

---

---

# Feasibility Study

## ELIZA Plus - Long-Term Memory

---

---

*by*

PRITAM SANGANI

SUPERVISED BY: DR. JAMES O'SHEA



School of Computing, Mathematics and Digital Technology  
MANCHESTER METROPOLITAN UNIVERSITY

A Feasibility Study document submitted to  
Manchester Metropolitan University, as part of a final  
year dissertation, in accordance with the requirements  
of the degree of BACHELOR OF SCIENCE in the  
Faculty of Science and Engineering.

OCTOBER 2018

## Course-Specific Learning Outcomes

Below are the Course-Specific Learning Outcomes for BSc (Hons) in Computer Science. Please note that only those outcomes which are relevant to this dissertation have been listed.

### On completion of the course students will be able to:

- use knowledge, abilities and skills for further study and for a range of employment in areas related to scientific and technical computing;
- interpret legislation appropriate to computer professionals and also be aware of relevant ethical issues and the role of professional bodies;
- analyse, design, and implement algorithms using a range of appropriate languages and/or methodologies;
- demonstrate effective communication, decision making and creative problem solving skills, and identify appropriate practices within a professional, legal and ethical framework;
- critically appraise and apply suitable artificial intelligence techniques for a variety of software systems.

## Background

In 1950, Alan Turing published a paper called 'Computing Machinery and Intelligence' (Turing 1950), in which he posed the question, "Can machines think?". Alongside this question, he also proposed the Turing Test - a test that would use conversation to answer this question. Since then, Computer Scientists have been researching, for decades, how to develop chatterbots that converse convincingly with humans. Research, in this field, intensified when in the mid-1960s, Joseph Weizenbaum, developed ELIZA (Weizenbaum 1966) - known to be the first chatter bot. ELIZA used an early-form of Natural Language Processing, where it would match patterns in the text input and substitute it with phrases, to create the illusion of understanding. In recent years, chatterbot developers have been trying to win the Loebner Prize - a modern day version of the Turing Test. This contest has been held since 1991 in which judges converse with chatterbots not knowing whether they are talking to a human or a chatterbot (Zdenek 2001). Chatterbots have also become very popular commercially, as businesses look to make their customer services more efficient, innovative and, most importantly, more personal, due to the rising demand of customers looking for fast resolutions to their problems and the increase in time they are spending online (Bakhshi et al. 2018). This report by Deloitte, also found that one of the market forces driving chatterbot development is the "technological advances in AI and NLP".

Figure 1 below shows an infographic, also from the Deloitte report, displaying the "different functions of the human brain" that chatterbots try to "mimic". For my thesis, I will be concentrating on the Natural Language Processing (NLP) and Entity Recognition parts of the infographic for Dialog Management. NLP is a branch of AI concerned with the research of interactions between humans and computers through natural language. Natural Language has been defined in a white paper on NLP as being "the most natural means of communication between humans, and the mode of expression of choice for most of the documents they produce" (Weischedel et al. 1989). Entity Recognition (also known as Named Entity Recognition[NER]) is a branch of NLP concerned with labelling "sequences of words in a text which are the names of things, such as person and company names, or gene and protein names." (*Stanford Named Entity Recognizer (NER)* n.d.)

A big part of developing chatterbots is actually evaluating the quality of existing chatterbots and

# Understanding Language & Context

Chatbots mimic different functions of the human brain.

