Final Report

40 Credits	CM3203 One Semester Project
Author	Spencer Thomas
Supervisor	Dr Alia Abdelmoty
Moderator	Professor Dave Marshall
Project Title	"To what extent is the user aware of the privacy Implications surrounding the Social Web and is there need for Personal Control of the issue?"

Abstract

The implications surrounding the provision of location information on social networking websites has become an increasing problem that many people are worried about. This is because there has been an increased presence of tools which allow for such information to be collected from a user and manipulated thereafter.

The project was tackled with the initial research into work which has already been undertaken regarding location information in social media platforms, in particular Facebook, Twitter and FourSquare. Nonetheless, given the general response from a questionnaire, of which had 42 responders; there was argument to be made for further work to be carried out. This is where the project aimed to develop a comprehensive definition of the problem surrounding these implications, and from there, a comprehensive set of activities could be outlined in order to successfully tackle the problem. These tasks were achieved through the application of Soft System Methodology tools; CATWOE analysis and a derivative root definition and a Conceptual Model, which was one of the key deliverables of the project and is contained in the appendices of the report.

A resulting set of initial requirements and prototype features were designed on the back of the analysis. These requirements and prototype designs are two of the key deliverable for the project and have been outlined within the report. The prototypes have also received user feedback through the medium of an interview, so that the prototypes could be initially tested in terms of the usefulness of it, the design of it and any changes that may be seen as necessary to the interviewee. These two steps are the first stages of tackling the problem, but the necessary future work of the project has been clearly defined in the "Future Work" section of the report.

Acknowledgements

This project was made possible because of Dr. Alia Abdelmoty input and knowledge and I would like to thank her for the initial input that she gave me at the beginning of the project, and for the continued support and feedback that she has provided throughout the duration of the year.

I would also like to give a thank you to the people that took the time out of their days to complete my questionnaire which gave me the material for which to base my user analysis. Also, I would like to give a special mention to those people who took the time to sit down with me and carry out either the pilot test of my questionnaire or the interviews for which I carried out to get user feedback on my prototype designs.

The final thank you is for my family and friends who continued to support and encourage me throughout my study at Cardiff University, which made the difficult times a lot easier.

Table of Contents

Table of Figures

Introduction

Geo-Social Networks have become the norm these days with the increasing popularity of websites such as Facebook and Twitter. These websites are defined as Geo-Social Networking application, since they used GPS-enabled technologies to be able to allow users to tag a post or a tweet with the location of where they are sending it from. The sites can also store the GPS information gathered from a sent tweet, which are produced through enabling things such as "Location Services" with an Apple Mobile device and enabling "Android Location Access" on Android mobile devices. This information may be stored even though it was not intentionally given away by the user as they may not be fully aware of the implication linked with enabling such services. Further categorisation of these websites show that Facebook and Twitter can be described as Location-Enabled Social Networks (LESN). The location information from these websites is available, but it is only a subset of the main information contained within a tweet or post. Whereas Location-Based Social Networks (LBSN) require location information for the primary source of their posts and this information is key for their services. A typical LBSN website is FourSquare, where a user of this site uses its "check-in" feature to state where they are checking-in so that they can share this with their friends. Therefore, FourSquare needs the location information of the user to be able to create the post for their friends to see. In comparison, on a website such a Twitter, the user can create a tweet with or without displaying the location of where it was created, meaning that the location information is available but is not necessary. However, the GPS information is created as a subset of sending a tweet even though it is not necessary.

This paper will focus mainly on the LESN websites, in particular, Twitter with the aim of determining what privacy implications surround providing location information to these websites. In recent years, there has been a huge increase in the amount of people using Social Network websites such as Twitter. In a study carried out by Statista (2014), the number of active users of Twitter per month has increased from 138 million users back in 2012 to over 240 million users as of the start of 2014. This growth has been a constant since its foundation back in March 2006 and seems set to be a constant upwards trend in the years to come with studies estimating, including one conducted by Fiegerman (2014) that the total number of users will likely reach the 400 million mark by 2018.

The paper will also be looking at the attitudes and concerns of Twitter users', with a clear plan of discovering what they think they know about their privacy and thereafter discovering what they think of their privacy using these sites after showing them the analysis which will been developed. Furthermore, the paper will highlight the need for control over user's geographic information, especially whether LESN websites are doing enough to give users control over this information in the first place, or if there needs to be features added to ensure that they have control over this sensitive information. Again, these features will be put to the opinions' of the users in form of an interview to see whether they feel more control over this information will make them feel more safe when using LESN.

The basic assumptions that the project is based on is that geo-tagging on GeoSN websites is not fully understood throughout general users of these sites, and there is a need for improved knowledge. Also, there can be a lot of improvement into what control is given to users with regards to what information they provide when using GeoSN websites.

The aim of the project is to take the necessary steps in order to raise user awareness of the implications relating to the provision of location information and also to increase the control users have over such information on social networking sites, such as Twitter. In order to achieve these aims, it will be a necessity to look into work already undertaken in trying to come up with solutions for these problems, from which a definitive problem can be defined with the use of soft system methodology. Supporting material which will help define the problem will be an analysis of various methods for which location information can be collected and manipulated. Also, there will be an analysis of user's underlying beliefs of the implications of providing location information through the use of a questionnaire. In order to take first steps to coming up with solutions for these problems, a set of requirements and a resulting set of prototype designs will be created and evaluated. From which further work can be carried forward to help achieve the overall aim of the project.

Related Work

This area of work is of increasing interest to those who study it, since there are a lot of unanswered questions that require answers. A lot of people are not fully aware of the privacy implications surrounding the provision of geotagging posts/tweets on their social media accounts, enabling technologies such as location services when using social media applications and linking third party applications to your social media account. This had led a lot research into the implications of privacy to users of Geo-Social Networking applications. Below, the report will be highlighting the findings from work carried out in the area of privacy implications on the social web, answering relevant questions which will be of importance to this project. Work including research into; users' attitudes and concerns of their privacy when using LESN websites, what is Twitter's stance on geographic information, how the collected geographic information is being used by others, in particular by Twitter, its users' and various third party entities and what sort of control can be feasibly added to these sites for users to gain control of the information they provide.

i.   Twitter Terms of Service and their Control of Geographic Information

A recent report from the Irish Examiner, conducted by O'Brien (2015), states that the European Data Protection Minister has gone on record of saying that there is a need to have control over their own personal data with the ever growing amount of a data people are willingly providing without knowing the full implication of their actions. A direct quote from that report, from Dara Murphy *"As individuals and users, we must educate ourselves on the implications of sharing more and more of our personal data online. We must look to increase individuals' control over their own personal data. "*

Twitter's main priority in terms of it users' are their privacy and control of their information when using "tweeting their location." To start, a user must opt-in to using this feature in the first place, so unless a user is aware that they have chosen to enable this feature, they should not worry. On top of this, there is an option available before every tweet is sent out that asks users whether or not they wish to enable the feature for that tweet. This means that even if the feature has been enabled by the user, it still needs another verification step so to say to enable the feature for a tweet. This means that users who may be unaware they have opted-in to use the feature accidentally can still be safe in the knowledge that this information is not being shared if they have not enabled it for a tweet. Also, Twitter's Help Center (2014) states that any location that has been attached to a tweet that was sent from an IPhone or Android device will include both the location label that they have put in themselves, and also the precise location from where they tweeted from. This information was found by looking into Twitter's terms of service solely for the purpose of this project. This

information would not be common among Twitter users, unless they took the time to look into the terms of service, which coupled with the fact that roughly 73% of people admit to not reading documents such as "terms of service" and "terms and conditions," (Glancy 2014) would imply that a lot of users are not really exposed to this type of information.

Twitter similarly states that it wants to be as transparent as possible to the user so that the user is aware of exactly what will be displayed in the tweet. It states in its "Geo Guidelines" that no explicit longitude or latitude information will be displayed. Instead, Twitter will show the user the names of the surrounding areas of which those co-ordinates are located within, and subsequently gives the user a choice of which one they selects. This means that the information provided is not entirely accurate, however clustering of this information will allow third parties to derive unwanted information from similar posts.

In terms of using location services with your Twitter account, it states that Twitter will use the information collected from enabling this feature when using Twitter's application to improve its services to its users. Twitter's Help Center (2014) states that these services are centred on the user geographic location at periodic time periods and building up this information to allow the users to tweet specifically where they are when sending a tweet or even tailoring the way they view trending tweets based off where they are. However, the user has full flexibility of their use of this feature, since it is prompted before using the application as to whether they wish to enable it or not. Moreover, they can then decide to disable this feature if they wish to, and Twitter has explained to its users exactly how to do this.

