UNIVERSITY OF
LINCOLN

Investigating Privacy Views with a Program to Collect and Analyse Information from Twitter

Being a dissertation submitted for the Project module in partial fulfilment of requirements for the degree of

BSc (Hons) Computer Science

By

Adam Walker

April 2017

WAL11357886 CMP3060M

Abstract

In today's age of instant communication and social media sharing, the amount of information being shared online by people every day is growing at an astronomical rate. This information is often public and easily accessible to people with a wide variety of intentions, from marketing research to criminals looking for easy targets. This therefore presents a problem, wherein people are sharing more and more information, without necessarily understanding the risk involved.

The literature shows that privacy is a contested concept, although certainly many researchers, experts and the general public have their own ideas about the precise nature of the right to privacy. The privacy paradox can cause problems for privacy researchers attempting to study the issue in detail. Furthermore, an increase in social media use, and fraud cases suggest this is a problem that may continue to grow for the foreseeable future.

The aim of this project was to determine whether an illustrative demonstration of the collection and visualisation of private data from a social media profile can change how people view their privacy. To achieve this a program was developed to analyse and extract information from a twitter account, and display the results to the user.

A study was conducted where participants were asked to rate how strongly they felt that their personal information could not be identified through their Twitter account. They were then asked to answer these questions again, having seen a demonstration of the program created for this project. Analysing the information gathered from these two sets of questions indicated the following results:

The use of a program to demonstrate collecting and analysing personal information from a Twitter account appears to have an effect on the privacy views of test participants, nudging their opinions further towards what they appeared to be thinking pre-demonstration.

Acknowledgements

"*Our work to improve privacy continues today*" – *Mark Zuckerberg*

I would like to thank my supervisor Dr Chris Headleand for his support and guidance in writing this dissertation.

I would like to thank my friends whom I subjected to testing and questioning over the course of this project. In particular Tom, Joe, Mary, Elli, Neville and Charlie for a whole bunch of different reasons that I don't fully remember but I appreciate you all for it.

I would also like to thank everyone who took part in my user study, I don't know most of you but you were very helpful.

I would like to thank Sir Tim Berners-Lee for creating the world wide web, because frankly without it we'd all be very bored and I'd have to find something else to do for my dissertation.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation and associated artefact are the result of my own work and have not been submitted for any other degree at the University of Lincoln or any other institution.

Signature: _____

List of Tables

Table 1.	Comparison of Social Media Sites	Page 12
Table 2.	Comparison between competing views of privacy	Page 14
Table 3.	Comparison of Programming languages	Page 23
Table 4.	Python libraries used	Page 24
Table 5.	Ethical concerns	Page 25
Table 6.	Twitter API rate limits per window	Page 33
Table 7.	Key observations from test participants	Page 40
Table 8.	Privacy protection methods	Page 42
Table 9.	Privacy confidence pre-demonstration	Page 45
Table 10.	Privacy confidence post-demonstration	Page 46

List of Figures

Fig 1.	Project management with Trello	Page 22
Fig 2.	Project repository on Github	Page 22
Fig 3.	Program display of account details	Page 35
Fig 4.	Breakdown of a regular expression	Page 37
Fig 5.	Matplotlib Pie chart	Page 39
Fig 6.	Matplotlib bar chart	Page 39
Fig 7.	Age results chart (Pre-Demonstration)	Page 41
Fig 8.	Gender results chart (Pre-Demonstration)	Page 41

Fig 9.	Twitter usage results chart (Pre-Demonstration)	Page 41
Fig 10.	Name results chart (Pre-Demonstration)	Page 42
Fig 11.	Name results chart (Post-Demonstration)	Page 42
Fig 12.	Age results chart (Pre-Demonstration)	Page 43
Fig 13.	Age results chart (Post-Demonstration)	Page 43
Fig 14.	Job results chart (Pre-Demonstration)	Page 43
Fig 15.	Job results chart (Post-Demonstration)	Page 43
Fig 16.	Gender results chart (Pre-Demonstration)	Page 43
Fig 17.	Gender results chart (Post-Demonstration)	Page 43
Fig 18.	Location results chart (Pre-Demonstration)	Page 44
Fig 19.	Location results chart (Post-Demonstration	Page 44
Fig 20.	Associated People results chart (Pre-Demonstration)	Page 44
Fig 21.	Associated People results chart (Post-Demonstration)	Page 44
Fig 22.	Interests results chart (Pre-Demonstration)	Page 44
Fig 23.	Interests results chart (Post-Demonstration)	Page 44
Fig 24.	Interest opinions results chart (Pre-Demonstration)	Page 45
Fig 25.	Interest opinions results chart (Post-Demonstration)	Page 45

Nomenclature

CFG	-	Context Free Grammars
Corpus	-	A large collection of text material used by some NLP techniques.
NE Recognition	-	Named Entity Recognition
NLP	-	Natural Language Processing
NLTK	-	Natural Language ToolKit
Phishing	-	Fraudulent activity where an attacker poses as a legitimate contact to learn sensitive information
POS Tagging	-	Parts of Speech Tagging
Privacy Paradox	-	Phenomenon where people say they care about their privacy, but also share private data
Regex	-	Regular Expression
Stop Words	-	The most common words in a language that aren't useful for NLP processes
Tokenisation	-	NLP technique to split text into small pieces, usually single words, called *tokens*
Vader	-	Sentiment analysis tool
VCS	-	Version Control System
XP	-	Extreme Programming

Table of Contents

Abstract **1**

Acknowledgements **2**

Author's Declaration **2**

List of Tables **3**

List of Figures **3**

Nomenclature **5**

Table of Contents **6**

Introduction **8**

Project aim 9

Research objectives 9

Literature Review **11**

Social Media 11

Privacy 12

Risks from social media 16

Risk mitigation 18

Natural Language Processing 19

POS Tagging 19

Sentiment Analysis 19

Grammars 20

Named Entity Recognition 20

Summary 20

Method **21**

Software Methodologies 21

Project Management 22

Tools 23

Languages 23

Environment 24

Libraries Used 24

Research Methods 25

Ethical concerns 25

Model Validation 26

User Study 27

Pilot Study 28

Design **28**

Prototype Testing 29

Requirements 30

Implementation **32**

Class Structure 32

Twitter API Rate Limits 32

Connecting to Twitter 33

Downloading Tweets 33

Identify personal information 34

Analysing data 34

Sentiment analysis 34

Keyword extraction 35

Location 35

Associated Users 35

Hashtag extraction 36

Output visual display of data 36

Results **38**

Observations 38

Demographic Questions 38

Privacy survey before demonstration 40

Privacy survey after demonstration 44

Evaluation **46**

Results Discussion 46

Supports hypothesis 46

Does not support hypothesis 47

Conclusion 47

Conclusion **48**

Future work 49

Critical reflection 49

Changes to the project objectives 49

References **50**

Appendices **54**

Introduction

The use of online social media, in the form of social networking sites, blogs and microblogging sites is becoming an increasingly large part of the daily lives of millions of people who use the internet around the world. Sites such as Facebook, Instagram, and Twitter reach 1.86 billion (Facebook, 2017b), 600 million (Instagram, 2017) and 313 million (Twitter, 2017) users respectively.

However it is not just for communication and sharing ideas that some people use these sites. The use of social media by cyber criminals as an easy way to find and steal personal information, to be used as part of crimes such as identity fraud or phishing scams, appears to be becoming more common (Nagunwa, 2014). Additionally, a report from fraud prevention non-profit organisation Cifas show UK fraud cases to be on the rise across all age groups in 2015 from 2014 (Cifas, 2016).

With this increase in the incident rate for fraud in addition to the sheer volume of users across different social networking sites, it raises the question of why are people willing to share so much information about themselves online. This *privacy paradox* is documented in several papers, which show that even when questioned on their privacy, people say they believe it to be important, yet their actions do not fit this view (Smith, Dinev and Xu, 2011; Baek, 2014; Kokolakis, 2015).

Another troubling concern raised is the question of whether the use of social media and the fraud rate are related. Whilst there is no clearly proven link between them, there have been papers hypothesising a link between online behaviour, including the use of social media, and an increased risk of victimisation for cybercrime (van Wilsem, 2013; Holtfreter, Reisig and Pratt, 2008).

A review of the relevant literature for this subject revealed research into differing areas of privacy and social media use, however a gap in the literature presented itself when there was little to be found regarding the use of technical implementations for extracting data from social media and their potential for educating users.

Therefore the purpose of this project was to investigate whether people are aware of how much personal information they may be sharing online using social networks. This was done by creating a program to collect and analyse information from a twitter account, and demonstrating the results to test participants as part of a user study.

Project aim

The aim of this project has been to investigate whether a demonstration of collecting personal information from a twitter account changes how people view their privacy. To achieve this a program was created which searches through a twitter account, downloads and analyses as much information as possible and presents the results to the user.

Research objectives

Conduct a comprehensive literature review

A comprehensive literature review of existing research into social media usage, privacy and cyber-crime involving identity theft will be conducted to ensure a thorough understanding of this area. This research will be used to help define the scope and guide the design and implementation of the project.

Conduct a prototype test to gather information and determine program scope

A manual prototype test should be conducted on a series of Twitter accounts to determine some useful information that could guide the development of the project. In addition the prototype can help define a set of criteria for what data can be found and what is likely to be present in other accounts.

Develop a system to extract and analyse information from a Twitter account

A system will be created to search through a Twitter account, taking any data it can find and analysing it. The system should produce a report detailing what information was found and display this for the user, demonstrating the amount of information they are sharing, and how it can be used to learn about them.

Conduct a user study to investigate the research question

A user study will be carried out, where test participants will use the program to collect and analyse their own data. The study will contain a survey to measure how the participants view their privacy and whether the program may have changed that.

Investigate ways to provide access to the program for ethical public use

A short investigation into the potential for releasing a public version of the developed program. The use of the tool should be ethical and not for malicious purposes.

Investigate further uses for extracted Twitter data

An investigation into potential further uses for any of the data extracted by the program developed for this research project.

Literature Review

Using a technological system to analyse the information from a social media account is a predominantly technical challenge. There are however some elements of sociology involved in the background context of social media and privacy for this project. Research into the work already performed in this area by academia has been conducted and the following section will present and discuss the results of this research and their bearing upon the project.

Papers cited as part of this literature review have been evaluated and selected as background contextual information, or for their relevance to the project directly. As part of the research into social media, references are made when necessary to the official support or developer pages of the site in question as the official source of information. The research conducted for this literature review has been divided into several distinct categories of: social media, privacy, risks, mitigation techniques and natural language processing (NLP).

Social Media

Our modern society makes use of social media as a tool for sharing and communicating ideas more so than at any point in history, with Ofcom figures showing 73% of internet users in the UK to have a social media profile, further Ofcom figures show there has been an increase in the amount of photos and videos uploaded online since 2014 (Ofcom, 2016), suggesting that social media is becoming a bigger part of the daily lives of most people. Overall these figures demonstrate a growth in the amount of both users, and content on social media.

The important factors for each site as they relate to this project have been compared (see *table 1* below) to provide a clear view of how suitable for the project each site may be. The user count refers to the total number of users each service has, and these figures have been self reported from the companies themselves (*Note:* not all sites have provided a total user count and instead provide monthly active user figures).

Knowing the kind of content which can be shared on each site allows provides a good view of the general purpose of the site as well as the kind of content that could be encountered using it. The purpose of the site is an extension of this, and as a result provides more information regarding the kind of behaviour of users and what they are looking to get from using the site.