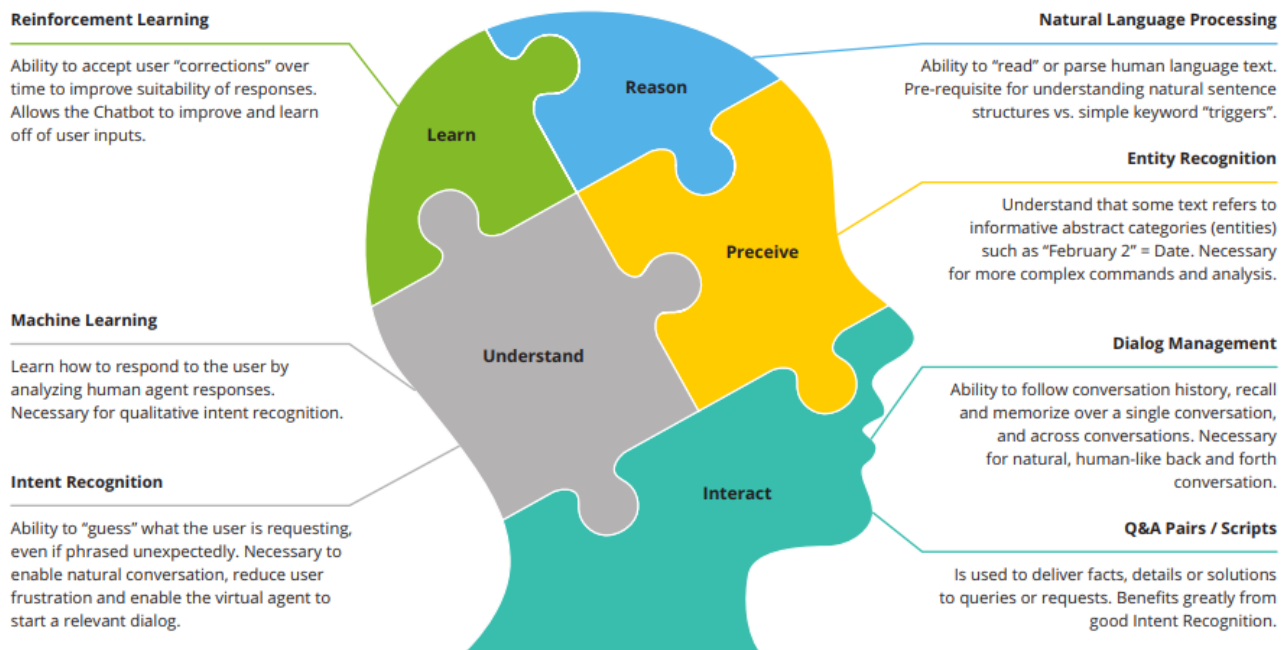


Figure 1: Infographic showing the different functions of a chatterbot architecture (Bakhshi et al. 2018)

how convincing they are at mimicking the functions of the human brain. Two recent studies have tried to answer two questions that must be answered when evaluating the quality of a chatterbot. Firstly, the current uses of chatterbots in society must be examined. Brandtzaeg and Følstad have published their study of the uses of chatterbots on various platforms and across various categories of uses. (Brandtzaeg & Følstad 2017) The vast majority of participants in the study reported using chatterbots to increase their productivity, by quickly retrieving information or accessing assistance. Rather surprisingly, the study also found that 12% of the participants reported using chatterbots for social or relational use. The study found that the human nature of chatterbots drove people to using them for this purpose - the responses stated that chatterbots were a way of "avoiding loneliness" and "improving their social and conversational skills". Another study tried to gather together all the attributes and features that can be used to assess the quality of chatbots. (Radziwill & Benton 2017) The study found that two attributes that measures the effectiveness of the chatterbot are the ability to "maintain themed discussion" and to deliver "convincing, satisfying and natural interaction". The study also found that one of the attributes that measures the satisfaction of the chatterbot is the ability to "detect meaning or intent".

A big part of this project is to add a "long-term memory" mechanism to the core of the ELIZA chatterbot. Long-term memory is the ability to refer back to earlier conversations and to bring the information back at relevant points in the current conversation. This would make the chatterbot more human-like and, as discussed, would improve the quality of the chatterbot. I plan to achieve this using NER and NLP to store relevant and linked information into a database to retrieve at a later point in the conversation.

## Research Question

The question I propose to answer with my research is as follows:

*Is it possible to implement convincing long-term memory into an existing chatterbot, such as Weizenbaum's ELIZA?*

## Hypotheses

From the research question above, I propose that the following hypotheses could be derived from my research:

- **H<sub>0</sub>** - it is not possible to add long-term memory to a chatterbot, at this stage, which is convincing enough to be comparable to human memory mechanism.
- **H<sub>1</sub>** - it is possible to add convincing long-term memory at a significant statistical level, if the number of participants in the research is high enough.

**Note:** The degree of convincingness would be determined by the analysis of conversation log files, to see how often the memory mechanism appeared to access what the user was referring to, and by a subjective analysis of a questionnaire put to the participants of a research experiment.

