FIT3162: Advanced Computer Science Project
Software Test/QA Report
Project Supervisor: Dr. Soon Lay Ki

Group: MCS15
Ooi Yi Sen – 30720699
Chan Wai Han - 31555373
Nawwaf Ali - 31261949
Yeonsoo Kim - 29584612

Table of Contents

Introduction

The report shall document the testing process of Team MCS15 for the Final Year Product of a Question-Answering System about Covid. The system shall be used by the public to essentially answer their queries regarding covid. The system consist of a few requirements, as follows:

- Access to the latest information about Covid. Users must be able to search for information about Covid, read the user guide on how to use our system, and see how to apply some of the information in real life
- Access to a repository where our user can perform self-browsing
- Access to reliable databases. So that our system can provide accurate information to our users
- Access to technical tools, languages, and software for our project.

Using programming languages such as Javascript, HTML and CSS, and Visual Studio Code as the development platform, requirements 1 and 2 have been satisfied in the form of a chatbot in our Q&A page and a self-browsing repository on the Repository page respectively. On the other hand, the Azure Data Studio was used to satisfy requirement 3.

The testing that will be performed on our system shall be focused around the 3 main requirements of our system, which are the first three requirements as listed above. Besides that, our team will focus on ensuring that our testing will satisfy the project's business objectives - to provide useful information relating to covid, as well as the project's social objective - to improve everyone's knowledge, acceptance, and perception of covid. Due to our team implementing the waterfall approach as the software development methodology for this project, the testing process will be simple and well-structured, as the requirements have already been set during the project inception phase.

The roles and responsibilities of the team members will be divided as such:

- Front-end testing: Yi Sen and Wai Han
- Back-end testing: Yeonsoo and Nawwaf

The testing platform that will be used is Visual Studio Code. Manual testing will be done to ensure the correctness of the code by running the program on the live server. Automated testing such as unit tests cannot be done due to the nature of the software requiring the use of a localhost live server.

The testing metrics, as defined by Hamilton (2022), that that our team will use are as follows:

- Passed Test Cases Percentage = (Number of Passes Tests/Total number of tests executed) x 100
- Failed Test Cases Percentage = (Number of Failed Tests/Total number of tests executed) X 100

The software testing methods that we will use are blackbox testing and whitebox testing. Control flow graphs will be drawn to enable whitebox testing. We will then discuss integration testing, system testing, our system's performance, scalability, and security, usability testing, the limitations of our software and testing process, and finally recommendations for the improvement to our system moving forward.

Black Box Testing

Black Box Testing, also known as Behavioural Testing, is defined as the software testing method where the software's functionalities are tested when there is no knowledge of the internal implementation of the code. The input and output of the software is the main focus, with the software's requirements and objectives being the primary concern (Hamilton, 2022). The type of black box testing that our team will focus on is functional testing, where the main priority lies in the functional requirements of the system. In other words, the most critical components of the software, which satisfy our project's requirements and objectives, will be tested.

Firstly, for our covid chatbot, the black box testing technique that was used was the Equivalence Class Testing technique, where possible inputs are divided into partitions, and only one example input from each group is tested (*Black Box Testing*, n.d.).

Black Box Testing 1
What is being tested: Covid Chatbot (responses.js)
How it is being tested: Equivalence Class Testing
What are the inputs to the code being tested: ● Invalid inputs that include "cq": "cq000" ● Invalid inputs that include "eq": "eq000" ● Invalid inputs that exclude "cq" and "eq": "naruto" ● Valid inputs that include "cq": "cq001" ● Valid inputs that include "eq": "eq001" ● Valid inputs that exclude "cq" and "eq": "temperature"
What are the expected outputs: ● "cq000" returns "We dont have this question try asking something else!" ● "eq000" returns "We dont have this question try asking something else!" ● "naruto" returns "We dont have this question try asking something else!" ● "cq001" returns the appropriate question id, question text, and answer text ● "eq001" returns a list of the appropriate question id, question text, and answer text ● "temperature" returns a list of the question id, question text, and answer text that include "temperature" in its keywords
Actual outputs observed: ● Input "cq000": PASSED (Appendix 1.1.1) ● Input "eq000": PASSED (Appendix 1.1.2) ● Input "naruto": PASSED (Appendix 1.1.3) ● Input "cq001": PASSED (Appendix 1.1.4) ● Input "eq001": PASSED (Appendix 1.1.5) ● Input "temperature": PASSED (Appendix 1.6)

Besides that, Random Testing was used. Random Testing is defined as the verification of the software's output using arbitrarily generated inputs. This testing method is not costly, unbiased, and can detect bugs quickly.

Black Box Testing 2
What is being tested: Covid Chatbot (responses.js)
How it is being tested: Random Testing
What are the inputs to the code being tested: ● "o" ● "aeiou" ● "1" ● "-69"
What are the expected outputs: ● "o" returns a list of the question id, question text, and answer text that include "o" in its keywords ● "aeiou" returns "We dont have this question try asking something else!" ● "1" returns "We dont have this question try asking something else!" ● "-69" returns "We dont have this question try asking something else!"
Actual outputs observed: ● Input "o": PASSED (Appendix 1.2.1) ● Input "aeiou": PASSED (Appendix 1.2.2) ● Input "1": PASSED (Appendix 1.2.3) ● Input "-69": PASSED (Appendix 1.2.4)

Next, for the repository, the first black box testing method that was used was the Equivalence Class Testing technique.

Black Box Testing 3
What is being tested: Repository (repo.html)
How it is being tested: Equivalence Class Testing
What are the inputs to the code being tested: ● Invalid inputs: "123" ● Valid inputs: "temperature" ● Empty input: ""
What are the expected outputs: ● "123" returns an empty table ● "temperature" returns a table containing the question text and answer text if the answer text includes "temperature" ● "" returns a table containing all the rows from the database
Actual outputs observed: ● Input "123": PASSED (Appendix 1.3.1) ● Input "temperature": PASSED (Appendix 1.3.2) ● Input "": PASSED (Appendix 1.3.3)

The second black box testing method used for the repository was the Random Testing technique.

Black Box Testing 4
What is being tested: Repository (repo.html)
How it is being tested: Random Testing
What are the inputs to the code being tested: ● "o" ● "aeiou" ● "1" ● "-69"
What are the expected outputs: ● "o" returns a table containing the question text and answer text if the answer text includes "o" ● "aeiou" returns an empty table ● "1" returns returns a table containing the question text and answer text if the answer text includes "1" ● "-69" returns an empty table
Actual outputs observed: ● Input "o": PASSED (Appendix 1.4.1) ● Input "aeiou": PASSED (Appendix 1.4.2) ● Input "1": PASSED (Appendix 1.4.3) ● Input "-69": PASSED (Appendix 1.4.4)

To populate the question and corresponding answers in our database, we developed a web scraping Python script to automatically get the answer from the Google search engine. To test the web scraping script, Equivalence partitioning technique is used. As the original script requires data fetched from the Database, not an user input, we modified the script to get a manual input for testing.

The script will output corresponding SQL statements if an answer to the input is found. Else, it will print out an error.

