are added to database and that they would have to be updated when any issues occurred. This was also bad because this list would have to be maintained and couldn't be automatically updated which would lead to it breaking down the line.

In the final section of the planned requirements **Further potential work** I have decided against integrating DemoHelper into AberLink as this would greatly increase it's complexity and make it much more difficult to maintain.

Chapter 4

Testing

As previously mentioned in section 1.3 I have used a Feature Driven Development (FDD) for this project to develop and test features as I developed the project. I worked in one week iteration windows to build and test the software as I went; this included all of the testing sections mentioned below apart from automated testing. This was completed at the end of the project instead.

The Discord.py Python framework doesn't contain any unit testing and there are no third party libraries which sufficiently perform this job. For the Django framework however there are some unit tests built into it so I have tried to use some of these and they are detailed further on this section. Django tests are used to cover both the website and the database.

4.1 Sample Data

I have considered using sample data for my project however deemed it to be impossible as due to the implementation of authentication. As stated previously the website uses the OpenID Connect [5] system to authenticate and log users in; as I don't have a dummy dataset that can authenticate with that system there is no other way for me to authenticate users and perform sample testing. The only workaround that I have found for this system has been to rely on automated and user testing.

4.2 Automated Testing

Automated testing has been difficult for this project as mentioned above. I was originally planning on including automated testing and building DevOps frameworks in Git, however as I ended up using the universities GitLab then there was no access to dedicated pipelines to build and run the code.

4.2.1 Unit Tests

4.2.1.1 Website & Database

Django provides some useful documentation (<https://docs.djangoproject.com/en/3.1/topics/testing/>) for creating unit tests and I have used these tests where I deemed most appropriate. All of the following tests can be found here `src\AberLinkAuthentication\login\tests.py`.

I've created two testing classes here; one called `TestUrls` and one called `TestModels`. These are responsible for testing that the urls are correctly found and that adding new users to the database works as intended. Included below is an example of each test:

```
def test_url_discord_redirect(self):
    url = reverse('Discord-response')
    self.assertEqual(resolve(url).func, views.
discord_oauth2_redirect)
```

Figure 4.1: Django URL render test

As seen above the test function simply gets the name of the url 'Discord-response' and then checks that it returns the correct Python function (view) to render that page

```
def test_model_openidc_user_staff(self):
    self.openidc_user2 = OpenIDCUser.objects.create(
        username="abc123",
        name="Bob Ross",
        email="abc123@aber.ac.uk",
        usertype="staff"
    )
    self.assertEqual(self.openidc_user2.username, "abc123")
    self.assertEqual(self.openidc_user2.name, "Bob Ross")
    self.assertEqual(self.openidc_user2.email,
"abc123@aber.ac.uk")
    self.assertEqual(self.openidc_user2.usertype, "staff")
    self.assertEqual(self.openidc_user2.is_admin, True)
```

Figure 4.2: Django database render test

This function creates a new user object and then checks to make sure that the user has been given the admin and has the usertype of staff.

4.2.1.2 Discord

Unit testing in `Discord.py` is impossible as there is no included framework for it and there are no external libraries capable of testing to check that it works.

4.2.2 Stress Testing

To test the websites capacity to deal with 100s of requests I used ApacheBenchmark [22] which is a built in Apache2 [3] library used to request a webpage and measure response time. To test this I temporarily disabled the OpenID Connect [5] authentication system so that HTML gets served to tge incoming requests. Due to the nature of my websites implementation I cannot query the main page because the user is not authenticated however I can query one of the subpages that doesn't require authentication. I used the following line of code to complete the request and uses 1000 requests with 100 concurrent threads to attempt to tank website performance:

```
ab -n 1000 -c 100 https://mmp-joa38.dcs.aber.ac.uk/privacy-policy
```

