Testing Document  
*UWA Biosecurity Game*  
By: Joel Dunstan  
Email: [jd427244@gmail.com](mailto:jd427244@gmail.com)  
biosecurity.are.uwa.edu.au

Table of Contents

[Objectives and Test Summary 3](#_Toc481344785)

[Testing Strategy 3](#_Toc481344786)

[Automated Testing 3](#_Toc481344787)

[Hardware and Manual Testing 5](#_Toc481344788)

[Test 1: The Frequency of Outbreaks 5](#_Toc481344789)

[How an Outbreak is Determined 6](#_Toc481344790)

[Test Specification 7](#_Toc481344791)

[Test Description 7](#_Toc481344792)

[Test Analysis Report 8](#_Toc481344793)

[Test 2 – Automated Command Line Testing 12](#_Toc481344794)

[Test Specification 12](#_Toc481344795)

[Test Description 12](#_Toc481344796)

[Test Analysis Report 13](#_Toc481344797)

[Test 3 – Automated Browser Bot Testing 17](#_Toc481344798)

[Test Specification 17](#_Toc481344799)

[Test Description 17](#_Toc481344800)

[Test Analysis Report 20](#_Toc481344801)

[Test 4 – Manual Browser Testing 21](#_Toc481344802)

[Test Specification 21](#_Toc481344803)

[Test Description 22](#_Toc481344804)

[Test Analysis Report 26](#_Toc481344805)

[Test 5 – Server Performance 28](#_Toc481344806)

[Test Specification 28](#_Toc481344807)

[Test Description 28](#_Toc481344808)

[Test Analysis Report 30](#_Toc481344809)

[Test Materials 32](#_Toc481344810)

[oTree\_Tags 33](#_Toc481344811)

[Glossary 34](#_Toc481344812)

[References 36](#_Toc481344813)

**Objectives and Test Summary**

This document presents a test of the biosecurity game system, this includes the website and the program underlying the website biosecurity.are.uwa.edu.au. Testing will focus the following

1) How outbreaks are determined and the amount of times outbreaks occur.  
2) The general usability and functionality of all the pages and how they work together  
3) The websites compatibility across systems, web browsers especially due to its use of HTML5 specific code used within the webpages, and its use of Django, a Python framework.  
4) The performance of the server, its backup strategies and recovery plans. Focusing on how many users the server can handle and creating a session and running a session with many participants.  
5) The export of the data from each game for experimental purposes and its readability.  
6) The use of the above components operating as a system.

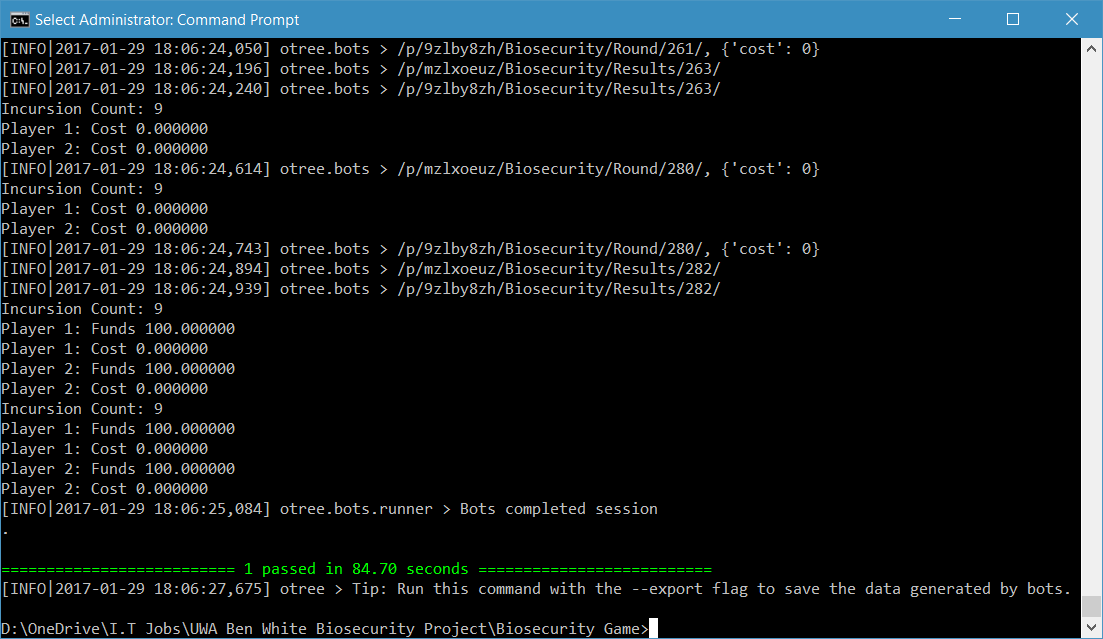
**Testing Strategy**

The overall strategy is a bottom-up approach test the raw code itself, using the command line (Refer to [Automated Testing](#_Automated_Testing) for more detail) The next stage uses browser bots to ensure that code works within a web browser (While checking its compatibility with different web browsers). Finally, the game is run manually across different browsers.

### Automated Testing

Automated Testing is used in most areas of testing using the tests.py from every application inside the oTree Project, the project was divided into three applications, Lottery Game, Biosecurity, and Results. These tests can be run from the command line using oTree test <session\_config\_name> to conduct tests with a different number of participants. An example of such a test would be *otree test basic\_biosecurity\_test 4* which means that the command line will use the test.py from the application Biosecurity and run the session configuration ‘basic\_biosecurity\_test’ using 4 participants.

The automated tests do verification checks to see if the game produces the correct results given costs, player approvals, etc. If the game completes a test, then it indicates that the game ran without an AssertionError which in oTree occurs whenever an assert statement failed or a game failed to finish due to oTree not progressing through the game (it doesn’t progress to the next page with a form). In the case of a test being completed successfully then this will appear on the command line at the end of the test:



For the Lottery Game, automated testing double checks that the quiz for testing people’s understanding of the game is working correctly giving an incorrect answer to see if it will fail, and then giving the correct answers. For the actual Lottery Game, itself, it randomly chooses A or B for the forms to make sure the data goes in through the form as there is no tangible way to verify that a person gets an exact result, due to the random selection of the winning game for every single session.

In the Biosecurity Game, automated testing applies different scenarios, the first being random where for every single form it produces a random value and performs checks on these values by replicating the same equations used in the code of the game and then comparing the results of the game with the results of the test using assert statements. To be specific the values that are tested throughout the game are the costs of protection for every single round, the funds at the end of each round, any group pledges that were performed, and finally the approvals by players. In the random scenario, we use random values for every single form required. The automated tests conduct other scenarios in the order mentioned below, which each scenario will show the values used for each of the values being tested above:

Quarter: 25% Max Protection for the cost of protection, funds are dependent on whether there was an outbreak or not and hence random and will not be mentioned further, group pledging is done randomly and the approvals are always -6.

Half: 50% Max Protection for the cost of protection, group pledging is done randomly and approvals are always 3.

Three-Quarters: 75% Max Protection for the cost of protection, group pledging is done randomly and approvals are always done at 6.

Full: Max Protection for the cost of protection, group pledging is random and approvals are always 0.

Half0HalfFull: Half of the bots do Max Protection, while the other half do no protection whatsoever, group pledging and approvals are random.

Bankrupt: No protection at all, group pledging and approvals are random.

There are default values used for every test throughout, the maximum amount of protection that can be used is always 15, the cost of upkeep is always 5, the revenue is always 25, the starting amount for every player is 25, the minimum probability that someone is not the source of the outbreak is 60% (refer to [Test 1: The Frequency of Outbreaks](#_Test_1:_The)), the number of participants is 4 and all participants are in a single group. Finally, a pledging round comes every 3 rounds and every round is when an approval by contribution takes place.

Finally, the Results has a survey to which each bot automatically completes and just waits for correct pages to come forth.

The same code that lies within the test.py for every application that does the command line testing also does the browser bot testing and therefore it’s safe to conclude that if a browser bot test completes without a fault, then the game’s logic is compatible with the browser.

### Hardware and Manual Testing

Manual testing involves running the game manually instead. Manual Testing is used during compatibility testing to verify that each form in each web browser works as expected. Manual Testing will include verifying the data produced by oTree every time a session is run, this will be the least tested part of the entire program as it simply loads data from a database which I don’t have much control of as oTree by design automatically puts the data into the database.

Hardware testing will be conducted to see how the server performs under load. To put strain on the server the game is run with aa number of participants that would far exceed the usual number playing the game which is less than 10.

## Test 1: The Frequency of Outbreaks

The frequency of outbreaks in the game is determined randomly by the joint probability of protection. This test involved changing the maximum amount of protection against biosecurity threats a player could provide in each round and the probability coefficient, or the minimum probability that a player is **not** the source of the outbreak. The determination of the probability of an outbreak is described in the following section.

### **How an Outbreak is Determined**

**MP** = **Max Protection =** (The maximum amount of protection a participant can put against biosecurity threats)  
  
**PC =** **Probability Coefficient =** (Essentially we moved the graph up so that $0 worth of protection doesn’t result in a 100% chance that someone is the source of the outbreak, e.g. if PC = 0.4, then there is a 40% chance that someone is **not** the source of the outbreak, or 60% chance that they’re the source of the outbreak)  
  
**C = Amount of protection a player put in that round** ($0 ≤ P ≤ MP)  
  
**P1, P2, P3, P4 = The probability that a player (4 players in this scenario) is NOT the source of the outbreak** (0 ≤ P1, P2, P3, P4 ≤ 0.999). Will be denoted as **Px** when referring to the function for P1-4.

**RAND =** A random number between 0 and 1 generated by the server

**U =** The cost of the crops per round  
  
**R =** Revenue which a player receives when there is no outbreak  
   
1) Get the Cost Factor:

**CF =**

2)Determine the probability that a player is the source of the outbreak:

**Px =**3)Determine the probability of no outbreak for the group (the set of players containing P1­ – P4):

**Probability of No Outbreak (PNO) = P1** x**P2** x**P3** x**P4**

4) Determine the probability of an outbreak for the group:

**Probability of an Outbreak (PO) = 1 – PNO**

5) Now generate a random number between 0 and 1 and follow the inequalities below:

**if PO > RAND, Outbreak Occurs**

**if PO < RAND, No Outbreak Occurs**

6) Now determine each player’s profits or losses as per the following functions:

If there was an Outbreak, then: **Current Funds = Current Funds – C – U**

If there was no Outbreak, then: **Current Funds = Current Funds – C – U + R**

### Test Specification

This test requires that the group’s actions proportionally affect how much the group gains from the game. That means:

*As a group if each player is cooperative, then we must see that on average, each player should receive a sizable profit compared with starting amount, while still having outbreaks to prevent a lack of engagement from the game being too easy and the result being independent to the group’s actions.*

*Inversely, as a group if each player is uncooperative, then we must see that on average, each player should receive a loss from the game compared with the starting amount, while still allowing players to ideally to finish with an amount more than $0 to prevent lack of motivation and engagement in the game***.**

It is worth mentioning too, that while the above is a requirement, it won’t always be met due to the nature of joint probability and the random determination of outbreaks. The game should allow cases where players are *cooperative* and outbreaks still occur often. Inversely, the game should also allow cases where players are *uncooperative* and outbreaks don’t occur often. Such cases in this game should be rare though to get the best data possible, as the player’s actions must determine the outcome as much as possible despite the randomness that exists in the game.

### Test Description

To perform this test, you will need a terminal (CMD, Bash) to perform the command line tests and export the data to a file (ideally, text, csv). The operating system you use, shouldn’t matter as long as you have a met these requirements (at the time of writing 02/04/2017):

oTree 1.2.8 (or above)  
Postgres 9.6.1 (or above)  
otreechat (latest version)  
psycopg2 (latest version)  
Python 3.6 (or above)  
  
Follow the instructions for oTree to use Postgres as per this page:  
<http://otree.readthedocs.io/en/latest/server/ubuntu.html#database-postgres>

Postgres here isn’t completely necessary to complete these tests, you could theoretically use MySQL, MSQL (Microsoft SQL) and many others, you must ensure it’s a Relational Database that can withstand multiple concurrent transactions without issue.

You should also refer to the notes on oTreee tags: [oTree\_Tags](#_oTree_Tags).

In the terminal (in my case PowerShell), you’ll need to put the following command to perform the test 20 times for the sake of getting a reasonable sample size to reduce the average being affected by any outliers.

for($i=1; $i -le 20; $i++) {

otree test basic\_biosecurity\_test 4 > “Path\To\File\file<i>.txt”

}  
  
This will run the basic biosecurity test 20 times which includes all the different case scenarios which were as follows (and in the order, they will be performed):

* Random
* Quarter
* Half
* Three-Quarters
* Full
* Half0HalfFull
* Bankrupt

For more details on what these scenarios include refer to [Automated Testing](#_Automated_Testing). The above code will also produce text files which you will need to record the incursion count and the funds each player has at round 5 and the end of the game. An Excel spreadsheet was used to calculate the average and mode of outbreaks (denoted as incursions in my documents as there was a change of terminology) and funds at rounds 5 and 15.