Attribute	Facebook	Twitter	Instagram [1]	LinkedIn [2]
User count	1.86 Billion per month	313 Million per month	600 Million +	467 Million +
Content type	Any	Any	Images & Video	Any
Purpose	Socialising + Communication	Socialising + Communication + News Aggregate[3]	Socialising + Sharing media	Networking + Career Progress
API Access	Yes	Yes	Yes	Yes

Table 1. Comparison of social media sites

Facebook is the largest social network in the world by user numbers, and therefore the largest amount of data will surely be there. However, when examining the Facebook privacy help pages (Facebook.com, 2017a) it becomes clear that there is significantly more individual control over privacy settings than the equivalent on Twitter (Twitter Help Center, 2017c). This additional access to settings, whilst good for a normal user is a problem regarding this project as it means any attempts to extract information may be significantly more difficult or impossible. The lack of complex privacy options on Twitter, where the protection appears to be completely on or completely off, means that non protected accounts are just broadcasting their information publicly for anyone who may be watching.

Privacy

There have been legal and ethical considerations regarding what privacy, if any, a person deserves online. Additionally there are many differing opinions on what a right to privacy even is or should be. Therefore in order to appropriately explore the potential privacy concerns related to the use of social media, there must first be a clear definition of what privacy is, in particular online privacy. To achieve this, several different definitions of privacy will be made and used to determine what privacy is for the purposes of this project.

1 Instagram user statistics cited from (Instagram.com, 2017)

2 LinkedIn user statistics cited from (LinkedIn.com, 2017)

3 Twitter is more than just a social network, it also acts as a news aggregate service (Kwak, 2010)

One definition of privacy can be given as “the claim of individuals, groups, or institutions to determine of themselves when, how, and to what extent information about them is communicated to others” (Chai et al., 2009). This definition introduces the idea that an individual should exert control over how information about themselves is shared with others. This idea is shared by Gupta and Dhami (2015), who state that privacy can be defined as 'control over the flow of one's personal information, including the transfer and exchange of that information'.

This idea of control is reiterated by Aïmeur et al (2016) who, when testing a model for better privacy policies, defined several important factors for a user to trust and share their information. These factors include, but aren’t limited to:

1. **Private Data Control**
    - The user has control and power over how their data is used.

2. **Management**
    - The user is able to decide for themselves between privacy and utility and what is most important to them

3. **Comprehension**
    - The user has a clearer and more comprehensive understanding of what information they are sharing and how it will be used.

Hughes (2015), argues that the driving factor for real life privacy claims is 'a concern on the part of an individual to prevent personal information about him/herself from being used to harm him/herself'. Hughes also argues that harm is ultimately the issue at the heart of privacy, that because privacy is “not an objectively identifiable thing that people can agree upon”, we should instead consider a privacy claim to be a privacy right only when the harm caused is not justifiable. This explanation of privacy is better, according to Hughes, because other attempts “fail to grasp what really matters in privacy because they have sought to identify the core of what is private in places, activities, or types of information” which inevitably fail because each of these situations have examples where the ‘right to privacy does not inhere’.

Additionally this harm can determine the difference between viewing privacy as intrinsically valuable or merely a tool of instrumental value, depending upon whether that harm can be justifiable. Whilst this argument by Hughes provides an interesting counterexample to the other ideas of privacy that have been discussed, the argument over the value of privacy does not fall within the scope of this project.

One commonality across each of these definitions is the idea that the individual should be able to maintain some level of control over their information. However as the focus of this project is more closely related to the risks of sharing information online, the definition of privacy must also include provisions for the knowledge of what information the user is sharing as well as how it will be used.

The definition given by Chai et al (2016) provides a clear and concise description of privacy with regards to user consent over how the information is used, however it makes no mention of user awareness about what information is being collected and how. This idea is suggested by Aïmeur et al (2016), and described as *comprehension*, which is a key criterion for privacy as part of this project.

To provide a clear definition for the concept of privacy to be used for this project, as shown in the previous section, is a difficult task. There are some common features across different ideas, however there are also some disagreements over the fundamental ways in which privacy rights can be determined. Therefore for this project, the issue of a privacy definition will be reframed to a more achievable definition of the criteria for good privacy.

Having more closely examined each of the definitions of privacy discussed so far, there have been four general concepts which can be shown in *table 2* below, and how they fit into these ideas of privacy.

Author / Paper	Knowledge of collection	Knowledge of use	Consent to transfer	Consent to use
(Chai et al., 2009)	~[4]	✓	✓	✓
(Gupta and Dhami, 2015)	X	X	✓	✓
(Aïmeur et al., 2016) [5]	✓	✓	~	✓
(Hughes, 2015) [6]	X	✓	X	✓

Table 2. Comparison between competing views of privacy

From this comparison, the ideas of privacy as provided by Chai et al (2009) are felt to cover most of these criteria, then additionally when incorporating the idea of *comprehension* as described by

4 ~ Indicates the definition could be interpreted as including this criterion but it is not certain

5 Aïmeur et al., also suggest allowing the individual to decide for themselves between privacy and utility.

6 Hughes suggests using harm as a way to determine between a privacy claim and a privacy right.

Aïmeur et al, (2016) the final criterion is definitely covered. Thereby a clear and open standard of criteria by which to judge a privacy claim is established, based upon abstracted concepts taken from a combination literature sources.

Criteria of privacy:

1. **Knowledge of collection**
    - An individual should have knowledge of what information they share will be collected and how.

2. **Knowledge of use**
    - An individual should know what their information, once collected, will be used for.

3. **Consent to transfer**
    - An individual should provide consent for their information to be transferred to a third party.

4. **Consent to use**
    - Following from knowledge of use, an individual should provide consent for their information to be used, especially in the case of unintended or additional use.

In applying these criteria to privacy issues with online social media use, it becomes more apparent that there is a conflict between the sharing of information which is often done very readily by users for even a small amount of utility (Kokolakis, 2017) and their awareness of what will happen to that information and the risks associated with sharing it so readily.

One hypothesis for why this may be the case suggests that it may be because often when people share information on these sites they are not necessarily aware that they have unsecured or publicly viewable profiles due to their privacy settings (Srinivasan, 2012).

Another hypothesis suggests that privacy policies, which are supposed to inform the use how their information will be collected and used, are widely ignored by users because of their 'lengthy verbose format' (Aïmeur et al., 2016). This author continues with a suggestion that due to a large amount of users simply not reading privacy policies, there is a lack of trust in the ability of these services to protect their private data. Srinivasan (2012) also suggests that privacy policies are contributing towards this lack of control, due to the regular changing and updating of the policies themselves, making it harder for users to know at any time what the policy actually is.

A significant problem related to privacy concerns and research is the privacy paradox. The privacy paradox is the phenomenon that people consistently say they are concerned about their privacy and care about protecting it, whilst simultaneously giving it up for little to no reward, often for small things such as drawing the attention of peers on social networks (Kokolakis, 2015).

Risks from social media

When using online social networks, there are potential risks for the user. Some potential risks may include “embarrassment, stalking and identity theft” according to Gupta and Dhami (2015). One of the primary concerns for this potential risk of identity theft, is that due to the large amounts of information being shared online, such crimes are becoming easier and more common to commit.

Identity theft can be defined as ‘involving the collection and use of an individual’s personal information, without his or her consent, for criminal purposes’ (Koops and Leenes, 2006). There has been a growth in the number of identity theft crimes recorded over several western countries in the past decade. According to Cifas (2016) there has been a 57% increase in cases of identity fraud in the UK across all age groups from 2014 to 2015. It has been suggested that due to a significant increase in the use of technology and the internet by people around the world, that this may be fueling the rise in identity theft incidents (White and Fisher, 2008). In support of this position, there are numerous studies which suggest a link between certain types of behaviour online and an increased risk of victimisation by criminals for phishing attacks, ID theft and a variety of other targeted attacks (Holt and Bossler, 2008; Holtfreter, Reisig and Pratt, 2008). It is however important to note that these studies are not indicative of a proven link between the use of social media and an increased risk of cybercrime victimisation.

A worry is that cyber criminals are exploiting information that they find on social media as part of their crimes. Information such as birthdates, mobile numbers, locations and friend's contacts are all useful information for a variety of crimes (Nagunwa, 2014). This concern is shared by van Wilsem (2011) who states that a person’s online activities can result in greater visibility to offenders, and that it is possible to ‘unconsciously send out signals of careless behaviour’ such as revealing personal details about themselves.

This concern for an uptake in identity theft and related crimes, is particularly prudent when considering the sheer number of social media users and the information they’re sharing online. The potential that there is a link between the use of social media and an increased risk of cybercrime victimisation should be reason enough to instigate a push towards better privacy online.

Some common uses for personal information that has been found by criminals include [7]:

1. **Identity Fraud**
    - Using personal information from another person to impersonate them and commit crimes in their name.

2. **Phishing scams**
    - Using personal information found, attackers can more easily commit believable phishing scams.
    - This can be known as 'spear phishing' when referring to specifically targeted attacks.

3. **Hacking**
    - Using information from online, it could be possible to try and answer or bypass standard security measures for other online services, such as security questions for resetting passwords.

According to a study of global security policies, 34% of IT professionals are not aware of, or do not understand the security policies of their organisations (Cisco, 2008). Although this report refers to corporate security policies and not individual users, it does promote a similar view of how people view policies as Aïmeur et al, (2016) provide when referring to the "widespread" ignoring of privacy policies. These studies indicate that using formal policies for educating users about the risks of social media may not be as effective as hoped. Therefore other methods of mitigating the risks of social media usage, and educating people may prove to be beneficial. This is especially important given that the great many risks to users across social media, and that the situation does not look likely to change anytime soon.

Risk mitigation

Some common ways of educating users on the subject of social media security and privacy is through the use of formal policy and training. Nonetheless as discussed in the previous sections, has its own problems as formal policies are often misunderstood or ignored by a fairly large amount of people.

Tayouri (2015) suggests that using 'complementary layers' of protection can provide better results for handling what he terms as the human factor in security. These layers, he continues, can be divided into two key aspects: education and training, wherein users are taught the dangers of

7 These crime examples from (Nagunwa, 2014)

social media and how to minimise their own risks; and technology, where the use of spam filters, firewalls and other technology can provide additional protection for users. Security policies are also recommended by He (2012), who suggests that the social media security policies should be updated and communicated regularly. Although this situation is specified more towards a corporate setting, this does further confirm the use of security policies as a commonplace measure for educating users.

In both papers, there is some element of utilising technology to provide protection for users. He (2012) suggests using firewalls and web monitoring to protect users from various risks online. This use of technology is also recommended by Tayouri (2015), who suggests using it as a 'complementary layer' of protection. This multiple 'layer' strategy is based on his suggestion to use multiple layers of defence to combat multi-layered attacks. This idea is interesting, however much like He, the use of technology is made to clearly be a separate element of protection than education.

This multi-layer approach demonstrates that the use of technology is clearly viewed as an independent approach to the problem, however because of this view there appears to be a gap in the literature. As far as the research conducted for this project is concerned, technology and education are well documented as separate entities for separate purposes. There does not appear to be anything documented about using technology as a tool to improve education.

Natural Language Processing

Learning information about the account user from their tweets is the primary objective of this project. As a result a large section of the program will be dedicated towards taking in text data from tweets, and extracting and analysing information from it will be an important technical challenge. Natural language is inherently subjective, ambiguous and difficult to understand in a meaningful way for a computer, and therefore requires the use of different techniques to provide structure, clarity and meaning to otherwise subjective text data. These topics fall under the category of natural language processing (NLP), which is an area of research that seeks to explore how computers can understand and use natural language for various applications (Chowdhury, 2005).