Storing the geographic data of its users' is another element which has been highlighted in these guidelines. Twitter states that it stores all geographic data collected from tweets where the feature was enabled, just like the rest of the data contained within the tweet. There is an option for users to clear their location history meaning Twitter will no longer have this data available to them. However the question raises again regarding how well known this tool is to its users without having to look into in more detail?

Nonetheless, there is an argument that could be made that the person whose information it is should be the person that decides what should be done with it, who is allowed access to it and who uses it. Snekkenes (2001) looked into this side of the argument, deciding that the user should have the final say on the policies affecting their own information, not Twitter or any third-parties. Snekkenes (2001) believes that these policies should be situational, depending on the circumstances, and that the users' themselves should be allowed to change the accuracy of the location information if the intended use of it is not what they were expecting or comfortable with. This point could be of relevance when deciding control aspects that could be implemented, since it puts the initiative in the user's hands.

ii.   How is the location information being collected and used by Twitter and others?

As highlighted previously, the main ways of Twitter collecting location information from its users are through GPS-enabling technologies used in mobile phone devices (such as "location services" for Apple IPhones), users actually geotagging their tweets by adding their location through the feature they provide its user, or through the user simply mentioning a location themselves in the text of the tweet. All three are viable pieces of information but each is more accurate than the one preceding it.

Global Positioning Systems (GPS) has become an emerging technology that is being implemented into many applications worldwide. The technology allows an application to get an approximation of the user's location. According to Wilson (2005) this approximation is based on the signals from at least three satellites above you, and your location in relation to these three satellites.

The primary reason of this project of looking into GPS is how it is used in mobile devices, in particular, what privacy implications are linked to the enabling of its services. Work carried out by Zickuhr (2013) found that almost three in four people enable location services on their mobile devices to allow them to obtain location information based on their current location.

Related work that has looked into how GPS and location information can cause privacy implications was carried out by Gruteser and Grunwald (2003). The main focus of their work was on how feasible it was to anonymously collect location information from location-based services. The findings of their work showed that the anonymous usage of location-based services bring about potentially "serious" risks. They state that due to the accuracy of GPS, in particular differential GPS which can provide accuracy of a location to better than 10 foot, that the amount of information that can be collected on someone, anonymously too, is serious.

Given that the user has GPS-enabling services enabled, then that data can be mined and used for various functions. Nikitha and Kumar (2014) used the GPS information contained within tweets, which is the co-ordinates from location of where the tweet was sent from, to create an event detection system. Similar to the situation that many find carrying out analysis in this field of work, Nikitha and Kumar (2014) found that only a small percentage of users actually geotag their tweets, therefore, most of their information came from the users who had "location services" enabled, which allowed for the collection of their location through GPS.

There are a number of third party entities that use these feature enabled by users for their own benefits. One of these entities is students such as Laurence Smith (2014) whose project was based around using Twitter's API to consume

tweets from users and analyse the resulting tweets in a variety of ways. A similar project was undertaken by "GNIP. " GNIP is a company which harnesses the power of social media platforms, such as Twitter, to allow its users to gain a competitive edge of its competition though the analysis of social data. Since Twitter has acquired GNIP, as reported on by Etherington (2014), it has become the standard application for many companies to use its facilities, including its analysis of geographic information, companies including Microsoft and IBM, which is shown GNIP's website. GNIP provides users with a similar application where it plots geotagged tweets onto a map to allow a visual reflection of where tweets are being published from in relation to keywords mentioned in the tweet. Moffitt (2014) reports that GNIP uses three types of metadata from a tweet to source the application; the exact location (longitude/latitude), a mentioned location from the tweet itself and any location specified within a user's profile. This metadata can then be filtered to whatever the user of the application is using it for by using GNIP operators.

Another application which is similar to GNIP is "Twilert". Twilert uses a similar concept to GNIP in terms of using the metadata of tweet to gather a number of tweets. This application differs as users of the application will first specify to particular location for where they wish to focus their collection. Then, any tweets that have been sent from that area will be added to the user's collection. It also has an alert function which alerts the user each time a new tweet mentioning or from that location has been sent by a user.

Alarmingly to users of such websites, the geographic information could be used to invade personal privacy by uncovering things such as their home locations or regular places they visit. Friedland and Sommer (No Date) studied what they described as "Cybercasing" by using a popular celebrity as an example of how it easy it is for people to access the information that it's available to gain perhaps unwanted personal privacy. Their work involved using this celebrity's various images which they had posted to Twitter, with the resulting images revealing where they had sent the tweet from, which incidentally included the celebrity's studio and their home.

Crandall et al. (2010) studied how geographic information can be used to infer social ties between two individuals. In this instance, Flickr was used as the social media platform, which is a LESN similar to Twitter, to collect both spatial and temporal information. The spatial information collected from a user can be used to infer the location of where the user had been while posting that information, while the temporal information relates to when the post was made. The results from the study highlighted that by using this information and deriving if two individuals had been in the same place at the same time, there was a high probability that the two people had a social connection. The report states that *"Rather than basing estimates on extensive high-resolution traces of individual behaviour, we ask what can be learned from an extremely small number of instances in which two people were proximate in time and space. This latter type*

of inference is arguably a greater privacy risk, because small quantities of such data are more easily exposed than detailed traces of physical comprehence." Being able to infer such events from such a small amount of data highlights the power of what can be achieved through mining LESN websites.

iii. What are peoples' current attitudes and general awareness of privacy implications using social media?

Raicu (2012) believes that the current attitudes of users of social media websites towards the privacy implications linked are deteriorating. His report includes an excerpt from research carried out by Madden (2012), from the Pew Research Center, which highlights that two in three online adults have made their social networking websites private so that only their friends have access to them. In comparison, only one in five adults admits that their profiles are completely public for others to see. Raicu (2012) suggests that users are becoming increasingly concerned about their privacy because of media coverage of the issue, not because of the actual dangers that are related to publicising your personal information.

Another survey was carried out in America by Software Advice's employee Humphries (2014), has its key findings surrounding the issues of privacy when using the social webs. One of the primary survey questions that were asked was whether or not users took the time to read the terms of services that needed to be agreed too. Of the survey, only 8% of users said that they always read the terms of services, meaning the majority of users either do not read them or sometimes read them. This is supported by the study carried out by Glancy (2014) of the Guardian which has already been previously highlighted within this report. Another question asked by Software Advice that is of relevance to this project is "Public Opinions on Data Collection." The results shows that 44% of respondents believe that this practise is an invasion of privacy, meaning that a lot of people are uneasy with idea of companies using personal data to use it for the company's own personal gain. The final question that was asked of Software Advice which relates to the collection of geographic data in a way is the "Top concerns with Data Collection Practises?" The results of this question show that 37% of respondents had no opinion of the matter, but alarmingly the rest had some type of issue with data collection, the main one being who exactly is accessing their personal data.

The trend of increasing concerns to personal privacy continues across other surveys carried out in recent times. A report conducted by Lai (2014) of Forrester's ConsumerVoices Market Research Online Community, carried out a survey on 600 adults again from America, to show that they have taken steps to protect their personal data. The results of that survey show that in June 2013 only 15 out of every 100 people surveyed took steps to protect their personal data, whereas a year later in June 2014 the number of people increased to 52 out

every 100 people took steps. One of these steps talked about in this survey was that of changing privacy settings their social networking accounts.

In terms of looking at users attitudes towards providing information such as their location information, the work done in this area in sparse. One study carried out in this area looked into the determinants of why a user would be reluctant to provide any information. Lederer et al. (2003) found that the main factor from people deterring to give away information would be because of the person inquiring for the information, rather than the situation. Another study carried out looked at the how comfortable people were about giving away their information. Ackerman et al (1999), cited in Barkuus and Dey (2003), found that people are more likely to be concerned with the type of information they are willing to give up. Along with that, Ackerman found that people would be more relaxed about providing information towards an application that was of use to them. These findings are of relevance to this project since users find that an application like Twitter can be deemed as trustworthy and useful to the general public, since it has such a large user-base.

With concerns to using Twitter or other social media platforms to access third party applications, a recent study carried out by Gigya (2012) found that privacy concerns with the feature were the underlying reasons why they opt out of using it. The privacy concerns which generated the greatest concern to users were of having to transfer personal information across different websites/applications and not knowing how their personal information will be used across the applications. This is of significant relevance to this report since users that have accessed third party applications through their Twitter accounts may then be vulnerable to the threat of their geographic information being used by this party. Because the information is now transferred between the two linked application, it gives different parties a user's geographic information, which will allow them use the information themselves.