From this command I received the following output which shows that the website is capable at easily handling this many requests simultaneously. The longest response time for serving the requests was 642ms which is still adequately fast for the average user and shows that Django handles requests reasonable well. It can also serve on average 174 requests per second and with the department I can't see more than 300 requests happening at the exact same time.

```
Server Software:      Apache/2.4.38
Server Hostname:      mmp-joa38.dcs.aber.ac.uk
Server Port:          443
SSL/TLS Protocol:     TLSv1.2,ECDHE-RSA-AES256-GCM-SHA384,2048,256
Server Temp Key:      X25519 253 bits
TLS Server Name:      mmp-joa38.dcs.aber.ac.uk

Document Path:        /privacy-policy
Document Length:      12448 bytes

Concurrency Level:     100
Time taken for tests:   5.719 seconds
Complete requests:     1000
Failed requests:        0
Total transferred:     12734000 bytes
HTML transferred:      12448000 bytes
Requests per second:   174.86 [#/sec] (mean)
Time per request:      571.882 [ms] (mean)
Time per request:      5.719 [ms] (mean, across all concurrent requests)
Transfer rate:         2174.50 [Kbytes/sec] received

Connection Times (ms)
      min    mean[+/-sd] median    max
Connect:    9    422 105.1    458    547
Processing:  7     86  61.1     67    404
Waiting:    6     83  60.1     64    381
Total:      42    508  96.3    531    642
```

Percentage of the requests served within a certain time (ms)

50%	531
66%	550
75%	560
80%	565
90%	574
95%	583
98%	593
99%	602
100%	642 (longest request)

4.3 User Interface Testing

Testing the user interface for the AberLink is mostly straight forward as Discord has full control over the UI so I've had to do no testing on that front.

When I came to building the UI of the website I decided to rely on the CSS library Bootstrap [21] as it is built from the ground up with responsive design in mind. This meant that website would scale very nicely and uniformly depending on screen size. I have however conducted testing and scaling of all the website pages on mobile devices, iPads and laptops with screens ranging from 11in to 40in using Safari, Chrome, Edge and Firefox. I have also tested the website on HTML validation website <https://validator.w3.org/> and fixed the warnings and errors where possible.

4.4 User Testing

Throughout this project I have worked with the mindset that you develop small sections of code and then review what effect they have on the system. This helps to catch errors as it is much easier to backtrack through small sections of code.

For user testing I asked in the comp sci server if students could try and login to the website and break it. This turned out to work very well and helps to find a few errors that I had previously missed in the code.

4.5 Integration Testing

Following the FDD approach I ran integration tests on this project as it contained many working parts. I used the Figure 2.1 as a reference guide when performing these tests.

For the AberLink Discord bot I only had two components to test; the bot requests and responses from the database for the commands I had. I used some user tests to verify that all these commands worked and attempted to account for all possible scenarios.

Chapter 5

Evaluation

Examiners expect to find a section addressing questions such 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?

Other questions can be addressed as appropriate for a project.

The questions are an indication of issues you should consider. They are not intended as a specification of a list of sections.

The evaluation is regarded as an important part of the project report; 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 in the work and aspects of the work that could be improved. 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.

In the latter stages of the module, we will discuss the evaluation. That will probably be around week 9, although that differs each year.

Annotated Bibliography

- [1] Discord Inc., "Discord," accessed January 2021. [Online]. Available: <https://discord.com/>

Discord is a voice, video and text communication platform.

- [2] Debian, "Debian 10 (Buster)," accessed April 2021. [Online]. Available: <https://www.debian.org/releases/buster/>

Debian 10 (Buster) is a Linux Distro and is used throughout this project

- [3] Apache Software Foundation, "Apache 2.0," accessed April 2021. [Online]. Available: <https://httpd.apache.org/>

Apache2.0 is a open source linux package for website hosting

- [4] Django Software Foundation, "Django," accessed April 2021. [Online]. Available: <https://www.djangoproject.com/>

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design.

- [5] OpenID Foundation, "OpenID Connect," accessed April 2021. [Online]. Available: <https://openid.net/connect/>

OpenID connect is a identity layer on top of OAuth 2.0 and is responsible for simple user verification. It provides a basic profile of the connected user to the client in a REST-like manner.

- [6] PostgreSQL Global Development Group, "PostgreSQL," accessed April 2021. [Online]. Available: <https://www.postgresql.org/>

PostgreSQL is a powerful, open source object-relational database system with over 30 years of active development that has earned it a strong reputation for reliability, feature robustness, and performance.

- [7] Danny (Rapptz) and Various, "discord.py," accessed January 2021. [Online]. Available: <https://github.com/Rapptz/discord.py>

Python API library for Discord currently available through Github.

- [8] N. Snooke, "AberVerify," Sept. 2020, accessed April 2021. [Online]. Available: <https://github.com/NealSnooke/Aber-Verify-Discord-Bot>

Discord bot for Aberystwyth user verification in discord servers using email accounts (e.g. joa38@aber.ac.uk) and discord IDs (e.g. Joelin-Rome#4893). These accounts are then linked together using JSON objects. This project is currently available through Github.