As NLP is a very large and complex field it is very difficult to adequately research specific topics with any degree of complexity within the constraints of a literature review. Therefore, several key topics related to NLP, which may prove useful to this project have been briefly defined and discussed below.

Tokenisation

To be able to perform NLP processes on text values, it is often important that the text be tokenised. This involves splitting the text into tokens, commonly individual words, sentences or paragraphs depending on the size of the text and the requirements of the program.

POS Tagging

The purpose of POS tagging is to label words into specific categories based on their parts-of-speech, such as *noun, verb*...etc (Nlp.stanford.edu, 2017). The technique is commonly used as a step among other stages of NLP techniques which rely upon the POS tags to function.

Sentiment Analysis

Sentiment analysis is used to determine the opinion expressed in a piece of text by the author. The technique works by comparing the words and phrases of the text against a given corpus which has been used as training data for the sentiment classifier (Bird, Klein and Loper, 2009).

The nature of extracting information from tweets may find this technique to be useful and important during the development of the project.

Grammars

A grammar is a set of rules that describes the structure of a language. They are used, among other things, as part of named entity recognition to determine the relationship between different the words in the text (Clark, Fox and Lappin, 2010).

Summary

The research conducted as part of this literature review has yielded the following key results that are directly applicable to this project.

Social media usage significant and widespread, with a variety of different social network sites offering different purposes for different audiences. These services provide privacy policies for users, but these are widely ignored or not understood.

The privacy risks associated with sharing personal information through these social networks appear to be understood and appreciated when people are asked, however their behaviour doesn't

seem to match their views in this regard. As a concept, privacy is hard to define, though there are some common factors across different definitions that have been taken and used as criteria for determining good privacy. There is a gap in the literature regarding risk mitigation, standard recommendations include using various methods of education and technology. There was however no mention of using technology to try and assist education.

There are many NLP techniques to handle and extract different information from unstructured text. These techniques will be used to help inform and guide the design and implementation sections of this project.

Method

Software Methodologies

The development of this project has aimed to follow the principles of an agile methodology, as it provides greater benefits than those offered by alternate methodologies such as waterfall. Due to the nature of this project being a completed as a sole developer, many of the benefits that different methodologies have for managing and dividing work among different people are irrelevant. Therefore, the selection of a methodology was based on the benefits it could provide to the development and testing stages of the project.

The requirements of this project (detailed in the *design* section) are the goals for the program to achieve in regards to the functional output of the program. However the specific details of how to achieve those goals or how they could be improved upon implementation are left flexible to allow more freedom in the development stage. It is this freedom and flexibility which enabled the development of better features by allowing feedback from rapid development and testing to improve the project.

Extreme programming (XP) is a methodology with principles that align with the agile manifesto, and focus on simplicity, “The core XP practices here are Pair Programming, Simple Design, Test-Driven Development, and Design Improvement” (Highsmith and Cockburn, 2001). With the exception of pair programming, these principles all align well with the nature of this project. A simple design (see *design* section for details) with a focus on completed functionality rather than specific methods for implementation, is in keeping with XP principles. The short timespan between development and testing of each feature, allowing improvements to be made, where appropriate, to the design and implementation of the feature also strongly follow the principles of XP methodology.

Project Management

To track the the current state of each task and overall progression of the project, the online software tool *Trello* was used. The software allows for a number of lists to be created and individual tasks may be added and moved between these lists (see *Fig 1.*), allowing for the representation of different states of progression for the project.

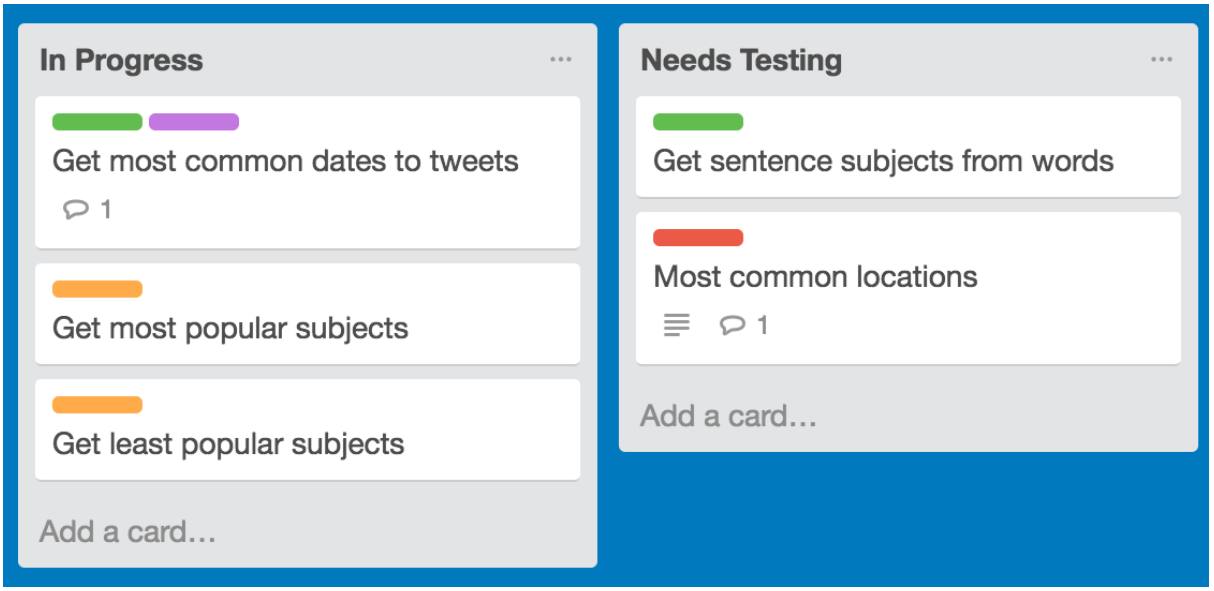

Fig 1. Project management with Trello

As per standard practice in modern software development, this project made use of a VCS (Version Control System) to backup and maintain the code in an online repository. In this case the repository was hosted on *github.com* which uses *git,* the most common VCS used by developers worldwide (Stack Overflow, 2015).

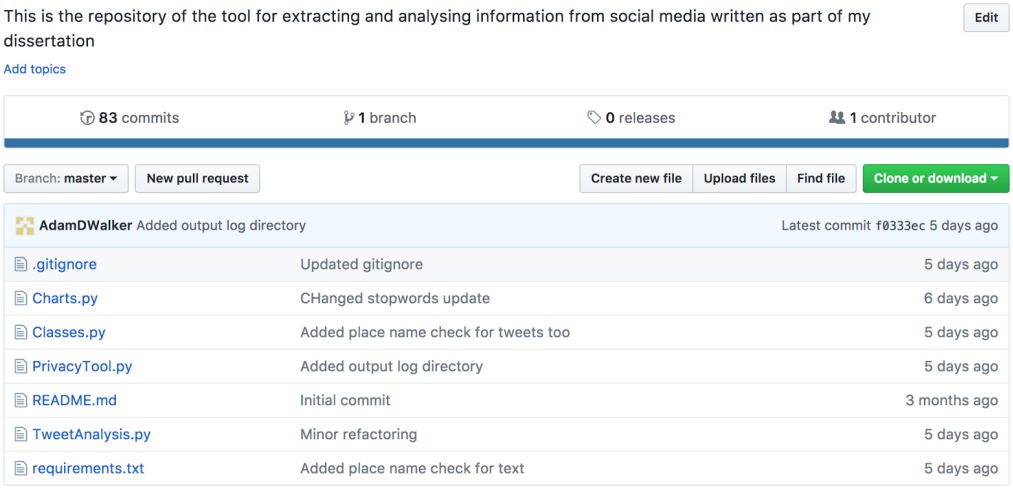

Fig 2. Project repository on Github

Throughout the project there have been supervisor meetings on a weekly or biweekly basis, depending upon the stage of the project. These meetings have served to provide feedback on current progress at that time, update targets for progress to be achieved by the next meeting and

generally discuss the project. These meetings also served as small milestones for progress and keep each step of the project in scope.

Tools

Languages

The choice of language for the development of this project has a significant effect on the structure of the program. Furthermore, the choice of language has several key factors which affect the decision, these factors are compared in Table 3 below.

Although it is a highly powerful and capable language, C++ was rejected simply because the power it offers over other languages is not required for this project. The program is not being developed to be the fastest most efficient implementation, but rather to be functional and for a study. This also applies to Java somewhat, though not as much.

Language	**Python**	**C++**	**Java**
Familiarity	Intermediate	Intermediate	Intermediate
Development Speed	Very Fast	Medium	Medium
Version	3.6	C++ 17	JDK 8
Interpreted / Compiled	Interpreted	Compiled	Compiled

Table 3. Comparison of programming languages by attributes

The choice of Python to develop this project was done so because of the speed and ease of development when compared to the other languages. Python, being an interpreted language can be written with just a text editor and environment to run the scripts (i.e. without the need for a compiler) which increases the speed at which testing and changes can be done.

Environment

The development environment for this project was a Python 3.6 environment, with the text editor Atom, which was chosen due to it being free and open source with community developed package integration providing custom functionality. Furthermore, as a text editor rather than IDE there is no

particular advantage over any other text editor and as such this functionality is beneficial, but the final choice was mostly due to familiarity with the software.

Libraries Used

Library	Purpose
Tweepy	Wrapper for making Twitter API calls Makes handling Twitter API requests easier
NLTK	Provides vast suite of NLP tools Used for Tokenisation, POS Tagging, Entity Chunking, Sentiment analysis
Twython	Not used directly Needed to add extra functionality for Vader sentiment analyser
Matplotlib	Generate graphs and charts to visualise data
Geopy	Convert longitude / latitude coordinates into an address
Numpy	Provides additional maths & stats functions. Used as part of NLTK and some functions used for Matplotlib
Spacy	Used to parse tweets and enable keyword extraction
Wordcloud	Generates wordcloud images from text data
Gender_Guesser	Estimates a gender from a first name
Geotext	Searches for city, country or nationalities in text

Table 4. Python Libraries used in the project

Research Methods

Ethical concerns

The nature of privacy concerns and collecting personal information as a matter of research raises several ethical concerns which ought be addressed here. As part of the *risks of social media* section

of the literature review it was discussed that, among other more nefarious concerns, there are still privacy concerns with regards to purposes for extracting and analysing personal information from social media.

In parts of the *privacy* and *risks of social media* sections from the literature review, it was discussed that various different third parties may be interested in harvesting data from social networks for their own purposes. This list of third parties included academic researchers who may be conducting studies upon the data they collect, which itself is a privacy concern. This is relevant as this concern applies to this project, which does collect and analyse personal information from twitter profiles. The fact that these Twitter profiles and associated information are in the public domain does not negate the need to ensure strong ethical standards are upheld and that every effort is made to minimise the risk of ethical problems.

Concern	Mitigated by
Participant wishes to leave the study	The participant is free to withdraw at any time and their data will be destroyed
Participant feels uncomfortable with the test	The participant is free to stop the test at any time
The participant does not understand how the test will affect them	The consent form details the structure of the test, and there is a time dedicated to ask questions before the test begins
Unable to get consent from people outside of the study whose data may be accessed during the study	No data is permanently stored anywhere, data is removed after the completion of the test by each participant. If data from someone is accidentally accessed, no details will be made of any of the information and the data will be removed ASAP.
Sharing datasets in prohibited in the Twitter API Terms of service.[8]	No data will be kept longer than the length of the study, and therefore no datasets will be shared with anyone.