Project Planning and Methodology

The plan behind the project was closely modelled off the waterfall software developmental process; however it could be seen as a waterfall-agile hybrid methodology since it differs from the traditional waterfall model. Following the waterfall process for this project would continue on after the specification of the requirements, development of the interface feature designs and so on, without looking into them further. Whereas the aim of this project is to be able to include the end-user in the designs of the end-product since they will be the ones either seeing the benefits or the drawbacks of the proposed solutions. Therefore, it will be of the best interest to include user feedback at each stage of the process so that the optimal solution can be delivered to the user. This can be best achieved through the development of a testable iteration at the end of each stage of the process which user can be involved in to voice their opinions.

Ideas behind choosing questionnaires

From looking into the best ways of planning a questionnaire, it was determined that the best practise for carrying out the process would be to fully understanding what was needed to achieved from the questionnaires first, before starting the design process. Therefore, the author came up with the main objectives for which he hoped that the survey would achieve with its results. The main objectives were:

a. *To obtain a feedback on how people use social media websites, what steps they take to ensure their safety and what are their current attitudes of the privacy implications surrounding the provision of geographic information*

   The thinking behind the first objective was to attempt to gain a basic understanding of what people understood about the privacy implications involved with using GeoSN websites. Through research and general discussion with users of Twitter, it was identified that there was not a lot of knowledge relating to the implications to their security when providing geographic information. Nevertheless, a lot of people had warmed to the idea that their information was being used more now than they had done previously. This was due to the new features on websites such as personalised advertisements popping up on users' Amazon accounts and features on YouTube such as "What to Watch" which suggests what videos to watch based off what you have been watching previously.

b. *To understand the users' actions while using the sites, to gain an understanding of what actions they are taking that makes them vulnerable to privacy violation from the collection of geographic information*

   As well as attempting to understand what the users believes they understand about GeoSN websites, the questionnaire aims to understand the actions taken by users which may be putting them at risk without them being aware of it.

Therefore, its purpose was to gauge an understanding of whether people actually use Twitter, how often do they use it, through what medium do they access the application and what actions do they take to ensure their own safety on the sites.

c. *To gain a basic understanding of what people think about the background research and analysis carried out into this topic and see if this changes their opinion on their vulnerability*

Having gauged a basic understanding about the user, the questionnaire would then present the recipients with the findings form the research and analysis to see if their opinions alter given the new information. By setting up the questionnaire in this way, it will allow to see whether the increased awareness of the potential privacy implications surrounding the provision of location information will make the users feel that they need more control over such information.

Design of Questionnaire

In an attempt to develop the correct questionnaire which would allow for the best retrieval of results from users, there was research done into the possible design choices of the questionnaire. There were a number of factors that needed consideration when designing the questionnaire.

a. *How the questions should be asked in terms of the language used?*

This project's topic can be seen as a technical topic; therefore, it was important to consider the basic understanding of the general twitter user's knowledge of the technical aspects, to ensure that the recipient of the questionnaire fully understands the questions being asked.

b. *What will the layout of the questionnaire be?*

There were a number of things to think about with regards to the layout of the questionnaires. Firstly, through looking into layouts closer, Pew Research Center (No Date) suggests considering the placement of open questions next to closed questions. Their reason being is that if the users is presented with an open-ended question after answering a closed-ended question, then their answers may not reflect their true opinion, but it could be influenced by the answers presented to them previously. An example of this instance occurring within this questionnaire would be asking the question of "What is your opinion of the privacy concerns of providing geographic information on your social network accounts?" next to the question "Which of the following applications of geographic information concerns you as a user the most?" If the user was previously unaware of the applications highlighted by the questionnaire, then they may not have been concerned about the privacy when providing such information. However, because of the questionnaire's design, the user's opinion

may then be altered because of the question asked. The other consideration was to include instructions for potentially difficult questions to understand, or for questions that required some work from the user. As mentioned previously, this area of work could prove to be too difficult to understand for the some users; therefore perhaps an explanation could be included beforehand to make sure that the question could be fully understood by most, if not all.

c. How should the questions be answered by the users?

The project required the results from the questionnaires to be presentable and measured against other data sets. Therefore, for questions which require an "either or" answer, the choices would be provided for the user to select from.

d. How long should the questionnaire last?

The idea of the questionnaire is to attempt to retrieve as much data as possible without losing the recipients interest with a drawn out questionnaire. Consequently, the length of questions and the length of the questionnaire must be considered when designing because anything too long may defer the recipient away from completing it.

e. Which application is the best to use in order to create the questionnaire?

Having spoken to people and through looking at user feedback, Google Forms seems to be the choice that many choose as their application for creating the questionnaire. The main reason behind the choice for this project was because of the fact that a Google Forms questionnaire can be distributed easily through social media platforms, like Twitter, which then obviously aims it at the target audience, and the answers will give a true reflection of different users and their opinions on the matter. Similarly, it can be distributed to other social media platforms like Facebook, which will be easier to distribute to a certain audience through uploading it to different group. Furthermore, Google Forms provides the user with a very user-friendly interface to use to create the questionnaires and for recipients to fill it in. One of its usability features is that it provides a variety of potential questionnaire style questions to be implemented, so it gives the creator a wide choice to choose from. Also, it allows the creator to generate a professionally looking finish to their questionnaires through the feature of being able to change the themes, which includes the heading of the document, the font style and colour and the background colour of the questionnaire itself. Finally, the results retrieved from completed questionnaires can then be used in a spreadsheet to generate easily readable results in the form of charts and graphs, which would be a useful tool for this project.

Other considerations that were taken were websites such as SurveyMonkey, Survey Planet and PollDaddy. However, the choice was made to use Google Form as per the reasons stated above and also with the familiarity of using

Google products and knowing the overall high quality that Google maintains with all its products, it seemed like the most logical choice for the author.

Reasons behind the Questions

Page One

- *Age and Gender, Access Twitter and Twitter Use*

The first 5 questions asked in the questionnaire are all included to gauge an idea of the type of user that is filling out the form. By attempting to find out things such as age or the gender of the recipient, analysis can be undertaken to see if the beliefs of people differ with age or depending on if they are a male or a female. M Madden et al. (2013) looked into the beliefs of teens that use social media and found that only 16% of teens set the accounts to automatically include their location information in their post. Also, when they carried out an online focus group with teens about the location privacy issues, many were cited as saying that they were against including their location information because of things such as strange people finding their home and also the annoyance of the feature itself. By understanding the ages of the recipients, hopefully the analysis can find a potential trend that can support this idea and also see what older people of social media websites think of the issue.

In terms of understanding how the recipients access the accounts, this will be of relevance to the questionnaire because there will need to be a differentiation between users who use their mobile phones and those who use other platforms to see which of the mobile phone users use "location services" also. Furthermore, by including the options of laptop, tablets and computers analysis should be able to show whether either of these methods of accessing Twitter will increase the likelihood of a person adding their location to a tweet. For example, adding location information to a tweet maybe more visible through an internet browser, rather than on Twitter's application.

Understanding a person's time spent using Twitter and sending tweets is also relevant because those who spend more time tweet could be of higher risk of increasing the visibility of their location information. As shown from Laurence Smith's analysis tool, a user's tweets can be collected and the location information can be plotted to a map. By having a high tweet count, there will obviously be more plots on that graph which gives a better picture of where the target is located. By understanding the usage information, there can be a better understanding of which user is more vulnerable to having their location information mined from their Twitter accounts.

- *Twitter Terms of Service and Tweets*

The next three questions are mostly related to Twitter's terms of service and trying to gauge an understanding of what the recipients know about them. This

will be an important finding because if the people have an understanding of twitter's terms of service and the content that is included in a basic tweet, then they would be aware of the potential privacy implications involved with sending a tweet with geographic locations. Hopefully, the analysis will be able to spot a trend between the people who have this knowledge and how safe they feel on Twitter with the knowledge they have or if their actions when using the site are different.

There is also a question asking whether people whether they feel the content of a tweet is explained to them well enough. By content, the question is implying that the user understands all the data that can be mined from them from sending out a single tweet. This question is hinting at the idea of whether the recipients feel that they need more awareness on their profiles and the results from this type of question will warrant whether or not recipients feel that more needs to be done about this.