Black Box Testing 5
What is being tested: get_documents.py
How it is being tested: Equivalence partitioning
What are the inputs to the code being tested: • Meaningless words = "hello", "question" • Hard questions to answer = "Why don't I get Covid-19?", "When will Covid-19 end?" • Appropriate questions = "Where Covid-19 started?", "Is COVID-19 a threat?"
What are the expected outputs: • All 'Meaningless words' class inputs will lead to errors • All 'Hard questions to answer' inputs will lead to errors • Corresponding SQL statements will be output for all 'Appropriate questions' input
Actual outputs observed: • Input "hello": PASSED (Appendix 1.5.1) • Input "question": PASSED (Appendix 1.5.2) • Input "Why don't I get Covid-19?": PASSED (Appendix 1.5.3) • Input "When will Covid-19 end?": PASSED (Appendix 1.5.4) • Input "Where Covid-19 started?": PASSED (Appendix 1.5.5) • Input "Is COVID-19 a threat?": PASSED (Appendix 1.5.6)

After much thought and analysis, our team has concluded that other common black box testing methods are not feasible for our project. For example, Boundary Value testing focuses on the boundary of the input values. Essentially, as Hamilton (2022) explains in his article, this technique is most suitable for systems that accept inputs with certain ranges. In other words, it is most suitable for programs that accept integer values. However, our chatbot takes in the input values as strings, thus being inappropriate to use this testing method.. Besides that, the Decision Table testing technique could not be used. As defined in Javatpoint (n.d.), it is a testing strategy that captures various input combinations into a table. This testing technique requires multiple input values. However, in our case, our program is a chatbot and a repository that only accepts one input, thus being inappropriate to use this testing method.

White Box Testing

Most execution paths in software are traversed correctly and produce sensible output/meaningful error messages (whitebox) (path coverage, condition coverage thingy)
- Special conditions
- Robustness (file exists before accessing or writing data to it)
- Functions tested separately - range of input parameters - boundary condition data

White Box Testing goes by various names, some being Clear Box Testing, Open Box testing, and Glass Box Testing. As the names suggest, Hamilton (2022) defines the testing technique as the verification of the internal components of the software. Being the counterpart of black box testing, this testing method prioritizes the end-user experience. Components such as paths in coding processes, the flow of inputs through the code, expected output, conditional loops, and code statements are involved in white box testing. Essentially, the working flow of the software is to be verified.

Control Flow Graphs (CFG) are a representation of the execution flow in a graph. It consists of basic blocks, decision blocks, and arrows or edges that connect these blocks. Through CFGs, the coverage of a function can be represented in the form of a graph. Consequently, decisions, branches, blocks, and lines can be understood easier. A CFG was created for the getBotResponse function in responses.js (Appendix 2.1.1), and for the script in repo.html (Appendix 2.1.2) and for the add_rows.py script (Appendix 2.1.3), which inserts datasets into database. With the CFGs, it will enable us to perform white box testing much easier, as there is now a clear visual representation of the main functions that satisfy the project requirements and objectives.

Firstly, for our covid chatbot, the white box testing technique that was used was the Condition Coverage Testing. Condition Coverage Testing, which is also known as predicate coverage, is defined as the testing of all conditional expressions in a program for all possible outcomes of the conditions (Asif, n.d.).

White Box Testing 1
What is being tested: Covid Chatbot (responses.js)
How it is being tested: Condition Coverage Testing
What are the inputs to the code being tested: ● "hello" ● "hi" ● "goodbye" ● "bye" ● "cq001" ● "cq0001" ● "covid" ● "temperature"
What are the expected outputs: ● "hello" returns "Hello there!" ● "hi" returns "Hello there!" ● "goodbye" returns "Talk to you later!" ● "bye" returns "Talk to you later!" ● "cq001" returns a list of the question id, question text, and answer text that include "cq001" in the question id ● "cq0001" returns "We dont have this question try asking something else!" ● "covid" returns a list of the question id, question text, and answer text that include "covid" in its keywords ● "temperature" returns a list of the question id, question text, and answer text that include "temperature" in its keywords
Actual outputs observed: ● Input "hello": PASSED (Appendix 2.2.1) ● Input "hi": PASSED (Appendix 2.2.2) ● Input "goodbye": PASSED (Appendix 2.2.3) ● Input "bye": PASSED (Appendix 2.2.4) ● Input "cq001": PASSED (Appendix 2.2.5) ● Input "cq0001": PASSED (Appendix 2.2.6) ● Input "covid": PASSED (Appendix 2.2.7) ● Input "temperature": PASSED (Appendix 2.2.8)

Next, for our Repository, the white box testing technique that was used was also the Condition Coverage Testing.

White Box Testing 2
What is being tested: Repository (repo.html)
How it is being tested: Condition Coverage Testing
What are the inputs to the code being tested: • "temperature" • ""
What are the expected outputs: • "temperature" returns a table containing the question text and answer text if the answer text includes "temperature" • "" returns a table containing all the rows from the database
Actual outputs observed: • Input "temperature": PASSED (Appendix 2.3.1) • Input "": PASSED (Appendix 2.3.2)

Our team decided to use Condition Coverage testing for the Covid Chatbot and the Repository because it is said to be more accurate than other methods such as Line Coverage and Statement Coverage, Block Coverage, Branch/Decision Coverage, and Path Coverage. For example, the line coverage of a one-line code, which is possible in Java, would always be 100%, as all lines are covered. Thus, it is not very accurate. Besides that, Branch Coverage may cover each branch in the code. However, when there are complicated conditions, which use multiple ANDs and ORs, not all possibilities are covered. Therefore, Condition Coverage splits the decisions into single conditions, rather than having one big decision block with the entire condition. Consequently, all conditions would be checked.

Thirdly, add_rows.py script, which automatically inserts JSON datasets into the database is tested. As the original script does not require any user input and was designed for our team's datasets only, we modified the script so that it can process any given files and output the result of the process so that we can actually see how the function handled the files. It was a script that corresponds to an already existing dataset only, but the test was conducted because it had some error handling codes in case wrong files are inputted to the script. As all the functions inside add_rows.py script share the same workings, we only tested the function to add 'document' datasets.

White Box Testing 3
What is being tested: add_rows_test.py (Modified add_rows.py for testing)
How it is being tested: Condition Coverage Testing

What are the inputs to the code being tested:
   ● A Normal 'document' JSON dataset, '000tfenb.json' (.json file)
   ● A text file, 'testing.txt' (.txt file)
   ● A csv file, 'testing.csv' (.csv file)

What are the expected outputs:
   ● The JSON file will output a query statement that inserts the data into the DB.
   ● The txt file will be ignored.
   ● The csv file will be ignored.

Actual outputs observed:
   ● Input "000tfenb.json": PASSED (Appendix 2.4.1)
   ● Input "testing.txt": PASSED (Appendix 2.4.2)
   ● Input "testing.csv": PASSED (Appendix 2.4.3)

Condition Coverage Testing is used for this add_rows.py because all the functions have only one if statement to check whether the file is JSON data or not. It means that a condition coverage testing for the if statement becomes the branch testing / path testing for the script as well due to the reason.

Performing Testing

A software system's features and functionality are not the only considerations. The performance of a software programme, such as its response time, reliability, resource utilisation, and scalability, is important. The purpose of Performance Testing is to reduce performance bottlenecks rather than to uncover flaws (Hamilton, 2022).