### Test Analysis Report

In the tests, there was two values that changed throughout the tests which were:

1. The minimum chance that a player wasn’t the source of the incursion, or the **Probability Coefficient (PC)** as per [How an Outbreak is Determined](#_How_an_Outbreak).
2. The Maximum amount of protection or effort that a player can put against biosecurity threats, or the **Maximum Protection (MP)** as per [How an Outbreak is Determined](#_How_an_Outbreak).

The reason that these values were adjusted is because these values adjust the frequency of outbreaks and how much funds each player finished with more than any other values. While one could adjust the revenue each player, $25 is a good value as a player can gain that amount and just as easily lose it with carelessness, any higher the game would be too easy, any lower too hard. As one will see from my results the same rings true for both **PC** and **MP**.

The defaults across all games were as follows:

* $5 for Upkeep **(U)**
* $25 for Revenue **(R)**
* 4 Participants
* $0 as the Minimum Amount of Protection
* 15 Rounds, thus we take the averages below as the average number of outbreaks occurring within 15 rounds.
* The Starting Funds is $25 for every player.

**The Average Outbreak Count**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario ↓ | **MP PC →** | **MP =** $15 **PC =** 40% | **MP =** $15 **PC =** 50% | **MP =** $10 **PC =** 60% | **MP =** $10 **PC =** 70% |
| Random | | 10.25 | 9.2 | 7.9 | 6.4 |
| Quarter | | 13 | 12.05 | 10.95 | 8.7 |
| Half | | 9.7 | 8.65 | 8 | 6.4 |
| Three-Quarters | | 5.6 | 5.45 | 4.45 | 3.65 |
| Full | | 0 | 0.05 | 0 | 0 |
| Half0HalfFull | | 12.75 | 11.35 | 9.7 | 7.8 |
| Bankrupt | | 14.8 | 13.6 | 13.1 | 11.9 |

**Mode – Outbreak Counts**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario ↓ | **MP PC →** | **MP =** $15 **PC =** 40% | **MP =** $15 **PC =** 50% | **MP =** $10 **PC =** 60% | **MP =** $10 **PC =** 70% |
| Random | | 10 | 10 | 9 | 6 |
| Quarter | | 12 | 13 | 12 | 9 |
| Half | | 8 | 9 | 7 | 7 |
| Three-Quarters | | 4 | 5 | 4 | 3 |
| Full | | 0 | 0 | 0 | 0 |
| Half0HalfFull | | 13 | 12 | 10 | 6 |
| Bankrupt | | 15 | 13 | 14 | 13 |

**Average Funds at Round 5 - Frequency of Outbreaks**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario ↓ | **MP PC →** | **MP =** $15 **PC =** 40% | **MP =** $15 **PC =** 50% | **MP =** $10 **PC =** 60% | **MP =** $10 **PC =** 70% |
| Random[[1]](#footnote-2) | | 5.26 | 11.30 | 43.51 | 47.93 |
| Quarter | | 0 | 3.75 | 23.75 | 48.75 |
| Half | | 6.25 | 7.5 | 40 | 36.25 |
| Three-Quarters | | 25 | 20 | 52.5 | 60 |
| Full | | 50 | 50 | 75 | 75 |
| Half0HalfFull[[2]](#footnote-3) | | -15 | -15 | 16.25 | 49.38 |
| Bankrupt | | 1.25 | 8.75 | 16.25 | 25 |

**Average Funds at Round 15 - Frequency of Outbreaks**

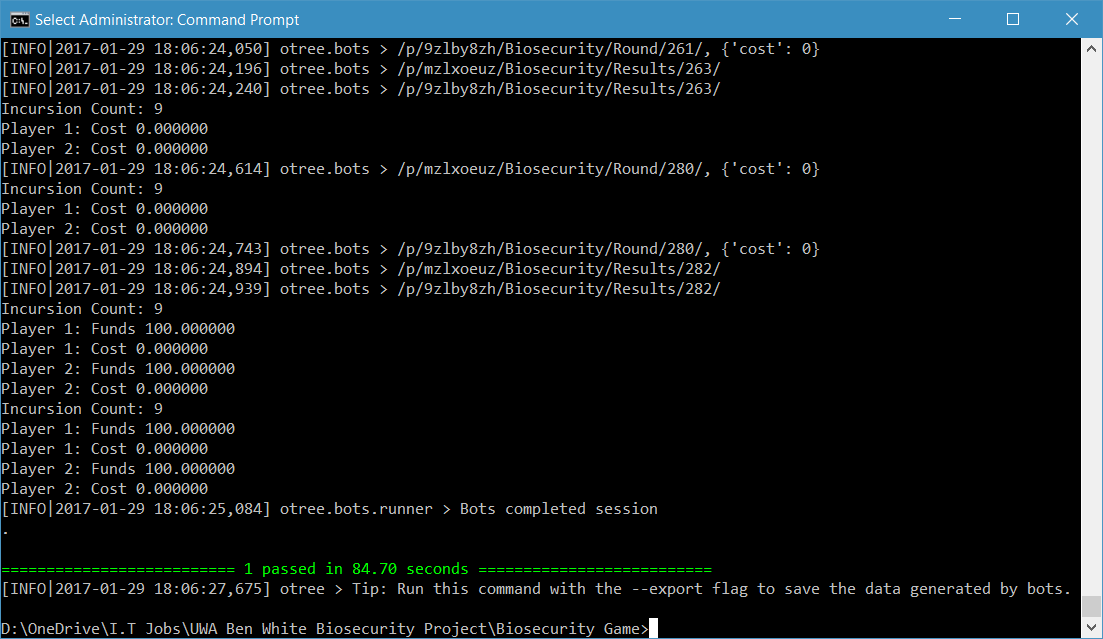
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario ↓ | **MP PC →** | **MP =** $15 **PC =** 40% | **MP =** $15 **PC =** 50% | **MP =** $10 **PC =** 60% | **MP =** $10 **PC =** 70% |
| Random[[3]](#footnote-4) | | -43.525 | -19.18 | 52.46 | 92.45 |
| Quarter | | -56.25 | -32.5 | 13.75 | 70 |
| Half | | -30 | -3.75 | 50 | 90 |
| Three-Quarters | | 16.25 | 20 | 101.25 | 121.25 |
| Full | | 100 | 98.75 | 175 | 175 |
| Half0HalfFull[[4]](#footnote-5) | | -106.25 | -71.25 | 7.5 | 48.13 |
| Bankrupt | | -45 | -15 | -2.5 | 27.5 |

The purpose of exploring the cost and probability parameter space was to determine a set of parameters that provided sufficient incentive to engage in the game over 15 rounds without making the game too easy to generate funds. The higher is the PC parameter, then the lower on average is the number of outbreaks with a random selection over the range from PC up to PMAX. PMAX was not varied, so PC became critical parameter. That determines the lower bound on the pobablity of production a PC4.

To conclude, the final configuration chosen for **MP** and **PC** were $10 and 60% respectively, the averages and modes indicate that with players who cooperate will *most likely* be rewarded and players who are uncooperative will *most likely* suffer. With this configuration, we can rely on the data to reflect the player’s actions and their thought processes behind those actions.

## Test 2 – Automated Command Line Testing

The purpose of this test is to verify that the logic of the game is stable and not capable of failure. To do this, tests are performed within a terminal which go through the scenarios as mentioned in [Automated Testing](#_Automated_Testing). To pass the test the game must be able to complete tests for all the different configurations of the biosecurity game. A test is complete when you observe the following inside a terminal:



### Test Specification

To complete the test and ensure the logic of the game is working, you must have a computer that has the following installed:

oTree 1.2.8 (or above)  
Postgres 9.6.1 (or above)  
otreechat (latest version)  
psycopg2 (latest version)  
Python 3.6 (or above)

The operating system should notmatter as Python and oTree do not have issues with compatibility across operating systems. Also ensure that you have followed the instructions as per [oTree Tags](#_oTree_Tags).

Upon the completion of this test, every single function and class written inside the program will have been tested multiple times, including the test code itself. The only requirement of this test is that it **must** complete every test.

### Test Description

The steps to perform the test are as follows:

1. Have a clone of the repository.
2. Inside a terminal and navigate to the directory where you stored the repository.
3. Run the following command:  
   otree test basic\_biosecurity\_game > basic\_biosecurity\_game\_cmd\_test.txt
4. Once completed wait for the test to finish and ensure the text file was made (it doesn’t have to be a text file, you could export it a different way if you have knowhow)
5. Now repeat 3) and 4) but with the following commands:  
   otree test freeform\_biosecurity\_game > freeform\_biosecurity\_game\_cmd\_test.txt

otree test monitoring\_biosecurity\_game > monitoring\_biosecurity\_game\_cmd\_test.txt

otree test pledging\_biosecurity\_game > pledging\_biosecurity\_game\_cmd\_test.txt

otree test mon\_pledging\_biosecurity\_game > mon\_pledging\_biosecurity\_game\_cmd\_test.txt

otree test aop\_biosecurity\_game > aop\_biosecurity\_game\_cmd\_test.txt

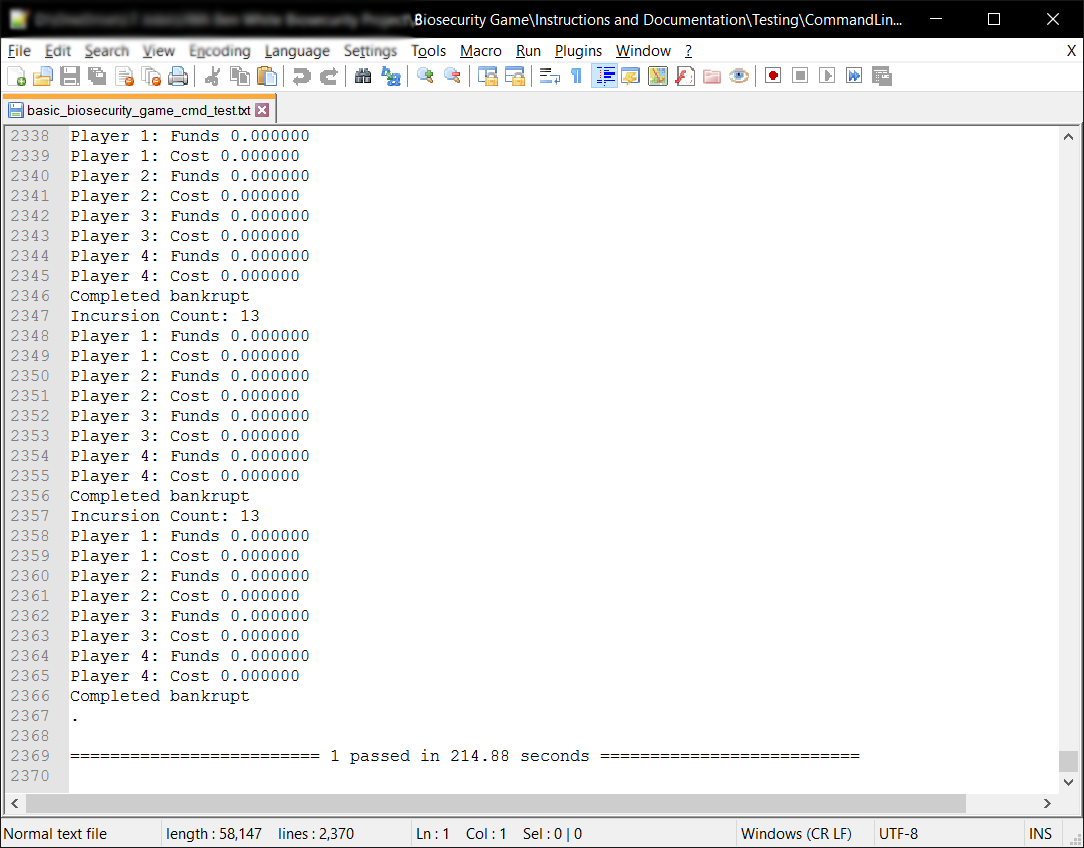
otree test aoc\_biosecurity\_game > aoc\_biosecurity\_game\_cmd\_test.txt