Table 5. Ethical concerns

8 Details from (dev.twitter.com, 2017c)

Model Validation

In order to perform this validation, there will be a selection of test data for the program to analyse. This data will be hand selected and then manually analysed to determine what information can and should be returned from the program. The program will then run using this test data and the output will be compared against the predicted output from the manual analysis. If these two outputs match with significantly similar results then the program is determined to have performed successfully, and the model is validated.

This validation stage is a predominantly qualitative research method, as it focuses on a smaller set of data that will be analysed for subjective qualities. The output of the tests will be manual descriptions of data the program should find, furthermore since this does not fit within the definition of quantitative research as the data is not analysed through statistical comparisons or detailed categorisations, the overall method can be considered to be qualitative in nature.

User Study

In order to answer the research question of this project, a user study was conducted to test the hypothesis that using a program to demonstrate what private data collection from a twitter account may prove to be surprising or educational to users. The dependent variable in this study is the participant's view of their own privacy, as this is what is being measured. The independent variable (i.e. what changes during the test) is the amount of knowledge about their own privacy that the participant has, which is changed due to the demonstration of the program.

The study involved 21 participants, most of whom were selected from the school of computer science student body. The criteria to take part in the study was that they must be a user of twitter and be over 18 (which as university students they all were anyway). The test would take approximately 15 minutes per person, and began with an explanation of the project and the ethical considerations therein, followed by a chance for the participant to ask any questions and sign a testing consent form (see *appendix C*). The study performed within group testing, as there was only one group of test participants, and the nature of the study had no need for more groups.

The study involved three main components for each participant. A short survey containing questions about the general background of the participant regarding Twitter. This was followed by a series of questions about the feeling the participant had towards their personal information and how

likely the believe it is that the particular piece of information could be learned from their twitter account, rated on a five point Likert scale.

The second stage of the test was a demonstration of the program developed for the study. This stage would involve the participant's twitter account name being provided to the program to perform its analysis on their account. The participant would then be shown the resultant information and charts generated by the program, with a short discussion of the results (see *pilot study* for details).

The final stage of the test involved the participant completing an identical second copy of the survey they completed before the demonstration, regarding their views of how likely it is that a particular piece of information could be learned about them from their Twitter. This second set of questions allows for a comparative analysis to be performed on the results to determine if the demonstration of the program has any affect on the participant's privacy views.

In addition to the quantitative data questions in the form of Likert scale interval data responses, there were also qualitative data questions asking the participant their reasoning behind each response. The purpose of using a quantitative data as the main data for this study is to provide objective, interval data which can be analysed and visualised more easily.

Pilot Study

To ensure that the test remained as consistent as possible for each participant, a short pilot study test was conducted to determine the structure and general flow of the test with regards to how information is presented and in what order this would occur. This was done so as to minimize any variation across tests and prevent any potential bias or changes from affecting the test (see *appendix E* for the specific test structure*)*.

During each test, some observational notes were taken regarding the reactions and thoughts of each test participant to the information they were being shown as part of the test, these observations will be presented in the *results* section and discussed in the *evaluation* section.

The survey (see appendix A) was divided into three sections for the participant to complete. The first section was a short section containing some questions regarding the participant's gender, average twitter usage and any methods they may use to try and protect their privacy on twitter. The second and third sections of the survey contained the same questions and were to be answered one before and one after the demonstration of the program was completed.

Design

The development of this project has followed the XP methodology, which focuses on a simple design and iterative improvements to be made over the course of development. This being the case, the design of the program has tried to remain open to flexibility and to allow changes to the design where improvements are possible. Therefore the design is predominantly based around functional requirements for the completed program to achieve, and these requirements have been established as they should enable the finished program to complete its purpose.

The design of the program for this project was done in an iterative manner, such that after initial prototyping was completed and functional requirements for the program had been defined, development work could begin. Therefore this section will discuss the prototype testing conducted, and what impact it had on defining the requirements of the program.

Prototype Testing

To gather some details on what kind of information is available on a Twitter account, a simple by-hand test was conducted on a select few Twitter accounts. This was done to determine what kind of information is immediately available, as well as to try and inform future development decisions about the functionality of the program.

The accounts used in this prototype stage were not chosen for any particular reason, and were simply used to examine the type of content that was posted, any noteworthy features of the account as well as numbers of tweets and other metrics. No data was stored from these tests, and only observational notes were kept. The key observations from this prototype testing are detailed below.

1. **Account description often contains very identifiable information**
    - The description information, when present, often contains information about the person themselves. Such as their job, location or just personality traits.

2. **Some accounts contains many thousands of tweets**
    - Though these vast amounts of content would provide more data to work with, the Twitter API limit of 3200 tweets still applies.

3. **Several ways of identifying locations**
    - Tweets can be checked for location coordinates
    - Tweets can be read for location names in the text
    - The account description can be checked for location names

4. **Liked and retweeted content may be less useful than normal tweets**
    - Tweets created by the user are more immediately relatable to that creator, rather than other content which may be less relevant.

Considering the different social networks for the development of this project, as discussed in the *social media* section of the literature review, it has become clear that for the purposes of this program a more general content style network is better. This allows for the maximum amount of different media to be shared online, which in turn means a broader range of information available for the program to access. In a choice between the social networks discussed, the top two generalised content sharing social networks by user numbers are Facebook and Twitter.

Of these two sites, Facebook commands the greatest user count by a very wide margin, though this does not necessarily equate to a better site for this project. In fact for the purposes of this project having a wider variety of potential content on the site may provide a better range of information available for the user to share, and by extension allow the program more to work with. The fact Twitter seems to have less sophisticated privacy control settings for users however is probably the key influential factor in this decision. Whilst Facebook and Twitter have many similar characteristics and features, Twitter users have less control over their privacy which is significantly more useful for this project.

Requirements

These requirements are expectations for the function of the completed program, and are deliberately flexible in their declaration to allow for changes to the design along the way. These requirements are defined from a combination of what the program must be capable of achieving for the project aim, observations made from the prototype tests (see above) and from information researched as part of the literature review.

Connect to Twitter

The program must utilise the Twitter API to authenticate the program with Twitter, and allow further access to data through the API.

Download Twitter data

Having connected to Twitter, the program must download a copy of any data it needs so as to perform local analysis on it.

Identify personal details such as:

- Name
- Gender
- Contact information
- Age
- Location

As described in the *privacy* and *risks of social media* sections of the literature review, pieces of information such as these are particularly attractive to criminals looking to commit ID fraud (Nagunwa, 2014). Furthermore, these are all pieces of information that were found to be pretty common in the immediate profile details on a Twitter account, and as a result these are all pieces of information to make a direct effort towards collecting.

Analyse tweets to find:

- Important topics
- Opinions on topics
- Locations
- People of interest

Sentiment Analysis

To perform sentiment analysis on the tweet data requires a sentiment analyser. It is likely that using a custom dataset to train a classifier specifically for this problem would result in the most accurate analysis tool for this project, however this is not practical for several reasons. Firstly, there needs to be a large dataset available to train a classifier tool, and this is something beyond the resources of this project. Secondly, the time and effort it would take to train and properly test a sentiment analyser would be significantly wasted, as this project is not attempting to find a better or more

efficient analyser tool, therefore there is simply no need for this when a pre-made tool will solve the problem far more quickly.

Vader is a sentiment analyser provided by *NLTK* and is trained specifically on social media text, therefore providing a more accurate result when applied to social media text data than alternative analysers (Hutto and Gilbert, 2014).

Output this information

- Charts
- Output log files

Implementation

This chapter will discuss and evaluate the technical considerations with regards to the implementation of this project, technical considerations and any significant issues encountered during development. This will follow the outline of requirements (as defined in the *design* chapter) and describe the process of implementing each of these features.

Class Structure

Making use of the OOP style allows for the program to encapsulate the extracted data into objects which fit the need of the program well. The program makes use of several scripts in order to maintain a clear structure to the code. In the *Classes* file there are two classes, firstly, a TwitterAccount class which is used to store information regarding the account user, such as real name or follower count. The second class is for tweets, this class is used to create a tweet object for each tweet downloaded from the target account. The class contains a number of different variables to store each piece of information relating to that specific tweet. These tweets are then stored in an array inside the TwitterAccount class.

Twitter API Rate Limits

The Twitter API makes use of rate limiting on a per user access token basis. These limits are the maximum amount of request to the API that can be made in any 15 minute interval. These rate limits (see *table 6)* were enough to be able to work without any real change in the way the program was tested or used in any demonstrations for the user study, however they were a constraint worth considering and with different program requirements could have been a problem for development.

Action	Rate Limit
GET status/user_timeline	900
GET status/retweets_of_me	75
GET followers/list	15
GET geo/id/:place_id	75

Table 6 - Twitter API rate limits per window (dev.twitter.com, 2017b)

Connecting to Twitter

In order to access Twitter data, it is first required to connect to Twitter. In order to achieve thishe Twitter API must be used, which requires OAuth authentication from the program trying to connect, the authentication requires API access tokens for OAuth verification to succeed. These tokens are generated and provided to developers by creating a Twitter application.

Having created the application on Twitter and generated the access tokens, these can be provided to the Twitter API OAuth call functions to authenticate the application and connect allow access to Twitter data. In order to connect with Twitter, the program makes use of a python package called *Tweepy* This package serves as a wrapper to simplify and handle Twitter API requests.

```
# OAuth process, using the keys and tokens
auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_token_secret)

# Creation of the actual interface, using authentication
api = tweepy.API(auth)
```

Downloading Tweets

To iterate through many pages of data, such as a timeline, on Twitter requires pagination. This pagination can require a lot of boilerplate code, therefore the Tweepy library provides a *Cursor* object to handle this and make accessing the data much simpler. This Cursor object is used to access the user timeline and download up to the maximum of 200 tweets per page, for up to 32 pages. These numbers are used as the Twitter API allows for access to a maximum of 3200 tweets, with a maximum of 200 tweets per page. Additionally, the page count is double the theoretical limit of 16 pages to account for potential overrun from pages which return less than 200 tweets.

```
for page in tweepy.Cursor(api.user_timeline,
                          id = user.screen_name,
                          count = 200,
                          include_rts = include_retweets).pages(32):
   page_list.append(page)

for page in page_list:
   for status in page:
```

```
        tweet = Classes.Tweet(status.text, status.created_at, status.coordinates)
```

Identify personal information

To identify personal information in the program takes several different stages, depending upon the information being gathered. The different pieces of information that are directly available from a Twitter account as directly defined in the *design* section, can mostly be accessed specifically from the account user details.

- Name
- Gender
- Job / occupation
- Description

To identify the immediate personal details as listed from the design requirements, this section focused mostly on extracting information stored as the account details rather than any specific tweets. Each Twitter account contains a *realname* value for the account, as well as a number of other attributes which can be accessed and displayed to the user (dev.twitter.com, 2017a). The results of this data are logged to the console window for the user to see (see *fig 3*).