Performance testing is carried out to offer stakeholders information on the speed, stability, and scalability of their application. More significantly, Performance Testing identifies what has to be changed before the product is released to the public. Without Performance Testing, the product is prone to drawbacks such as slowness when several people use it at the same time, inconsistencies across different operating systems, and poor usability (Hamilton, 2022).

We have undertaken the following sorts of performance testing:
1. Stress testing
2. Spike testing

For Stress Testing, we put our software through extensive testing to determine how it handles high traffic and data processing. The goal is to determine a software's breaking point. To do the testing, we entered 5 queries into the application successively to see if the application could handle them or it crashes.
1. Test case 1: Input 1 query (pass)
2. Test case 2: Input 3 queries (pass)
3. Test case 3: Input 5 queries (pass)

Moreover, we conducted Stress Testing for our Database as well because our database is located in the U.S. and the database is a free version of Azure SQL Database. To do the testing, we executed multiple SELECT query statements consecutively in the 'Question_Answer' table which are mainly used in the front-end software.

1. Test case 1: Input 3 SELECT query (pass)
2. Test case 2: Input 5 SELECT queries (pass)
3. Test case 3: Input 7 SELECT queries (pass)

For Spike testing, we test the software's response to unexpectedly big surges in user load. To do the testing, we spam multiple queries simultaneously into the application to see if the application could handle them or it crashes.
1. Test case 1: Spam 5 queries (pass)
2. Test case 2: Spam 10 queries (pass)
3. Test case 3: Spam 15 queries (pass)

The similar tests are conducted for the database again to test if the database can handle all the queries when the front-end websites repeatedly and simultaneously query the database in short-time. To reproduce multiple queries being executed in the database, multiple 'SELECT TOP 3' query files were executed in the 'Question_Answer' table simultaneously at the same time.

1. Test case 1: Running 5 SELECT TOP 3 queries at the same time (pass)

2. Test case 2: Running 10 SELECT TOP 3 queries at the same time (pass)
3. Test case 3: Running 15 SELECT TOP 3 queries at the same time (pass)

The methodology used for performance testing might vary greatly, but the goal of performance testing stays constant. It can assist in demonstrating that our software system fulfils pre-defined performance standards. It can also assist in identifying sections of our software system that are degrading its performance (Hamilton, 2022).

Scalability

The scalability of our software was considered when we were in the early planning stages of our project. In order to ensure the scalability of our software, our team performed preprocessing in the backend. The preprocessing ensures that only the required columns are put into the Azure database. Therefore, when the database information is retrieved in JavaScript to be stored into the json file, it is extremely straightforward. Essentially, if more rows are added into the data, it will always be preprocessed in the backend before being retrieved. In other words, the JavaScript code will only retrieve the dataset, regardless of its size, and will not need to perform any processing.

Security

Website security refers to any action or software put in place to guarantee that website data is not exposed to cybercriminals or to prevent the website from being abused in any way. These activities aid in the protection of sensitive data, hardware, and software on a website against the numerous sorts of threats that exist today.

Since we are using a local host rather than online hosting, website security is not a primary concern for our project. We picked local hosting because it allows us to do private testing on our computer without transmitting files over the internet, which is more secure and will not expose your website to the public (Juviler, 2021).

HTTPs protocol is used by the website from where we obtained our datasets. The use of HTTPs avoids interceptions and disruptions while the material is in transit. We also verify that the website has an SSL Certificate to ensure the internet connection is safe. SSL is another essential site protocol that encrypts data to prevent others from accessing it while it is in transit. It prevents individuals without sufficient authorisation from accessing the data (Hendrick, n.d.). Hence we can conclude that our dataset's security is safe and secure.

Security of the back-end Database strictly depends on the security of the Microsoft Azure SQL Database itself. It is because Microsoft Azure SQL is a cloud-based database which is actually managed by Microsoft, which means that the database is not opened and configured by individuals. The only security measures we can do are to set up firewall rules, classify user/admin roles and prevent the admin account password leakage.

But our team thinks this is an advantage, not a disadvantage. Security measures are often not within the reach of developers, and are largely dependent on security professionals. Because our team is also just students, not security experts, we believe that using cloud-based DB, run by large corporations is bound to be more secure than DB, running personally.

Optimization

Our system went through a fair amount of optimization, particularly in the back-end of the program. When the database is downloaded online, it is pre-processed by the back-end team before being inserted into the Azure database. Besides that, certain questions with inappropriate or missing answers will be filled in with online answers. Essentially, answer snippets from Google would be retrieved and inserted into the answer text column for those particular questions. All the preprocessing is done in the back-end before the data is inserted into the database. Therefore, when the data is retrieved from the database and stored into the json file, time complexity is significantly reduced as only retrieval is done, without any need to process or filter.

Usability Testing

Usability Testing is essentially testing by a third party to assess the ease of completing a certain task using the program's interface. Our team took the user guide and software and gave it to some volunteers to use.

Volunteer 1: Leong Pui Lum Occupation: Third Year Student at Monash University
Feedback: Pui Lum was aware of the amount of tabs available in our software, through the navigation bar at the very top. She was able to navigate through each page with ease by clicking on each tab. Pui Lum also knew what the purpose of each tab was, as they were clearly labelled at the top with a short and simple keyword. However, Pui Lum was not aware of the Quote of the Day at the bottom of the Home page, albeit it not being essential to our project's main requirement and objective. In terms of the main requirements of our software, upon reading the user guide, Pui Lum was able to easily understand how to use the Covid Chatbot. She knew how to click onto the FAQ button to view the dropdown list of questions, and commended the legend on the left side which enabled her to quickly understand each question. Furthermore, Pui Lum knew how to use the chatbot, as the interface was similar to other chatbots, with a simple input message bar and a send button. She was satisfied with how the bot listed out the answers to her queries. On the other hand, upon reading the user guide and the tutorial on the left side of the Repository page, Pui Lum immediately knew how to use the repository to self-browse. In no time, she was searching for various types of things related to covid. She also knew how to view our entire database by clicking on the search icon without any input keywords.

Volunteer 2: Goh Kai Yuan Occupation: Second Year Student at Monash University
Feedback: Kai Yuan seemed to have no problems utilising our programme after reading through the end user documentation. He was able to navigate through all of the webpages and use all of the features that our team had created. Kai Yuan had no problem using the Covid Chatbot, he knew that he had to query using keywords and choose the questions from the display. He also knew that the legend on the left side of the screen are the questions to the frequently asked questions (FAQs). He tested the repository website after evaluating the covid chatbot page. He then clicks on the navigation bar to go to the repository page. Kai Yuan tested the repository by searching for keywords in his head. He was unaware that our repository allows comprehensive browsing of our datasets, but with the assistance of our tutorial on the left side of the screen, he immediately understood that he could query without writing anything. He was really satisfied with how user friendly our system is after testing all of its functions.