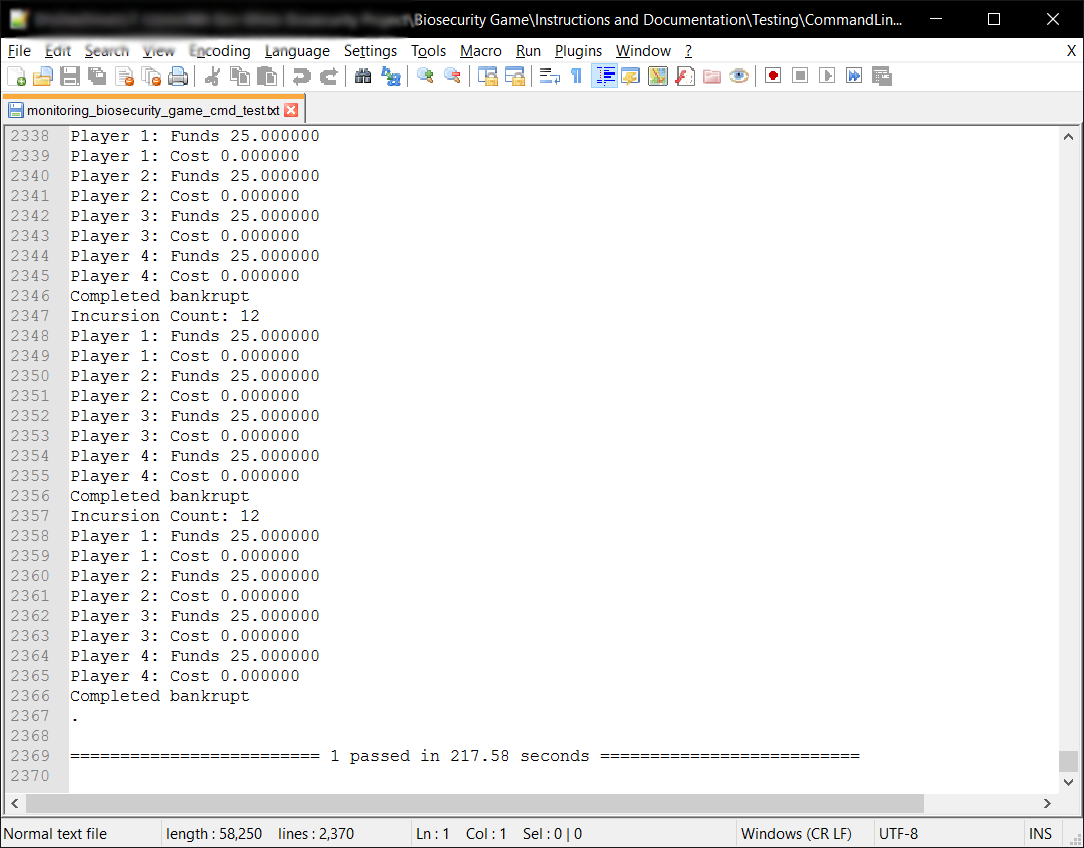
Once you have completed these steps and assuming you received no errors in your testing, then the testing is complete. If the testing failed, then it will most likely be one of 3 errors:

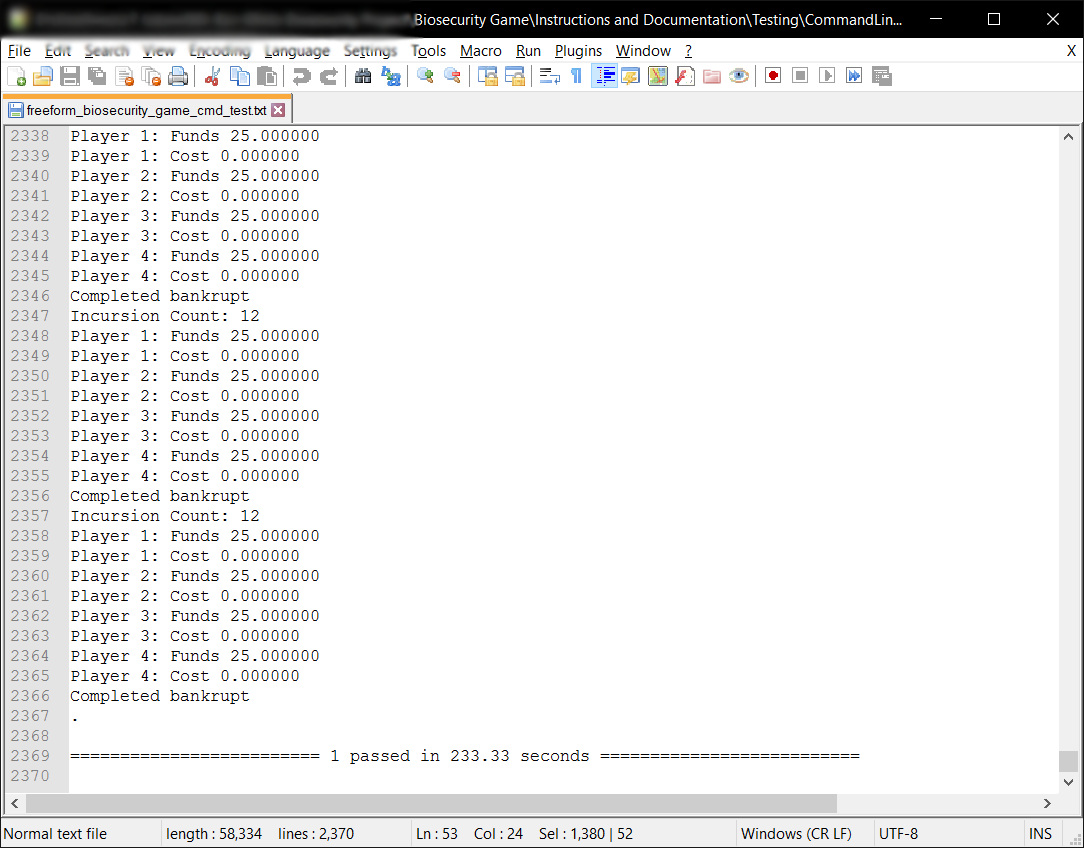
1. You haven’t followed the instructions in [oTree Tags](#_oTree_Tags)
2. It has generated an Assertion Error, in the traceback it should reveal which assert statement failed. Upon finding out which assert failed, debug the code to see what could have caused the Assertion Error.
3. You are running a future version of Python or oTree, its possible some features being used in this program might be deprecated and therefore no longer work with the version of Python or oTree that you are running.

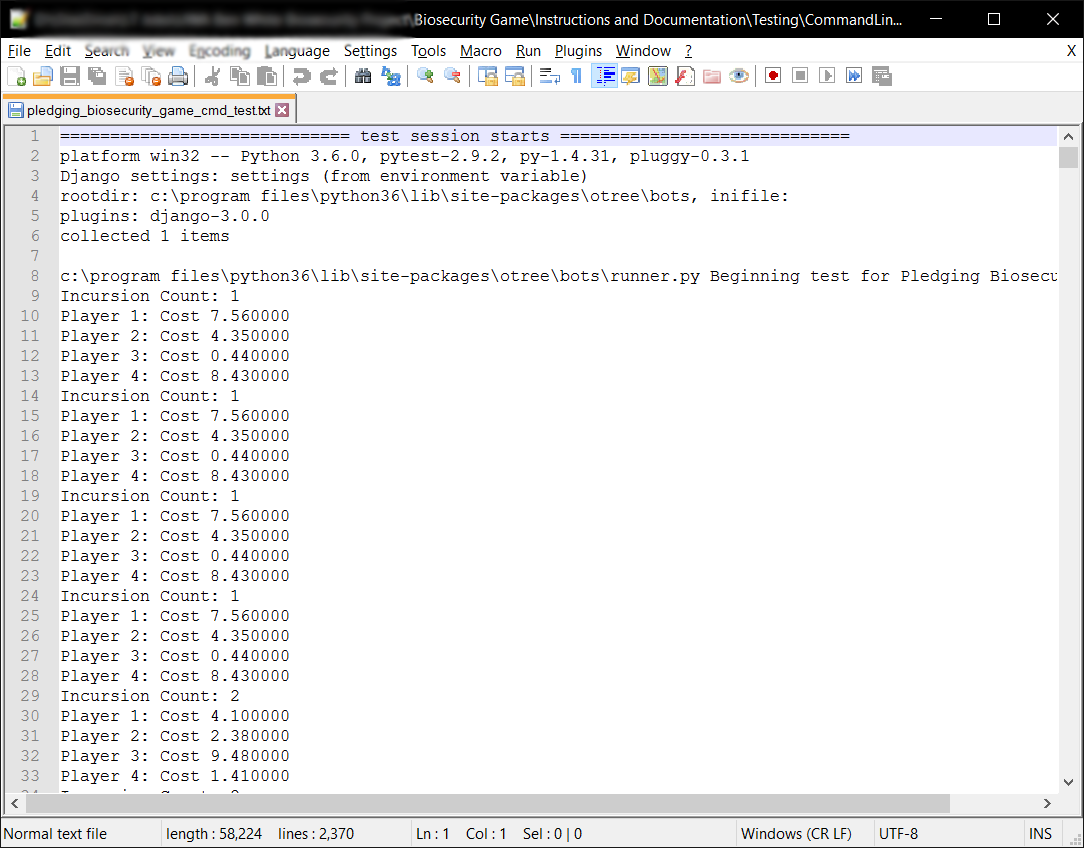
### Test Analysis Report

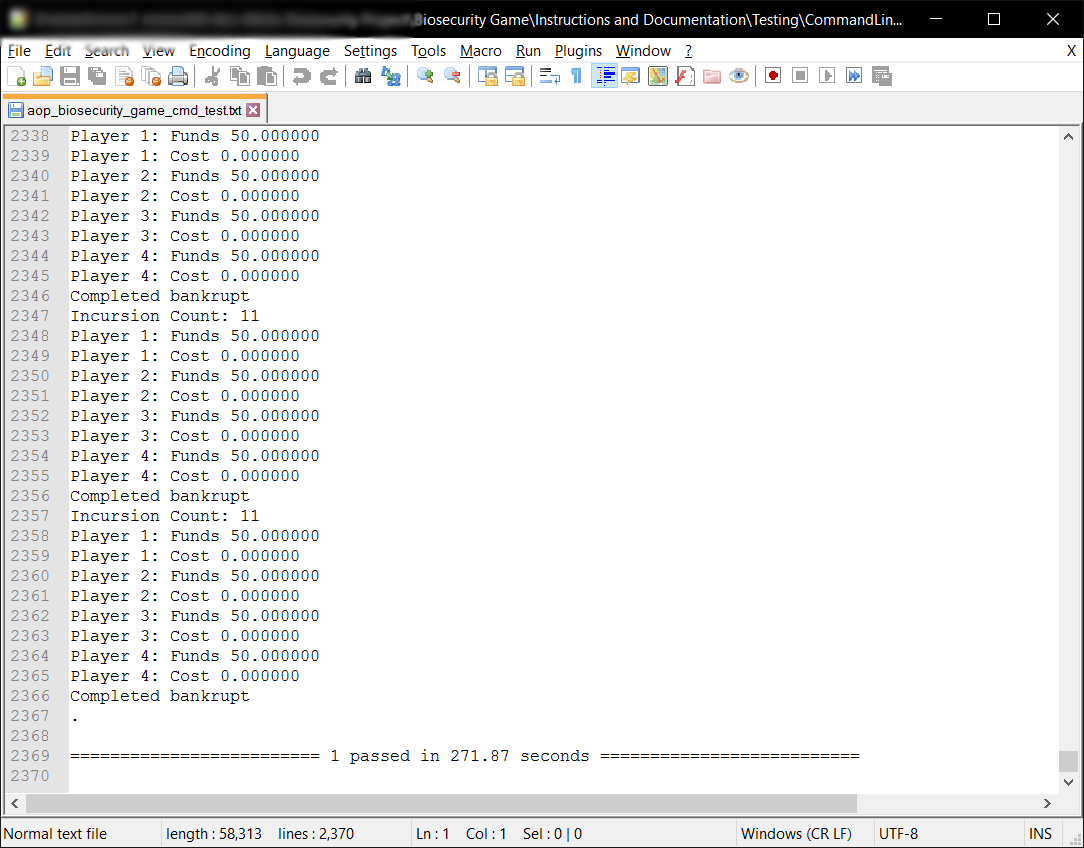
My testing was successful, my proof is below and inside the text files I made. The images below are from the text files, which have recorded that the test passed and the amount of time it took to pass the test.

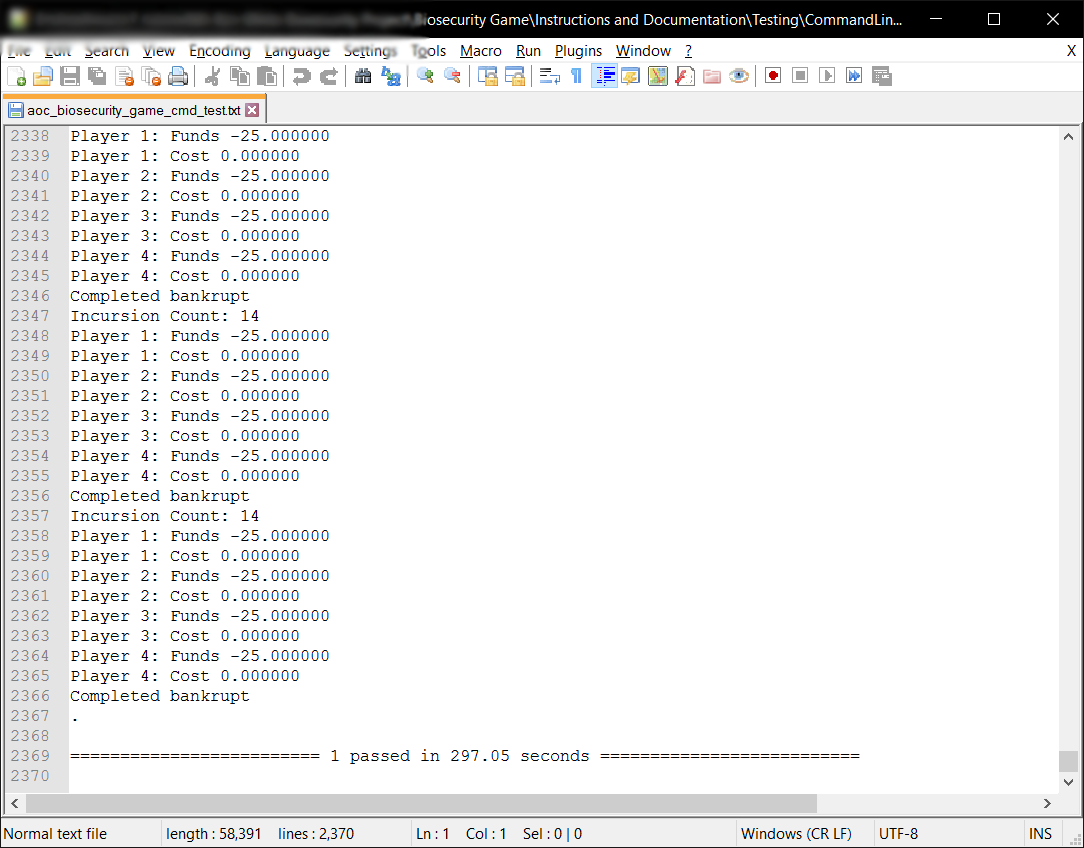












As the above proves, every single test passed and completed without issue.