- *User Safety*

  The following two questions are asked to gauge an idea of how safe the user feels when using Twitter and what are their main concerns, if any, that they have when using the site. This question was purposely asked at the beginning of the questionnaire because of the need to get the recipients' raw opinions without any information being provided to them. This way, there can then be a similar question included towards the end of the questionnaire, and see if any of their opinions alter about their safety while they have undertaken the questionnaire. Hopefully, given what is shown to them in this questionnaire relating to the privacy implications surrounding GeoSN websites like Twitter, their minds may broaden to thinking that they are not as safe as they first thought. Again, this is important because if users' opinions change then the idea of introducing control features or introducing feature which will increase the awareness of the information that they are given away in a tweet can be looked into, and it will have a substantial backing behind it.

- *Location and Third Party Understanding*

  The final four questions are specifically related to geographic location and also there is a question related to third party applications. The questions are put in to understand how people use their Twitter accounts in relation to giving away geographic location information, and to see just how many people do give it away. These questions are needed because they will be trying to gauge how many people use geographic enabling features such as location services or add location information to tweets so that it can be seen exactly how many people are vulnerable to the potential threats linked with the provision of this information. Also, the question regarding third party applications was included to determine exactly how many of the recipients used their Twitter accounts to access accounts on other applications, without the knowledge of what doing that means to the information they are providing when using their Twitter accounts.

Page Two

- *Privacy Implications*

The user will be presented with 12 statements and a choice of opinions about how they feel about each statement. Each statement is included in the questionnaire after it was determined, through research and analysis, that there is a possibility of each happening if someone has access to your location information from accessing your tweets. It was determined that these are possible and they have been included in order to gauge the recipient's opinion on each. The statements used are possibly alarming, but also realistic statements so they show the recipients just how vulnerable they are from giving away their location information. Through analysis of the results it should be interesting to see if some users may not be aware that such fine levels of detail are possible from the tweets they send out and that because of this question, and that their opinions on the matter change as a result.

- *Control Aspects*

The user will be presented with 8 possible control features or awareness raising features and their choice of whether they feel each is necessary or not. The features are basic add-on ideas that could be implemented onto Twitter's home page that will either give the user more control over their personal information, such as location information, or will help raise awareness of the issue. These features, although pretty basic as ideas, can be elaborated on in future work, given that the recipients of the questionnaire think that they are a necessary feature. From the results of this question, it will be possible to draw up a set of requirements, and subsequently, a prototype design for the ideas that the recipients feel are of most necessity to implement. By gauging the user's opinion on what control features are necessary, it will allow the author to have a better understanding of how to model a possible feature that could be implemented onto Twitter.

- *Repeat Questions*

There has been an inclusion of two questions on the second page, which have been repeated from the first page of the questionnaire. The first question is "Do you currently feel safe using Twitter?" and the second is "Would you like to have more control of the information that you are giving out in your tweets?" It was decided to ask both these questions again **after** asking the questions on the second page to see if the recipient's basic understanding that they outlined from the first page has altered having answered the questions from the second page. The predicting outcomes is that the research and analysis carried out previously in this study, to find the information to include in the second page of questionnaires, will open the recipients eyes to what is possible from collecting their location information and that their opinions will change after being presented with this information.

Analysis

This part of the analysis was then carried out on datasets produced by previous work carried out by Cardiff University students from previous years. The first work was carried out by Laurence Smith, who generated a dataset and the resulting analysis as an example to display the functionality of his Twitter analysis tool. The second work was carried out by Fatma S. Alrayes and this was based on the social media platform FourSquare, rather than Twitter. Even though this project has focused on Twitter and the privacy implications related to geographic location provision on Twitter, Fatama's analysis is of relevance because it shows the depth of analysis that can be carried out based off the derivation of geographic location information.
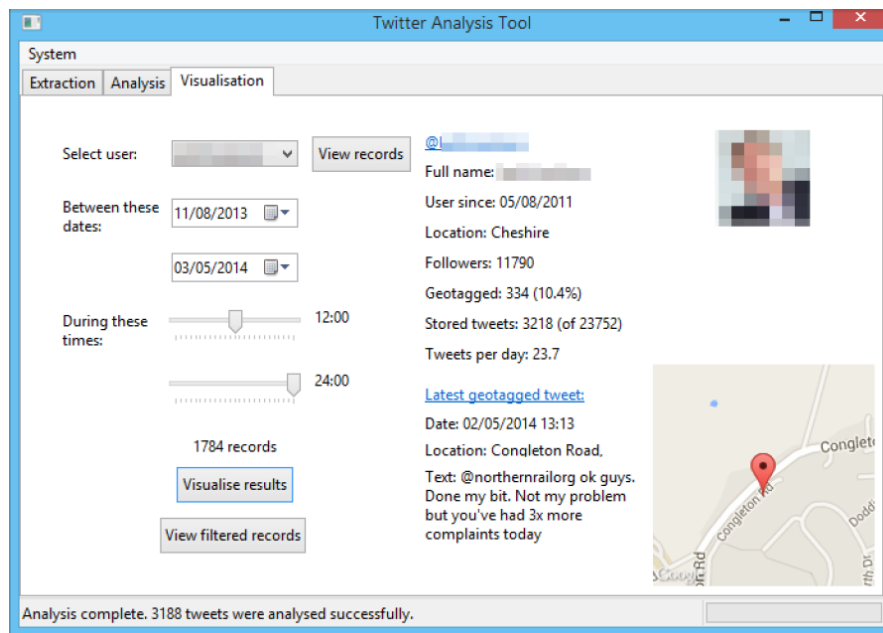
Twitter Application Analysis

Laurence carried out a use case on his Twitter analysis tool. The tool allowed him to lookup a specific username and subsequently, the tool retrieves up to 3,200 of that user's last tweets. These tweets were then stored in a database, and from which he could then apply various analysis tools to this set of tweets to derive the information he desired from them. The information of interest to this project was related to what geographic information could be collected.

Figure 1: Tool allowing a look-up of users through an association with a place name
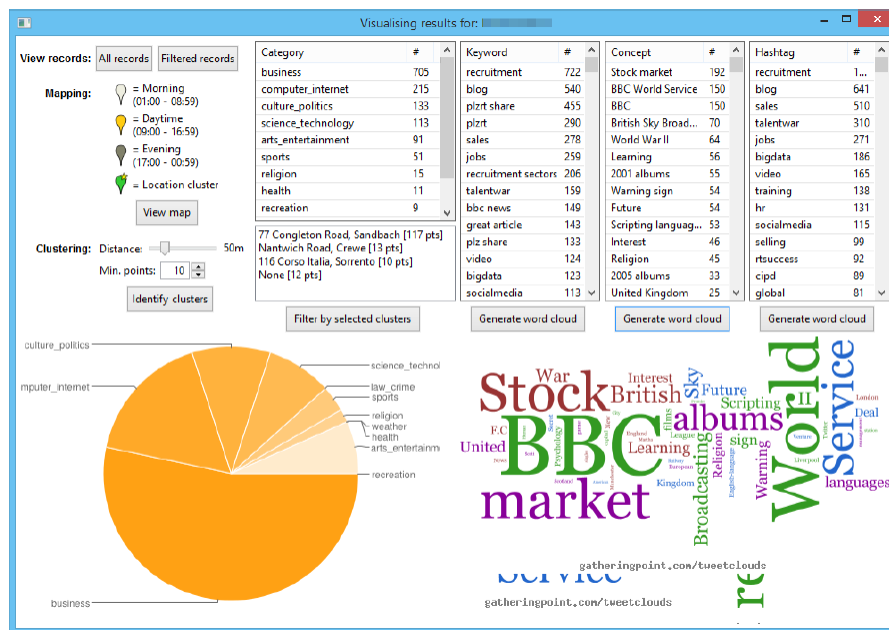

The first significant tool that Laurence used was that he could find out a username of any user that had any relation to a specified location that he had manually inputted into his form. In his example, he specified that he wished to find users from Cardiff and any area within 1km radius of Cardiff, and the results found that 57 usernames had some association with this location. Also as shown in the image, the tool could also use more specific geographic information, a user's co-ordinates, to find them. This is of relevance to the user's that use location services when accessing their Twitter accounts, since their co-ordinates of their tweets will be available for collection once they have posted a tweet from a mobile application.