## Aim

The aims of this thesis are as follows:

- **Integrate Long-Term Memory mechanism to the ELIZA chatterbot core.**
- **Evaluate how convincing the long-term memory mechanism is, by way of research methods determined by research done in literature review.**

## Objectives

The problem is a complex one to be solved and so I have broken the problem down into small steps and approaches that I will take to solve the problem. These are as follows:

1. Complete a Literature Review by reading research papers and articles on the topic of Natural Language Processing and Named Entity Recognition from the Internet and textbooks.
2. Find source code for the ELIZA chatterbot in Python and start understanding how the code works.
3. Compare the different tools for Named Entity Recognition readily available for use in Python.
4. Design an overview of the software and a plan to implement.
5. Implement the software according to the plan, whilst incrementally testing where appropriate.
6. Carry out a final testing of the software.
7. Evaluate the quality of the software by carrying out a study with participants.
8. Write up the research findings in a report.

As well as the objectives listed above there are a number of interim deliverables that need to be met. These are as follows:

1. Prototype Report
2. Prototype Software
3. Final Software
4. Report Outline
5. Showcase Event
6. Final Report

## Problems

As with any project, problems can arise and it is important to resolve them quickly. To mitigate the effect of problems during this project, I am considering the problems that could arise and how I would overcome each problem.

Firstly, a problem that is likely to occur is in the implementation of the product. As I have not worked with Natural Language Processing before, I am likely going to have some trouble, initially, implementing the Named Entity Recognition algorithms. To resolve any issues in implementation, I will first look at the documentation for any modules I use. If this does not solve the problem, I can use sites, such as Stack Overflow, to search for possible solutions.

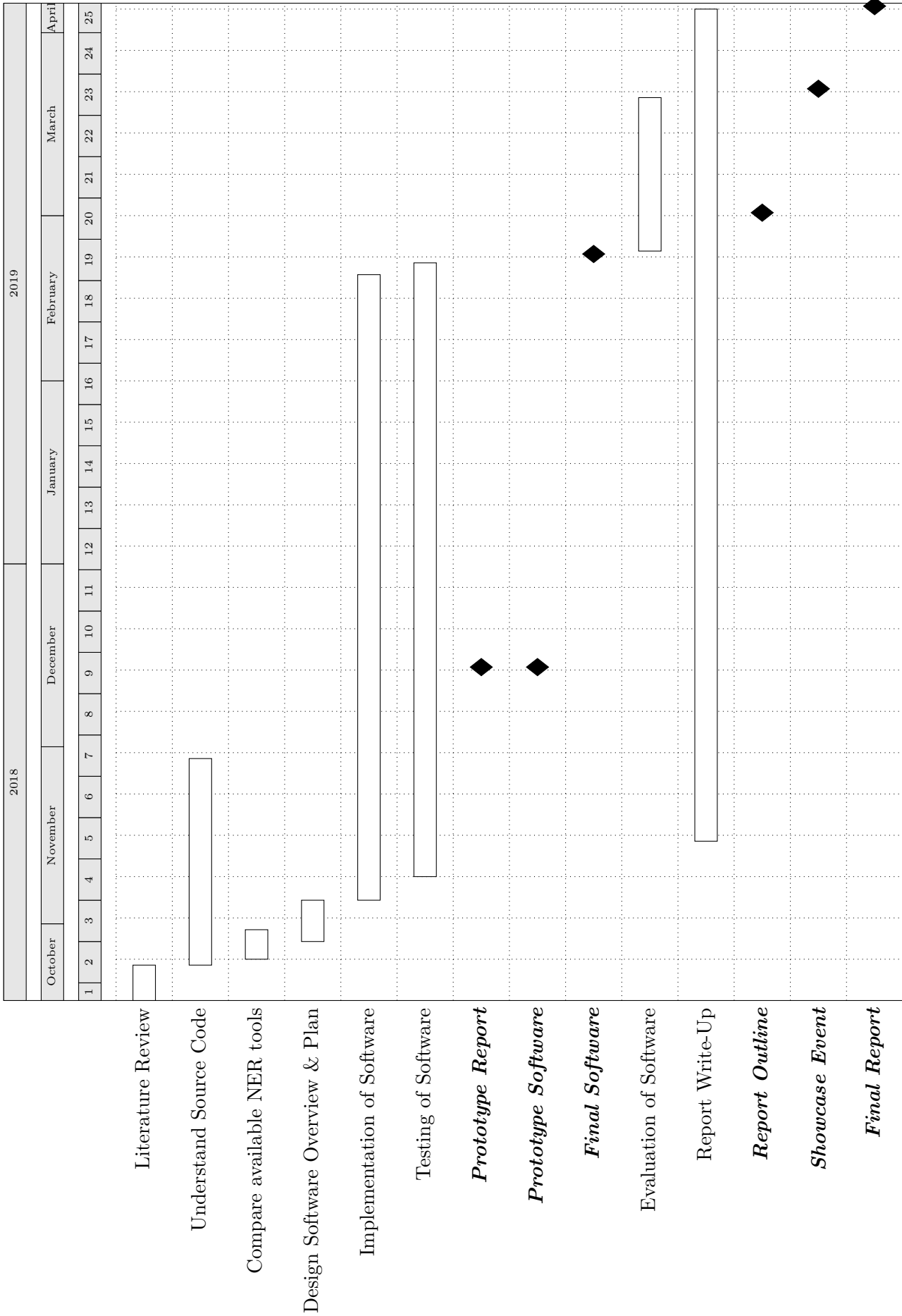
A second problem, that is not likely to occur but could still occur, is that my code could be lost, either by my laptop becoming corrupt or stop working. To mitigate this problem, I will use a version control system called Git and I will host my code on GitHub, which is a cloud-based storage service for code repositories.