- [9] —, “I am here,” Sept. 2020, accessed April 2021. [Online]. Available: <https://github.com/NealSnooke/IAmHere-Discord-Bot>

Discord bot for marking Aberystwyth students’ attendance during a practical. A student writes a message and the bot adds them to a list of students that have attended the practical and emails the lecturer at the end. This project is currently available through Github.

- [10] F5 Company Leadership, “NGINX,” accessed April 2021. [Online]. Available: <https://www.nginx.com/>

NGINX is a open source linux package for website hosting

- [11] Joel Adams, Nathan Williams, “Demohelper,” accessed April 2021. [Online]. Available: <https://github.com/AberDiscordBotsTeam/demoHelperBot>

DemoHelper is a Discord bot created to manage demonstrating online for students on Discord.

- [12] Blaine Cook, “OAuth 2.0,” accessed April 2021. [Online]. Available: <https://oauth.net/2/>

OAuth 2.0 is the industry-standard protocol for authorization.

- [13] Various authors, “OpenID Connect authentication module for Apache2,” accessed April 2021. [Online]. Available: <https://packages.debian.org/sid/libapache2-mod-auth-openidc>

mod-auth-openidc is a authentication library for Apache2 that uses OpenID Connect to authenticate users.

- [14] —, “Pipenv,” accessed April 2021. [Online]. Available: <https://pypi.org/project/pipenv/>

Pipenv is a package manager for Python that uses a Pipfile to store a list of dependencies that need to be installed. These dependencies are then installed inside of a virtual environment.

- [15] graffatcolmingov, Lukasa and nateprewitt, “Requests: HTTP for Humans™,” accessed April 2021. [Online]. Available: <https://docs.python-requests.org/en/master/>

Requests is a Python library used to make queries to API endpoints or websites. It is open source and is available via PyPI.

- [16] Various authors, “Pyscopg2,” accessed April 2021. [Online]. Available: <https://www.pyscopg.org/docs/>

Psycopg2 is a Python library responsible for connecting to PostgreSQL databases. It is built in C as a libpq wrapper.

- [17] Evan Wallace & Dylan Field, "Figma," accessed April 2021. [Online]. Available: <https://www.figma.com/>

Figma is a design website used to build application mockups.

- [18] JetBrains, "PyCharm," accessed April 2021. [Online]. Available: <https://www.jetbrains.com/pycharm/>

A professional IDE for Python built by JetBrains.

- [19] Microsoft, "Visual Studio Code," accessed April 2021. [Online]. Available: <https://code.visualstudio.com/>

Free, open source code editor with a wide range of extensions to support compiling code in many languages.

- [20] Electronic Frontier Foundation (EFF), "Certbot," accessed April 2021. [Online]. Available: <https://certbot.eff.org/>

Certbot is an open-source tool used to create a Let's Encrypt certificate to allow a website to use HTTPS.

- [21] Various authors on Github, "Bootstrap 5.0," accessed April 2021. [Online]. Available: <https://getbootstrap.com/>

Bootstrap is a responsive plugin to help with styling webpages using HTML, CSS and JavaScript plugins.

- [22] Apache Software Foundation, "Apache HTTP server benchmarking tool (ApacheBench)," accessed April 2021. [Online]. Available: <https://httpd.apache.org/docs/current/programs/ab.html>

AB is a tool used to stress test a website's capacity to deal with 1000s of requests at once.

- [23] stub, "pytz," accessed April 2021. [Online]. Available: <https://pypi.org/project/pytz/>

Used for accurate, cross-platform timezone calculation. It is open source and is available via PyPI.

- [24] aalbrecht, "sqlparse," accessed April 2021. [Online]. Available: <https://pypi.org/project/pytz/>

Python library for parsing, splitting and formatting Non-validationg SQL libraries. It is open source and is available via PyPI.

- [25] Various authors, "ASGI," accessed April 2021. [Online]. Available: <https://asgi.readthedocs.io/en/latest/>

Asynchronous Server Gateway Interface is the successor to WSGI used as an interface for Python web servers and frameworks.