Figure 2: Tool shows a break-down of the user's location related activities. Tool also showing the possibility of spatiotemporal analysis with use of timestamps

After analysis was carried out on a user's tweets, the results were then published in the form of the image above. The right side of the page is information related to the user that has been generated through analysis of their tweets. As shown, information such as the number of geo-tagged tweets and when and where the last time the user geo-tagged a tweet can be derived from the collection of tweets. This image also shows that the location information can be mapped to a Google Maps interface. This has relevance to the analyser of the tweets as it shows where the tweet was posted from in relation to their specified location. Another feature which has been included for the user to use is the ability to change the times and dates for which the tweets can be determined by. This could be of particular use for a companies who wish to target their advertisement more specifically to the time of day in which their product or service is being used, in relation to their location. For example, a chain of pubs may want to use the information from this application to derive when their pubs are being mentioned more than usual throughout the course of a week. They could then see this evidence from the users' location information and the timestamp on the tweet. This time could then be used for promotional purposes such as reduce prices on drinks or food being served with the idea that more people seeing this deal will want to tell others about it, thus bringing about more custom to their pubs. Nonetheless, this type of manipulation of information could be of cause for concern to users who may only see the negative to people having such freedom of access to this level of information. They could think that this information could be used to track them overtime to spot trends in the patterns of their lives to determine the times of day in which they are most vulnerable.

Figure 3: Level of analysis possible from the tweets collected by the tool
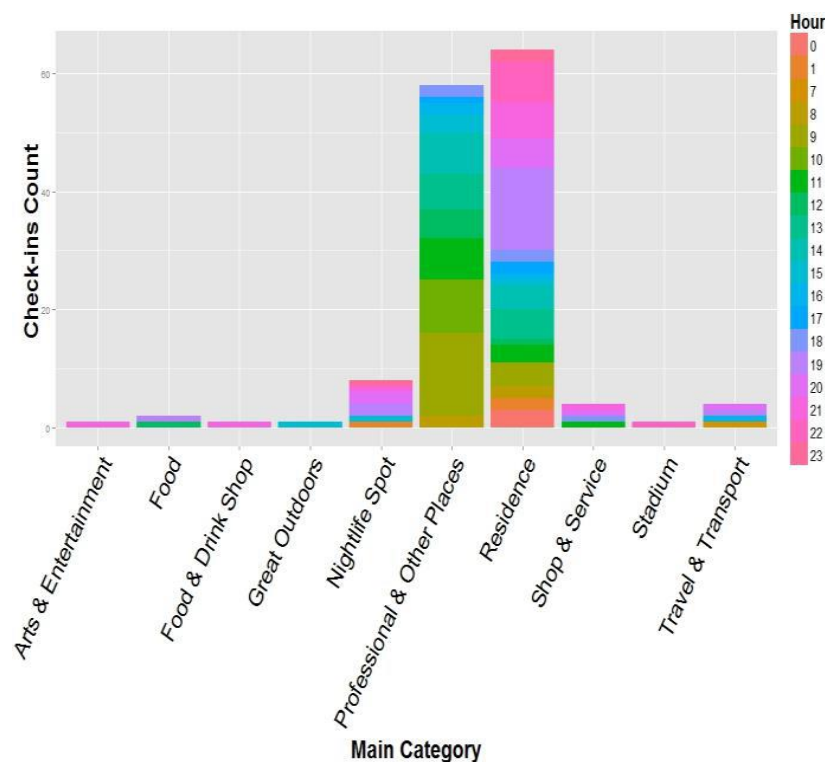

The user of the applications then has the ability to visualise the analysis in different forms, as shown in *figure 3*. The visualisation tool of most interest to this project is the mapping tool which is located in the top left of the figure. The tool allows the user to map the tweets onto a Google map image, and see where each tweet was tweeted from relating to where they were sent from. This can allow the user to map out frequent locations that a user may visit to gauge an understanding of what places (e.g shops, restaurants) are popular with different people. The user analysing this data can then use this information for various reasons, one being advertisement for an example, as they can identify the trends within users' daily activities. This is of interest to the project as a tool like this is potentially seen as being dangerous if put in the hands of the wrong people. Users may feel that they could be seen as vulnerable with such a fine level of detail about themselves readily available to whoever can access it. On top of the mapping, Laurence has also set-up the tools so that the person analysing the data can also roughly see when a user has visited these places. This could potentially only add to the concerns of some people as it may put the idea in their minds that people could be potentially stalking them having all this information available to them, and consequently could be at unease with giving location information in future.
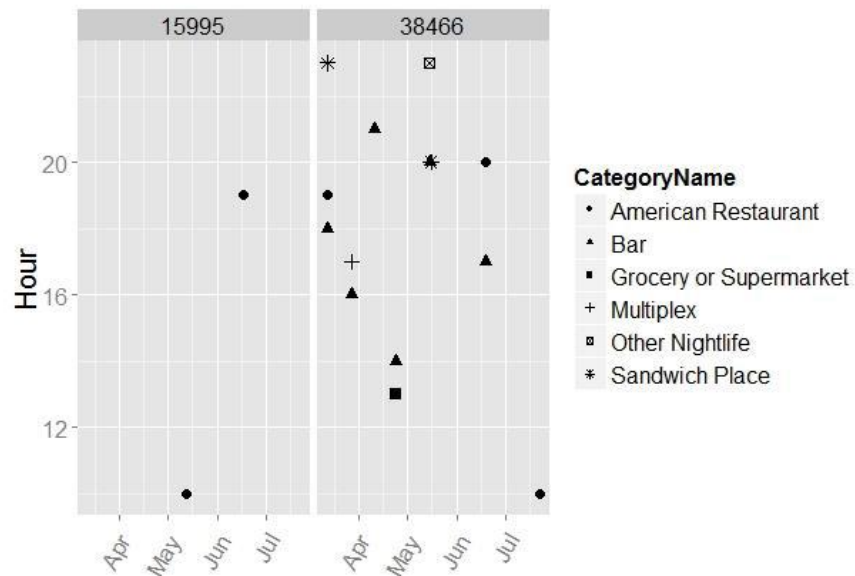
Foursquare Analysis

Fatama carried out a study on three separate users of the FourSquare application. From these three users, the intention was to identify three unique patterns of the users' activity in relation to their location, and by doing this with the categorisation of users into the frequency of their posts. The three categories were a moderate user, a frequent user and a hyper-active user.

Figure 4: Graph showing the number of check-ins to a location, against the time the check-in took place. Graph shows a high check-in count for home and work
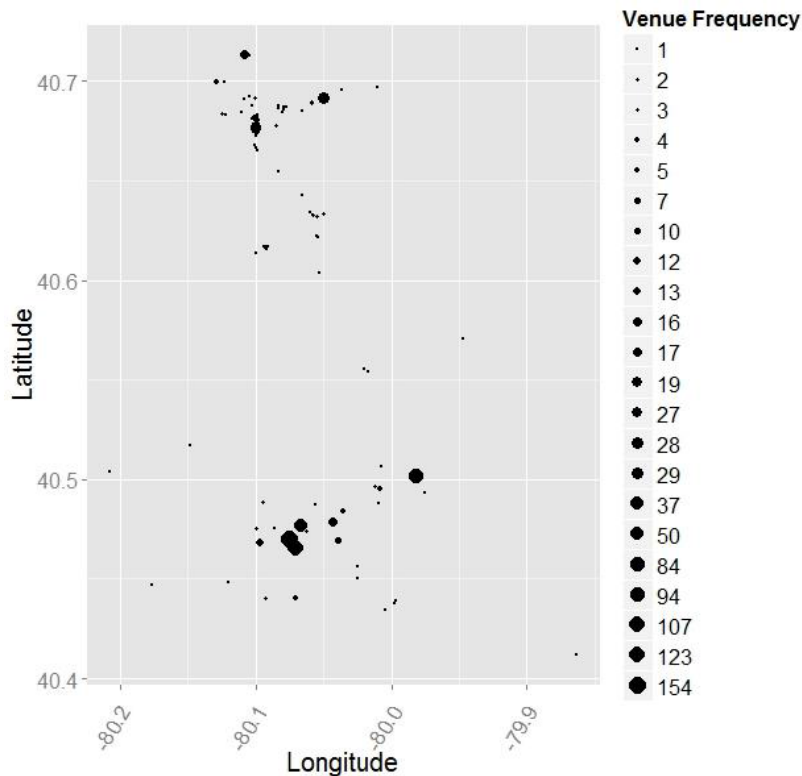

The first task was to be able to identify the relationship between a user and their association to a particular place, which is what is being shown *figure 4*. This was particular study was carried out on the moderate user, and the biggest take away from these findings is that Fatama was able to derive some highly sensitive information from the user. Such information that was derived is that the user had checked into their home and office at 44% and 36% respectively. This means that from analysing the location information that was accessible from FourSquare, it was possible to identify places such as a user's home and work location, just from manipulating their data over a period of time.

