Project Plan

Group 9785-23-06

ITS Capstone Project

E-COMpliance:

Detecting Non-Compliant Therapeutic Goods on E-commerce Websites

Audience: *Project Sponsors*

Purpose: *This document outlines the project we will be completing for the ITS capstone unit.*

Authors:	**Version:**	**Date:**
Dylan Henderson Swikriti Chapai Parbat Khadka Matthew Borowski	1.3	10/03/2023

Approval by Team Members:

Parbat Khadka	Dylan Henderson
Matthew Borowski	Swikriti Chapai

CPA 1/2005 Group: #, Group Name> Group Logo>

TABLE OF CONTENTS

1 INTRODUCTION

The detection of non-compliance in the online sales of therapeutic products is a significant concern for governing agencies worldwide. Although numerous of these e-commerce websites operate lawfully and in accordance with the rules established by their individual nations, there are many products on these platforms and non-compliant goods can slip through the cracks. This problem occurs because products can be approved for use in other countries (like the US) but contain ingredients that are still deemed poisonous by the Therapeutic Goods Administration (TGA) of the Australian government. Only products that meet Australia's criteria for therapeutic goods should be available for sale in Australia. Products that do meet these criteria are able to be identified by an ARTG (Australian Register of Therapeutics Goods) product code which is printed on the product's label. Non-compliant products can cause serious harm or endanger public health.

A therapeutic goods fall into one of these three categories:

- Medicines - including prescription, over the counter and complementary medicines, such as paracetamol and echinacea.
- Biologicals - something made from or containing human cells or tissues, such as human stem cells or skin.
- Medical devices - including instruments, implants and appliances, such as pacemakers and sterile bandages.

2 PROJECT NAME

E-COMpliance: Detecting Non-Compliant Therapeutic Goods on E-commerce Websites

3 PROJECT DESCRIPTION

The current situation is that there is no efficient system in place to detect non-compliant goods, which can lead to serious health consequences for consumers and the current solution is completed manually by members in the TGA, which is time-consuming and costly.

This project aims to solve the issue by developing a system that can automatically collect products from Australian e-commerce websites and detect which products are non-compliant under TGA criteria.

The system will work by web scraping e-commerce websites and collecting data on product advertisements, product descriptions, images, and ingredients and any other relevant information on the product page. The collected product data will then be analysed for compliance with TGA regulations, including:

- Poisonous ingredients check.
- Therapeutic claims intentions.
- ARTG number register.

The system will then check to see if any of the criteria above are not met to a satisfactory standard by the product or its associated product advertisement page. The output of the system will include metrics on the compliance of the e-commerce websites as well as a .csv file of non-compliant products, and a final executable. We believe that this system will provide an efficient and effective solution to detect non-compliant therapeutic goods on e-commerce Platforms in Australia.

4 PROJECT PERSONNEL

4.1 Sponsors

Name	Type	Contact Information
Dat Tran	Internal	Dat.Tran@canberra.edu.au
Shuangzhe Liu	Internal	Shuangzhe.Liu@canberra.edu.au
Julie Doan	External	Julie.doan@health.gov.au

4.2 Project manager and key team members

Name	Role	Contact Information
Matthew Borowski	Project Manager	u3214653@uni.canberra.edu.au
Parbat Khadka	Team Member	u3229244@uni.canberra.edu.au
Swikriti Chapai	Team Member	u3234545@uni.canberra.edu.au
Dylan Henderson	Team Member	u3203758@uni.canberra.edu.au

5 SCOPE

5.1 Business rules

The system will only operate within the legal framework of the Australian Therapeutic Goods Administration (TGA) and adhere to its guidelines and regulations.

The system will only analyze products that are sold on specific e-commerce websites operating within Australia.

The system will only analyse products that are written in english and no other languages.

The system will only analyze therapeutic goods, as defined by the TGA, and not other types of products sold on e-commerce websites.

The system will not perform any actions that may interfere with the security of e-commerce websites.

The system will only collect and store data required for analysis and will not store any personal or sensitive information.

The system will provide accurate and reliable results, with a low false positive rate, to ensure that compliant products are not flagged as non-compliant.

The system will provide a clear and easy-to-understand output, with reasoning for products flagged as non-compliant.

The system should be designed to be easily updated to ensure compliance with any changes to the TGA guidelines and regulations, as well as any changes to e-commerce websites' functionalities.