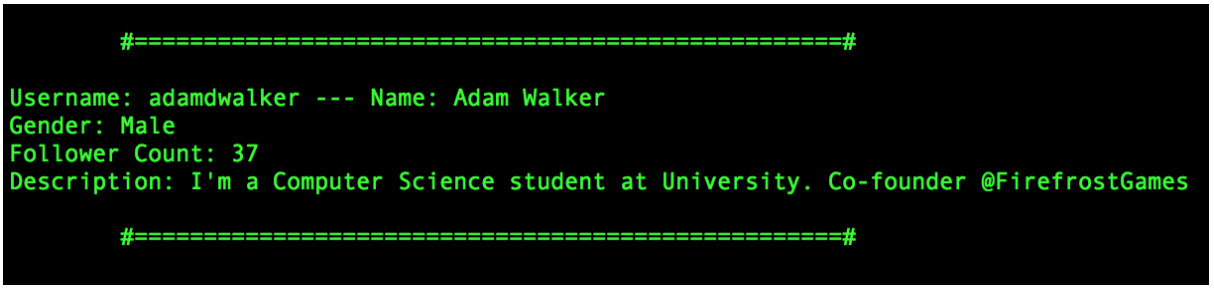

Fig 3. Program display of account details

Analysing data

Sentiment analysis

Extracting the sentiment values from each tweet was done using the Vader sentiment analyser tool from NLTK. Vader is a pre-trained sentiment analyser, which has been trained especially for social media data. In order to return sentiment scores for a given piece of text, firstly an instance of the sentiment analyser object is created. This object has the function *get_polarity_scores()*, which will

return a dictionary of positive, neutral and negative scores as well as a compound score which represents the total sentiment of the text.

Keyword extraction

To collect the important topics and keywords out of the tweet text data, keyword extraction must be performed. The first attempt at extracting keywords was a failure and did not correctly function. The second attempt, utilising a slightly different method however succeeded with a reasonable degree of accuracy. The first step in extracting keywords is to get the text data into a sufficiently structured format so as to make the task easier. This is accomplished by tokenising the text into individual words, and removing an undesirable parts of the text (i.e. stripping punctuation, removing stop words and any usernames or hashtags).

The first method of extracting keywords that was attempted was using parse trees and named entity recognition. The way this attempt worked was to use grammars and NE recognition to parse the text into a tree structure and then determine the various relationships and entities in the text. This could then be used to find the subject of the sentence, as well as key nouns and verbs based upon the relationships in the parse tree.

Unfortunately this approach suffered a series of technical problems with the grammar and NE recognition approach and was abandoned in favour of an alternative approach using a more intuitive library. *Spacy* is a Python library which handles the same techniques as described above but with its own grammars and parsing contained in simple wrapper functions. This therefore allowed the words to be categorised, and the final output keywords are returned nicely without the manual trouble of the first attempt.

```
def getKeywords(text):
   keywords = []
   text = nlp(text)
   for sentence in text.sents:
       for word in sentence:
           if(word.dep_ == "ROOT"):
               keywords.append(str(word))
           elif(word.dep_ == "nsubj"):
               keywords.append(str(word))
           elif(word.dep_ == "dobj"):
               keywords.append(str(word))
   return keywords
```

Location

When the tweet is downloaded, the tweet object can contain a geo attribute which is used to store the coordinates of location that tweet was sent from, provided it is location enabled. If this is the case, the tweet object will contain longitude and latitude coordinates. When using the *geopy* python package, it is possible to perform a reverse lookup on these longitude and latitude coordinates to return an address value for that location.

There is also a secondary implementation of location detection in the program. Using the *geotext* Python library, the program searches through the account description and the text of each tweet to compare against a database of locations.

Associated Users

To find a specific name in a larger string of text, with the only definable characteristics being that it starts with an @ symbol and can contain any letters, numbers or an underscore symbol (Twitter help center, 2017b). This would normally mean writing a complex piece of code to iterate through the string and make lots of checks for characters. The exact same outcome however, can be achieved through the use of regular expressions.

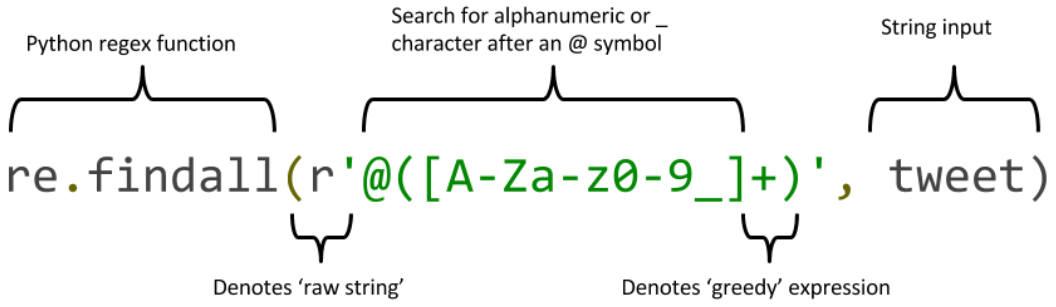

Fig 4. Breakdown of a regular expression

This regular expression is used to search for any username within the text of a tweet. It uses the python *re* package to allow for the use of regular expressions. The *findall()* function takes a regex pattern and input string as parameters, and will return all non-overlapping matches of the pattern in the string (Docs.python.org, 2017). This is also a *greedy* expression, which means it will return the longest match it can find rather than a *lazy* expression which will return the shortest match.

Hashtag extraction

The use of hashtags on Twitter allows users to index their tweets with keywords, starting with the # symbol (Twitter Help Center, 2017a). These hashtags are extracted with another regular expression, modified from the username extraction expression.

$$re.findall(r'\#([A-Za-z0-9]+)', tweet)$$

Output visual display of data

To display the results of the program in a clear and coherent way requires several different implementations, to account for the differing formats of data and how best to represent them.

The program contains a *Charts.py* file which contains a selection of functions to generate and display different kinds of graphs. These output graphs are generated using the *matplotlib* library, as a way to display the information gathered by the program in a more clear and visually appealing way to the user.

```
def generatePieChart(n, data, labels, explode, title, filename):
   fig = plot.figure(n)
   ax = plot.axes([0.1, 0.1, 0.8, 0.8])

   ax.pie(data, explode, labels, autopct='%1.1f%%',
          shadow=False, startangle=90)
   ax.axis('equal')  # Equal aspect ratio ensures that pie is drawn as a circle.
   ax.set_title(title)
```

These functions are available so as to enable the program to scale up if required, such that if new data is collected it need only be passed to these functions and new charts will be added to the output directory.

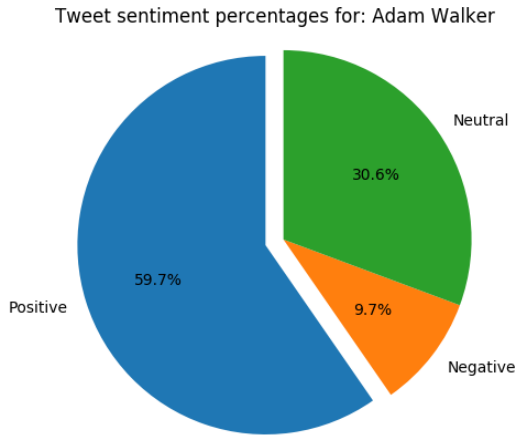

Fig 5. Example pie chart

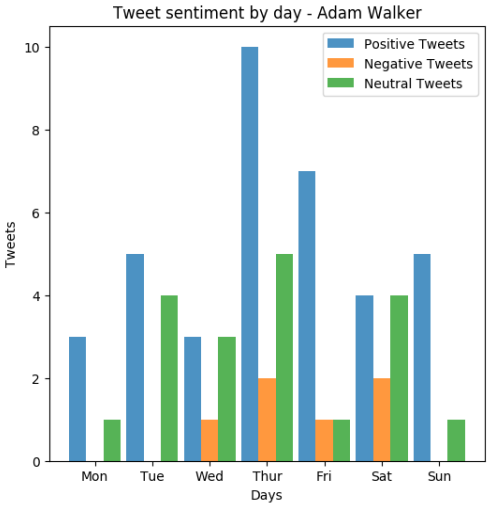

Fig 6. Example bar chart

Testing

Testing was completed in an iterative process as each feature was developed. The process involved working on a feature until it was complete, with periodic reviewal and testing of each small code change in keeping with XP principles.

This method of repeated small tests and refactoring code as development was still fully underway enabled the debugging process to be much simpler when an issue did arise, as the total amount of code to inspect and test was much smaller. Moreover the quality of the code was felt to be higher in general due to this process.

Results

Observations

This is a selection of key observations made by the researcher during the test, focused on the reactions of the participant towards the results of the program upon first seeing the data. Observations are noted as short bullet points and any relevant quotes from the test, see *appendix B* for the full set of observations.

Test No	Observations
4	- Surprised by the amount of data - *"It's not private at all is it?"*
10	- Everything is publicly available - Job + Youtube exposure mean the participant was sharing significant amounts of information - *"Someone could steal my life if they wanted to"*
12	- Participant was ok with what they share on Twitter - *"I imagine that I don't have much privacy though"*
15	- Participant was aware of the information they were sharing but still very surprised at how much data was available, and what could be inferred from that information.
16	- Participant had very low confidence in their privacy

Table 7. Key observations from test participants

Demographic Questions

In order to draw some comparisons from the results of the primary survey questions, demographic questions were also asked of the participants so as to provide some alternative data for analysis.

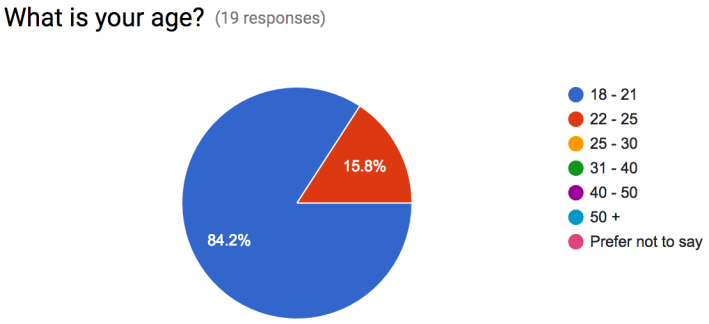

Fig 7. Age results chart (Pre-Demonstration)

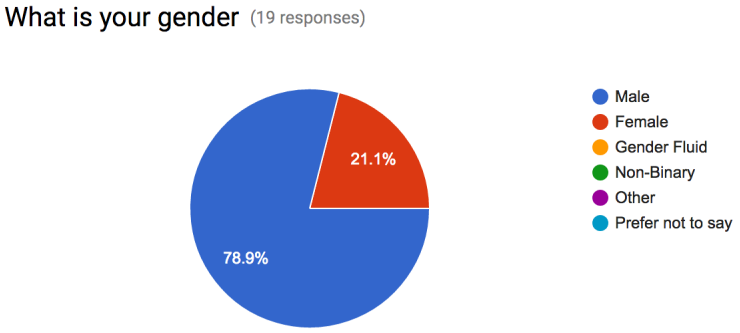

Fig 8. Gender results chart (Pre-Demonstration)

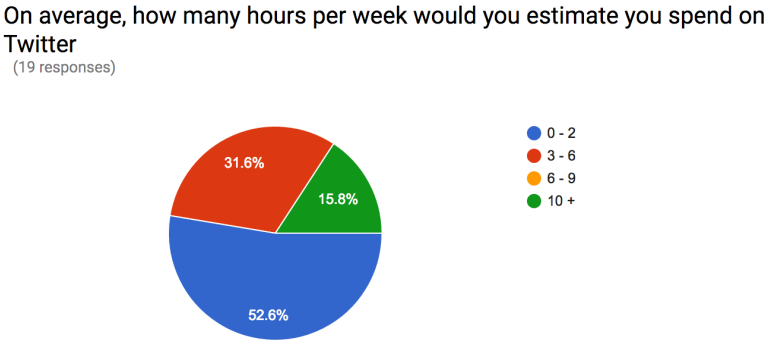

Fig 9. Twitter usage results chart (Pre-Demonstration)

Please describe what, if any, methods to protect your privacy you use on Twitter
- *"None"*
- *"N/A"*
- *"Pseudonym for actual username. Don't use it often. Don't post about personal stuff"*
- *"Post and then delete later if the post could be misconstrued as something else"*
- *"Don't use my real name I use an online alias"*
- *"Literally none"*

Table 8. Privacy protection methods results

Privacy survey before demonstration

This survey asked each participant to rate on a five point Likert scale, how likely they felt it was that the piece of information in question could be learned about them from their Twitter account. Each of these questions was followed by an optional description about their reasoning, and some of these quotes have been included where appropriate.