Figure 5: Graph showing the co-occurrence of a check-in from the user at the same time as another user, in the same place

The next focus was on the degree of association with other users, which is what is being highlighted in the graph above. Through analysing the information from the check-ins, it was identified that the frequent user had checked into a location at the same time as another user on 36 occasions, with further analysis showing that of those 36, of which 26 co-location were shared with one user in particular. On top of this, it was also identified that the user had 16 spatiotemporal co-occurrences with another user, 14 of which were with the same user which had been identified previously. It was denoted that from these results that both user must have shared a relationship with one another, having so many check-ins together in the same places. This just proves the fact that this level of manipulation could be achieved on any two people, and that level could be quite alarming if the common perception of users changed to consider these possibilities, or even see them.

Figure 6: Graph showing the co-ordinates of places within Pittsburgh, USA against the frequency a user checked into those co-ordinates. Graph shows more defined spots for frequently visited locations.

The final focus was on the relationship between the frequencies of a check-in over a period of time, against the co-ordinates of that location. As you can see from *figure 6*, the more defined speckles on the graph represents a higher frequency of check-ins to that place by a user. Therefore, by analysing these two components of a FourSquare check-in, Fatama was able to derive exactly where that user had visited the most during a select time period and was able to see, to the co-ordinate, where this place was within the given area from which the users were being studied from, which incidentally was Pittsburgh, USA.

Questionnaire Analysis

Analysis of Respondents Background Knowledge

The questionnaire has been targeted at Twitter users that the author was able to reach out to through different mediums such as social networking sites, work emails and through passing and sharing the questionnaire into other social circles. The response of only 42 respondents completing the questionnaire was not what was expected having hoped for more response to be able to identify clearer trends in the responses that may not have shown with this level of response. As stated previously, the questionnaire was designed using Google Forms while the resulting data retrieved was analysed using Microsoft Excel. Of the 42 participants, the gender was evenly split with 57% of the recipients being male and 43% being female. The age was more or less evenly spread across ages ranging from 19 and up, with the majority of the respondents falling within the age range of 19 and 39 as 55% of respondents fell within this range. Since the majority of the people targeted were of similar age to the author, coupled with the fact that the average age of Twitter user's has been identified as 26.58 years, as reported on by Statista (2013), this is why the majority of respondents fell within this range. Having initially expected the older respondents to be more cautious when it comes to their safety when using Twitter, the results show that safety was not dependent on age since there were respondents from all age ranges that stated they had insecurities about using Twitter. Similarly, there was no difference in opinions of safety with respect to gender, as again there was no pattern to spot that suggested a certain gender had more insecurities than the other.

Of the 42 respondents, all 42 said that they use their mobile phone to access Twitter, or a combination of that with another device. However, less than half of those people (47%) enable the location services feature that can be enabled when using Twitter's app, so they are less likely to be at risk to the implications linked with the provision of their location information. Nonetheless, of the 47% of the respondents who enable location services, their knowledge of what the implications surrounding this are seem to be sparse, since their responses suggest they are quite concerned with what is possible. Of those respondents who said they enable location services when using Twitter, 77% were either worried or uncomfortable with the scenarios which are possible from deriving and manipulating location information, which will be fully explained later in the analysis. This suggests that even though they use this feature, their awareness of the implications is still low considering they do not really feel at ease with some of the possibilities of the implications. This coupled with the fact that of the respondents, the majority (64%) stated that they never read up on things such as terms of services (which Twitter uses) or terms and conditions. This would mean that unless this information was passed on to them from their friends or family, they would be completely unaware of the implications surrounding the provision of their location information. Although it could be have been brought up within the questionnaire, the reasons behind why these people tend not to read such documents were never discovered through this questionnaire or the

background research; but the obvious reason behind this would be the length of such documents and the total amount of irrelevant information contained within those documents that users do not need to waste their time on ciphering through it.

Another two issues were raised within the questionnaire to gauge the recipients' knowledge. Firstly, the questionnaire was trying to gauge an understanding of the user's underlying understanding of a tweet, and how much they understood about what is contained in a tweet and if they felt that this knowledge was portrayed well enough to them by Twitter. The results from the questionnaire were somewhat contradictory since of the 64% of respondents that felt they were safe enough about using Twitter, just under half (48%) of those respondents felt that Twitter does not do a good enough job of explaining what is included within a tweet. In general, a majority (57%) felt the same way, so this is obviously part of the problem in the eyes of the user. This needs to be addressed by Twitter since it's clear that stating these things only in long documents such as a "terms of service" is not sufficient for a lot of users.
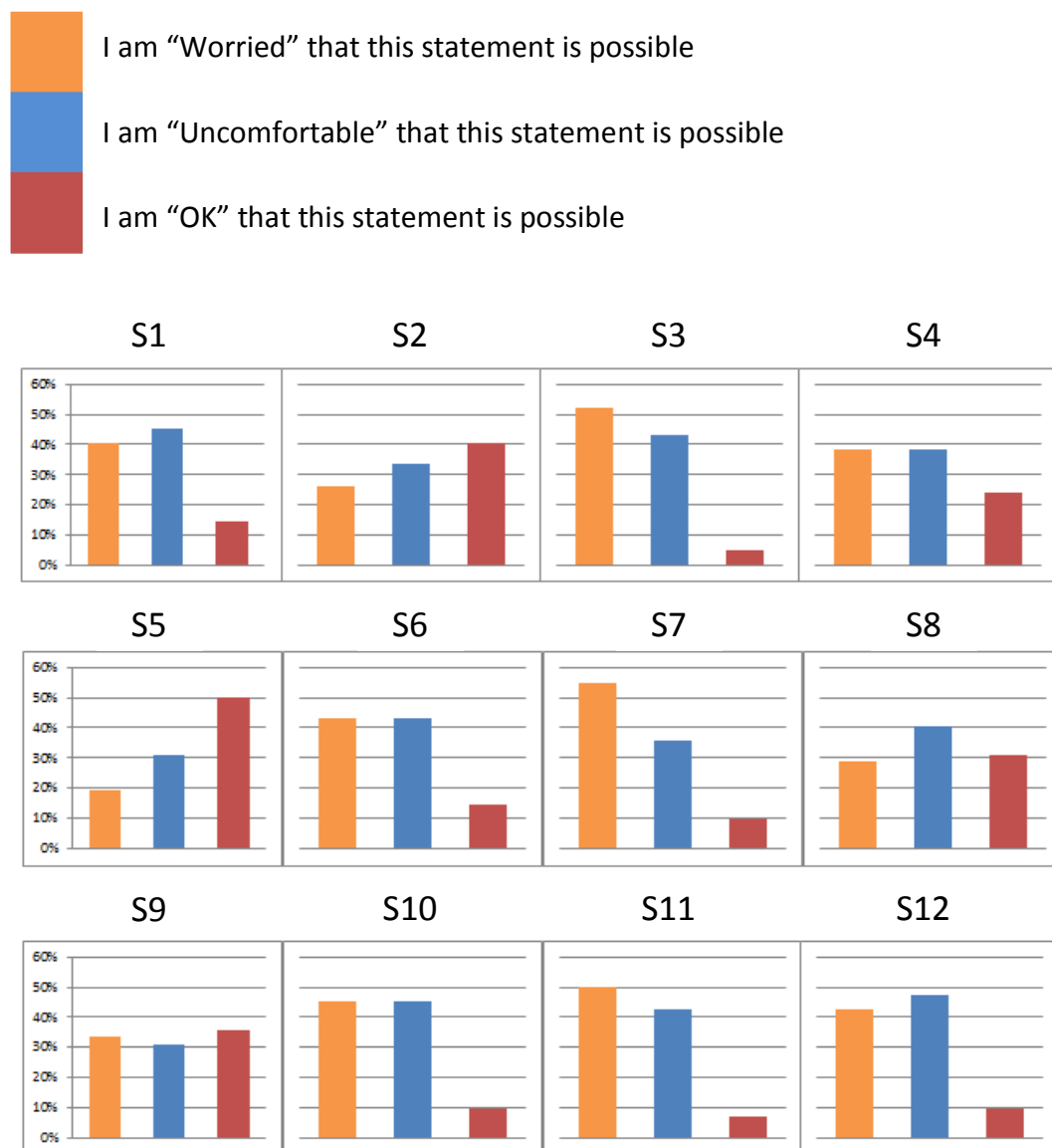
Analysis of Respondent's Awareness

As highlighted previously, there had been a set of 12 scenarios which had been created, through prior research and analysis, which are possible from the derivation and manipulation of user's location information. These scenarios were then presented to the recipients of the questionnaire in an attempt to gauge their opinion on each. The predicted outcome of the respondents was for those users' awareness to be low and that they would be concerned with the scenarios that were presented.