5.2 MoSCoW Prioritisation

Must Have:

- Web scraping of two Australian e-commerce websites.
- collect product information from e-commerce pages.
- check if it contains ingredients on the ARTG poison registry, if the product has an ARTG number and if the product sales page makes any unsubstantiated health claims.

Should Have:

- System to read a list of keywords to search the e-commerce sites instead of being hardcoded into the program.

Could Have:

- Have compatibility for the system to be run on an apple computer.

- A dashboard created to view the results in a more interactive format.

Will Not Have:

- Any form of reporting system implemented, eg. the system won't automatically report product listings to the respective e-commerce websites.
- Support for languages other than English.

5.3 Assumptions

We are assuming that we can make an educated guess at whether a product is compliant or not based on some simple factors so we can have a metric to check against so we can determine how our system is working.

We assume that products on Australian e-commerce websites will be in English and not in other languages.

5.4 WBS (Work Breakdown Structure)

1. Web Scrape
    a. e-commerce website
        i. input search term into search bar.
        ii. collect results from the page.
        iii. go to the next page (and when it gets to the last page).
        iv. repeat for the next search word.
    b. collect product data.
        i. collect text components.
        ii. collect image information.
    c. bulk output results for further processing
    d. adapt components in part a to work for alternative e-commerce website
    e. Testing
2. Data Analysis
    f. Poison check
        i. format poison list from ARTG website
        ii. parse product ingredient list
        iii. compare ingredient to poison list
        iv. flag problem as poison or not
        v. Output decision of poisonous or not
    g. Therapeutic claim
        i. parse information of product description
        ii. determine if it contains therapeutic claims.
        iii. Output decision on if claims are compliant or not
    h. ARTG number
        i. check product information for ARTG number
        ii. check ARTG number online
        iii. Output if its got a number or not
    i. Compliance Check
        i. check poison status
        ii. check therapeutic claim status
        iii. check ARTG