## Test 3 – Automated Browser Bot Testing

The purpose of this test isn’t necessarily to test the python code or the logic of the game *per se*, it is to test the compatibility of Django, Python and oTree across multiple browsers and different devices. This test will reveal any issues the game has with specific web browsers and operating systems. In theory, oTree, Django and Python, and therefore the game should be able to run on any web browser, on any device, or is this the case or not?

### Test Specification

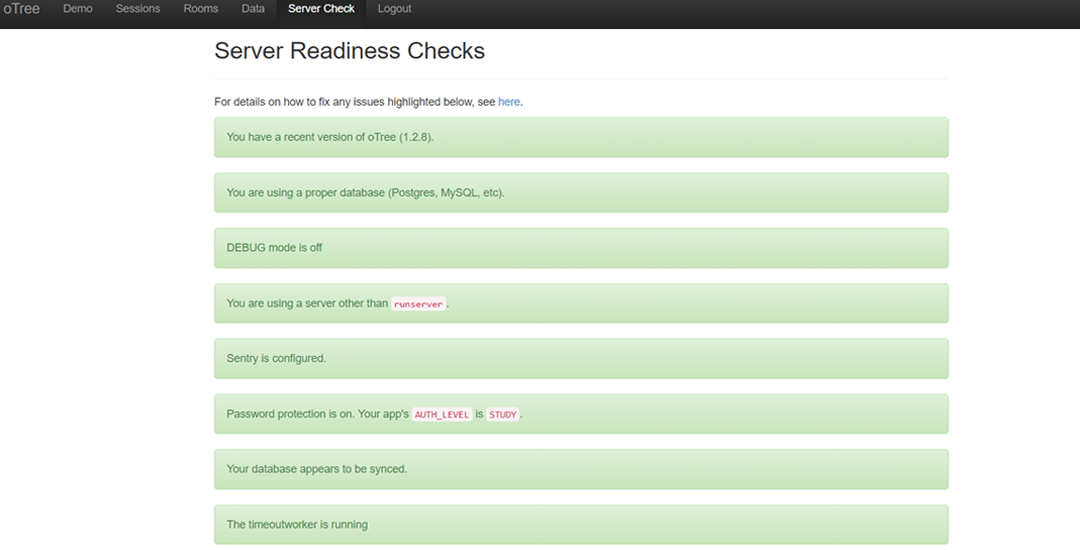
To perform this test, you need to have the following:

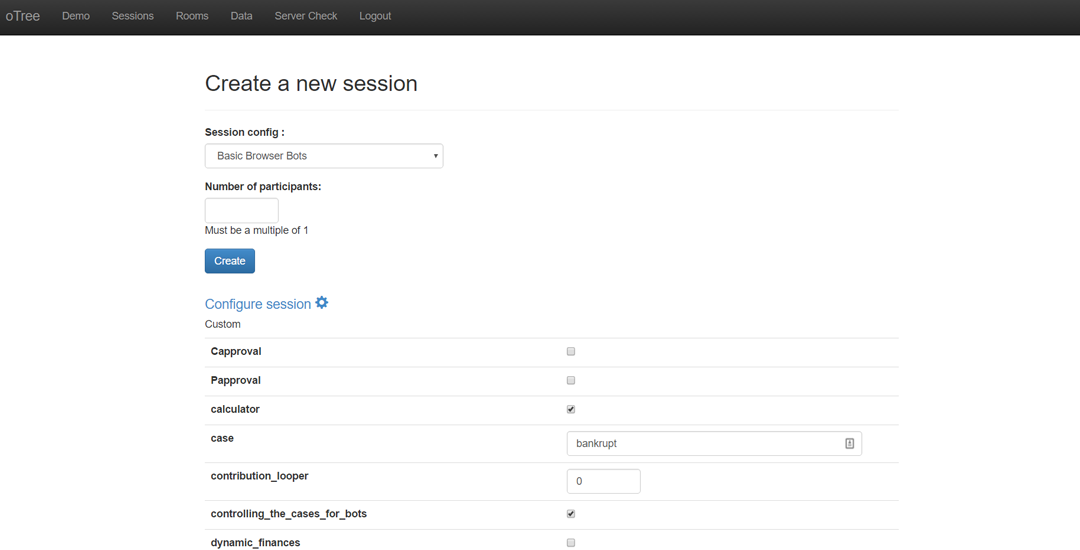
1. A web server that has completed the following test [Automated Command Line Testing](#_Test_2_–) to its requirements. This web server needs to be running the oTree program via otree runprodserver -p 80 (This will run the program and a Daphne web server that will put all input and output through port 80, or HTTP). You can, once you have gained my permission, run this test on biosecurity.are.uwa.edu.au.
2. Ideally you want to have as many different devices as possible, with each device having as many web browsers as you think appropriate. For my testing, I used a Windows 10 PC with Google Chrome, Mozilla Firefox, Opera, Internet Explorer and Microsoft Edge; a Nexus 6P an Android smartphone running Nougat 7.1.1 with Google Chrome, Mozilla Firefox and Opera; an iPad Mini 2 running iOS10.2.1 with Safari, Google Chrome, Mozilla Firefox and Opera. My test covered at least 75% of the market, as Android, iOS and Windows Users use the above web browsers with Google Chrome being the most used browser of them all (Anon., 2017; Anon., 2017).

To complete this test successfully, you ideally want to have tested all the browsers you have chosen using the current market share of your time to decide what browsers to test. It is not a requirement of the test that it work on **every** web browser, the **only** requirement is that the test works on the browsers **you wish to support**. In my case I want to support Google Chrome in any version, Safari, Firefox and Microsoft Edge, thus it was a requirement that these web browsers are compatible with the game.

### Test Description

The steps to complete the test are:

1. Navigate to the web site that you are using to test this program, or use biosecurity.are.uwa.edu.au. You could run this locally, in which case navigate to 127.0.0.1:8000, unless you used the -p argument with otree runprodserver in which case navigate to whatever port you assigned for it.
2. Once here, go to ‘Server Check’ located at the top of the administration page and ensure that you are running using the most recent version of oTree (it was oTree 1.2.8 in my case), using a proper database, not using runserver (using otree runprodserver when launching the server), the database must be synced and timeoutworker is running. Observe the following image below, the green details as labelled by red arrows are the required settings to be configured:
3. Now that the server is ready to go, navigate to ‘Sessions’ in the top menu. Here you will be greeted with a list of all the sessions you have created and a button *‘Create New Session’*. Click this button, now select a *Session Configuration*, choose ‘*Basic Browser Bots’* and set the number of participants to 4.
4. Here comes the most important part for testing, click *‘Configure session’*, this will reveal parameters you can change. For testing you must change the case parameter, the case will determine what test scenario it runs, for example, bankrupt will run a case scenario where every player contributes $0 of protection per round in the biosecurity game. Do *‘bankrupt’* as your first case scenario.



4

1. Create the session and open all the links which will conduct the automated tests in your selected browser. Let them continue until they reach a page where it simply says “*Bot Completed”*.
2. To complete testing for the ‘Basic Browser Bots’, you must repeat steps 3-5, except you must run all these case scenarios successfully (in no specific order and **case specific,** you cannot do **bankrupt** as **Bankrupt**):
   1. bankrupt – Every player does $0 protection every round. You should have just completed this one.
   2. half0halffull – Half of the players do maximum amount of protection, half do $0 every round.
   3. full – Every player does maximum amount of protection every round.
   4. threequarters – Every player does 75% of maximum protection every round.
   5. half – Every player does 50% of maximum protection every round.
   6. quarter – Every player does 25% of maximum protection every round.
   7. random – Every player does a random amount of protection ($0 – Max Protection) every round.
3. Now complete steps 3-6, except choose the following *Session Configurations*:
   1. Freeform Communication Browser Bots
   2. Approval on Pledges Browser Bots
   3. Approval on Contributions Browser Bots
4. Now that you have completed all the Session Configurations, save the data using the Data Export Page. Navigate to ‘Data’ on the top menu in the oTree administration page, and take the excel or csv sheet per app or through all-apps.

The test is completed to its requirements once you have conducted steps 1-8 on the browsers and the systems those browsers are running. This test will take a fair amount of time to complete, here’s some tips to complete it faster:

* You can run one browser bot session and complete the test for multiple browsers by opening (or copying) the links in different browsers across different devices. This is a lesson I learned in doing the manual tests, as iOS only allows 2 active tabs at the same time, meaning that to perform an automated test on an iPad you must constantly switch tabs. This resulted in much manual work for a supposedly automated test, which at the time meant that I didn’t complete the automated testing for iOS and resorted to Manual Testing for the iPad.
* You could create all the sessions first with their appropriate cases, and then proceed to run the browser bots for different cases simultaneously, although, be warned this will cause the web server to slow down. In my testing, which will be examined later, I could run around 20-30 participants or bots simultaneously.
* If the browsers can’t even complete the *Basic Browser Bots* test successfully, chances are they will not complete the rest of the Session Configurations successfully as the other configurations put more stress on the web server and the web browser (and the device its running on) alike.

### Test Analysis Report

In my testing, Windows and Android all completed their tests successfully, iOS had some difficulties as mentioned [above](#_Test_Description), namely that it could only have 2 active tabs. This limitation by iOS on web browsers meant that every time you opened 4 links, only 2 of those links could be doing something in their respective processes. This resulted in me having to switch tabs constantly on iOS, in my case this was too much time to complete the test this way and had to move on to manual testing.

The below tables go through Windows, Android and all the web browsers I tested on these platforms, a tick will indicate that the browser passed the test, and cross will indicate it didn’t. I will explain why each didn’t complete the test appropriately below the tables. As a reminder, the only requirement for the test was for the program to work on any web browsers that the tester (you) wishes to support. To meet this requirement, it needs to *complete every single scenario for each Session Configuration* as mentioned above in the [Test Description](#_Test_Description). Thus, in the tables below, a tick means it completed all the scenarios, in that Session Configuration.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Windows 10 | Basic | Freeform | Approval on Pledges | Approval on Contributions |
| Chrome | ✔ | ✔ | ✔ | ✔ |
| Internet Explorer 11 | 🗙 | 🗙 | 🗙 | 🗙 |
| Microsoft Edge | ✔ | ✔ | ✔ | ✔ |
| Firefox | ✔ | ✔ | ✔ | ✔ |
| Opera | ✔ | ✔ | ✔ | ✔ |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Android 7.1.1 | Basic | Freeform | Approval on Pledges | Approval on Contributions |
| Chrome | ✔ | ✔ | ✔ | ✔ |
| Firefox | ✔ | ✔ | ✔ | ✔ |
| Opera | ✔ | ✔ | ✔ | ✔ |

In these tests, I found all the browsers except Internet Explorer 11 to be compatible with the biosecurity game. The main problem with Internet Explorer 11 was that it occasionally crashed or halted during the test.

As mentioned, in the [Test Description](#_Test_Description), I didn’t do any automated tests on iOS due to the limitation on web browsers put on by iOS. The limitation was that iOS didn’t allow any more than 2 active tabs at a time. On iOS, this makes sense as it increases battery life with less processor usage, and reduces the amount of data used by the device, meaning people on 3G/4G (and potentially 5G) plans in the future will not suffer large costs. Unfortunately, in this context, it was not helpful and to complete an automated bot test, the tester would have to sit there constantly switching tabs on the device, spending just as much time (or longer) as completing a manual test. To be clear I will support iOS and the web browsers Safari and Chrome, as one will see from the Manual Testing below, they are successful and are quite good at running the program.