The nine initial statements were in relation to Twitter itself and what is possible from deriving information from the users. The final three statements were related to information being used from Twitter by third party applications, which is possible when users agree to link their Twitter accounts to other applications. The statements were flanked by three possible responses of either "Worried", "Uncomfortable" or "I'm OK with it." Although this does not exactly warrant showing the user's awareness of the statements, it may indicate that they had not been aware of these things if they did in fact state that they were either worried or uncomfortable with them being possible. Here are the scenarios that were presented to the questionnaire recipients:

    S1. I can guess where you live

    S2. I can guess where you work

    S3. I can guess when you are away from home

    S4. I can guess when you are off work

    S5. I can guess when you have been with your friends

    S6. I can guess where you have been and at what time of day

S7. I can map out where you have been

S8. I can guess where you are after you send out a tweet

S9. By entering a place name or a set of co-ordinates, I can potentially find your tweets

S10. Third-party applications have access to this information also

S11. Third-party applications can guess where you are at any given time

S12. Third-party applications can guess when you have been somewhere

Figure 7: Results from the questionnaire regarding the 12 statements relating to the users' concerns. Key complimenting the set of results

I am "Worried" that this statement is possible

I am "Uncomfortable" that this statement is possible

I am "OK" that this statement is possible

The results in from *figure 7* show the respondents' opinions to the scenarios that were provided to them.

Firstly, the results regarding the first nine scenarios show that respondents were either worried or uncomfortable with the majority of scenarios, with scenarios one, three and seven gathering a total negative response of 95%, 95% and 91% respectively (negative response being a combination of "worried" and "uncomfortable" responses.) These seem like the most personal scenarios that would affect a user, which would suggest why these scenarios would gain the most amounts of negative responses. On the contrary, the scenarios which respondents would find that would have less impact on their privacy were scenarios two and five, collecting a majority number of responses favouring the opinion of being "OK" with the scenario occurring. Overall, the results show that most of the scenarios, except for scenarios two and five, gathered a majority negative response to the scenarios presented to them which would suggest that they may have been unaware of these being possible.

Secondly, the results regarding the final three scenarios, which are concerning third party applications, show that the respondents generally had a negative response towards the idea of other applications accessing information from their Twitter accounts and using it potentially to affect them. For each of the scenarios relating to third party applications (scenario ten, eleven and twelve,) the responses towards being OK with each scenario playing out gathered either a 10% or less (10%, 7% and 10% respectively) percentage of the response, meaning that very few are OK with idea of other applications having access to this information.

Analysis of Respondent's attitude towards Control

Carrying on from understanding respondent's attitudes to Twitter's current situation, the questionnaire next attempted to understand the respondent's attitudes towards the possibility of increasing the awareness of location information on Twitter and the possibility of gaining more control over such information. This would then allow for a better understanding as to what users would like to see added to Twitter in order for them to become more at ease with providing their location information, so that they do not feel as vulnerable to the scenarios which they had previously been either worried or uncomfortable with.

The recipients of the questionnaire were presented with a set of eight features, each either being related to increase of user's awareness or giving them more control over their information. They were then asked whether they felt that each feature was necessary to be included or not. Here are the features that were presented to the questionnaire recipients:

F1. A new feature on Twitter to raise a user's awareness of what information they are giving away with each tweet

F2. A new feature on Twitter to hide your personal location information when tweeting from places like home or work

F3. A new feature allowing you to control what information you give away when you tweet

F4. A new feature allowing you to control who can access your location information

F5. A video within Twitter's "Help" menus explaining what is included in a tweet

F6. A video within Twitter's "Help" menus explaining how others can use the data from tweets

F7. A video within Twitter's "Help" menus why people use the data from tweets for their own work and projects etc.

F8. Warning messages before agreeing to use a third-party application with your Twitter account that will potentially give away personal information

From analysing the feedback for each of these features, each seemed to have an overwhelming positive response from the respondents. Five of the eight features gathered an 85% or higher response for being a necessary feature to be included to allow Twitter users to have more control and an increased awareness over their location information. On top of that, four of those fives features even gathered a 90% response for being a "necessary" feature. Therefore, it would seem a necessity to think about including these types of features when drawing up a set of requirements for a change to Twitter's interface. Nevertheless, three of the features that were options in the questionnaire did not gain as much of a positive response from respondents as the rest, with the three gaining necessary votes of around 70% each. These three features were ideas which included tutorial-type videos that would provide users with an in-sight into their tweets and how they can be manipulated by others for their own use, and to give them an idea of how this could be potentially used in an adverse way against them. These three features gained a notably less positive response meaning that the respondent possibly did not feel the need for these to be included in the possible changes to Twitter's interface.

Analysis of Respondents General Opinion

The final question within the questionnaire was aimed at getting the respondents' open opinions on the topic of controlling their own information, without the restrictions of having to select pre-determined answers. The responses were all along a similar pattern of being concerned at the level of depth that can be achieved through manipulating information that they are making available by sending out tweets. However, there were a variety of answer which was of interest to the project and which were helpful towards deriving a set of requirements for Twitter's interface.

One respondent was quoted on saying "*I think that users should be made aware of possible implications (such as burglaries through guessing when away from home) of using Twitter without fully understanding what can happen.*" This respondent felt that the best approach to achieving an increase of awareness and control to the user would not be through "scare tactics", but to imply that potentially dangerous things are possible without actually mentioning them. Nevertheless, when answering the questions regarding the features they felt would be necessary to include, they stated that the idea of having videos to help users raise their awareness would be unnecessary. Although he did not mention an alternative, by him quoting this, he must have felt that there would be an alternative method of achieving an increase in the public awareness of privacy concerning location information. Other alternatives that were conceivably possible could be that the user may be alluding to the idea of education users to the dangers of providing location information on social networking websites early on in their lives. Although there is an age restrictions set to social media websites, which is 13 on a lot of them including Twitter which has been reported on by Bennett (2014) of the Social Times. This is still a very immature age to be allowed to openly provide such sensitive information to practically the whole world. Therefore, this respondent could be implying to include these issues into children's education since nowadays social networking sites are becoming so prominent throughout the general public.

Another respondent was quoted on saying "*After going through the questionnaire I have been made more aware of what information is being given out without me knowing. I do not feel you could fully eradicate the problem due to the world we live in, but it could be nullified. In addition, although I have said that it is worrying that information is being released about Twitter users in general, I am personally not that bothered about my information being available to others.*" This respondent makes two points regarding the issue. The first one point is that they feel due to the way the world is today there is no way to fully eradicate the issues surrounding the provision of location information, assuming that they mean that people always manage to find other ways of getting this information from people once the last way has been quelled. This is a good point that is made, and is part of the problem that needs to be addressed since there is not really one universal way to prevent the privacy issues arising without taking everybody off social networking sites altogether. Therefore, the process needs to be monitored, by users and by Twitter themselves so that they can see that the issue is being tackled effectively and that users are becoming more aware, and in the process more secure with providing such information. If the problem is not being addressed then there can be control action taken so that the necessary steps are being taken to, and to quote the respondent, "nullify", the problem. The second point he raises is about not being concerned about providing this information to others. This is a point that could be made for a lot of users, and even those who carried out this questionnaire but did not state this opinion as the question was not a required part of the questionnaire. But for every person who has this opinion, there is another person with the opposite opinion which is concerned with their information being given away. Therefore, this could be outlined within the requirements that features should be made optional so that

those who feel that the need them included within their interface can have them, but those with less concerns can carry on using Twitter as it was.

Another respondent from the questionnaire was quoted on saying "*I do not post on twitter just use it to follow people mainly celebrities and never turn on location services.*" This person would fall under the same category as the person who was quoted from the previous paragraph, as these features would not be necessary to them since "I never tweet" or "I never use the location services feature". Nonetheless, this respondent does not specify their actions on Twitter with respects to their actions when interacting with these celebrities. Many users may be unaware that when they are tweeting at a celebrity in response to one of their post, or even re-tweeting a tweet they have sent out, then these actions are also making the user vulnerable to having their sensitive information being available to others. Users may think that they are safe because their level of knowledge about a tweet, or which has been highlighted in this study a number of times, their awareness about what is included in a tweet is low. As shown from the analysis of Laurence Smith's application earlier in the study, it is possible to modify your search terms to look for re-tweeted tweets, meaning the information from those who have re-tweeted that tweet is now available. Therefore, these people are still as vulnerable as those who are constantly sending out tweets.