iv. Output a final decision on compliance

j. Testing

3. Output and Delivery

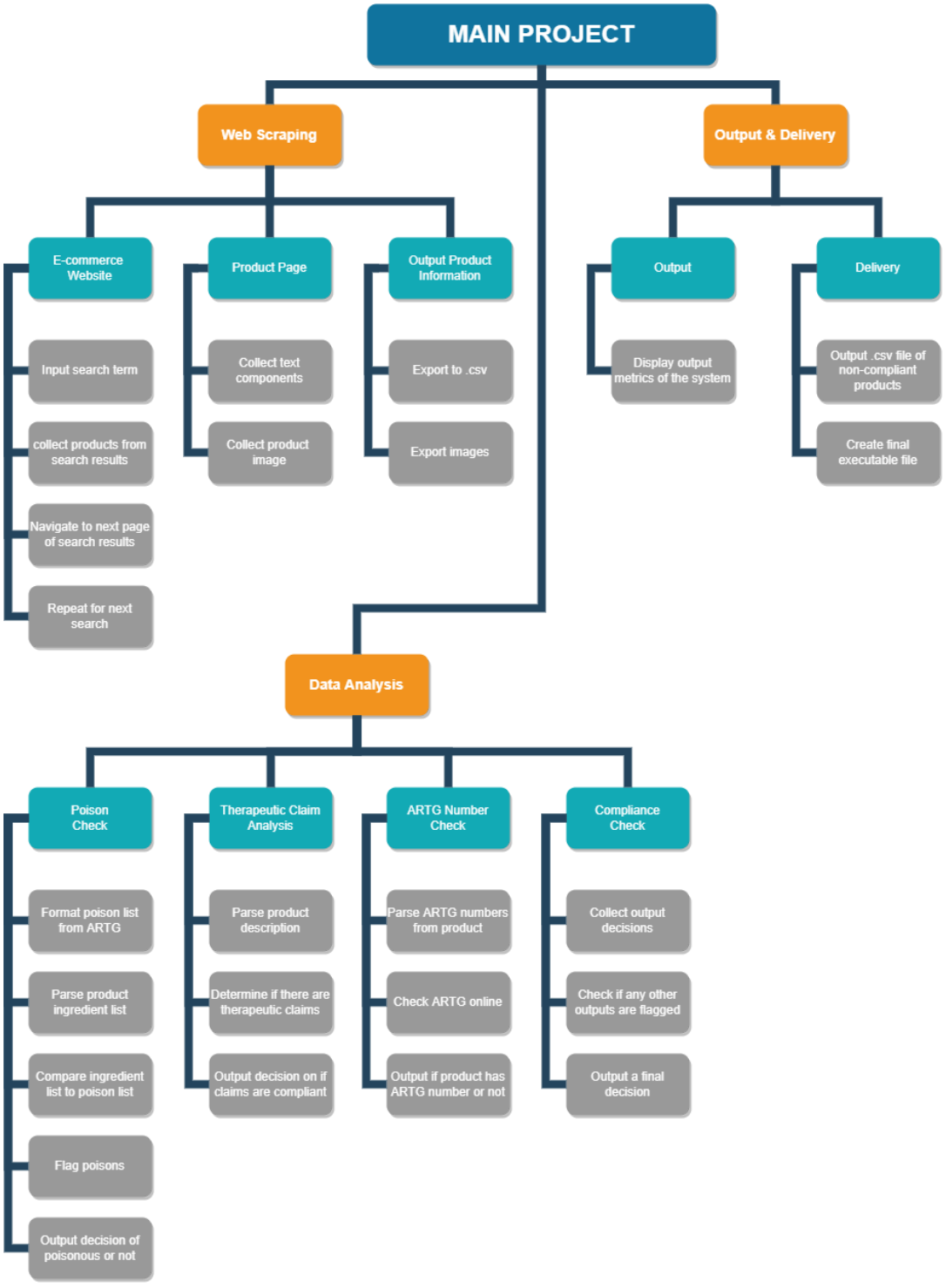
k. Output

i. Display metrics of system

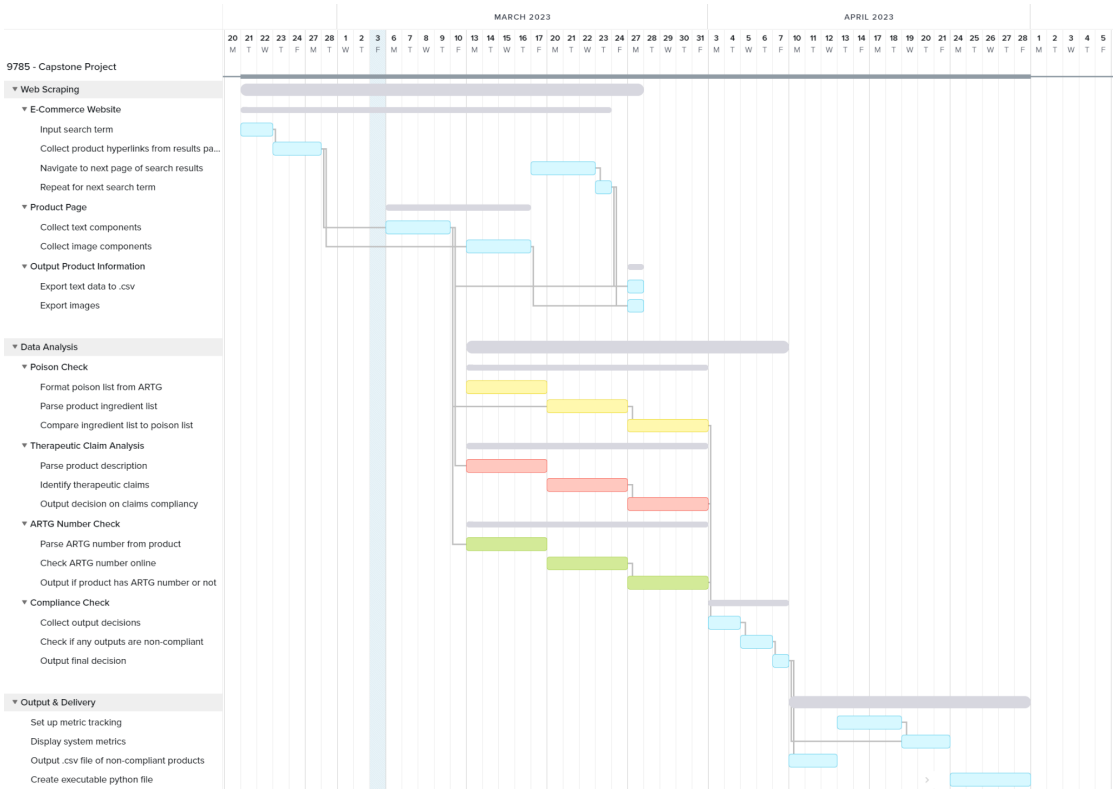
l. Delivery

i. Output csv file of just non-compliant products

ii. Create final executable



5.5 Gantt Chart



6 PROJECT DELIVERABLES

6.1 Deliverable List

- An automated tool for identifying and flagging non-compliant therapeutic goods listings on e-commerce websites.
- A .csv file of non-compliant therapeutic goods identified during the monitoring process, which can be used to track compliance over time and inform future compliance monitoring efforts.
- A detailed report on the non-compliant therapeutic goods identified on e-commerce websites, including the number and type of non-compliant listings, the frequency of non-compliance, and the reasons for non-compliance.
- Maintenance and Support Plan : A maintenance and support plan details the continuing assistance and upkeep procedures necessary to maintain the system current and working efficiently. This has to include regular monitoring and reporting, as well as protocols for dealing with issues and upgrades.
- Project Report: A summary of the project's goals, deliverables, and deliverables along with any difficulties encountered along the way and suggestions for future work.
- An output .csv file MUST contain:

- Manufacturer
- Product Name
- ASIN
- Seller name
- Searching keywords (they are all line up with search keywords such as: Weight Loss, Weight Loss Supplements, weight loss Vitamins, and Diet Supplements)
- ARTG (Yes/No)
- Therapeutic Claim (Yes/No)
- Contained Ingredients – leave it blank if not available on the list
- Image attached on the list
- Collected date
- Compliant/non-compliant predictive outcome

6.2 Key Performance Indicators