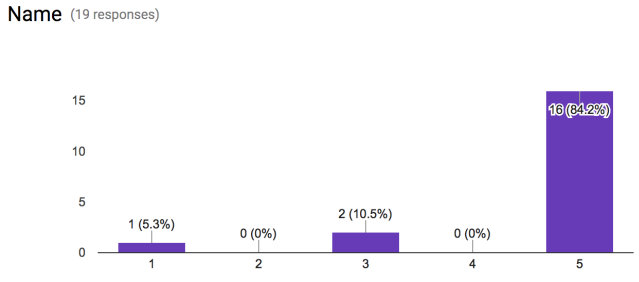

Fig 10. Name results chart (Pre-Demonstration)

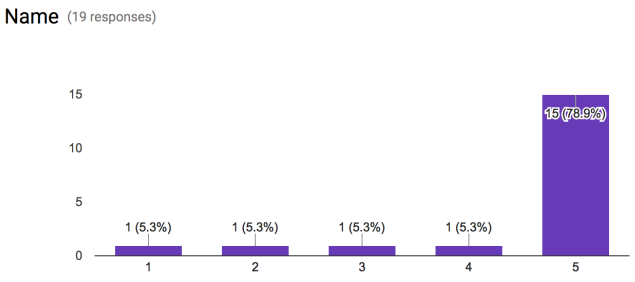

Fig 11. Name results chart (Post-Demonstration)

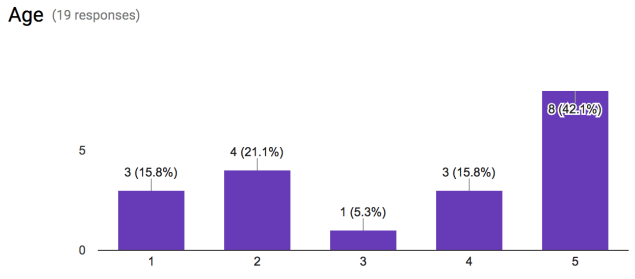

Fig 12. Age results chart (Pre-Demonstration)

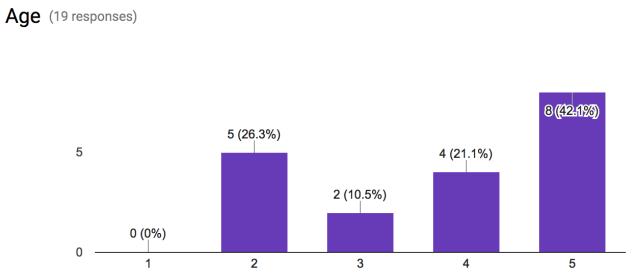

Fig 13. Age results chart survey (Post-Demonstration)

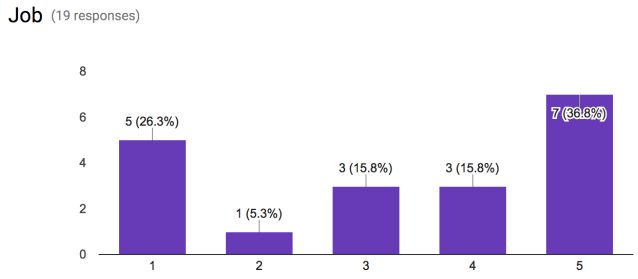

Fig 14. Job results chart (Pre-Demonstration)

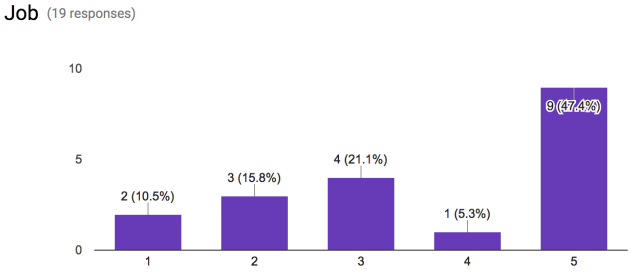

Fig 15. Job results chart (Post-Demonstration)

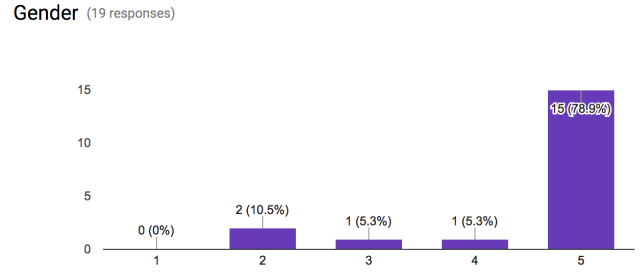

Fig 16. Gender results chart (Pre-Demonstration)

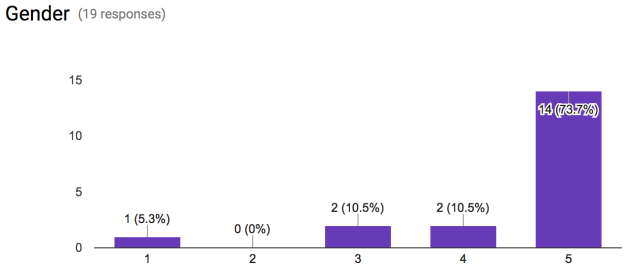

Fig 17. Gender results chart (Pre-Demonstration)

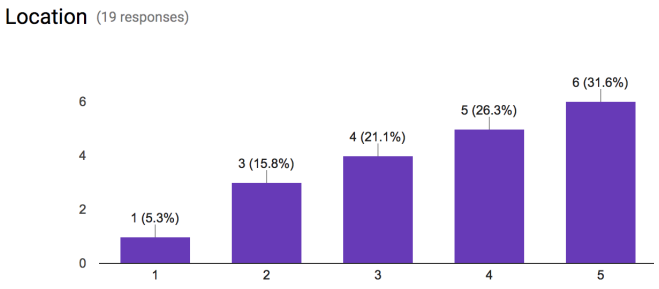

Fig 18. Location results chart (Pre-Demonstration)

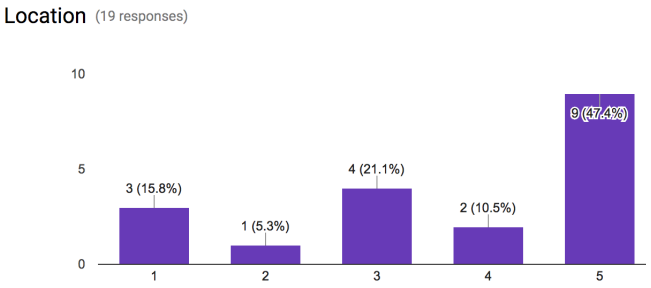

Fig 19. Location results chart (Post-Demonstration)

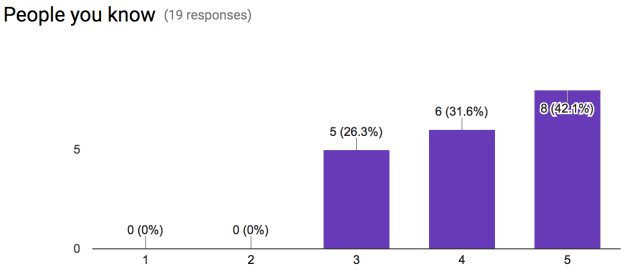

Fig 20. Associated people results chart (Pre-Demonstration)

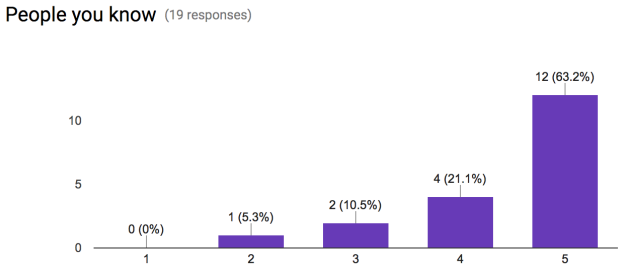

Fig 21. Associated people results chart (Post-Demonstration)

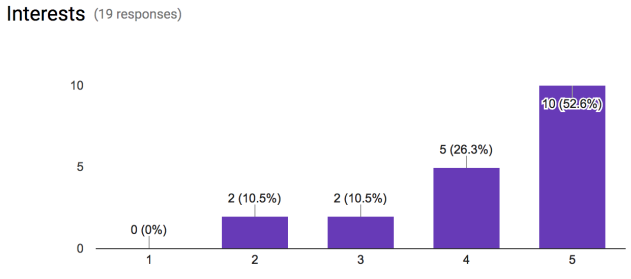

Fig 22. Interests results chart (Pre-Demonstration)

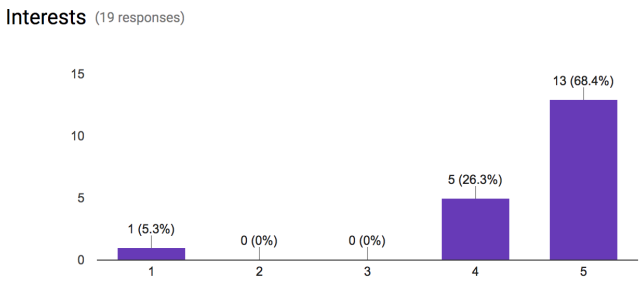

Fig 23. Interests results chart (Post-Demonstration)

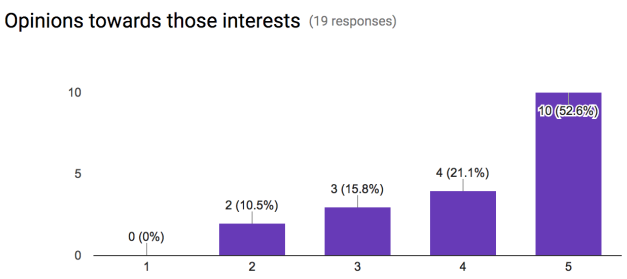

Fig 24. Interest opinions results chart (Pre-Demonstration)

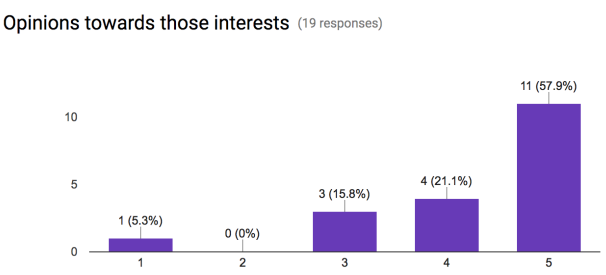

Fig 25. Interest opinions results chart (Post-Demonstration)

Overall how would you describe your confidence in your privacy on Twitter? [9]
- *"I feel secure about the information I choose to share."*
- *"I'm OK with what I post being in the public domain. It's fine."*
- *"Not very confident"*
- *"I have very little confidence, though I consider this to be my own fault"*
- *"Relatively confident"*
- *"Not very private but that's all social networks"*
- *"I know there's a lot of information out there but I don't think anything too private is out there."*
- *"Uncertain. Could be compromising information"*
- *"Not confident, I have plenty of information on display."*

Table 9. Privacy confidence on Twitter results (Before Demo)

9 Responses may have been corrected for spelling and grammar mistakes. Not all responses shown, duplicates have been removed.