- [26] Internet Security Research Group (ISRG), “Let’s Encrypt,” accessed April 2021. [Online]. Available: <https://letsencrypt.org/>

A nonprofit Certificate Authority providing TLS certificates to HTTPS.

- [27] Ashald and ncbi, “libpq-dev,” accessed April 2021. [Online]. Available: <https://packages.debian.org/sid/libpq-dev>

Debian package required for building C header files for the psycopg2 library.

- [28] Guido van Rossum, “libpq-dev,” accessed April 2021. [Online]. Available: <https://packages.debian.org/buster/python-dev>

Debian package required for building Python 2 modules such as psycopg2.

Appendices

Appendix A

Third-Party Code and Libraries

I declare that all the following libraries are unmodified.

- **Pipenv** [14] - This is a Python package manager that is used throughout this project to create and maintain virtual environments. It uses Pipfiles (text files) to store a list of dependencies that are then downloaded and stored in a virtual environment. In this project I have two Pipfiles; one for the website (`src\AberLinkAuthentication`) and one for the Discord bot (`src\AberLinkDiscord`). This library uses the MIT license so is free to use.
- **Django** [4] - This is a Python web framework used for building websites. It is used to display the webpages, sign users in and help users and staff manage their data. This library is open source and provided by the Django Software Foundation.
- **Apache2** [3] - This Debian [2] package is responsible for hosting a HTTP web server. In this case it hosts the Django [4] framework so it can be accessed on the web. This library is released under the Apache License, Version 2.0.
- **libapache2-mod-auth-openidc** [13] - Debian [2] package used to authenticate users when they first connect to the site. It authenticates users against their Aberystwyth University account and then creates a session where they can access content on the website. This library is released under the GNU Free Documentation License
- **certbot** [20] - This package was only briefly required to setup HTTPS on the website using Let's Encrypt [26]. Once SSL certificates have been made cerbot then runs a cronjob in the background to update it. This library is licensed under the Apache 2.0 license
- **psycopg2** [16] - This library is used for talking to PostgreSQL databases in Python. The version used here is called psycopg2-binary and it requires two additional dependencies to be built libpq-dev [27] and python-dev [28]. This library is licensed under the GNU Lesser General Public License.
- **Discord.py** [7] - This Python library is used to create and write the Discord bot. It is basically an interface that allows me to write bots without making API requests straight to Discord. This library is licensed under the MIT License.

- **Requests** [15] - It is used to make HTTP requests to the web. In this project it is used in the website to get information from the Discord login and in the Discord bot it is used to send and receive data from the attendance API endpoint <https://integration.aber.ac.uk/joa38/submit.php>. This library is licensed under the Apache Software License (Apache 2.0)
- **Pytz, SQLParse & ASGIref** [23] [24] [25] - These are dependencies required to work with Django. pytz helps with managing timezones and timezone data. SQLParse handles SQL queries, formatting and splitting. ASGIref is an interface used to talk between Apache2 [3] and Django [4] for web server hosting and frameworks. These libraries are all licensed under the MIT License.

Appendix B

Ethics Submission

Application Number: **18847**

Assessment Details

AU Status

Undergraduate or PG Taught

Your aber.ac.uk email address

joa38@aber.ac.uk

Full Name

Joel Adams

Please enter the name of the person responsible for reviewing your assessment.

Neil Taylor

Please enter the aber.ac.uk email address of the person responsible for reviewing your assessment

nst@aber.ac.uk

Supervisor or Institute Director of Research Department

CS

Module code (Only enter if you have been asked to do so)

CS39440

Proposed Study Title

Discord bot for verification and attendance linked to OpenID Connect

Proposed Start Date

25 January 2021

Proposed Completion Date

1st June 2021

Are you conducting a quantitative or qualitative research project?

Mixed Methods

Does your research require external ethical approval under the Health Research Authority?

No

Does your research involve animals?

No

Does your research involve human participants?

Yes

Are you completing this form for your own research?

Yes

Does your research involve human participants?

Yes

Institute

IMPACS

Please provide a brief summary of your project (150 word max)

I'm creating a system that links student's discord accounts to their uni accounts. This is done through a website that allows users to log in with their Discord accounts and their student account using Discords login system and OpenID Connect respectively so no password information is ever stored. This information is then saved in a backend postgresql database stored in the Department of Computer Science's servers so it is protected from remote attacks. A discord bot is then used on the front end to verify if a student has a verified Discord account and what student account that is linked to. It is also responsible for marking students' attendance in practicals.