- The system must accurately determine what products are therapeutic goods from search results with a greater than 95% success rate.
- The system must correctly identify non-compliant products >90% with less than 5% false positive rate.

7 ROLES AND RESPONSIBILITIES

Project Manager
The project manager plays a critical role in completing the project within the timetable and financial constraints, as well as satisfying its objectives. Managing relationships with contributors and stakeholders, as well as ensuring that initiatives receive adequate finance.

Matthew Borowski has been appointed as the project manager in charge of our project.
Responsibilities:

- Making a project plan.
- Ensuring deliverables are met for each milestone.
- Choosing the project management methodology
- Task delegation based on team members strengths.
- Creating a project schedule and identifying each step
- Updating Gantt chart as we meet or fail to meet task deadlines.
- Interacting with stakeholders
- Scheduling sponsor and mentor meetings

Team Member
Our entire team is equally liable for every step of the project. Once the project manager allocates tasks, the team members adhere to it. Team members should aim to adopt the right techniques and processes that will benefit the project.

Responsibilities:

- Contribution to the overall project goals.
- Completing individual allocated tasks.
- Learn any skills that will help complete a task.
- Communicate effectively with the project manager if tasks are taking too long or they become stuck and need guidance.

Project Sponsors
The project sponsor collaborates closely with the project manager and validates the project's goals. Sponsor's may also assist in resolving conceptual issues, as well as signing off at key milestones in the project.

Responsibilities:
- Making critical project strategic decisions.
- Communicating any changes in deliverables.
- Signing off on milestone deliverables.