Limitations of Software

Our team managed to identify a few limitations to our software. The main limitation is that users should try to query with only keywords. This is because for our Covid Chatbot, if users query with a single alphabet, such as the letter "o" (Appendix 1.2.1), the chatbot will return a list of the question id, question text, and answer text that include "o" in its keywords. This is because the letter "o" is included in the keywords of those rows. Ultimately, such a query would not benefit the user. Besides that, single alphabets as such would typically not be a common user query. Therefore, in order to properly use the chatbot, the user should ideally search a single keyword or phrase. Besides that, users should avoid querying a sentence. For example, "what is temperature" (Appendix 3.1) "temperature" (Appendix 3.2) and returns different results, despite both the queries containing the main keyword "temperature" in it. Therefore, in conclusion, users should only search using the main keyword.

Moreover, our software cannot answer some detailed or non-popular questions. It is because our software does not automatically generate answers whenever it receives the users' questions, instead our software uses stored and pre-processed questions and answers from the Database. For this reason, to increase the amount of questions that can be answered, we need to increase the amount of questions and answers that are stored in the database. Currently, all the famous questions and answers are stored in the database, so most of the ordinary and non-detailed questions are answerable.

Another limitation is the database. Because we are currently using Microsoft's paid cloud-based database rather than using a database that we have set up and opened ourselves, it is also a cost burden to keep the database. If we do not keep the database, the software will work as well. But it is not a problem with the software itself, so it is not a serious problem currently.

Recommendation for Improvement

Our team was able to identify a few recommendations for improvement to our software moving forward. Firstly, from the Usability Testing, Pui Lum had mentioned that she could not see the Quote of the Day section at the Home page. Therefore, we could either bring the Quote of the Day up to make it more obvious, or we could tell the user of its existence in the user guide.

Besides that, during the Equivalence Partitioning Testing, we realized that the Consumer Questions (CQ) and Expert Questions (EQ) could be better categorized for the testing. This led us to thinking of a special input for the users. For example, if the user inputs "cq" into the chatbot, it will then list out all the CQ. On the other hand, if the user inputs "eq" into the chatbot, it will then list out all the EQ. Subsequently, from there, the user can select the question that he or she is interested in. This would allow the user to view all CQs and EQs in one list.

Limitation of Testing Process

The main limitation of the testing process for this software is that we could not use automated unit tests to test each function. This is because our software requires the localhost live server to run. It is unable to output in the terminal of our development platform, Visual Studio Code. Therefore, the only way to test our software is to run it on the live server and manually test it by inputting the values one by one.

There were also many problems when testing scripts for data population. Although all scripts are designed to ignore or safely stop without affecting the database when errors occur, the scripts are not suitable for testing separately and it is the another reason why we could not use automated unit tests as well. Eventually, in order to proceed with the test, we created test files by modifying the existing scripts so that the testing files can print the output and can get separate inputs, rather than fixed inputs.

Integration Testing

Our software is not a live software. Essentially, it retrieves the relevant information from the database on the back-end, and then gets retrieved by the front-end. Therefore, it is actually just a manual retrieval of data using the front-end code. As shown in the white box testing and black box testing sections above, the front-end is able to successfully retrieve data from the back-end. On the other hand, the back-end is able to successfully send data to the front-end. It can be concluded that the integration testing works as both sides are able to communicate and achieve their intended task.

Testing Metrics

The calculation for the testing metrics of our software is as follows:

Passed Test Cases Percentage = (Number of Passes Tests/Total number of tests executed) x 100
Passed Test Cases Percentage = (36/36) x 100
Passed Test Cases Percentage = 100%

Failed Test Cases Percentage = (Number of Failed Tests/Total number of tests executed) X 100
Failed Test Cases Percentage = (0/36) x 100
Failed Test Cases Percentage = 0%

Conclusion

Since project inception, our team has managed to identify the main requirements of our project:

1. Access to the latest information about Covid. Users must be able to search for information about Covid, read the user guide on how to use our system, and see how to apply some of the information in real life
2. Access to a repository where our user can perform self-browsing
3. Access to reliable databases. So that our system can provide accurate information to our users
4. Access to technical tools, languages, and software for our project.

Our software's system testing is a success. Requirements 1 and 2 were implemented in the Covid Chatbot and Repository respectively. Through this test report, it is evident that the main requirements of our software are satisfied, without any fatal errors. The testing methods used were the Black Box testing method and the White Box testing method. For the Black Box testing method, Equivalence Testing Method and Random Testing were performed. On the other hand, for the White Box Testing method, Condition Coverage testing was done. Control Flow Graphs were drawn to help visualize the tested functions better. Consequently, the project's business objectives - to provide useful information relating to covid, as well as the project's social objective - to improve everyone's knowledge, acceptance, and perception of covid have been satisfied.

References

Asif, K. (n.d.). *What is condition coverage testing?* Educative.io.
　　　https://www.educative.io/answers/what-is-condition-coverage-testing

Black Box Testing. (n.d.). Imperva.
　　　https://www.imperva.com/learn/application-security/black-box-testing/

Black box testing. (n.d.). Javatpoint. https://www.javatpoint.com/black-box-testing

Hamilton, T. (2022, August 27). *Software Testing Metrics: What is, Types & Example*.
　　　Guru99. https://www.guru99.com/software-testing-metrics-complete-tutorial.html

Hamilton, T. (2022, September 3). *What is BLACK Box Testing? Techniques, Example &*
　　　Types. Guru99. https://www.guru99.com/black-box-testing.html

Hamilton, T. (2022, September 16). *What is, Techniques, Example & Types*. Guru99.
　　　https://www.guru99.com/white-box-testing.html

Hamilton, T. (2022, September 16). *Performance Testing Tutorial – Types (Example)*.
　　　Guru99. https://www.guru99.com/performance-testing.html

Hendricks, D. (n.d.). 10 Essential Steps To Improve Your Website Security.
　　　https://www.computer.org/publications/tech-news/trends/10-essential-steps-to-improve-your-
　　　website-security

Juviler, J. (2021, October 14). What Is Localhost? (And How Can You Use It?).
　　　https://blog.hubspot.com/website/what-is-localhost

Appendix

Appendix 1: Black Box Testing

Appendix 1.1: Black Box Testing 1

Appendix 1.1.1: Input "cq000"

Appendix 1.1.2: Input "cq000"

Appendix 1.1.3: Input "naruto"

Appendix 1.1.4: Input "cq001

Appendix 1.1.5: Input "eq001"

temperature

Here are the information that we have related to *"temperature"*.
Please type in the Question ID that you are interested in into the chat to view the full answer to the question.

Question ID: EQ002
Question Text: how does the coronavirus respond to changes in the weather

Question ID: EQ005
Question Text: what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?

Question ID: CQ002
Question Text: how does the coronavirus respond to changes in the weather

Appendix 1.2: Black Box Testing 2

Appendix 1.2.1: Input "o"

o

Here are the information that we have related to **"o"**.
Please type in the Question ID that you are interested in into the chat to view the full answer to the question.

Question ID: EQ001
Question Text: what is the origin of COVID-19

Question ID: EQ002
Question Text: how does the coronavirus respond to changes in the weather

Question ID: EQ005
Question Text: what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?

Question ID: EQ007
Question Text: are there serological tests that detect antibodies to coronavirus?

Appendix 1.2.2: Input "aeiou"

aeiou

We dont have this question try asking something else!

Appendix 1.2.3: Input "1"

Here are the information that we have related to **"1"**.
Please type in the Question ID that you are interested in into the chat to view the full answer to the question.