## Evaluation

As with all projects of this nature, an evaluation of the produced software must be undertaken. I intend to do this by running an experiment where participants would converse with the chatterbot for two sessions. There will be two sessions as this project involves implementing long-term memory. Therefore, in order for the participant to properly evaluate the quality of the chatterbot, in particular its ability to retrieve information from previous conversations, I would need the participant to have two sessions with the chatterbot. Once the participant has concluded the experiment, I would have them answer a questionnaire to gauge their responses and feelings about the chatterbot's quality. I would then subjectively analyse the responses, as well as, look at the conversation logs to determine the degree of convincingness of the chatterbot.

## Schedule

Below is the proposed week-by-week schedule for the project as outlined by the tasks/objectives outlined in the Objectives section. I have displayed the objectives as bars, whilst the deliverables are classed as milestones, which are displayed using diamond symbols.



## Required Resources

The resources required to carry out this research project is as follows:

- A laptop or PC capable of smoothly running:
  - Python 3.x
  - Google Chrome
  - PyCharm - an IDE, by JetBrains, to write and compile Python code
  - Visual Studio Code - an IDE to write code in any programming language (HTML, CSS and JavaScript for this project)

All resources listed above are all available to download for free. (Note: PyCharm has a paid version, but, for the scope of this project, the free community edition will suffice.) Furthermore, I will not be relying on the University Laboratory PCs as I can run all required resources on my personal laptop.

## References

- Bakhshi, N., van den Berg, H., Broersen, S., de Vries, D., El Bouazzaoui, H. & Michels, B. (2018), Chatbots point of view, Technical report, Deloitte The Netherlands.
- Brandtzaeg, P. B. & Følstad, A. (2017), Why people use chatbots, *in* 'International Conference on Internet Science', Springer, pp. 377–392.
- Radziwill, N. M. & Benton, M. C. (2017), 'Evaluating quality of chatbots and intelligent conversational agents', *arXiv preprint arXiv:1704.04579*.
- Stanford Named Entity Recognizer (NER)* (n.d.), Available at <https://nlp.stanford.edu/software/CRF-NER.shtml> (Accessed: 17/10/2018).
- Turing, A. M. (1950), 'I.—computing machinery and intelligence', *Mind* **LIX**(236), 433–460.
- Weischedel, R., Carbonell, J., Grosz, B., Lehnert, W., Marcus, M., Perrault, R. & Wilensky, R. (1989), White paper on natural language processing, *in* 'Proceedings of the Workshop on Speech and Natural Language', HLT '89, Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 481–493.
- Weizenbaum, J. (1966), 'Eliza - a computer program for the study of natural language communication between man and machine', *Commun. ACM* **9**(1), 36–45.
- Zdenek, S. (2001), 'Passing loebner's turing test: A case of conflicting discourse functions', *Minds and Machines* **11**(1), 53–76.