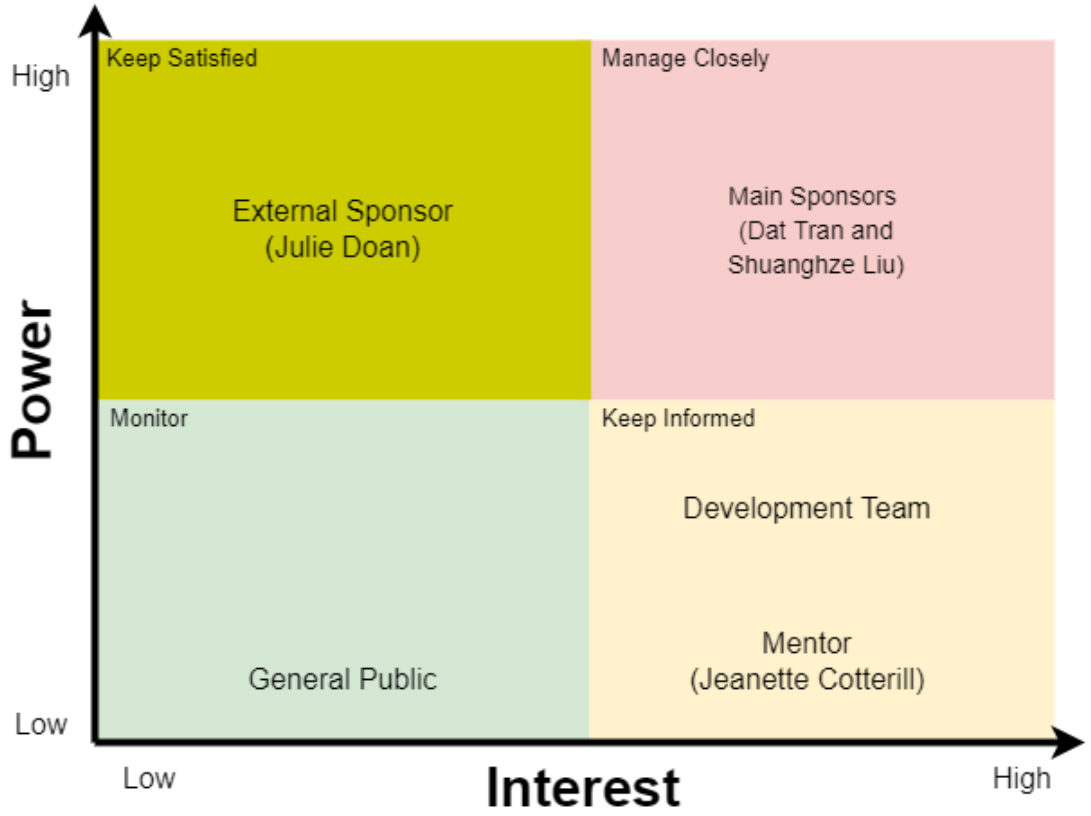
8 PROJECT MILESTONES

Milestone 1
Completing the web scraping from the two e-commerce websites is the first major milestone. This milestone means that all the information that we need to collect is all in one place and exported to a .csv file, ready to be analysed.

Milestone 2
Data analysis of the four major checks have been completed, those being: ARTG approved product number, determine if therapeutic claims are genuine, products contain poisonous ingredients and a final compliant or non-compliant decision. This milestone represents that a majority of the programming is completed and the product is nearly ready for delivery. At this milestone, the team can also look at possible extra features and determine if time permits.

Milestone 3
Project delivery is the final milestone, this represents the end of the project. The project reaches this milestone once a final output is created and the system can be executed by running a python script.

9 STAKEHOLDER ANALYSIS



10 TECHNICAL APPROACH

10.1 Collecting Data

Scraping the data from e-commerce websites will be difficult as not all e-commerce websites are formatted the same, so different approaches need to be developed to collect the information from each one. This problem is compounded when you need to ensure all the information conforms to the same headings.

10.2 Data Integrity

Keeping the generated .csv files of product information in a centralised location like GitHub can be used for the initial testing and development of the system as having all programmers have access to the same information to use for testing different aspects of the system.

10.3 Multiple Programmers

Using multiple python scripts for development means all programmers can work concurrently on their respective tasks without worrying about any incompatibilities or creating bugs between each other's code. This approach also helps debugging when all parts are called from the main script.

11 RESOURCES

11.1 Hardware

Name	Description
Computer	A computer with at least a 2.4GHz processor running Windows 10 or newer.
Internet Connection	Stable internet connection to access e-commerce etc. websites to scrape information from. It is also used for our version control and documentation needs.

11.2 Software

Name	Description
Python IDE	We have chosen to use vs code for the majority of programming tasks but other python IDEs would suffice.
GitHub	GitHub is an online version control platform that allows the team to keep our codebase up to date and all in one easily accessible place. This will only be used for the programming side of the project.

12 MANAGEMENT APPROACH

12.1 Project Management Style

It was decided that this project would be created using an agile project management approach. This methodology was chosen as it offers the greatest amount of flexibility to task allocation and scheduling can be changed around after “sprints” to better utilise the time and resources available to the team. Tasks that are taking longer than anticipated can be put back into the “backlog” and other important tasks can be put into the next sprint. These short iterative sprints really lend themselves to the changing landscape of prototyping and small scale development that this project is. A strong focus on documentation is required to ensure there is no wasted time/resources in sprints.