Question ID: EQ001
Question Text: what is the origin of COVID-19

Question ID: EQ005
Question Text: what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?

Question ID: EQ011
Question Text: what are the guidelines for triaging patients infected with coronavirus?

Question ID: EQ013
Question Text: what are the transmission routes of coronavirus?

Question ID: EQ020

1

Appendix 1.2.4: Input "-69"

-69

We dont have this question try asking something else!

Appendix 1.3: Black Box Testing 3

Appendix 1.3.1: Input "123"

Repository	123 🔍

q_text	answer_text

Appendix 1.3.2: Input "temperature"

Repository	temperature 🔍

q_text	answer_text
how does the coronavirus respond to changes in the weather	Temperature and relative humidity are major factors determining virus inactivation in the environment. Temperature and relative humidity have been shown to impact the kinetics of inactivation of coronaviruses. Increased temperatures have been shown to increase the rate of the inactivation, and decreased relative humidity have been associated with a reduction of coronaviruses inactivation rate on surfaces. Inactivation rates were lower in suspensions compared to surfaces in studies that tested both suspensions and surfaces at similar temperatures. Hence, the prediction of the persistence of SARS-CoV-2 on fomites is essential to investigate the importance of contact transmission.
how does the coronavirus respond to changes in the weather	It is not yet known whether weather and temperature affect the spread of COVID-19. Some other viruses, like those that cause the common cold and flu, spread more during cold weather months but that does not mean it is impossible to become sick with these viruses during other months. There is much more to learn about the transmissibility, severity, and other features associated with COVID-19 and investigations are ongoing.

Appendix 1.3.3: Input ""

Repository	Search.. 🔍

q_text	answer_text
what is the origin of COVID-19	Sars-cov-2 is thought to be originated from an animal coronavirus that successfully adapted to humans. The species of origin of sars-cov-2 has not been fully identified, but the virus seems to be related to sars-cov and other coronaviruses found in bats and other mammal species. The first cases of the novel coronavirus associated disease (covid-19) have been traced to the Chinese province of Hubei in early December 2019. Although the actual index case is not really known, the first sequence of the novel coronavirus was produced within weeks from the emergence of the disease.
how does the coronavirus respond to changes in the weather	Temperature and relative humidity are major factors determining virus inactivation in the environment. Temperature and relative humidity have been shown to impact the kinetics of inactivation of coronaviruses. Increased temperatures have been shown to increase the rate of the inactivation, and decreased relative humidity have been associated with a reduction of coronaviruses inactivation rate on surfaces. Inactivation rates were lower in suspensions compared to surfaces in studies that tested both suspensions and surfaces at similar temperatures. Hence, the prediction of the persistence of SARS-CoV-2 on fomites is essential to investigate the importance of contact transmission.
what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?	Favipiravir is a nucleoside analog with an ability to inhibit RNA-dependent polymerase and was approved for marketing in Japan in 2014. In vitro experiments have shown that favipiravir is effective for COVID-19 and that its EC 50 is 61.88 M. To date, some clinical trials of favipiravir in the treatment of COVID-19 have been carried out in China. Recent clinical studies have found that compared with the antiviral drug arbidol, the clinical effect of favipiravir is more significant. Nucleic acid positive-to-negative time, mean antipyretic time and cough remission time were all better than those of the arbidol group. Remdesivir(GS-5734) Remdesivir was first used to treat Ebola virus, and it has completed phase 2 clinical trials. As a nucleoside analog, it can interact with RdRp, and the triphosphate form of remdesivir will compete with adenosine triphosphate, leading to delayed chain termination and inhibiting viral replication and transcription. number of in vitro studies have shown that remdesivir has inhibitory effects on a variety of human and animal coronaviruses.
	The serological tests that have recently been developed employ different methods and target either IgG or IgM or both. In an attempt to fill the knowledge gap, this systematic review summarized evidence from 38 studies involving 7848 individuals. Although the US Food and Drug Administration (FDA) has approved ELISA, LFIA, and neutralization

Appendix 1.4: Black Box Testing 4

Appendix 1.4.1: Input "o"

q_text	answer_text
what is the origin of COVID-19	Sars-cov-2 is thought to be originated from an animal coronavirus that successfully adapted to humans. The species of origin of sars-cov-2 has not been fully identified, but the virus seems to be related to sars-cov and other coronaviruses found in bats and other mammal species. The first cases of the novel coronavirus associated disease (covid-19) have been traced to the Chinese province of Hubei in early December 2019. Although the actual index case is not really known, the first sequence of the novel coronavirus was produced within weeks from the emergence of the disease.
how does the coronavirus respond to changes in the weather	Temperature and relative humidity are major factors determining virus inactivation in the environment. Temperature and relative humidity have been shown to impact the kinetics of inactivation of coronaviruses. Increased temperatures have been shown to increase the rate of the inactivation, and decreased relative humidity have been associated with a reduction of coronaviruses inactivation rate on surfaces. Inactivation rates were lower in suspensions compared to surfaces in studies that tested both suspensions and surfaces at similar temperatures. Hence, the prediction of the persistence of SARS-CoV-2 on fomites is essential to investigate the importance of contact transmission.
what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?	Favipiravir is a nucleoside analog with an ability to inhibit RNA-dependent polymerase and was approved for marketing in Japan in 2014. In vitro experiments have shown that favipiravir is effective for COVID-19 and that its EC 50 is 61.88 M. To date, some clinical trials of favipiravir in the treatment of COVID-19 have been carried out in China. Recent clinical studies have found that compared with the antiviral drug arbidol, the clinical effect of favipiravir is more significant. Nucleic acid positive-to-negative time, mean antipyretic time and cough remission time were all better than those of the arbidol group. Remdesivir(GS-5734) Remdesivir was first used to treat Ebola virus, and it has completed phase 2 clinical trials. As a nucleoside analog, it can interact with RdRp, and the triphosphate form of remdesivir will compete with adenosine triphosphate, leading to delayed chain termination and inhibiting viral replication and transcription. number of in vitro studies have shown that remdesivir has inhibitory effects on a variety of human and animal coronaviruses.
	The serological tests that have recently been developed employ different methods and target either IgG or IgM or both. In an attempt to fill the knowledge gap, this systematic review summarized evidence from 38 studies involving 7848 individuals. Although the US Food and Drug Administration (FDA) has approved ELISA, LFIA, and neutralization

Appendix 1.4.2: Input "aeiou"