## Undergraduate and PGT Application

### START HERE - Basic Information

This form must be completed for all student projects.

#### Before you proceed

Some activities inherently involve increased risks or approval by external regulatory bodies, so a proportional ethics review is not recommended and a full ethical review may be required.

These may include:

- i. Approval from an external regulatory body (including, but not limited to: NHS (HRA), HMPPS etc.);
- ii. Misleading participants;
- iii. Research without the participants' consent;
- iv. Clinical procedures with participants;
- v. The ingestion or administration of any substance to participants by any means of delivery;
- vi. The use of novel techniques, even where apparently non-invasive, whose safety may be open to question;
- vii. The use of ionising radiation or exposure to radioactive materials;
- viii. Engaging in, witnessing, or monitoring criminal activity;
- ix. Engaging with, or accessing terrorism related materials;
- x. A requirement for security clearance to access participants, data or materials;
- xi. Physical or psychological risk to the participants or researcher;
- xii. The project activity takes place in a country outside of the UK for which there is currently an active travel warning issued by the authorities (see info button);
- xiii. Animals, animal tissue, new or existing human tissue, or biological toxins and agents.

**If any of these activities are fundamental to your project, please contact your supervisor to determine if a full application is required.**

This form must be completed for each research project which you undertake at the University. It must be approved by your supervisor (where relevant) PRIOR to the start of any data collection.

In completing this form, please consult the University's [ACADEMIC ETHICAL FRAMEWORK](#) for ethical research.

A1 Please confirm that you will abide by the University's Academic Ethical Framework in relation to this project.

- ☒ Yes  
☐ No

A2 Are you submitting this application as a learning experience, for a unit which already has ethical approval? (please confirm with your supervisor)

- ☐ Yes  
☒ No



A3 Student details

Title	First Name	Surname
<input type="text"/>	<input type="text" value="Pritam"/>	<input type="text" value="Sangani"/>
Email	<input type="text" value="pritam.sangani@stu.mmu.ac.uk"/>	

A4 Supervisor

Title	First Name	Surname
<input type="text" value="Dr"/>	<input type="text" value="James"/>	<input type="text" value="O'Shea"/>
Faculty	<input type="text" value="Science and Engineering"/>	
Telephone	<input type="text" value="0161 247 1546"/>	
Email	<input type="text" value="j.d.oshea@mmu.ac.uk"/>	

A5 Which Faculty is responsible for the project?

Science and Engineering

A6 Course title

A7 Project title

A8 What is the proposed start date of your project?

A9 When do you expect to complete your project?

A10 Please describe the overall aims of your project (3-4 sentences). Research questions should also be included here.

This project has two aims. The first aim is to "Integrate Long-Term Memory mechanism into the ELIZA chatterbot core." The second aim is to "Evaluate how convincing the long-term memory mechanism is, by way of research methods determined by research done in the literature review." I also have a research question that I aim to answer. This is "Is it possible to implement convincing long-term memory into an existing chatterbot, such as Weizenbaum's ELIZA?"

#### A11 Please describe the research activity

I will start my project by completing a short literature review by reading research papers and articles on the topic of Natural Language Processing and Named Entity Recognition, which are the topics my project is based upon. I will obtain these from the internet and textbooks. At the same time, I will also find open-source code for the ELIZA chatterbot and start looking through the code to gain an understanding of how the software works. I will obtain the code from GitHub, a cloud-based storage service for code repositories. I will then research and compare the different tools available for implementing Named Entity Recognition. Once this is done, I will begin planning for the implementation stage, by designing an overview of the software and creating a plan of features that need to be implemented, thus breaking down the implementation into small tasks. During the implementation stage, I have decided to incrementally test the software, where appropriate, by conversing with the chatterbot - I will personally be doing the testing. Once the implementation stage is complete, I will be carrying out a final testing of the software to make sure it is ready to be evaluated by participants - again I will personally be doing the testing. Once testing is complete, I will be ready to ask participants to evaluate the quality of the chatterbot by having them converse with the chatterbot. I will be keeping a log of the conversation history for each participant, which I will analyse once the study is complete. I will also ask the participants to fill out a questionnaire to get their responses and feelings about the quality of the chatterbot. I will subjectively analyse the questionnaires and then write up my findings in a report.