12.2 Quality Management

Project Quality Measurements		
Key measurement	**Acceptable level definition**	**Comments**
Project plan completion	Project plan is completed and is well defined	Early draft checked by sponsors and mentors. Final draft is proofread by multiple sources to confirm correct flow.
Meeting optimisation	Meetings will take place with all required stakeholders. All topics will be discussed and all questions clarified.	All meeting times, questions, and topics will be established in advance. Additionally, all meetings will be documented for missing team members and future reference through Trello.
Web scraping	Web scraping will be able to scrap 99% of required information from Amazon and Ebay	Required information will be ingredients, ARTG number, and screenshot of product.
Poison Checking	All ingredients scraped will automatically get checked for poisons listed on the ARTG poison registry. At least 90% of products with illegal ingredients will be denied.	Any false positives or negatives will be identified from the list of products given, and the success rate will be determined.
Therapeutic claim analysis	All therapeutic claims will be identified from product description. At least 90% of false claims will be denied.	Any false positives or negatives will be identified from the list of products given, and the success rate will be determined.
ARTG number check	At least 90% of ARTG number checks will be able to correctly identify ARTG numbers.	Any false positives or negatives will be identified from the list of products given, and the success rate will be determined.
Compliance check	Overall success rate of at least 70% for non-compliance detection	Stratified sampling by taking from non-compliant and compliant products and manually checking if products are correctly categorised.

Key Quality Responsibilities		
Activity	**Team Member**	**Comments**
Release approval	Matthew Borowski	Will determine whether the final product is ready for handover to sponsors.
Acceptance criteria	Dylan Henderson	Will determine the acceptance criteria for each measurement.
Web Scraping tests	Matthew Borowski	Responsible for carrying out web scraping tests and ensuring that the function adheres to the acceptance criteria
Poison checking tests	Swikriti Chapai	Responsible for carrying out poison check tests and ensuring that the function adheres to the acceptance criteria
Therapeutic claim tests	Dylan Henderson	Responsible for carrying out therapeutic claim tests and ensuring that the function adheres to the acceptance criteria
ARTG number check tests	Parbat Khadka	Responsible for carrying out ARTG number tests and ensuring that the function adheres to the acceptance criteria
Compliance check	Matthew Borowski	Will carry out compliance check tests to ensure the final product

		adheres to the acceptance criteria.
Implementation Checklist		
TRUE	Is there a shared understanding of what project management is and how it will be achieved?	
TRUE	Are responsibilities clearly defined and divided up in the context of quality management?	
TRUE	Has acceptance criteria been created for all required functions?	
TRUE	Is the project being managed through project management software?	

12.3 Risk Management

A risk management plan has been implemented to identify, evaluate, and control any potential risks that are associated with the project, and to ensure that all stakeholders involved are satisfied. Our risk management plan looks at cost/time, schedule, and performance risks that may hinder the progression of the project. It addresses these risks by implementing strategies to prioritize and mitigate any potential problems.

Risk	Event description	Probability/ Likelihood	Impact	Rating	Mitigation Strategy
The project taking too long to complete	The project runs the risk of getting behind schedule if there are any issues along the way that aren't properly met. As a result, this could lead to additional cost if the project was to be delayed.	Low	High	High	We address this by clearly defining the project scope before beginning, and making sure that the entirety of the project is capable of being completed within the set timeframe.
Unclear or unrealistic expectations.	An unclear understanding between all parties on what is involved in the project. Additionally, if unrealistic expectations are set then project goals may become unobtainable or may be completed at a lower quality standard.	Low	Medium	Low	To mitigate either of these risks, we will ensure that careful planning of estimation of costs, time to complete milestones, and communication is fully established.
Missed deadlines and deliverables.	Deliverables are missed before the required deadline. This can put the whole project behind schedule and potentially not completing the project on time.	Medium	High	High	To mitigate the chance of any missed deadlines and deliverables we will implement prioritisation of tasks. Tasks that are crucial to the project will take the majority of resources and will be given the highest priority, while those of lowest priority will be given the least number of resources. Additionally, dependencies will be determined, and tasks will be completed synchronously wherever possible.