Finally, to discuss and compare the behaviour of browsers during these tests, on Windows 10, Chrome and Microsoft Edge were the fastest to complete their automated tests by far. With Opera and Firefox being comparable in speed to each other, but not as fast as Edge or Chrome. Thus, on Windows 10, I will support all browsers that passed all the automated tests, favouring Chrome or Edge. On Android, unsurprisingly, Chrome was the fastest browser, due to Android being an OS developed by Google, Chrome’s optimisation on Android would undoubtedly high priority. Thus, I will support Android’s Chrome Web browser, however I will advise, that people should not play this game on devices smaller than 7”, for an explanation of this refer to the [test report for the Manual Testing](#_Test_Analysis_Report).

To conclude, Google Chrome is the best web browser for the game to run on, the Manual Testing supports this, as below.

## Test 4 – Manual Browser Testing

The purpose of manual testing is to observe how the game behaves in real time, to see if the manual testing reflects the results of the automated browser bot testing. The tests also include any actions coming from administrators like moving the slowest players forward, checking the data as games progress. Mainly this test is about how the game looks and works on many different devices and browsers. It is this test that is the final stage of testing the game itself, ensuring that its working in all areas as it should.

### Test Specification

This test has the same specifications as [Automated Browser Bot Testing](#_Test_Specification), like that test you need to ensure it passes the test on the browser and devices you wish to support. To pass the test, the game must perform the following the requirements:

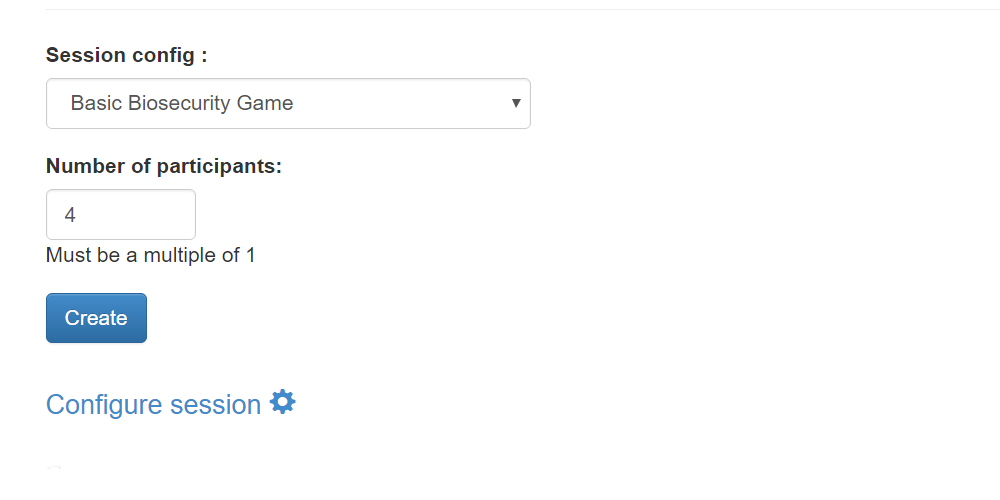
1. Complete the Lottery Game, perform at least 5 rounds of the Biosecurity Game, and display the results and Questionnaire properly at the end, with ***absolute success***, there must not be a single issue by this point.
2. The administrator must be able to use the ‘move slowest players forward’ functionality for at least 5 rounds of the Biosecurity Game.
3. The administrator can check the data while the game continues and the game dynamically updates the data appropriately, including the Payments Page.
4. To have met the previous 3 requirements in the Session Configurations:
   1. Basic Biosecurity Game
   2. Freeform Communication Game
   3. Approval on Pledges Game
   4. Approval on Contributions Game
5. Finally, have met requirements 1-4 on each device and all the browser you wish to support on that device

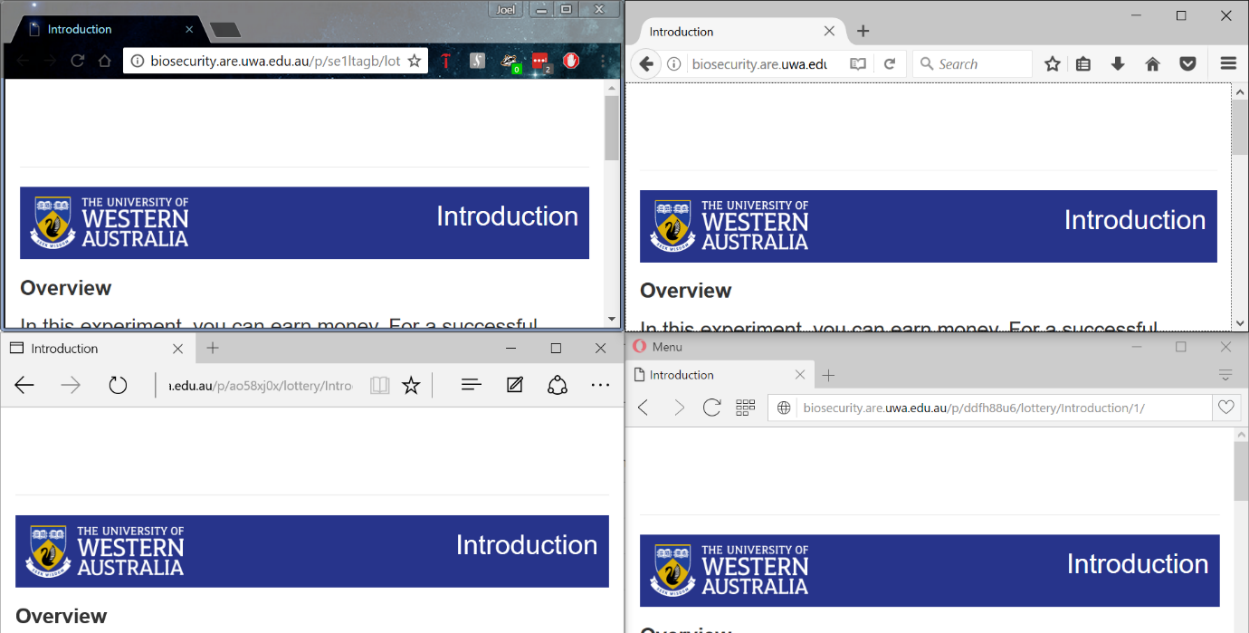
The browsers and devices I tested on are the same as the [Automated Browser Bots Tests](#_Test_Specification).

### Test Description

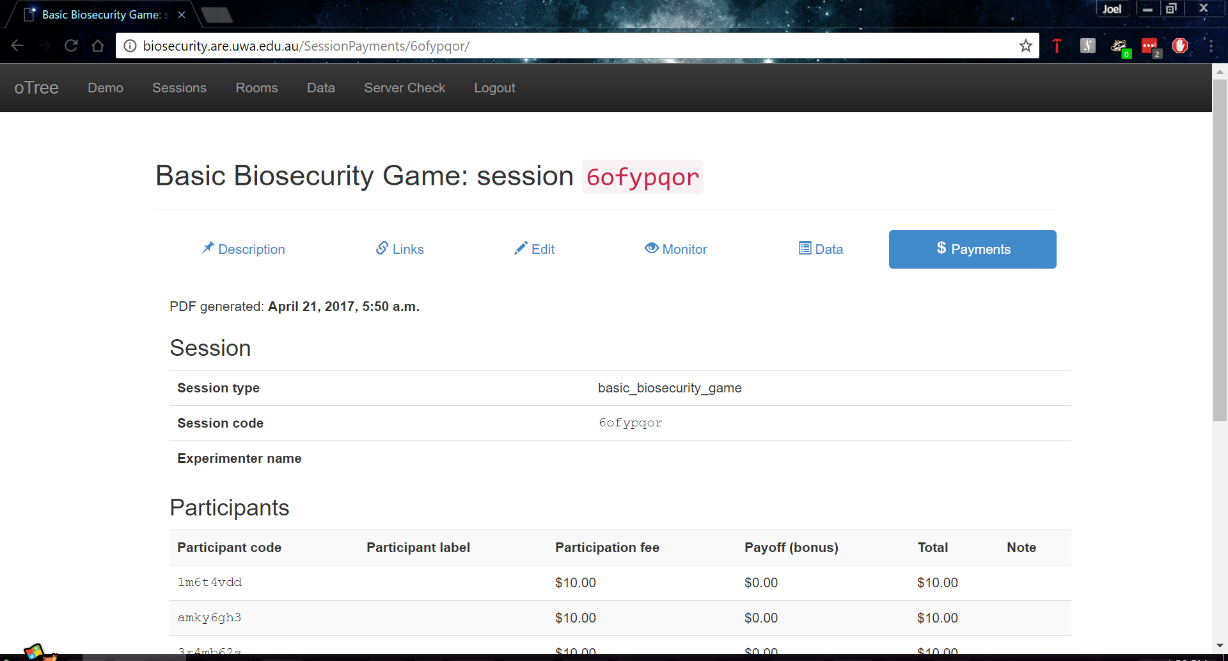
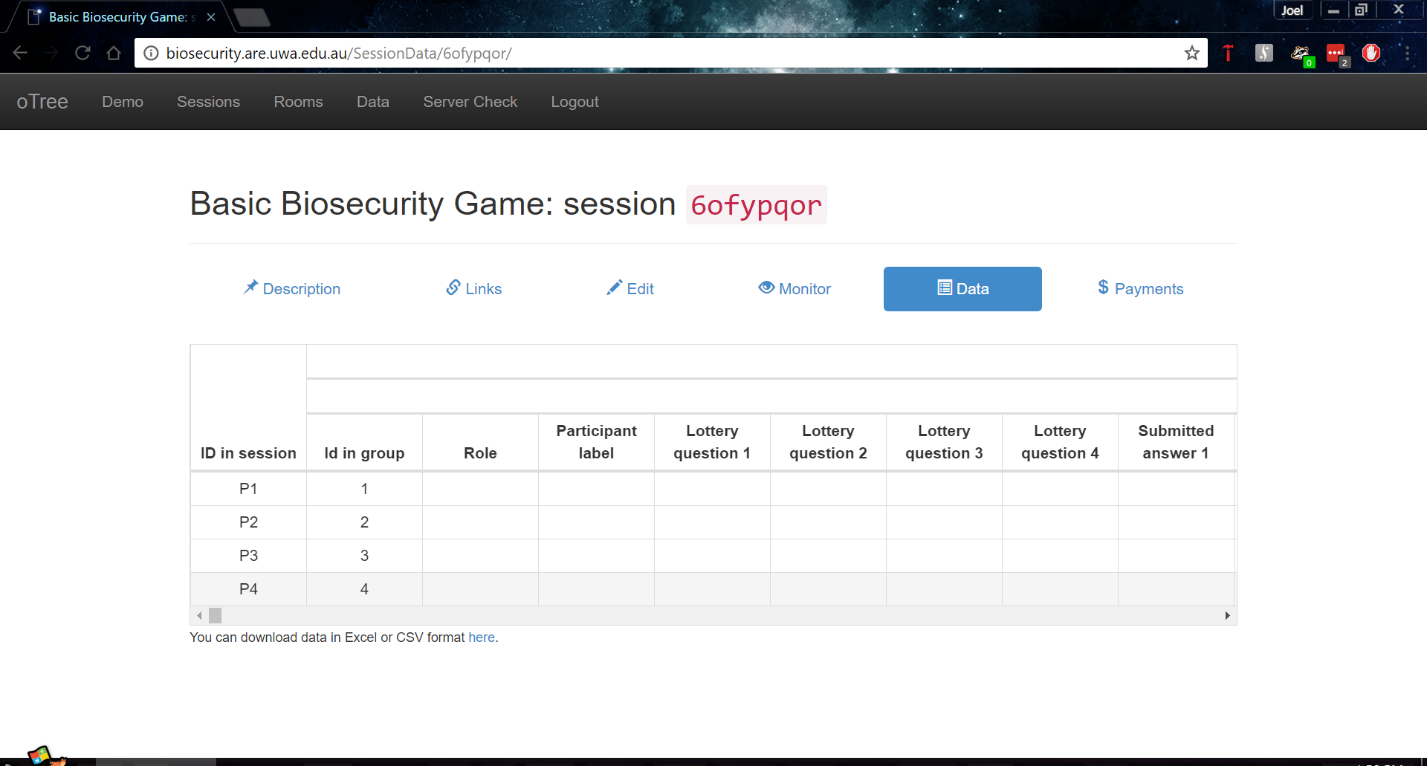
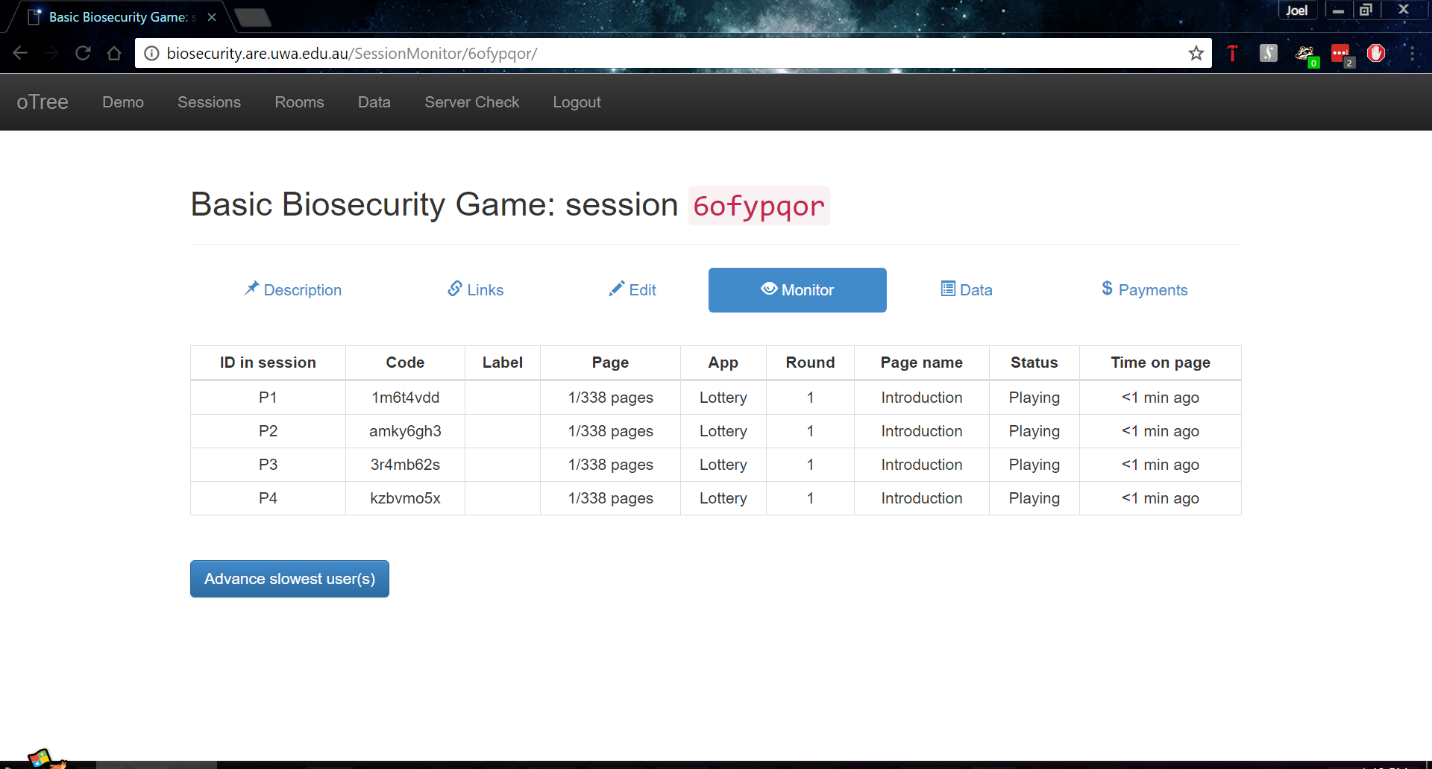
The steps to complete this test are as follows:

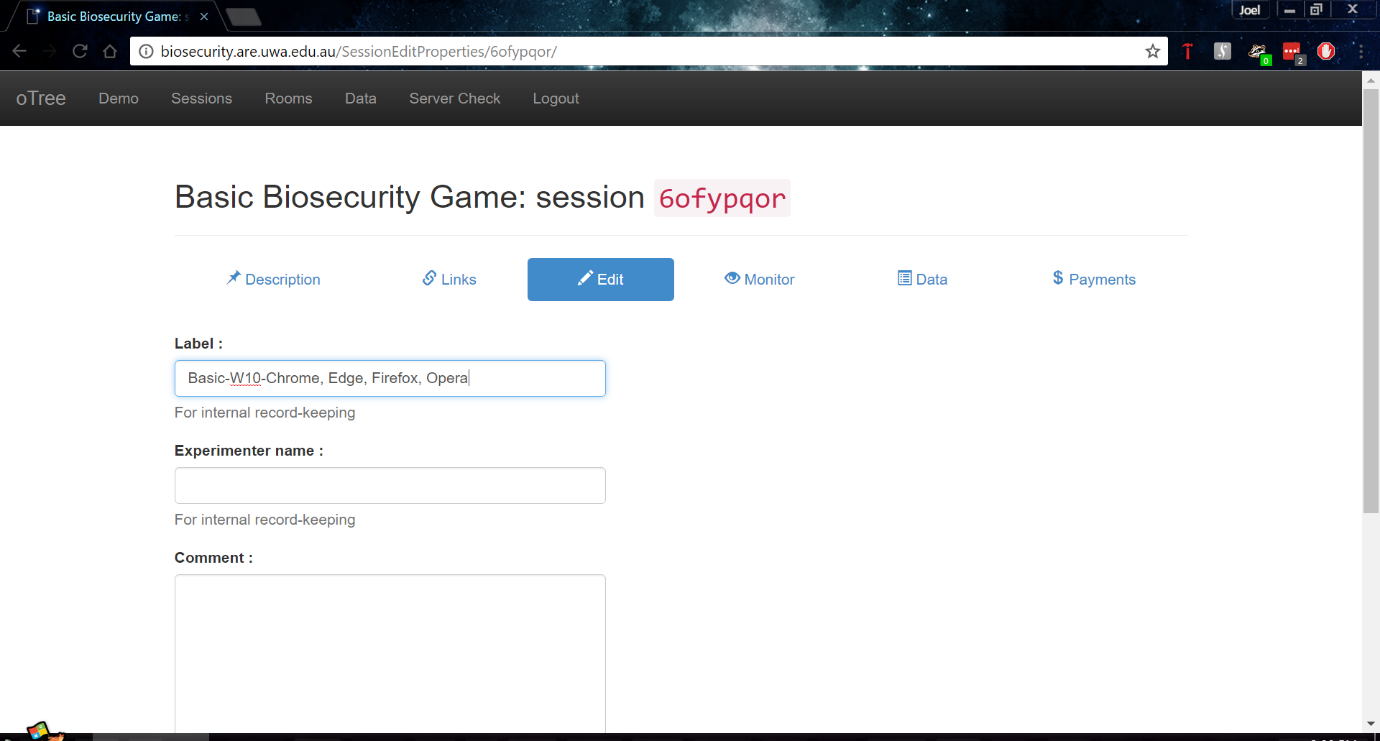
1. Navigate to the same website/server that you used to perform the Automated Browser Bot Testing, in my case this was biosecurity.are.uwa.edu.au. Also like the browser bot testing, ensure the server is still healthy as per [step 2 here](#_Test_Description).
2. Navigate to ‘Sessions’, and click ‘Create a new session’. Now in the list of Session Configurations choose ‘Basic Biosecurity Game’. Like the browser bot testing, put 4 as the number of participants and click create. Here there is no need to configure the session like the browser bot testing as we do not need to change any variables to test the game in its ‘default’ states as per the Session Configurations.

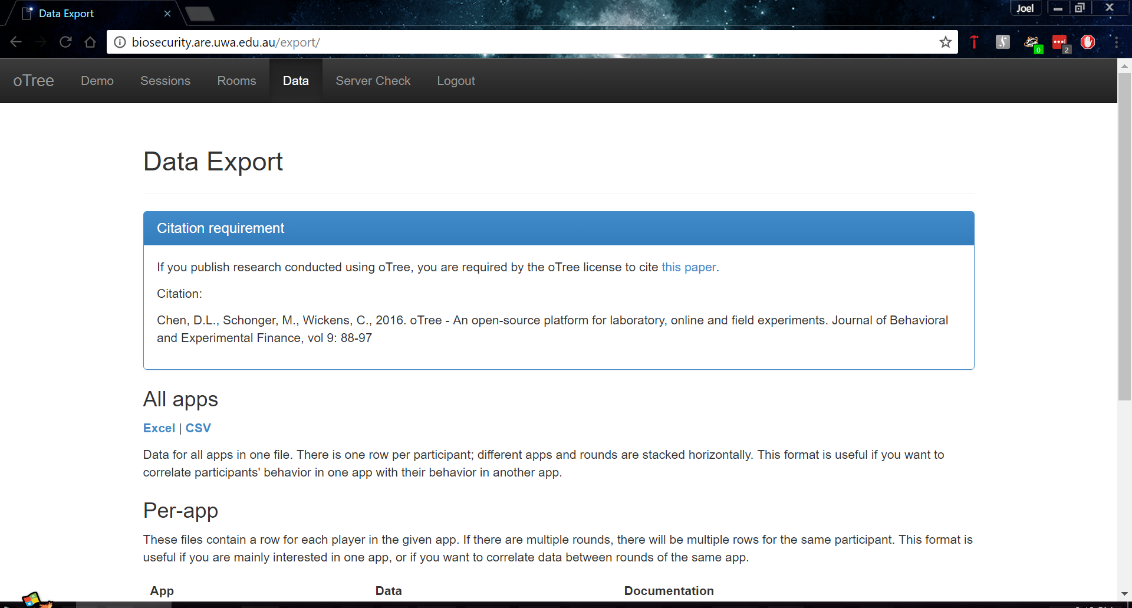


1. Now, here is when you have options as to how to perform the test, each will give you the same result even if there are issues. The first option is on each device and each browser, open the 4 links for every participant in new tabs on that one browser, meaning you will test each browser one by one with a full game. This is very time consuming, so my suggestion is to do a new browser for every player as per the following example:  
   

This method will allow you to complete your tests more quickly and more efficiently, with a bonus of testing how browsers and devices interact with each other in a real world like scenario.

1. Now play the Lottery Game, and the first 5 Biosecurity Game rounds yourself, in each round you must use different numbers in a random sequence to identify that the functions and algorithms work. To ensure that the calculations are working perfectly, use the first round of the Biosecurity Game as a chance for every player to do $0 or the maximum amount of protection to get expected and verifiable results. Test every slider, every button, observe how the elements look on screen (does font become unreadable on certain devices? Are the pictures too big for certain browsers or devices? Do the sliders work properly alongside the JavaScript that accompanies them? Etc.) ensure that the right information appears based on how you made the players play the game, for example in the instructions we don’t want chat instructions appearing when we aren’t in the Freeform Communication Game session configuration.
2. Once you have completed 5 rounds yourself, check the ‘Monitor’, ‘Data’ and ‘Payments’ page, ensure that each page displays the correct information. The ‘Monitor’ page should show which page each player is currently on and how long they’ve been on the page. The ‘Data’ page should show the data for the game as it progresses, ensure the data for the Lottery Game is there and that the first 5 rounds of the game have been recorded on the data page. The ‘Payments’ should show how much each player is owed as per the conversion rate and participation fee for the game, this page also dynamically updates with the game. Observe the following images below as to what the pages look like.  
   

1. Navigate back to the Monitor page, you should see that everyone has probably advanced a few pages while you were performing step 5. If they’re all still at the same page you left them and didn’t advance even a single page, then the timeouts might not be working as expected. If they did advance then continue the test by clicking ‘Advance slowest user(s)’ as seen on the ‘Monitor’ page, assuming you have completed all tests properly so far, clicking this will advance every user to the next page. You are to keep advancing every player till every single player reaches round 11, you will notice that not every player seems to be on the same page according to the monitor page, don’t worry, the monitor page is slower to update than the game, chances they are on the same page, the monitor page just needs to catch up.
2. Now that you’re at round 11, perform the same checks as you did in step 5, ensure that the Data and Payments pages have updated accordingly.
3. Now complete the game youself, complete the questionnaire and when it reaches the final page that shows the result for the entire game, verify that the results reflect what it says on the Data page.
4. Label the session with the session configuration it was and what devices and browsers you used to test it, this will show up when you extract all the data at the end. For example:  
   
5. Now repeat steps 1-9 till you have tested all the session configurations as mentioned in the [Test Specification](#_Test_Specification_1) on every device and browser you wish to support.
6. Finally, once you completed steps 1-10 now extract the data using the ‘Data Export’ page and clicking ‘Excel’ inside the ‘All apps’ section. This will download an excel file with every single session configuration you tested, and if labled your sessions you will have proof that you completed the tests on your devices, and browsers.



### Test Analysis Report

As mentioned in the test specification, the test must pass with absolute success to pass the test. Thus, a tick in the tables below indicate that it completed the Lottery, Biosecurity and the Results completely without any issues. The issues I refer to here are ones that don’t allow someone to complete the game or to play the game as efficiently as other participants on other browsers.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Windows 10 | Basic | Freeform | Approval on Pledges | Approval on Contributions |
| Chrome | ✔ | ✔ | ✔ | ✔ |
| Internet Explorer 11 | ✔ | ✔ | ✔ | ✔ |
| Microsoft Edge | ✔ | ✔ | ✔ | ✔ |
| Firefox | ✔ | ✔ | ✔ | ✔ |
| Opera | ✔ | ✔ | ✔ | ✔ |

While in the automated tests I noticed, there were issues in Internet Explorer 11, these issues did not reveal themselves during the manual tests. However, Internet Explorer 11 seems to have inconsistent behaviour, sometimes working perfectly, other times not at all, combining this with the lack of support in the future for the browser, I will not support Internet Explorer 11.

In this test, Chrome and Opera worked the best in terms of stability: there were no errors and they performed well relative to other browsers.

Microsoft Edge was *almost* the best, by appearances it was performing faster than Chrome and Opera, however the only aspect that let it down was its strange behaviour with the sliders. While dragging the slider it would work, while clicking on a different place on the slider wouldn’t always work.