#### A12 Please provide details of the participants you intend to involve (please include information relating to the number involved and their demographics; the inclusion and exclusion criteria)

I will try and involve a minimum of 5 participants to a maximum of 15. The participants will all be students and I will try and have a mixed demographic of participants. (a mix of gender and race)

### Project Activity

#### B1 Are there any Health and Safety risks to the researcher and/or participants?

- ☐ Yes  
☒ No

#### B2 Please select any of the following which apply to your project

- ☒ Aspects involving human participants (including, but not limited to interviews, questionnaires, images, artefacts and social media data)  
☐ Aspects that the researcher or participants could find embarrassing or emotionally upsetting  
☐ Aspects that include culturally sensitive issues (e.g. age, gender, ethnicity etc.)  
☐ Aspects involving vulnerable groups (e.g. prisoners, pregnant women, children, elderly or disabled people, people experiencing mental health problems, victims of crime etc.), but does not require special approval from external bodies (NHS, security clearance, etc.)  
☐ Project activity which will take place in a country outside of the UK  
☐ None of the above

#### B2.4 Is this project being undertaken as part of a larger research study for which a Manchester Metropolitan application for ethical approval has already been granted or submitted?

- ☐ Yes  
☒ No

### Informed Participation/Consent

C1 Will participants be given accessible information about:

- a) the general purpose of the project
- b) what is expected from them in the project
- c) their right to refuse or withdraw at any time
- d) how their data will be used and managed, and their relevant legal rights

- ☒ Yes  
☐ No

C1.2 Please describe how you will do this

I will give the participants all information regarding the project and what is expected of them in written form (probably a leaflet). I will also outline how they can withdraw from the project at any time, as well as how their conversation logs will be used and kept strictly private.

C2 Will you ask for informed consent from all participants?

- ☒ Yes  
☐ No

C3 Will any participants be legally unable to consent, and require you to obtain informed consent from a legal representative?

- ☐ Yes  
☒ No  
☐ Not Applicable

C3.1 Please give details why this will not be necessary or relevant

Study will only involve day-to-day conversation and will not involve any information which is confidential or requires any legal consent from authorities.

C4 Will participants have an opportunity to ask questions prior to agreeing to participate?

- ☒ Yes  
☐ No

C4.1 Please describe how participants will be able to ask questions prior to agreeing to participate

I will provide possible participants with my student email which they can use if they have any questions to ask.

C5 Will participants receive any payments, reimbursements of expenses or any other benefits or expenses for taking part in this project?

- ☐ Yes  
☒ No

C6 Have relevant authorities (gatekeepers) given their permission for project activities to take place on their premises (e.g. shop managers, service managers, head teachers, classroom lecturers)?

- ☐ Yes  
☐ No  
☒ Not Applicable

C7 Could your past or present relationship with the potential participants give rise to a perceived pressure to participate?

- ☐ Yes  
☒ No

C9 Will any participants be identified through posters, leaflets, adverts, social media or websites?

- ☐ Yes  
☒ No

## Anonymity and Confidentiality

D1 Please describe how you will protect participants anonymity

I will remove all identifying data, such as name, from the conversation logs and also I will inform the participants that they should not enter any personal information when conversing with the chatterbot.

D2 Please describe how you will ensure that individuals cannot be identified indirectly (e.g. via other information that is collected)?

No personal details will be collected before, during or after the study

D3 Please describe how you will protect participants confidentiality?

Conversation logs will be cleared of any identifying details before being written up in the report.

## Debriefing

E1 Will participants have the opportunity to obtain feedback or the results after the project has ended?

- ☐ Yes  
☒ No

## Data

F1 How and where will data and documentation be stored?

Data will be collected in the form of conversation logs and will be kept on university servers which are protected by security that the university employs

F2 Will you be collecting personal data or sensitive personal data as part of this project?

- ☐ Yes  
☒ No

## Additional Information

G1 Do you have any additional information or comments which have not been covered in this form?

- ☐ Yes  
☒ No

G2 Do you have any additional documentation which you want to upload?