Risk	Event description	Probability/ Likelihood	Impact	Rating	Mitigation Strategy
Poor communication.	Poor communication between team members and stakeholders. Team members will be unable to perform optimally without a clear direction and end goal.	Low	High	High	To mitigate the risk of poor communication, all stakeholders will be made aware of the best communication channels. Team members will give weekly updates on the progress of their allocated parts to the project manager, and sponsors will be made aware of progress regularly. This will ensure that all employees are on the same page and there is no miscommunication.
The final product not being what it was intended to be.	The final product is not what the sponsors intended it to look like. The means additional costs and time to rectify the problem, or the project being regarded as a complete failure	Low	Medium	Medium	To mitigate this risk, we will clearly define the requirements for the project, and will seek clarification from the sponsors before commencing. Additionally, we will continually give updates to the sponsors on the progress of the project, this will be through fortnightly meetings and will include demonstrations wherever possible. Any clarifications will be dealt with during these meetings to ensure team members are completing their tasks according to the requirements.
Scope creep.	Additional requirements added late into project development. This means more time may be required, or the final product may be of lower quality standard.	Medium	Medium	Medium	To mitigate this risk a rigorous change control process will be established. The change control process will require signing off and agreement from multiple stakeholders.
Final product unable to perform required tasks.	The final product is unable to perform required functions/tasks.	Low	High	Medium	To mitigate this risk we will ensure that all “must have” parts are fully realised before beginning and completed before the end of the final deadline. Additionally, rigorous testing of the product will be done regularly so that everything works correctly during the final demonstration.
Corrupted, damaged, or missing work.	Involves any work on the project becoming corrupted, damaged, or going missing. This would cause major setbacks, causing time delays and additional costs.	Low	Low	Low	To mitigate this risk, all files will follow a backup and recovery plan incase of a disaster. Work is to be backed up to the cloud regularly and autosave wherever possible.

Risk	Event description	Probability/ Likelihood	Impact	Rating	Mitigation Strategy
Team member absences	Team members are unable to perform their assigned tasks due to major sickness, injury, or departure from the team.	Low	High	Medium	To mitigate this risk extra time has been allocated to the project schedule in the event a team member is unable to continue their tasks. Additionally, the communication plan will allow other team members to be fully aware of what each person is doing and take over if need be.

12.4 Communication Management

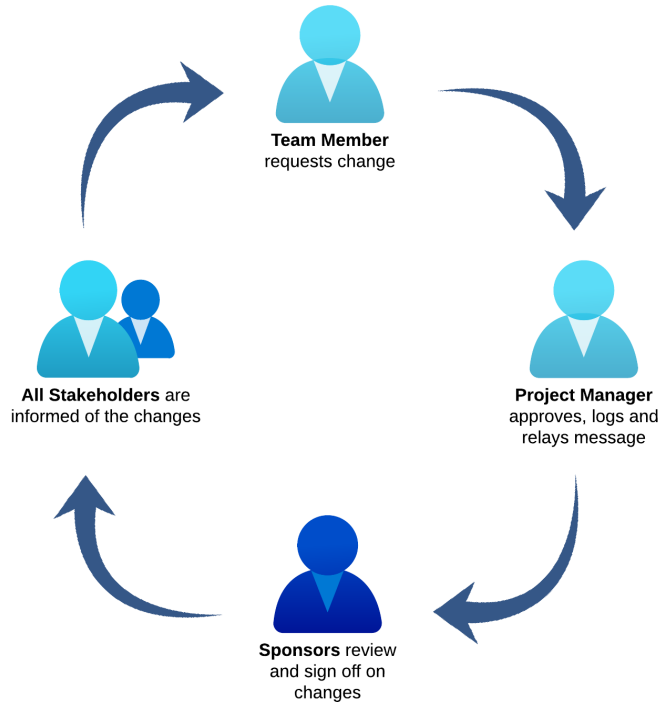
The communication management will include fortnightly meetings with the project sponsors updating them on progress and clarifying any outstanding issues. Fortnightly meetings have been chosen to allow adequate time to make progress on the project, and to therefore, have something meaningful to show stakeholders each time. The meetings are held online through Microsoft teams. Additionally, any urgent matters that need to be communicated to the sponsor is done through email. Finally, any documentation surrounding the project is completed through a Google document.