I can confirm that the study does not involve vulnerable participants including participants under the age of 18, those with learning/communication or associated difficulties or those that are otherwise unable to provide informed consent?

Yes

I can confirm that the participants will not be asked to take part in the study without their consent or knowledge at the time and participants will be fully informed of the purpose of the research (including what data will be gathered and how it shall be used during and after the study). Participants will also be given time to consider whether they wish to take part in the study and be given the right to withdraw at any given time.

Yes

I can confirm that there is no risk that the nature of the research topic might lead to disclosures from the participant concerning their own involvement in illegal activities or other activities that represent a risk to themselves or others (e.g. sexual activity, drug use or professional misconduct).

Yes

I can confirm that the study will not induce stress, anxiety, lead to humiliation or cause harm or any other negative consequences beyond the risks encountered in the participant's day-to-day lives.

Yes

Please include any further relevant information for this section here:

Where appropriate, do you have consent for the publication, reproduction or use of any unpublished material?

Not applicable

Will appropriate measures be put in place for the secure and confidential storage of data?

Yes

Does the research pose more than minimal and predictable risk to the researcher?

Not applicable

Will you be travelling, as a foreign national, in to any areas that the UK Foreign and Commonwealth Office advise against travel to?

No

Please include any further relevant information for this section here:

Is your research study related to COVID-19?

No

If you are to be working alone with vulnerable people or children, you may need a DBS (CRB) check. Tick to confirm that you will ensure you comply with this requirement should you identify that you require one.

Yes

Declaration: Please tick to confirm that you have completed this form to the best of your knowledge and that you will inform your department should the proposal significantly change.

Yes

Please include any further relevant information for this section here:

N/A

Appendix C

Code Examples

For some projects, it might be relevant to include some code extracts in an appendix. You are not expected to put all of your code here - the correct place for all of your code is in the technical submission that is made in addition to the Project Report. However, if there are some notable aspects of the code that you discuss, including that in an appendix might be useful to make it easier for your readers to access.

As a general guide, if you are discussing short extracts of code then you are advised to include such code in the body of the report. If there is a longer extract that is relevant, then you might include it as shown in the following section.

Only include code in the appendix if that code is discussed and referred to in the body of the report.

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 [?].

```
#define IM1 2147483563
#define IM2 2147483399
#define AM (1.0/IM1)
#define IMM1 (IM1-1)
#define IA1 40014
#define IA2 40692
#define IQ1 53668
#define IQ2 52774
#define IR1 12211
#define IR2 3791
#define NTAB 32
#define NDIV (1+IMM1/NTAB)
#define EPS 1.2e-7
```

```

#define RNMX (1.0 - EPS)

double ran2(long *idum)
{
    /*-----*/
    /* Minimum Standard Random Number Generator      */
    /* Taken from Numerical recipies in C             */
    /* Based on Park and Miller with Bays Durham Shuffle */
    /* Coupled Schrage methods for extra periodicity   */
    /* Always call with negative number to initialise  */
    /*-----*/

    int j;
    long k;
    static long idum2=123456789;
    static long iy=0;
    static long iv[NTAB];
    double temp;

    if (*idum <=0)
    {
        if (-(*idum) < 1)
        {
            *idum = 1;
        }else
        {
            *idum = -(*idum);
        }
        idum2=(*idum);
        for (j=NTAB+7; j>=0; j--)
        {
            k = (*idum)/IQ1;
            *idum = IA1 *(*idum-k*IQ1) - IR1*k;
            if (*idum < 0)
            {
                *idum += IM1;
            }
            if (j < NTAB)
            {
                iv[j] = *idum;
            }
        }
        iy = iv[0];
    }
    k = (*idum)/IQ1;
    *idum = IA1*(*idum-k*IQ1) - IR1*k;
    if (*idum < 0)
    {

```

```
        *idum += IM1;
    }
    k = (idum2)/IQ2;
    idum2 = IA2*(idum2-k*IQ2) - IR2*k;
    if (idum2 < 0)
    {
        idum2 += IM2;
    }
    j = iy/NDIV;
    iy=iv[j] - idum2;
    iv[j] = *idum;
    if (iy < 1)
    {
        iy += IMM1;
    }
    if ((temp=AM*iy) > RNMx)
    {
        return RNMx;
    }else
    {
        return temp;
    }
}
```