Overall how would you describe your confidence in your privacy on Twitter?
- *“Same as before I believe most of the data I want to keep secure is secure. However, I am now more scared of location data.”*
- *“It's good enough for me. I don't post anything that I wouldn't want people to know.”*
- *“Nothing is private”*
- *“Not very private at all”*
- *“Still pretty confident”*
- *“I knew there was a lot of information out there, but I didn't quite expect that amount of information could be learned from it.”*
- *“My information is not entirely secure. Can quite easily be inferred.”*
- *“Very confident, moreso, as I didn't know that I refuse location information per tweet. I'm aware of what I'm letting public, and I'm fine with this.”*
- *“My twitter is an open book, there is almost no privacy.”*

Table 10. Privacy confidence on Twitter results (After Demo)

Evaluation

Results Discussion

The results for the user study have been presented in the *results* section, and this section will contain a discussion and analysis of these results.

The specific details and justifications for this study are provided in the *research methods* section. The purpose of the freeform questions asking for the participant's reasoning for their answers was to provide a more qualitative set of data, which is interesting to examine but also backed up by quantitative data in the form of Likert scale responses. These Likert scale responses represent how likely the participant felt their information could be learned through their Twitter, with a one representing not at all, and five being definitely.

The total number of test participants was 21, however the data from two participants was removed from the results as these test participants had their Twitter accounts set onto private mode. This private mode prevents the program from viewing any information from the profile, and therefore since it was unable to perform a standard test and the results could not be fairly compared against the other results, they were removed leaving 19 survey responses.

Studying the data seems to suggest that the participants fell into one of two groups with their responses to the questions in section two and three of the study. Some participant answers from before the test show that when asked, they conclude they don't have as much privacy as they might like. This is apparent in the *figs 20 and 22* which have a strong leaning towards the higher end of the scale, stating they feel it is more likely the information will be found. Furthermore, looking at the freeform qualitative data gathered from the study, asking why participants responded the way they did suggests the participants are aware of their privacy in these instances:

- *" Tweeted people I know, spoke about others anonymously "*
- *" Talk to close friends publicly on twitter "*

Across multiple charts there appears to be a small trend towards higher scores on the second (after demonstration) questions over the first questions. This trend can be seen in the

difference between *figs 12 & 13* and *figs 14 & 15,* indicating that the demonstration has changed their views on their privacy.

There are also some graphs which indicate a negative trend in privacy views. Such that some participants have gained confidence upon seeing that pieces of their information are not found in the demonstration.

The results in graphs from before the demonstration indicate this quite strongly in some instances, *fig 10* shows that the majority of participants feel their names can definitely be found from their Twitter profile. Comparing this to the responses from the same question after the demonstration showed a smaller percentage of people compared to before who believed their name could be found.

The data seems to suggest that the people who when asked to think about their privacy tend to feel that they probably have less than they would like, are the most surprised by the amount of data collected. Though this was not the case for most participants, as most of the data seemed to indicate that there was a pretty good awareness of what information was being shared on their accounts.

It is worth noting that the study was conducted with 19 responses (two rejected participants from 21) which severely limits the ability to make detailed analysis or draw many conclusions from the data. Furthermore the participants were all university students under the age of 25, an age range and education level that is by no means representative of the general public and especially being young there is likely to be an increased awareness of social media usage in general, which may have skewed the results of the study. Finally it is also important to address the fact that a large majority of the test participants were from the school of computer science, and as such are likely to have greater awareness of privacy and security issues than non computer science students.

Conclusion

Across most charts there is a small trend towards higher values after the demonstration was completed. This indicates a general trend among participants towards less confidence in their privacy. However as some of the charts contain a slight negative trend this is suggestive that some participants gained confidence in their privacy upon seeing the demonstration.

These two trends suggest that the demonstration has a polarising effect on participants. Where those with less confidence in their privacy tend towards even less confidence, and the

participants with more confidence in the privacy feel vindicated in their views if the program does not collect much information from them.

The fact that the views expressed by participants post demonstration tended to be polarised when compared to the pre demonstration views indicates that a sufficient change on view has occurred to the participants, and therefore the project is considered successful on this basis.

Conclusion

Critical reflection

The aim of this project was to investigate whether using a program to demonstrate the collection and analysis of personal information from a social media account would have an affect on the privacy views of social media users. To achieve this aim, a set of objectives were set to complete the project, and each of these has been completed to some degree of success.

To carry out this research, a comprehensive literature review of existing work in the area was conducted, and it was found that there was little in the way of existing research into using technology for extracting and analysing social media data as a tool for educating users about their social media privacy. This gap in the literature helped to inform the design of the user study so as to gather some additional qualitative data for providing some interesting context behind the quantitative data. This literature review was originally intended to be complete significantly earlier than it was in reality, this is partly due to relative inexperience with academic research and therefore the timeframe for completing was underestimated quite drastically.

The prototype for this project was generally a success in providing relevant information and guiding the design and scope of the program. The method of conducting the prototype test however was, on reflection, not particularly well structured and relied somewhat heavily on random observations. Were this project to be done again, this is something that would be worth reconsidering the approach and finding a clearer more efficient way of examining twitter accounts for relevant observations and notes. Designing the test in advance and producing a set of expectations for a successful outcome would have added focus and improved the quality of the test.

The *methods* chapter discusses model validation to ensure that the program works correctly as intended. This model validation was defined early and as part of testing for the completed program, however due to the way in which the program was developed, it became unnecessary to engage in formal testing and validation of this sort, as the testing had taken part over the entire development stage of the project.

The analysis of the main results, when focused on the change from before and after the demonstration was successful and covered most of the details. However due to time constraints and

technical problems with performing statistical analysis on the data, it was not possible to compare the specific demographic responses to the changes in privacy views. This is an unfortunate result but the key results were completed and the research question answered from this so the testing is overall a success, despite the fact it would have been preferable to analyse the data further with specific demographic breakdowns.

For each of the objectives specified for this project there were two which stood out as less relevant towards the achieval of the overall project aim, and instead focused more upon potential exploration of future work for this project.

These two objectives were an investigation into the potential future use of the program and any data collected. These objectives were unfortunately neglected as part of this project due to their relative irrelevance towards completing the final program and study.

Future work

To continue with this project or similar work in the future, it would be interesting to conduct another study but with a much larger number of participants and from a wider more diverse audience so as to address the key failings in the user study for this project.

Development of the program would also be something to change for future work, in order to take advantage of experience with NLP techniques, there would be an opportunity for greater depth to the program in potential future work. This in turn would allow for a more detailed analysis from any user studies performed, as with a more in depth program should be able to produce more detail for test participants to be shown.

Any future work would benefit hugely from a more detailed time management plan with more strict adherence to the schedule. Perhaps the single largest issue with this project has been sticking to the original timeline established in the project proposal.

References

Ahern, S., Eckles, D., Good, N., King, S., Naaman, M. and Nair, R. (2007). Over-Exposed? Privacy Patterns and Considerations in Online and Mobile Photo Sharing. In: *CHI 2007*. San Jose: CHI.

Aïmeur, E., Lawani, O. and Dalkir, K. (2016). When changing the look of privacy policies affects user trust: An experimental study. *Computers in Human Behavior*, 58, pp.368-379.

Baek, Y. (2014). Solving the privacy paradox: A counter-argument experimental approach. *Computers in Human Behavior*, 38, pp.33-42.

Bird, S., Klein, E. and Loper, E. (2009). *Natural Language Processing with Python*. 1st ed. Sebastopol, CA: O'Reilly Media, pp.10 - 20.

Chowdhury, G. (2005). Natural language processing. *Annual Review of Information Science and Technology*, 37(1), pp.51-89.

Cifas (2016) Cifas - Criminals Target UK Youth As Identity Fraud Rises. [online] London: Cifas. Available from https://www.cifas.org.uk/press_centre/criminals_target_UK_youth_as_dentity_fraud_rises [Accessed 21 February 2017].

Cisco Systems (2008). *Data Leakage Worldwide: The Effectiveness of Corporate Security Policies*. Cisco Systems Inc.

Chai, S., Bagchi-Sen, S., Morrell, C., Rao, H. and Upadhyaya, S. (2009). Internet and Online Information Privacy: An Exploratory Study of Preteens and Early Teens. *IEEE Transactions on Professional Communication*, 52(2), pp.167-182.

Chan, H., Wang, X., Lacka, E. and Zhang, M. (2015). A Mixed-Method Approach to Extracting the Value of Social Media Data. *Production and Operations Management*, 25(3), pp.568-583.

Christofides, E., Muise, A. and Desmarais, S. (2012). Hey Mom, What's on Your Facebook? Comparing Facebook Disclosure and Privacy in Adolescents and Adults. Social Psychological and Personality Science, 3(1), pp.48-54.

Clark, A., Fox, C. and Lappin, S. (2010). *The handbook of computational linguistics and natural language processing*. 1st ed. Chichester, West Sussex: Wiley-Blackwell.

dev.twitter.com. (2017a). GET users/lookup - Twitter Developers. [online] Available at: https://dev.twitter.com/rest/reference/get/users/lookup [Accessed 7 Mar. 2017].

dev.twitter.com. (2017b). Rate Limits Chart - Twitter Developers. [online] Available at: https://dev.twitter.com/rest/public/rate-limits [Accessed 11 Mar. 2017].

dev.twitter.com. (2017c). *Twitter Developer Agreement & Policy*. [online] Available at: https://dev.twitter.com/overview/terms/agreement-and-policy [Accessed 13 Mar. 2017].

Docs.python.org. (2017). *6.2. re — Regular expression operations — Python 3.6.1 documentation*. [online] Available at: https://docs.python.org/3/library/re.html [Accessed 4 Apr. 2017].

Facebook.com. (2017a). *Basic Privacy Settings & Tools | Facebook Help Centre | Facebook*. [online] Available at: https://www.facebook.com/help/325807937506242 [Accessed 10 Mar. 2017].

Facebook (2017b). *Company Info | Facebook Newsroom*. [online] Menlo Park: Facebook. Available from: https://newsroom.fb.com/company-info/ [Accessed 26 Mar 2017].

Fang, Q., Sang, J., Xu, C. and Hossain, M. (2015). Relational User Attribute Inference in Social Media. *IEEE Transactions on Multimedia*, 17(7), pp.1031-1044.

Gupta, A. and Dhami, A. (2015). Measuring the impact of security, trust and privacy in information sharing: A study on social networking sites. *Journal of Direct, Data and Digital Marketing Practice*, 17(1), pp.43-53.

He, W. (2012). A review of social media security risks and mitigation techniques. *Journal of Systems and Information Technology*, 14(2), pp.171-180.

Highsmith, J. and Cockburn, A. (2001). Agile software development: the business of innovation. *Computer*, 34(9), pp.120-127.

Holt, T. and Bossler, A. (2008). Examining the Applicability of Lifestyle-Routine Activities Theory for Cybercrime Victimization. *Deviant Behavior*, 30(1), pp.1-25.

Holtfreter, K., Reisig, M. and Pratt, T. (2008). Low Self-Control, Routine Activities, and Fraud Victimisation. *Criminology*, 46(1), pp.189-220.