Communication	Method	Frequency	Goal	Owner	Audience
Project Introduction and planning	Meeting	Beginning of project	To understand the project objectives and requirements. It is also an introduction to the team and discussion on the management style approach that will be taken	Project Manager	Project team
Team project progress update	Meeting	Weekly	To deliver weekly updates on progress of individual parts, including any queries or concerns that need to be discussed.	Project Manager	Project team
Sponsors project progress update	Meeting	fortnightly	To deliver fortnightly updates on the progress of the project, including demonstrations, and any clarification needed.	Project Sponsors	Project team, Sponsors
Urgent clarifications/ queries	Email	When required	Any clarification or queries that require immediate attention, or can't wait until a progress meeting can be discussed through email.	Project Manager	Sponsors, Project team

Communication	Method	Frequency	Goal	Owner	Audience
Milestone review	Meeting	At the end of each milestone	At the end of each milestone the team will discuss what has been completed and what needs to be completed moving forward	Project Manager	Project team
Project testing and error evaluation	Meeting	Everyday/ Weekly	Testing will be performed everyday, with any outstanding errors discussed during the weekly meetings.	Project Manager	Project team
Post completion meeting	Meeting	After completion of project	Was the project successful? How can it be improved? What worked well. This meeting is about discussing the key takeaways, and looking through the final deliverables.	Project Manager	Project Team

13 CHANGE CONTROL

We will follow a change control process that involves all of the stakeholders in the project. We define change as any modification that deviates from the original project plan, or impacts the project in any capacity. Team members can ask for potential changes to the project, however, the project manager will be responsible for logging and relaying change requests to the project sponsors. All these changes will be fully documented and will need to be signed off by the project sponsors. The project sponsors have the ability to approve or deny these changes. When changes have been made team members will be communicated to about the changes as soon as possible.



ISSUES AND PROBLEMS

When creating an end-to-end product that scans e-commerce websites and discovers non-compliant marketing for therapeutic goods, there are a number of potential concerns and problems that could occur. A few of these are:

- Difficulties in Data Collection: Due to the amount and diversity of data sources, it can be difficult to gather reliable and pertinent data from e-commerce websites. Data gathering may be more challenging on some websites due to their unique formats, organizational, or access limitations.
- Data Accuracy: The accuracy of the data on e-commerce websites might vary greatly. It may be difficult to extract meaningful features since the product descriptions may be deficient or inaccurate.
- Database Usability: The .csv used to store the information. There may come a time where the amount of products is too much to load in a reasonable amount of time, so other solutions may need to be implemented.
- Legal and Regulatory Issues: Different interpretations of legal definitions can lead to some confusion in what is compliant and non-compliant but with an external sponsor we are able to get an opinion from an expert who can guide us in the right direction.
- Limited Resources: It takes a lot of resources, including computing power, data storage, to develop and operate a system for identifying non-compliant medicinal items. The system's effectiveness may suffer as a result of data processing delays brought on by a lack of resources.

Finally, it should be noted that identifying non-compliant medicinal products on e-commerce websites is a challenging task that may encounter problems with data availability, data quality, legal and regulatory issues and resource constraints. A multidisciplinary strategy involving legal, technical, and regulatory knowledge is necessary to overcome these obstacles.

14 QUALITY ASSURANCE

1. Process of Document Development

This document was constructed from

- A review of text book Schwalbe K (2004) Information Technology Project Management Thomson Learning"
- Meetings with our sponsors and mentor.

2. Traceability

- Refer to stakeholders' requests and requirements; and charter to determine descriptions of the products involved in the project, project constraints, and project assumptions.
- Work Breakdown Structure for breaking down a lot of work into manageable pieces.
- Gantt Chart for timing of task completion as well as dependencies for tasks.

3. Verification

This document was signed off by sponsors indicating that aspects of this project are approved.

4. References

Schwalbe, K. (2004), Information Technology Project Management, 3rd edn, Course Technology, Canada

Cervone, H.F. (2011), "*Understanding agile project management methods using Scrum*", OCLC Systems & Services: International digital library perspectives, Vol. 27 No. 1, pp. 18-22. Available at: *https://doi.org/10.1108/10650751111106528 (*Accessed: 28 February 2023).

Therapeutic Goods Administration (2020) What are 'therapeutic goods'?. Available at: https://www.tga.gov.au/about-tga/what-we-do/what-are-therapeutic-goods (Accessed: 28 February 2023).

5. Document History

Changes from Previous Version:	**Source of Change:**
Version 1.1 Changes based on tutorial feedback.	Tutorial insight gained.
Version 1.2 Changed writing style in places to match document specification.	See mentor meeting #2 minutes.
Version 1.3 Updated some of the language to increase cohesion between different member's contributions.	Planned from the start as it was a foreseeable issue.