Firefox worked well on Windows 10, no issues appeared at all and while the browser performed well, I had superior performance on Chrome, Opera and Edge. The only issue with Firefox is the inconsistency of the web browser across multiple operating systems, even those made by the same company. For example, in Windows 7 (which is what the lab machines run on in UWA behavioural psychology labs). Firefox refused to show the instructions on these machines despite the instructions relying on basic HTML and CSS for their appearance. As I will go into in Android and iOS tests below, Firefox while performing well in the manual tests showed too much inconsistency.

To conclude my testing for Windows 10, I will highly recommend using the latest version of Chrome to run the game and the administration of the game, it was the most stable across all tests and performed consistently. If Chrome isn’t allowed due to some restrictions on the OS or device itself, alternative browsers I will recommend are Opera and Microsoft Edge, just be aware of Edge’s strange slider behaviour.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Android 7.1.1 | Basic | Freeform | Approval on Pledges | Approval on Contributions |
| Chrome | ✔ | ✔ | ✔ | ✔ |
| Firefox | 🗙 | ✔ | ✔ | ✔ |
| Opera | ✔ | ✔ | ✔ | ✔ |

Running the game on Android with different browsers gave variable results, the performance and their behaviours varied between them frequently. Chrome as expected ran perfectly without issue with impressive performance. Opera was the same as Chrome except for being slightly slower.

Firefox on the other hand had inconsistent performance and inconsistent issues across my manual testing. These issues appeared in the Lottery game, at one stage, Firefox didn’t let me use the radio buttons to select the options in the play panel which is troubling. Also, Firefox had inconsistent performance throughout the test, sometimes performing extremely well, other times being slower than Internet Explorer 11 on Windows 10 which is an impressive feat for the wrong reasons. I did notice this during my automated testing however I only took note of it and blamed it on Firefox’s lack of optimisation with tabs in Android, however the manual testing revealed that it wasn’t just the number of tabs open that was the problem in the automated testing but the web browser itself.

As a result, I will support Chrome primarily on Android, and the alternative I can recommend is Opera. Firefox, I just cannot recommend, it needs to improve on its performance and stability in general before I can support it.

As a final note in my Android testing in general I got to see how the game appears on smaller devices. The game doesn’t look great on smaller devices, the lottery game for example becomes unreadable in portrait mode, due to test being aligned to the same spot. This behaviour was observed in the biosecurity game during certain stages too. While yes, you can play the game on Android devices, I suggest you play the game on a tablet, its appearance is too inconsistent on smaller devices. If you must play the game on smaller devices like a smartphone then play in landscape and prepare to do lots of scrolling, pictures and certain tables can appear quite large on a smartphone.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| iOS 10.2.1 | Basic | Freeform | Approval on Pledges | Approval on Contributions |
| Chrome | ✔ | ✔ | ✔ | ✔ |
| Firefox | ✔ | ✔ | ✔ | ✔ |
| Opera | ✔ | ✔ | ✔ | ✔ |
| Safari | ✔ | ✔ | ✔ | ✔ |

Finally, iOS 10.2.1 Chrome and Safari worked perfectly and their performance was exceptional. Firefox and Opera here performed well and worked without issue. iOS’ issues with multiple tabs didn’t show up during manual testing as we only needed one tab open to play, in fact this time, the ‘feature’ actually worked to its advantage as other tabs weren’t taking up system resources meaning it could allocate them to the game. I will fully support Chrome and Safari on iOS 10.2.1, while offering Opera and Firefox as an alternative to them, however the chances of someone using Opera and Firefox in this case is extremely low as Chrome and Safari just work so well.

To conclude my testing of the game in general, the most recommended web browser across every device is the latest version of Google Chrome, the most popular web browser in the world (Anon., 2017; Anon., 2017). The only other highly recommneded web browser compared to Google Chrome at this time is Safari on iOS 10.2.1. Other browsers can be used, just they might have issues from time to time and may have inferior performance to Chrome.

**Test 5 – Server Performance**

The purpose of this test is to investigate the server’s performance when playing the game and to see how far the server could be pushed before it got to a point where it couldn’t run the game anymore.

**Test Specification**

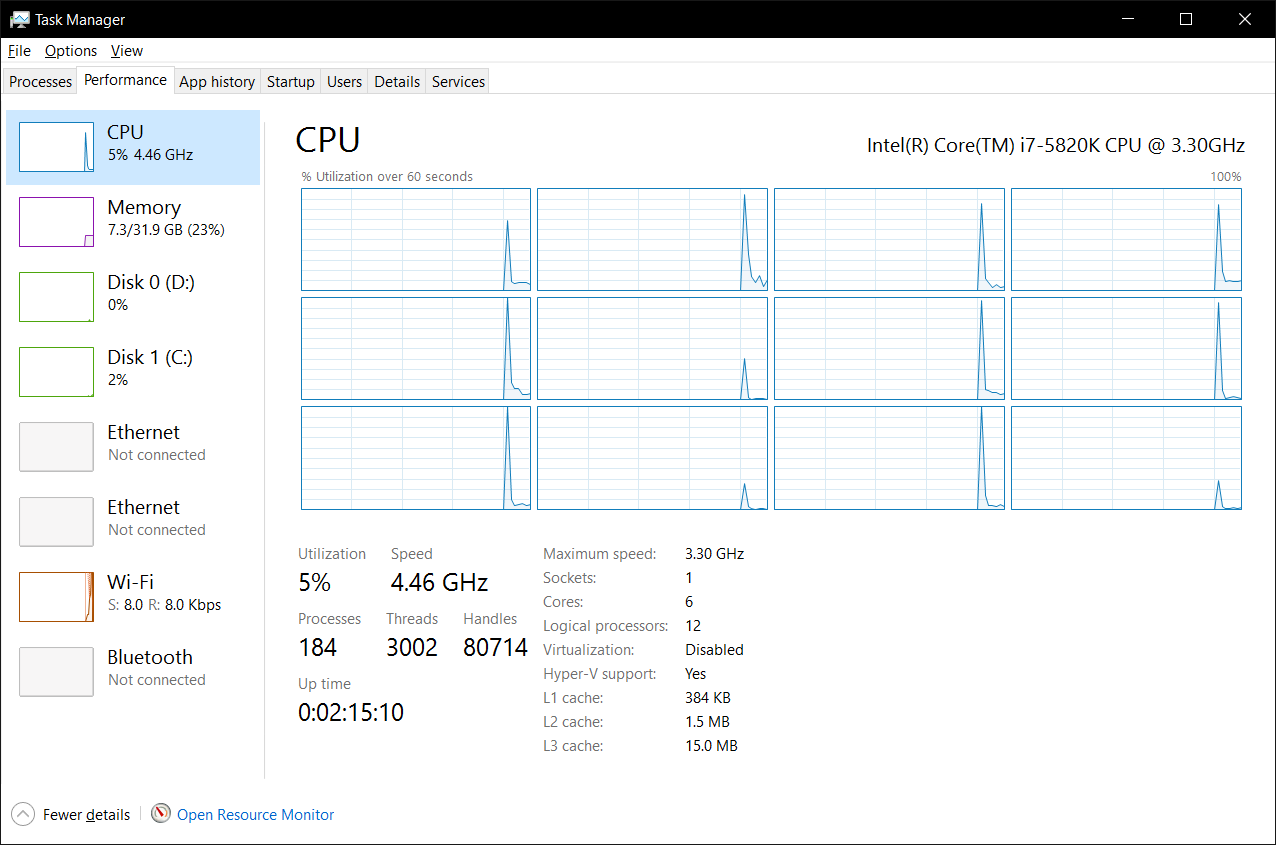
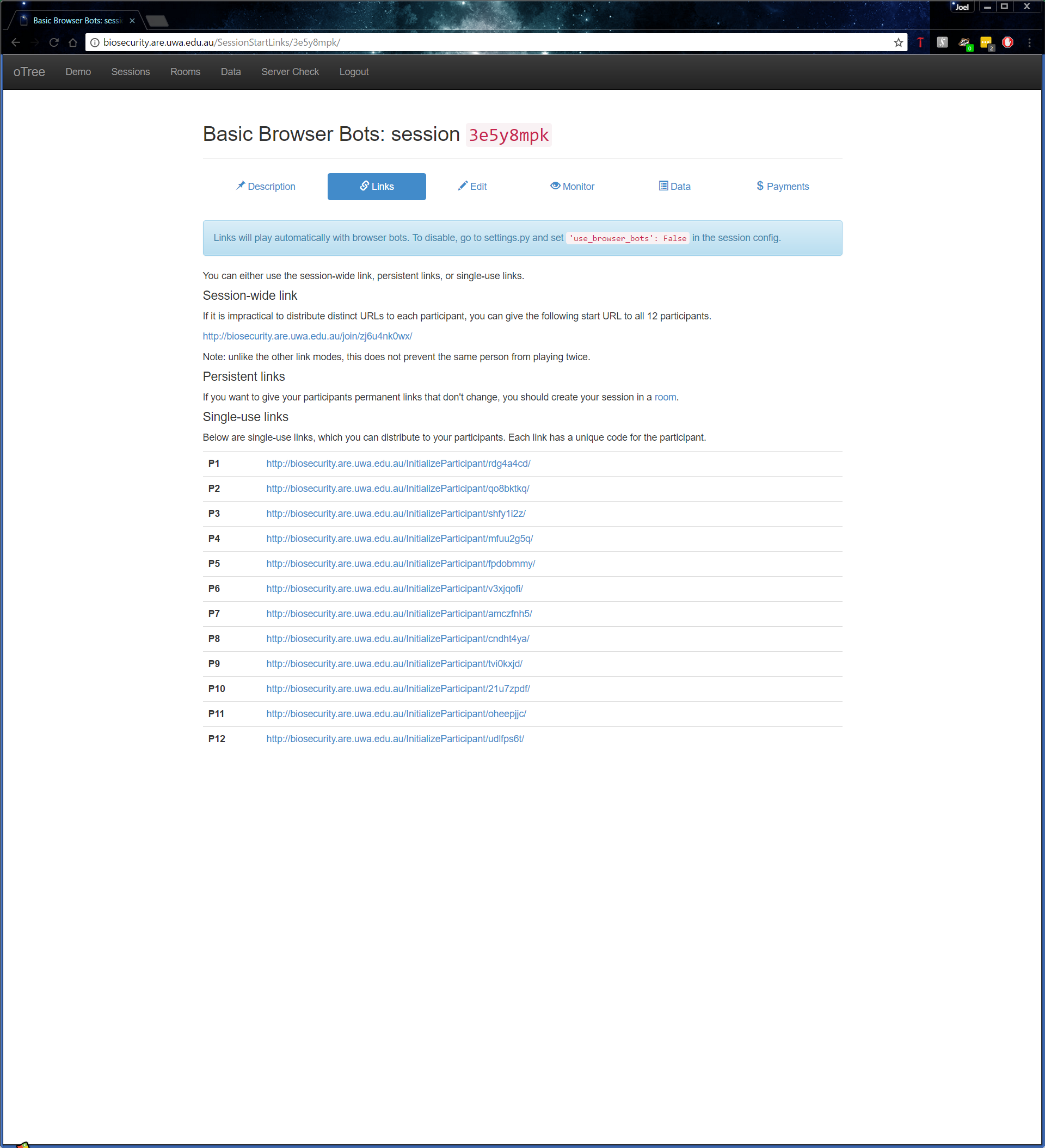
The idea of this test wasn’t to meet a certain requirement, but rather to see how far the server could go. During this time, experiments are only happening within the lab and thus don’t need to support more than 20 players at a time, while this number may increase in the future that will be a case of adding more resources when the time comes.

Currently the server has 2 cores and 4 threads due to hyperthreading, 16GB RAM and is running Windows Server 2012 within VMware as a virtual operating system. To perform this test, you want a computer that can handle a substantial number of tabs within Google Chrome (an excess of 200+ tabs) or multiple devices running Chrome to have up to 200 tabs open across them. I performed my test on my own desktop at home with an Intel i7-5820k a processor 6 cores, 12 threads combined with 32GB RAM running at 2666Mhz ensuring that any bottlenecks caused by the web browser and/or client device will be non-existent.

Also, as we will be timing in this test make sure you have a stopwatch available, most smartphones have one built in the clock settings so getting one is easy.