Repository

aeiou

q_text answer_text

Appendix 1.4.3: Input "1"

q_text	answer_text
what is the origin of COVID-19	Sars-cov-2 is thought to be originated from an animal coronavirus that successfully adapted to humans. The species of origin of sars-cov-2 has not been fully identified, but the virus seems to be related to sars-cov and other coronaviruses found in bats and other mammal species. The first cases of the novel coronavirus associated disease (covid-19) have been traced to the Chinese province of Hubei in early December 2019. Although the actual index case is not really known, the first sequence of the novel coronavirus was produced within weeks from the emergence of the disease.
what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?	Favipiravir is a nucleoside analog with an ability to inhibit RNA-dependent polymerase and was approved for marketing in Japan in 2014. In vitro experiments have shown that favipiravir is effective for COVID-19 and that its EC 50 is 61.88 M. To date, some clinical trials of favipiravir in the treatment of COVID-19 have been carried out in China. Recent clinical studies have found that compared with the antiviral drug arbidol, the clinical effect of favipiravir is more significant. Nucleic acid positive-to-negative time, mean antipyretic time and cough remission time were all better than those of the arbidol group. Remdesivir(GS-5734) Remdesivir was first used to treat Ebola virus, and it has completed phase 2 clinical trials. As a nucleoside analog, it can interact with RdRp, and the triphosphate form of remdesivir will compete with adenosine triphosphate, leading to delayed chain termination and inhibiting viral replication and transcription. number of in vitro studies have shown that remdesivir has inhibitory effects on a variety of human and animal coronaviruses.
what are the guidelines for triaging patients infected with coronavirus?	COVID-19 illuminates the need for a tiered diagnostic approach to rapidly identify clinically significant infections and reduce disease spread. Without the ability to efficiently screen patients, hospitals are overwhelmed, potentially delaying treatment for other emergencies. A multi-tiered, diagnostic strategy incorporating a rapid host immune response assay as a screening test, molecular confirmatory testing and rapid IgM/IgG testing to assess benefit from quarantine/further testing and provide information on population exposure/herd immunity would efficiently evaluate potential COVID-19 patients. Triaging patients within minutes with a fingerstick rather than hours/days after an invasive swab is critical to pandemic response as reliance on the existing strategy is limited by assay accuracy, time to results, and testing capacity. Early screening and triage is achievable from the outset of a pandemic with point-of-care host immune response testing which will improve response time to clinical and public health actions.

Appendix 1.4.4: Input "-69"

Repository

-69 🔍

q_text	answer_text

Appendix 1.5: Black Box Testing 5

Appendix 1.5.1: Input "hello"

```
hello - !!No Main Answer Found for this question!!
---------------------
```

Appendix 1.5.2: Input "question"

```
question - !!No Main Answer Found for this question!!
---------------------
```

Appendix 1.5.3: Input "Why don't I get Covid-19?"

```
Why don't I get Covid-19? - !!No Main Answer Found for this question!!
---------------------
```

Appendix 1.5.4: Input "When will Covid-19 end?"

```
When will Covid-19 end? - !!No Main Answer Found for this question!!
---------------------
```

Appendix 1.5.5: Input "Where Covid-19 started?"

```
INSERT INTO DOCUMENT (doc_id, doc_title, doc_url) VALUES ('70cc7edf4a3a11ed8eb6b0227aed472f',
INSERT INTO QUESTION_ANSWER (q_id, answer_text, doc_id) VALUES ('Where Covid-19 started?', '
---------------------
```

Appendix 1.5.6: Input "Is Covid-19 a threat?"

```
INSERT INTO DOCUMENT (doc_id, doc_title, doc_url) VALUES ('8995b9434a3a11edb3fcb0227aed47
INSERT INTO QUESTION_ANSWER (q_id, answer_text, doc_id) VALUES ('Is COVID-19 a threat?',
```
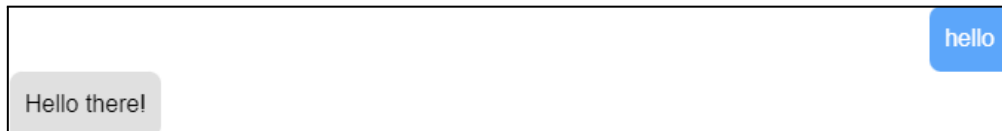
Appendix 2: White Box Testing