Hughes, R. (2015). Two concepts of privacy. *Computer Law & Security Review*, 31(4), pp.527-537.

Hutto, C. and Gilbert, E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. In: *International AAAI Conference on Weblogs and Social Media*. Ann Arbor, MI: ICWSM, pp.1-4.
Instagram.com. (2017). *About Us • Instagram*. [online] Available at: https://www.instagram.com/about/us/ [Accessed 3 Mar. 2017].

Internet Live Stats (2017) Twitter usage statistics - Internet Live Stats. [online] Delaware: Worldometers Licensing. Available from http://www.internetlivestats.com/twitter-statistics [Accessed 3 March 2017]

Kokolakis, S. (2015). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, 64, pp.122-134.

Koops, B. and Leenes, R. (2006). Identity theft, identity fraud and/or identity-related crime. *Datenschutz und Datensicherheit - DuD*, 30(9), pp.553-556.

Kwak, H., Lee, C., Park, H. and Moon, S. (2010). What is Twitter, a Social Network or a News Media?. In: *Proceedings of the 19th international conference on World wide web*. New York, NY: ACM, pp.591 - 600.

Nagunwa, T. (2014). Behind Identity Theft and Fraud in Cyberspace: The Current Landscape of Phishing Vectors. *International Journal of Cyber-Security and Digital Forensics*, 3(1), pp.72-83.

Nlp.stanford.edu. (2017). The Stanford Natural Language Processing Group. [online] Available at: https://nlp.stanford.edu/software/tagger.shtml [Accessed 5 Mar. 2017].

Ofcom, (2016). *Adults' media use and attitudes*. Adult's media use and attitudes. [online] Ofcom, pp.73 - 88. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0026/80828/2016-adults-media-use-and-attitudes.pdf [Accessed 4 Feb. 2017].

Osatuyi, B. (2013). Information sharing on social media sites. *Computers in Human Behavior*, 29(6), pp.2622-2631.

Press.linkedin.com. (2017). *About LinkedIn*. [online] Available at: https://press.linkedin.com/about-linkedin [Accessed 3 Mar. 2017].

Reyns, B. and Henson, B. (2016). The Thief With a Thousand Faces and the Victim With None. *International Journal of Offender Therapy and Comparative Criminology*, 60(10), pp.1119-1139.

Ridley-Siegert, T. (2015). Data privacy: What the consumer really thinks. *Journal of Direct, Data and Digital Marketing Practice*, 17(1), pp.30-35.

Smith, J., Dinev, T. and Xu, H. (2011). Information Privacy Research: An Interdisciplinary Review. *MIS Quarterly*, 35(4), pp.989-1015.

Srinivasan, S. (2012) Lack of privacy awareness in social networks. Information Systems Audit and Control Association, 6, 21 - 25.

Stack Overflow. (2015). *Stack Overflow Developer Survey 2015*. [online] Available at: http://stackoverflow.com/insights/survey/2015 [Accessed 5 Apr. 2017].

Statista. (2017). *Instagram: active users 2016 | Statista*. [online] Hamburg: Statista GmbH Available at: https://www.statista.com/statistics/253577/number-of-monthly-active-instagram-users/ [Accessed 26 March 2017].

Steijn, W. (2014). A developmental perspective regarding the behaviour of adolescents, young adults, and adults on social network sites. Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 8(2).

Tayouri, D. (2015). The Human Factor in the Social Media Security – Combining Education and Technology to Reduce Social Engineering Risks and Damages. Procedia Manufacturing, 3, pp.1096-1100.

Tweepy.readthedocs.io. (2017). Tweepy Documentation — tweepy 3.5.0 documentation. [online] Available at: http://tweepy.readthedocs.io/en/v3.5.0/index.html [Accessed 12 Mar. 2017].

Twitter (2017) Compay | About. [online] San Francisco: Twitter. Available from http://about.twitter.com/company [Accessed 26 March 2017].

Twitter Help Center. (2017a). *Using hashtags on Twitter*. [online] Available at: https://support.twitter.com/articles/49309 [Accessed 6 Mar. 2017].

Twitter Help Center. (2017b). *Why can't I register certain usernames?*. [online] Available at: https://support.twitter.com/articles/101299 [Accessed 29 Mar. 2017].

Twitter Help Center. (2017c). *Protecting and unprotecting your Tweets*. [online] Available at: https://support.twitter.com/articles/20169886 [Accessed 11 Mar. 2017].

van Wilsem, J. (2011). 'Bought it, but Never Got it' Assessing Risk Factors for Online Consumer Fraud Victimization. *European Sociological Review*, 29(2), pp.168-178.

White, M. and Fisher, C. (2008). Assessing Our Knowledge of Identity Theft. *Criminal Justice Policy Review*, 19(1), pp.3-24.

Appendices

Appendix A - Study Survey

Section 1 of 3

Privacy Study Survey Part 1

This study is designed to gather information on how you view your online privacy and how much information you believe you're sharing online.

Please enter a unique ID

This will be used to remove your data if you choose to withdraw from the study

Short-answer text

What is your age? *

- 18 - 21
- 22 - 25
- 25 - 30
- 31 - 40
- 40 - 50
- 50 +
- Prefer not to say

What is your gender *

- Male
- Female
- Gender Fluid
- Non-Binary
- Other
- Prefer not to say

On average, how many hours per week would you estimate you spend on Twitter *

- 0 - 2
- 3 - 6
- 6 - 9
- 10 +

Please describe what, if any, methods to protect your privacy you use on Twitter *

For example, using a pseudonym instead of your real name or setting your account to private mode

Long-answer text

Section 2 of 3

Privacy survey section 2

For each question please rate from 1 - 5 (1 being not at all, 5 being definitely) how likely you feel it is that you're sharing the piece of information in question on your twitter account.

Name *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Age *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Job *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Gender *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Location *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

People you know *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definite

Please state a reason for your answer

Long-answer text

Interests *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Opinions towards those interests *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Overall how would you describe your confidence in your privacy on Twitter? *

Please add any thoughts you have on your privacy when using Twitter

Long-answer text

Section 3 of 3

Privacy Study Survey Part 3

Now that you have seen the results of an analysis on your Twitter account, please answer the same questions again. This test is to determine if your opinions have changed due to this demonstration.

Name *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Age *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Job *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Gender *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Location *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

People you know *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definite

Please state a reason for your answer

Long-answer text

Interests *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Opinions towards those interests *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Definitely

Please state a reason for your answer

Long-answer text

Overall how would you describe your confidence in your privacy on Twitter? *

Please add any thoughts you have on your privacy when using Twitter

Long-answer text

Appendix B - Study Observations

Test No	Observations
1	- Participant surprised about recognising gender
2	- Participant was up front about how much information they shared, and therefore not surprised.
3	- Surprised by location information
4	- Surprised by the amount of data - “It’s not private at all is it?”
5	- Participant was very shocked by the amount of data
6	- Found the amount of information to be about what they expected - Didn’t realise what kind of things could be learned or inferred from it however.
7	- Participant noted parts of their information were either very easily found, or not mentioned at all.
8	- Some information was very explicitly found, such as location. - Other information not found such as job.
9	- Results removed from study due to protected account
10	- Everything is publically available - Job + Youtube exposure meant sharing lots of information - “Someone could steal my life if they wanted to”
11	- Participant tends to share things publicly to followers anyway so isn’t surprised by what’s available
12	- Participant was ok with what they share on Twitter - “I’d imagine that I don’t have much privacy”
13	- Surprised by the amount of things you can infer from their

	information
14	- Pretty confident in their privacy - Didn't share a lot of information, deliberately to protect their privacy
15	- Participant was aware of how much information they were sharing but very surprised at how much data was available and what could be inferred from it.
16	- Participant had very low confidence in their privacy
17	- Removed due to private account
18	- Expected most information to be found, and was mostly correct.
19	- Had much lower confidence after the test than before
20	- Expected most information to be found - Was surprised by what detail could be inferred however
21	- Did not expect location information to be found

Appendix C - Consent Form

Twitter Privacy Study

Researcher: Adam Walker
Email: 11357886@students.lincoln.ac.uk

Supervisor: Chris Headleand
Email: cheadleand@lincoln.ac.uk

In this study, you will be asked to complete a short survey regarding your awareness of the information you're sharing online (Twitter) and how likely you think it is that this information can be used to learn about you. You will then be shown a short demonstration of a program which will attempt to extract information about you from your twitter account. This will be followed by completing the same survey again to see if your opinion has changed at all due to this demonstration.

If you have any questions regarding the nature of the program or study, you are encouraged to ask the researcher at any time if you are unsure of something.

The participant is free to withdraw from the study at any time and their recorded responses will be discarded and destroyed in such an instance.

	Please tick
1. I confirm that I have read and understand the information above and have had the opportunity to ask questions.	☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving reason.	☐
3. I agree to take part in the above study.	☐
4. I agree to the use of anonymised quotes in publications.	☐

_____	_____	_____
Name of Participant	Date	Signature
_____	_____	_____
Name of Researcher	Date	Signature

Appendix D - Study Participant answer reasons

Test	Reasoning
Name	- *"My account is public and includes my name"* - *"I don't directly have my name on Twitter"*
Age	- *"Birthday on profile"* - *"I do not have my age on my twitter publicly in any way"*
Job	- *"Social media is used as an interface for my job..."* - *"Could easily work out that I am a [JOB TITLE REMOVED] [10]"*
Gender	- *"Haven't posted or made any direct inference towards it"* - *"My gender is in my public account info"*
Location	- "*I don't usually have location enabled posts but I'm not sure"* - *"Probably the city I'm in, but not more exact than that."*
People you know	- *"Talk to close friends publicly on twitter"* - *"I don't mention people very often but I follow people i know"*
Interests	- *"I often tweet about what i'm doing"* - *"I talk about my interests a lot"*
Opinions towards interests	- *"I post about thing i like and hate"* - *"I happily express my opinions"*

Participant reasons for their views (Before Demo)

10 Identifiable information redacted.

Test	Reasoning
Name	- *"This was as expected, so nothing has changed."* - *"Never gave it"*
Age	- *"Could work out a range easily"* - *"Didn't get my age but I think from what I've seen that it would be possible if I tweeted more extensively"*
Job	- *"Got stuff I mentioned a lot but I don't think it could give a confident statement"* - *"Can be inferred from tweets"*
Gender	- *"I think you could guess from my tweets"* - *"It got it right"*
Location	- *"Degree of location accuracy is too spooky."* - *"No location enabled tweets"*
People you know	- *"I only frequently talk to one or two people"* - *"People I tweet at often come up a lot"*
Interests	- *"Pretty clear what my interests are from my tweets"* - *"Keywords are consistent and apparent. My interests can easily be tracked"*
Opinions towards interests	- *"Positive and negative analysis is very interesting."* - *"keywords showing what i express make my opinion apparent"*

Participant reasons for their views (After Demo)

Appendix E - Test Structure

Step	Stage	Detail
1	Consent Form	- Participant would have the test explain to them - Any questions would be answered - Complete and sign the consent form
2	Conduct first survey	- Participant would conduct the first of the two surveys
3	Run Program	- The program would run on the participant's Twitter account

4	View/discuss results	- The results would be shown and discussed in the same order each test:
		  - Console output overview
		  - Pie chart stats
		  - Word clouds
		  - Location Log
		  - Other users
		  - Sentiment logs
		  - Total data log

5	Complete second survey	- The participant would complete the second survey