- ☒ Yes  
☐ No

G2.1 Please attach a copy of any other materials relevant to this application

Type	Name	File Name	Date	Version	Size
Additional Documentation	Risk Assessment	RA_Project_SoftwareDevelopment_270918.pdf	18/10/2018 12:00:00 AM	1	369.0 KB

## Signatures

H1 I confirm that all information in this application is accurate and true. I will not start this project until I have received Ethical Approval.

- ☒ I confirm  
☐ I do not confirm

H2 Please notify your supervisor that this application is complete and ready to be submitted by clicking "Request" below. Do not begin your project until you have received confirmation from your supervisor - it is your responsibility to ensure that they do this.

**Signed:** This form was signed by Jim O'Shea (J.D.Oshea@mmu.ac.uk) on 19/10/2018 11:04

H3 By signing this application you are confirming that all details included in the form have been completed accurately and truthfully.

**Signed:** This form was signed by Pritam Sangani (pritam.sangani@stu.mmu.ac.uk) on 18/10/2018 15:02



# Research Insurance Checklist



## ADMINISTRATIVE DETAILS

**Lead Investigator Name  
(Title/Forename/Surname)** Pritam Sangani

**Contact Email Address** pritam.sangani@stu.mmu.ac.uk

**Full Title of the Research** Eliza – Long Term Memory

## SECTION 1 – TECHNIQUES, TESTING AND INTERVENTIONS

Does your research study involve:

☐ **Physically invasive techniques?**

This refers to any test in which the skin of the participant is broken or an implement is inserted into any opening of the human body (e.g. eyes, ears, nose, mouth, lungs, stomach, rectum, vagina and urethra) or involves the taking of body samples such as saliva, hair, urine, faeces, sputum, skin, nails, or taking biopsies of any form for any purpose, or any form of scanning such as DEXA scans, Ultrasound scans, MRI, fMRI, CT, or PET scanning.

☐ **Ingestion of food stuffs or drugs?**

This refers to the consumption of any substance which may impact on psychological or physical state. Substances may include but are not limited to food, beverages or drugs.

☐ **Physical testing?**

This refers to any test in which a participant must perform an action resulting in the use of any muscle of the body and/or involves the use of scanning procedures, eye-trackers, mounted body cameras, sensors or electrodes, or the taking of swabs from any cavity of the body, respiratory challenge testing or recording of peak flows, EEG, ECG, Exercise ECG, Treadmill work.

☐ **Psychological intervention?**

This refers to any test which purposely alters the mood of the participant or involves administering personality inventories, or any other form of psychological test.

**OR**

✓ **I confirm that my research does not fall into any of the above categories (*please go straight to Section 3*)**



## SECTION 2 – CLINICAL TRIALS INSURANCE

Please complete this section only if you ticked one of the boxes in Section 1.

Does your research study involve:

- ☐ Pregnant persons as participants with procedures other than blood samples being taken from them?
- ☐ Children aged five or under with procedures other than blood samples being taken from them?
- ☐ Activities being undertaken by the lead investigator or any other member of the study team in a country outside of the UK? *If 'Yes', please refer to the 'Travel Insurance' guidance on Page 1 of this form.*

OR

- ☐ I confirm that my research does not fall into any of the above categories

## SECTION 3 – OTHER HAZARDS

Does your research study involve:

- ☐ Working with Hepatitis, Human T-Cell Lymphotropic Virus Type iii (HTLV iii), or Lymphadenopathy Associated Virus (LAV) or the mutants, derivatives or variations thereof or Acquired Immune Deficiency Syndrome (AIDS) or any syndrome or condition of a similar kind?
- ☐ Working with Transmissible Spongiform Encephalopathy (TSE), Creutzfeldt-Jakob Disease (CJD), variant Creutzfeldt-Jakob Disease (vCJD) or new variant Creutzfeldt-Jakob Disease (nvCJD)?
- ☐ Working in hazardous areas or high risk countries? *Please refer to the 'High Risk Countries' guidance on Page 1 of this form.*
- ☐ Working with hazardous substances outside of a controlled environment?
- ☐ Working with persons with a history of violence, substance abuse or a criminal record?

OR

- ☒ I confirm that my research does not fall into any of the above categories