Appendix 2.1: CFG

Appendix 2.1.1: CFG for the getBotResponse function in responses.js
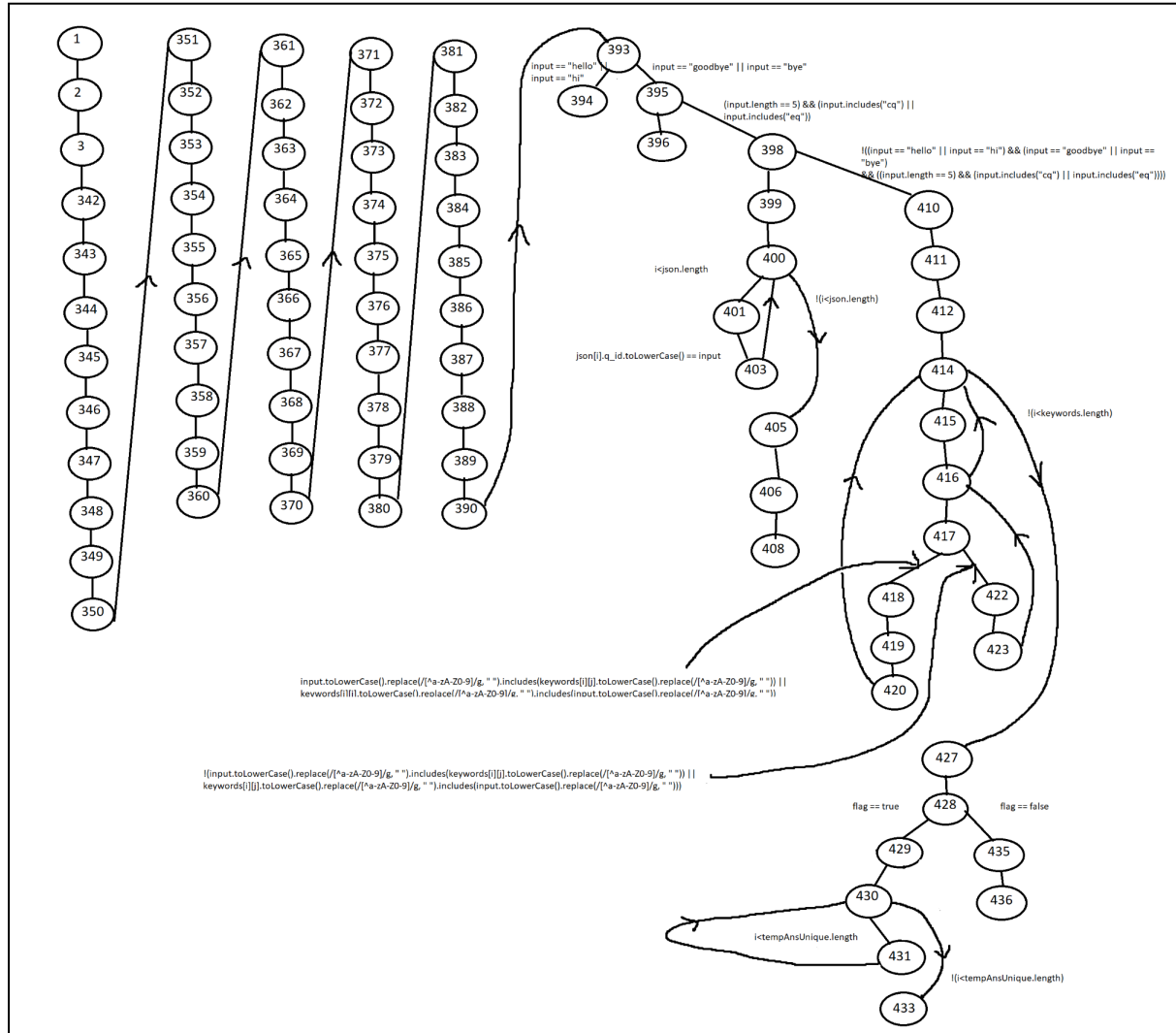
Appendix 2.1.2: CFG for the script in repo.html
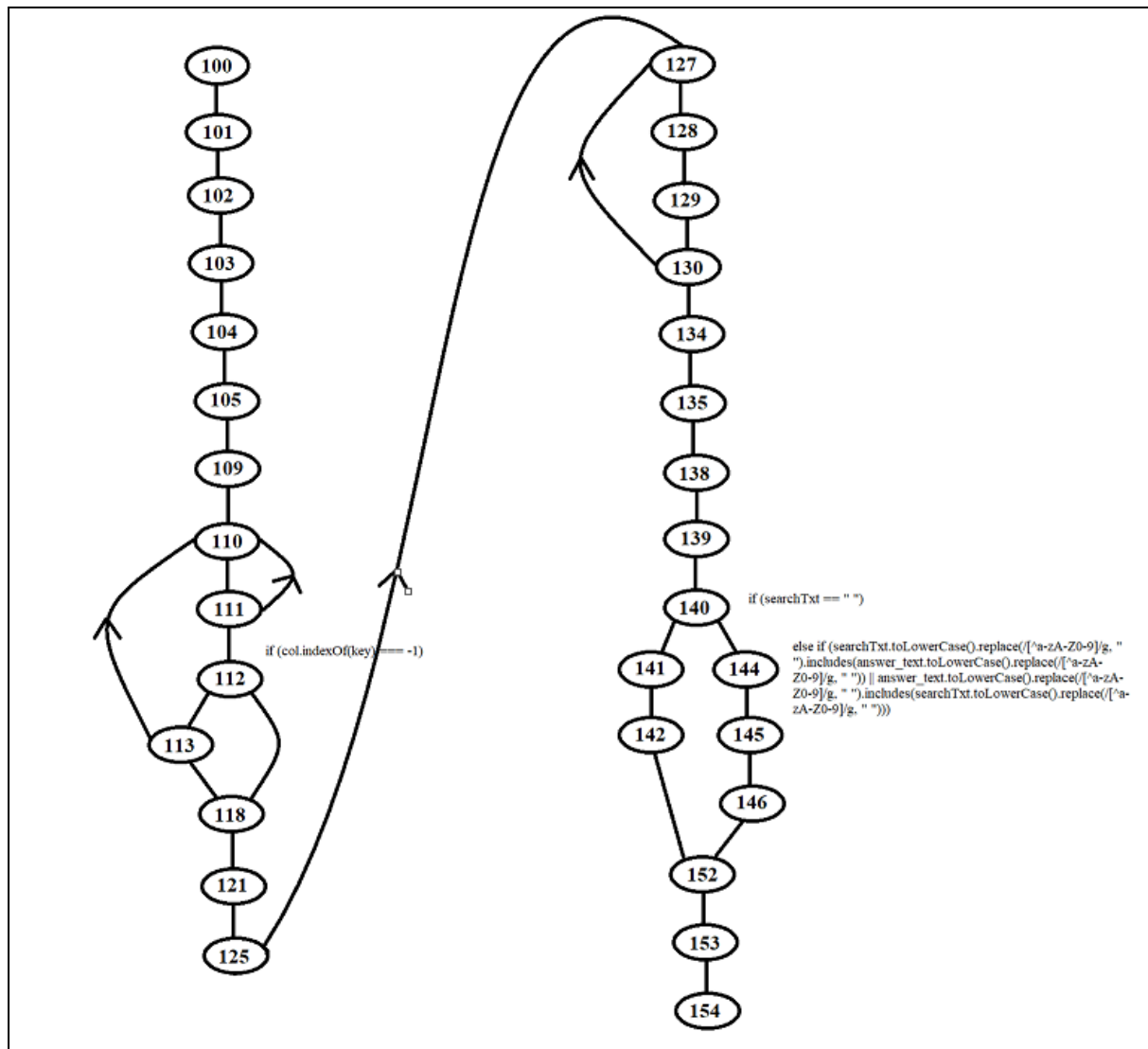
Appendix 2.1.2: CFG for the add_rows_test.py script

27

28

29   if file.endswith('.json')

30   39

37
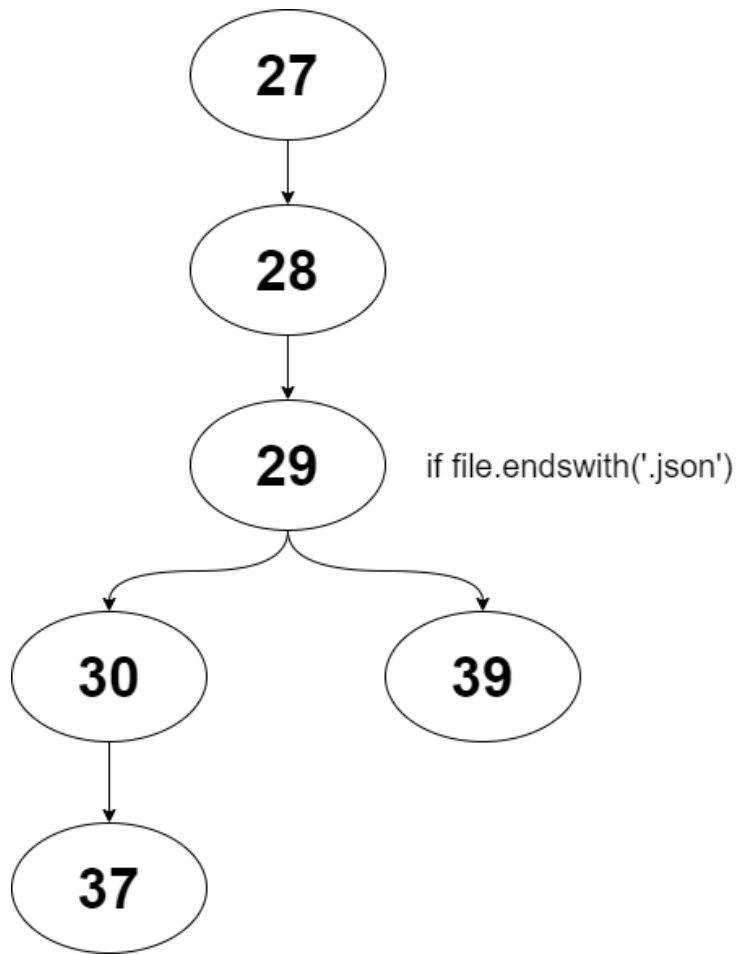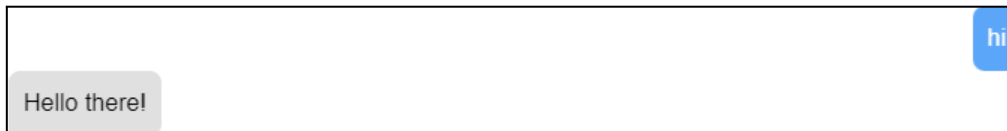
Appendix 2.2: White Box Testing 1

Appendix 2.2.1: Input "hello"