**Test Description**

To perform this test, we will follow the steps below[[5]](#footnote-6):

1. Navigate to biosecurity.are.uwa.edu.au and login. You will need the username and password from an Administrator.
2. Do a server check by going to the ‘Server Check’ page and ensuring everything on that page is green to ensure everything is working as expected including Sentry which allows Administrators such as myself to receive the tracebacks for errors.
3. On another computer/device (I used my laptop), but you can do it on the same one, access the server using remote desktop software and connecting the university using uniconnect if outside UWA, instructions on how to do this are here: http://www.bits.uwa.edu.au/it-help/access/uniconnect/connecting. The server’s IP address is 130.95.176.155.
4. Once you have remote access to the server, open Task Manager and go to the ‘Performance’ tab and make sure you’re observing CPU, like below:  
   
5. Now, going back to the computer (which was my desktop) which you have the website up on, open the ‘Sessions’ in the main administration page. Now create a new ‘Basic Browser Bots’ session with 4 participants.
6. Open each link in a new tab starting the timer when the last tab was opened when doing 4 participants, when performing this test with more than 4 participants you need to open every single player link that isn’t a multiple of 4 first e.g.
7. Start the timer and begin opening the remaining tabs, record the time between opening the first tab that is a multiple of 4 to the last tab that you open. This will be your error margin, this will get significantly larger as you have more tabs to open.
8. Allow the test to keep going until every tab has said its finished and record the time it took to complete the test, this may be time consuming as you want to try to keep an eye on it in the case any errors do occur.
9. Complete steps 1-8 with the number of participants at 4, 12, 32, 60, 80, and 100
10. Complete steps 1-9 with the Session Configurations:
    1. Basic Browser Bots
    2. Approval on Contributions Browser Bots (AOC Browser Bots)

Now you may now try to improve the times recorded, if you attempt to do this ensure you follow the same strategy each time to keep consistent results, error margins and participants. You don’t have to follow my strategy exactly, you could start the timer from the time you open the first participant link and get the same conclusions.

**Test Analysis Report**

Before looking at the times it took the server to complete the game, I also observed how many participants the server could take when creating a session. At first, I found the limit to be 24 participants, however I found after debugging that I had assigned the names per session rather than per group. Since there are only 25 names to choose from it wouldn’t allow any more than 24 players at once when creating the session. After fixing this I found that the server could create a session with 1000+ participants well exceeding what the server could handle when participants are playing the game concurrently.

Now the results of the server performance tests are below, to remind you, the error margin is the amount of time it took to open the first tab that’s a multiple of 4 to the last tab you open. The time is given is MM:SS ± Error Margin.

|  |  |  |
| --- | --- | --- |
| Number of Participants | Basic Browser Bots | AOC Browser Bots |
| 4 | 01:15 | 02:05 |
| 12 | 02:02 ± 00:05 | 03:52 ± 00:10 |
| 32 | 04:33 ± 00:20 | 08:10 ± 00:25 |
| 60 | 09:05 ± 00:20 | 14:10 ± 00:50 |
| 80 | 12:30 ± 01:00 | 20:00 ± 01:10 |
| 100 | 13:45 ± 01:15 | 26:00 ± 02:00 |

Looking at the time’s alone isn’t enough, I did record the CPU utilisation and when games started to finish to get a better idea of things might go if all the participants started at the same and how the game affects the server. Throughout the testing of the ‘Basic Browser Bots’ I noticed a CPU utilisation somewhere around between 50-75% on average occasionally spiking to 80-90% when the game went to 32+ participants. On average the first game to finish was well before the actual finishing time with participants finishing a minute or two before the rest of them completed the test.

Going from the times and the utilisation and from what I observed during the tests, I believe the server can handle an absolute maximum of 50 participants concurrently with the amount of resources it currently has. At 50 participants, the game would be slow but verging on unplayable.

There is one limitation to this test, this test uses bots which push the server to constantly take in http requests, timeout requests and other requests, much more than what would happen if people were playing. I believe this to be so, as players will always take longer than the bot to decide their actions, thus when 50 players are playing concurrently, it doesn’t necessarily mean it’s taking every single player’s data once a second, its taking the data of 50 players over a minimum time of around 10-20 seconds (or longer) and a maximum of 90 seconds. This could result in the game being much smoother with 32 or 60 players in a real-world experiment than it was in testing.

I did attempt to see if I could improve the times above by adding more runworkers to run with the command otree runprodserver, unfortunately I saw no difference on my local machine or running on the server, each time being produced being within margin of error. It did make the administration page and the creation of sessions perform better, but not enough to warrant doing it. The main bottleneck and limitation to the game itself performing better is the use of Daphne as the HTTP server. Daphne as observed by others [here](https://groups.google.com/d/msg/otree/oWoypNd4O0c/CkhRfH3LAwAJ) needs some load balancing to properly utilise multiple cores and therefore increase performance of the game.

## Test Materials

The test materials required will be a PC (Linux, Windows or OSX, Windows preferably as the other devices have browsers can have those that OSX and Linux can have, while Windows has IE and Edge), Android Smartphone or tablet, and an iOS phone or tablet. With these 3 devices one will be able to test web browsers that will be used by most participants including Chrome, Firefox, Internet Explorer 11, Safari, Microsoft Edge and Opera. The server will need to be on during the testing period, however one should not have to worry about the server being on due to it being on 24/7 except for about 5 mins in the middle of the night once a week.  
  
Theoretically, the game should work no matter what OS you’re running on your own personal device as the Server runs all the python code and produces the Django templates which your web browser views as HTML. However, to be thorough, each device should have as much web browsers as possible for testing.

## oTree\_Tags

For this program to work you will need to adjust the oTree package files, as this program requires an adjusted otree\_tags file. There are two ways you can do this:

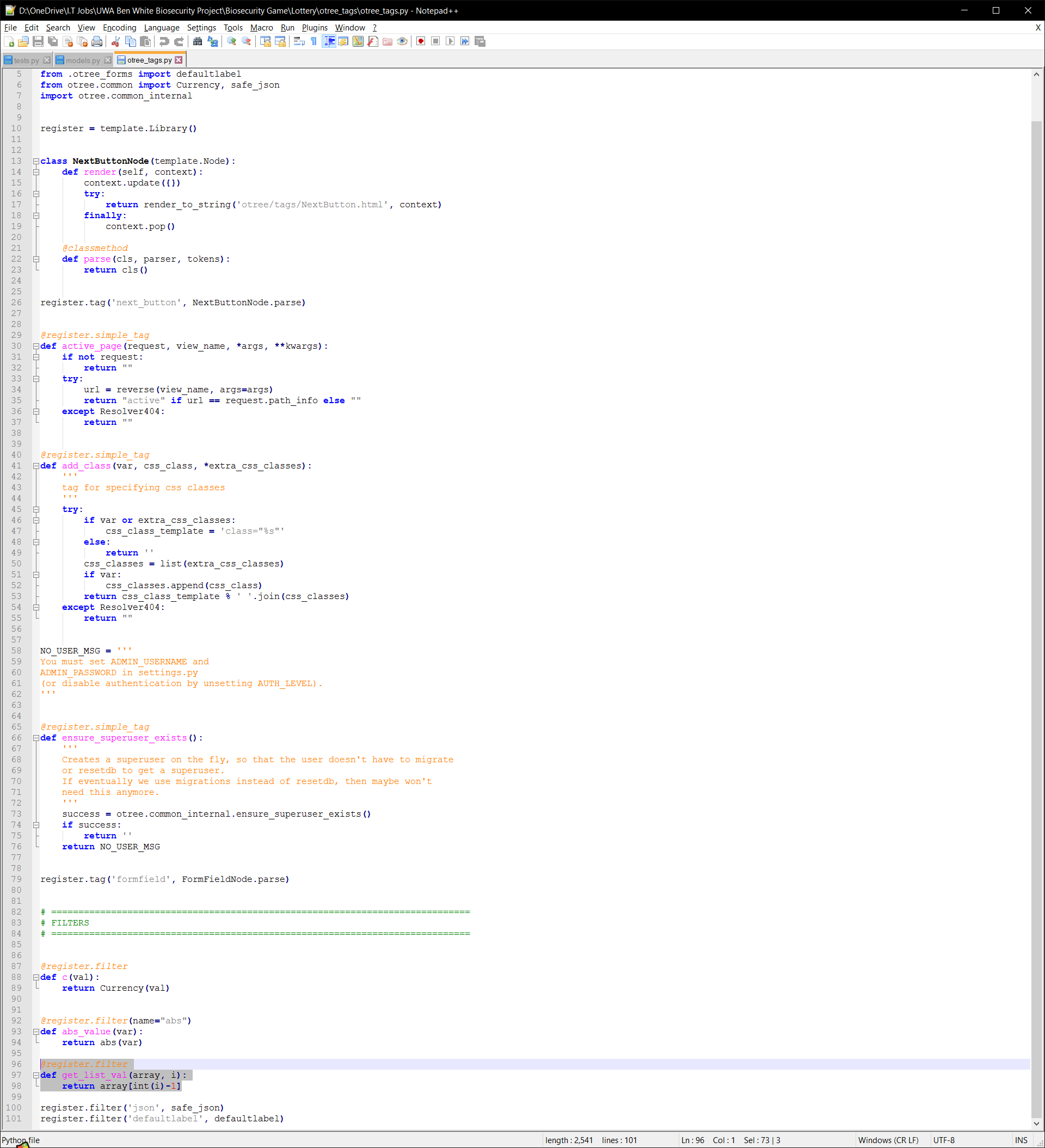
1) Assuming you cloned the repository or have the program files available to you, inside the Lottery app folder named *Lottery*, there should be a directory named otree\_tags, go in there.  
Copy the otree\_tags file inside a paste the file inside this directory:

<Home>\Python<version>\Lib\sitepackages\otree\templatetags\  
  
<Home> refers to your account files and folders which will be located differently depending on your OS, please refer to the Python documentation and how you installed Python on your machine to find the directory we’re looking for. Also, keep in mind the folder will be named differently depending on your Python version, so refer to what version you have as well to fill in <version>.

2) Instead of copying the file in the repository, find this file below (refer to above to understand what <Home> and <version> mean):

<Home>\Python<version>\Lib\sitepackages\otree\templatetags\otree\_tags.py

In this file add the code as per the image below:



# **Glossary**

* oTree: “oTree is a framework based on Python and Django that lets you build:
  + Multiplayer strategy games, like the prisoner’s dilemma, public goods game, and auctions
  + Controlled behavioural experiments in economics, psychology, and related fields
  + Surveys and quizzes” (oTree, 2016)
* Python: A high-level programming language that allows you to design and implement computer programs on any machine.
* Django: A web framework written in Python used to create complex, database driven websites.
* HTML: Stands for **H**yper**T**ext **M**arkup **L**anguage, it’s the language used to produce text, images, colour, graphics and links on the world-wide web. Every single page you access on the internet uses HTML.
* Web Server: The computer/device/machine that runs the necessary programs and services to allow people to access the website on the internet and run the oTree project using Python and Django on a web browser without a client’s computer needing Python, Django or any of the necessary programs to run the game.
* Web Browser or Browser: An application that is used to retrieve and present information and graphics sent by a web server.
* CSS: Stands for **C**ascading **S**tyle **S**heet, it is a language used to format how HTML elements will appear on a web browser.
* Relational Database: A database that properly recognises many to many relationships and is structured in a way to store information efficiently and effectively.
* Postgres: A Database Management System (DBMS), to which a DBMS allows one to create and manage databases.
* HTTP: Stands for **H**yper**T**ext **T**ransfer **P**rotocol, a high-level protocol designed to take and receive requests for HTML pages on the internet, it is a major part of communicating on the world-wide web.
* Operating System: Often used as an acronym **OS**, low-level software that allows a computer to manage basic functions like scheduling tasks, providing API’s for other programs etc. Windows, Android, OSX and iOS are examples of operating systems.
* IP Address: Stands for **I**nternet **P**rotocol **Address**, it is your “location” in a network, you can think of it like your home address, a unique set of numbers (and letters if using IPv6) that identifies your computer.
* Hardware: A physical part, component of a computer
* Software: A virtual program that can run a computer/device/machine created using a programming language, and designed for a certain purpose.
* Session: In the context of oTree, “a session is an event during which multiple participants take part in a series of tasks or games.” (oTree, 2016)
* Protection: In the context of the game, this is the amount of effort a player contributes during a round in the biosecurity game in terms of probability of the player **not** being the source of the outbreak. This is not to be confused with Cost which is still the amount of effort a player contributes, but represented in dollars.
* Outbreak: In the context of the game, this is an event when pests have infected the player’s crop for that round and therefore get no revenue.

# References

Anon., 2017. *Net Market Share.* [Online]   
Available at: https://www.netmarketshare.com/  
[Accessed April 2017].

Anon., 2017. *W3 Counter.* [Online]   
Available at: https://www.w3counter.com/globalstats.php  
[Accessed April 2017].

Chen, D., Schonger, M. & Wickens, C., 2016. oTree - An open-source platform for laboratory, online and field experiments.. *Journal of Behavioral and Experimental Finance,* Volume 9, pp. 88-97.

oTree, 2016. *oTree Documentation.* [Online]   
Available at: http://otree.readthedocs.io/en/latest/  
[Accessed April 2016].

1. These averages were the combined average of the 4 players at Round 5, for more detailed results on how each player performed refer to the Excel Spreadsheets or perform your own. [↑](#footnote-ref-2)
2. Same as 1 [↑](#footnote-ref-3)
3. These averages were the combined average of the 4 players at Round 5, for more detailed results on how each player performed refer to the Excel Spreadsheets or perform your own. [↑](#footnote-ref-4)
4. Same as 3 [↑](#footnote-ref-5)
5. The test instructions will assume you have done Tests 1-4 and have the administration instructions available to you. [↑](#footnote-ref-6)