Appendix 2.2.2: Input "hi

Appendix 2.2.3: Input "goodbye"
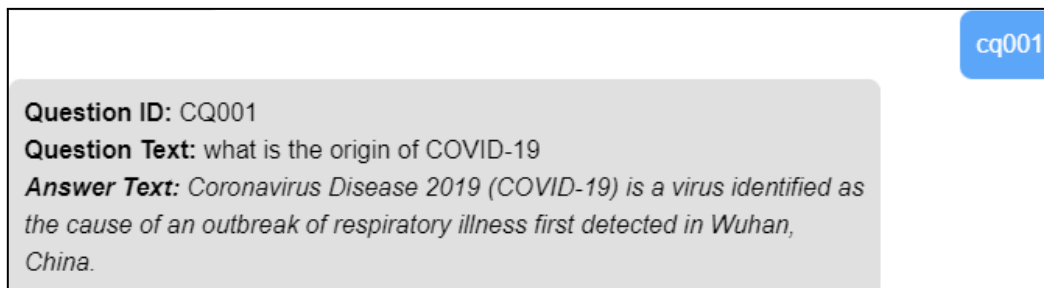
Appendix 2.2.4: Input "bye"
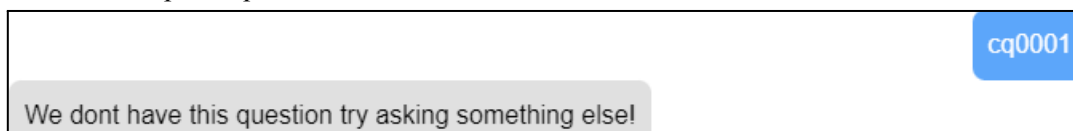
Appendix 2.2.5: Input "cq001"

Appendix 2.2.6: Input "cq0001"

Appendix 2.2.7: Input "covid"

covid

Here are the information that we have related to **"covid"**.
Please type in the Question ID that you are interested in into the chat to view the full answer to the question.

Question ID: EQ001
Question Text: what is the origin of COVID-19

Question ID: EQ005
Question Text: what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?

Question ID: EQ011
Question Text: what are the guidelines for triaging patients infected with coronavirus?

Question ID: EQ013
Question Text: what are the transmission routes of coronavirus?

Question ID: EQ020

Appendix 2.2.8: Input "temperature"

temperature

Here are the information that we have related to **"temperature"**.
Please type in the Question ID that you are interested in into the chat to view the full answer to the question.

Question ID: EQ002
Question Text: how does the coronavirus respond to changes in the weather

Question ID: EQ005
**Question Text:** what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?

Question ID: CQ002
Question Text: how does the coronavirus respond to changes in the weather

Appendix 2.3: White Box Testing 2

Appendix 2.3.1: Input "temperature"

Repository	temperature 🔍

q_text	answer_text
how does the coronavirus respond to changes in the weather	Temperature and relative humidity are major factors determining virus inactivation in the environment. Temperature and relative humidity have been shown to impact the kinetics of inactivation of coronaviruses. Increased temperatures have been shown to increase the rate of the inactivation, and decreased relative humidity have been associated with a reduction of coronaviruses inactivation rate on surfaces. Inactivation rates were lower in suspensions compared to surfaces in studies that tested both suspensions and surfaces at similar temperatures. Hence, the prediction of the persistence of SARS-CoV-2 on fomites is essential to investigate the importance of contact transmission.
how does the coronavirus respond to changes in the weather	It is not yet known whether weather and temperature affect the spread of COVID-19. Some other viruses, like those that cause the common cold and flu, spread more during cold weather months but that does not mean it is impossible to become sick with these viruses during other months. There is much more to learn about the transmissibility, severity, and other features associated with COVID-19 and investigations are ongoing.

Appendix 2.3.2: Input ""

Search.. 🔍

q_text	answer_text
what is the origin of COVID-19	Sars-cov-2 is thought to be originated from an animal coronavirus that successfully adapted to humans. The species of origin of sars-cov-2 has not been fully identified, but the virus seems to be related to sars-cov and other coronaviruses found in bats and other mammal species. The first cases of the novel coronavirus associated disease (covid-19) have been traced to the Chinese province of Hubei in early December 2019. Although the actual index case is not really known, the first sequence of the novel coronavirus was produced within weeks from the emergence of the disease.
how does the coronavirus respond to changes in the weather	Temperature and relative humidity are major factors determining virus inactivation in the environment. Temperature and relative humidity have been shown to impact the kinetics of inactivation of coronaviruses. Increased temperatures have been shown to increase the rate of the inactivation, and decreased relative humidity have been associated with a reduction of coronaviruses inactivation rate on surfaces. Inactivation rates were lower in suspensions compared to surfaces in studies that tested both suspensions and surfaces at similar temperatures. Hence, the prediction of the persistence of SARS-CoV-2 on fomites is essential to investigate the importance of contact transmission.
what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?	Favipiravir is a nucleoside analog with an ability to inhibit RNA-dependent polymerase and was approved for marketing in Japan in 2014. In vitro experiments have shown that favipiravir is effective for COVID-19 and that its EC 50 is 61.88 M. To date, some clinical trials of favipiravir in the treatment of COVID-19 have been carried out in China. Recent clinical studies have found that compared with the antiviral drug arbidol, the clinical effect of favipiravir is more significant. Nucleic acid positive-to-negative time, mean antipyretic time and cough remission time were all better than those of the arbidol group. Remdesivir(GS-5734) Remdesivir was first used to treat Ebola virus, and it has completed phase 2 clinical trials. As a nucleoside analog, it can interact with RdRp, and the triphosphate form of remdesivir will compete with adenosine triphosphate, leading to delayed chain termination and inhibiting viral replication and transcription. number of in vitro studies have shown that remdesivir has inhibitory effects on a variety of human and animal coronaviruses.
	The serological tests that have recently been developed employ different methods and target either IgG or IgM or both. In an attempt to fill the knowledge gap, this systematic review summarized evidence from 38 studies involving 7848 individuals. Although the US Food and Drug Administration (FDA) has approved ELISA, LFIA, and neutralization

Appendix 2.4: White Box Testing 3

Appendix 2.4.1: Input 'Normal JSON document data'

```
Testing 000tfenb.json
INSERT INTO DOCUMENT VALUES (000tfenb, Prevalence and genetic diversity analysis
```

Appendix 2.4.2: Input 'testing.txt file'

```
Testing testing.txt
Ignored
```

Appendix 2.4.3: Input 'testing.csv file'

```
Testing testing.csv
Ignored
```

Appendix 3: Limitations of Software

Appendix 3.1: Input "what is temperature"

what is temperature

Here are the information that we have related to **"what is temperature"**.
Please type in the Question ID that you are interested in into the chat to view the full answer to the question.

Question ID: EQ002
Question Text: how does the coronavirus respond to changes in the weather

Question ID: EQ005
Question Text: what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?

Appendix 3.2: Input "temperature"

temperature

Here are the information that we have related to **"temperature"**.
Please type in the Question ID that you are interested in into the chat to view the full answer to the question.

Question ID: EQ002
Question Text: how does the coronavirus respond to changes in the weather

Question ID: EQ005
Question Text: what drugs have been active against SARS-CoV or SARS-CoV-2 in animal studies?

Question ID: CQ002
Question Text: how does the coronavirus respond to changes in the weather