There were a few more general comments which expressed the opinions which were initially predicted from the outset, having presented the respondents with all the possibly new information to them. Respondents were quoted as saying things such as "*very concerned that so much information about me is readily available to possibly the wrong person,*" "*I am probably more worried for the kids as I am already wary enough to watch what I am sending*" and "*Having used Twitter since I was younger I've always felt safe because nothing bad has happened to me from using twitter. But knowing all these things are possible has made me more cautious about things I do when I'm on twitter.*" All these opinions were expressed from a wide age range of users, ranging from 18 and under demographic up to the over 55's. As expressed earlier in the analysis, these quotes reinforce the idea that privacy concerns seem to of some concern to all age groups, be it that they are just starting to understand the implications having grown up a bit to realise what is possible, or be it from a parental stand-point where they are concerned for the well-being of their children.

Soft System Methodology

Reasons behind CATWOE Analysis and Root Definition

A CATWOE analysis was outlined as a necessary tool to use for this project to be able to identify the problem faced with being able to increase control and awareness of location information on Twitter. An online article developed by Pandey (2011) quotes that the term "CATWOE" was defined by Peter Checkland as an acronym for the process taken to identify each of the following key components of the tool itself; clients, actors, transformation, Weltanschauung or worldview, owners and environmental constraints. This tool is commonly used by business analysts to identify exactly "how" the system will work in order to combat the problem identified.

Based off of the research, analysis and questionnaire results that will be produced, it will be possible to identify the two main components required to carry out the CATWOE analysis, the transformation and the worldview. Having found some of the possibilities from derivation of information from location information, and highlighted this in the questionnaire, it has been found that the majority of the replies found that the facts are worrying. From this analysis, the author has been able to derive the transformation component of the CATWOE analysis, as the transformation component is included to identify what changes the system will bring about in relation to what is required from the system. The recipients of the questionnaire have also identified the need for control measures and an increased awareness of information on Twitter's user interface. This has helped with the identification of the worldview, since this is included in the CATWOE analysis to show exactly what people believe should be put in place in order to achieve the changes of needed from the transformation component. The clients must be identified as the beneficiary of the system, which in this case will obviously be the Twitter users as they will be the ones hopefully benefiting from the increased presence of their location information on the interface and the inclusion of more control features which will allow them to have more control over their location information. Additionally, the Twitter users have been identified as the actors behind the system as they will have to be the ones that push the issue to the owners of the system, which is obviously Twitter, because as things stands Twitter does not seem to have an issue with its users have problems with its current location privacy settings. Another actor which has been identified in this system is the Twitter development team, as once the issue has been pressed to Twitter by its user, it must then be designed and implemented by the relevant developers employed by Twitter. The final component that was identified for the CATWOE analysis was the possible environmental constraints of the system. Since user's privacy is of paramount importance to Twitter, there should be no financial constraints to the system because if the problem of privacy issues highlighted has substantial backing from its users then it should be dealt with whatever the cost. Therefore, the environmental constraints refer more towards Twitter and its user's opinions of the system. The first constraint highlighted was Twitter's possible un-willingness or reluctance to

change their location privacy settings. JW Owens [1] has identified that the satisfaction rate of Twitter's user towards its interface is currently pretty average, with a mere 38.46% score out of 100% in usability satisfaction. Therefore, by having to include more control features or cluttering the interface with more location information will possibly make the interface more complex or confusing to use, therefore, it may detract or turn user away from using Twitter. With this in mind, the system must be able to deliver the relevant control features and increased information presence without affecting user's usability of the interface. The second environmental constraint that has been identified is the possibility of user being unwilling or reluctant to learn to understand location privacy concerns even with the inclusion of the new features. Even though the system may have all these new features included in on Twitter to improve the awareness of the concerns, users may still be unwilling to look at this information or use the features that are in place because they do not feel that the concerns are of much relevance to take the time to use the new features included.

CATWOE Analysis

C – Twitter Users

A – Twitter's Development Team, Twitter Users

T – Increase and continue developing Twitter user's awareness of the issues surrounding location privacy when providing geographic information on Twitter

W – Providing users with the necessary means which will help increase their awareness and increase their control of location privacy on Twitter and continue to keep them in-the-know with future privacy concerns

O – Twitter

E – User's reluctance or un-willingness to learn to understand location privacy concerns, Twitter's reluctance or un-willingness to change user interface because of fear of negatively affecting the user's experience while using Twitter's interface

Root Definition

A system owned by Twitter, where their development team can increase and maintain the awareness of Twitter users in relation to the issues surrounding location privacy by means of providing users with the necessary means to which they can improve their own awareness and control over their personal geographic information and the provision of such information, while giving consideration to the fact that users may have a reluctance or un-willingness to learn to understand issues surrounding location privacy and Twitter itself may have a reluctance and un-willingness to change user interface because of fear of negatively affecting the user's experience while using Twitter's interface.

The Design behind the Conceptual Model

Although the root definition explains what the system will do in order to solve the problem, it does not outline how the system will do it. This is the reason for the inclusion of the conceptual model, as it explains what parts of the system need to be put in place in order for the system to be able address the problem. The model is designed with the root definition in mind, to ensure that the problem identified within the root definition is addressed. The conceptual model includes activities from which have been outlined from the root definition, and it explains the relationships and interactions between the activities and how they work to address the problem. In this conceptual model, the activities which have been outlined were derived from looking into background research and analysis of the problem and a set of questionnaire results which help back-up the case for both the research and the analysis.

The conceptual model has been designed with the following notations in mind:

- A square box represents an activity that needs to be carried out by an actor of the system, which has been identified from the CATWOE analysis as a member of the Twitter development team and Twitter users. An example of an activity from the conceptual model is "Decide how to give users more control over their location information." This is activity which needs to be carried out in order for the system to function. Each activity has its own purpose within the system to take steps to addressing the problem identified.
- An arrow represents the flow or path required to take, in sequence, to eventually arrive at point within the system which can be seen as an activity which, when carried out by an actor, is seen as addressing the problem. An arrow is directional in the sense that the arrow follows the flow of activities from start to end.
- The notation "Performance Info" identifies points within the system where the performance of that activity and all of the prerequisite activities that came before it are evaluated or monitored to see exactly how the activities are being carried out in order to address the problems. Once the activities have been monitored, the actors of the system can the decide whether to leave the activities carry on as normal, if they are achieving their intended objectives, or to take control actions, which will be explained below.
- The notation "C.A" represents a "Control Action." Control actions are put in place within the system at points where the system is seen as addressing the problem. These actions are located directly after the "Performance Information" activities, which as stated, are used to identify whether the activities are addressing the problem identified. The purpose for the control actions are so that if the problem is not being correctly addressed by that activity, then the actors can then step in to ensure that the system changes in order to address the problem as it was intended to. For example, within the conceptual model it has identified that one of the activities for the development team to carry out is to make sure that the Twitter users are satisfied with the control over the location information. At this stage, a control action can be made by the team to ensure that if the satisfaction levels are to up to what were expected, then something needs to be changed.

Mapping the Conceptual Model

This section is set to outline why each set of activities have been included within the conceptual model and why they have been identified as necessary activities in order to address the problem or the "Transformation" identified in the CATWOE analysis. The "Transformation" being able to increase and continue developing Twitter user's awareness of the issues surrounding location privacy when providing geographic information on Twitter.

The first two key chains of activities revolve around the provision of control of location information and the increase of presence of location information on Twitter's interface. Having researched current attitudes of users (Raicu 2012),

location privacy is of some concern to users since they feel that the collection of their location information is seen as an invasion of privacy. Coupled with the fact that the results from my questionnaire suggest that users would be extremely open to these ideas being implemented; then these chains of activities were of necessity to the system. The sequence of activities will need to begin with the identification of how to implement these ideas onto the interface, followed by the actual implementation, followed by a monitoring process which needs to take place before any actions are taken to alter the effect of the activities. Further on in the project, there will steps taken to address the first step in the sequence of activities, by identifying ways in which the ideas can be implemented onto Twitter's interface.
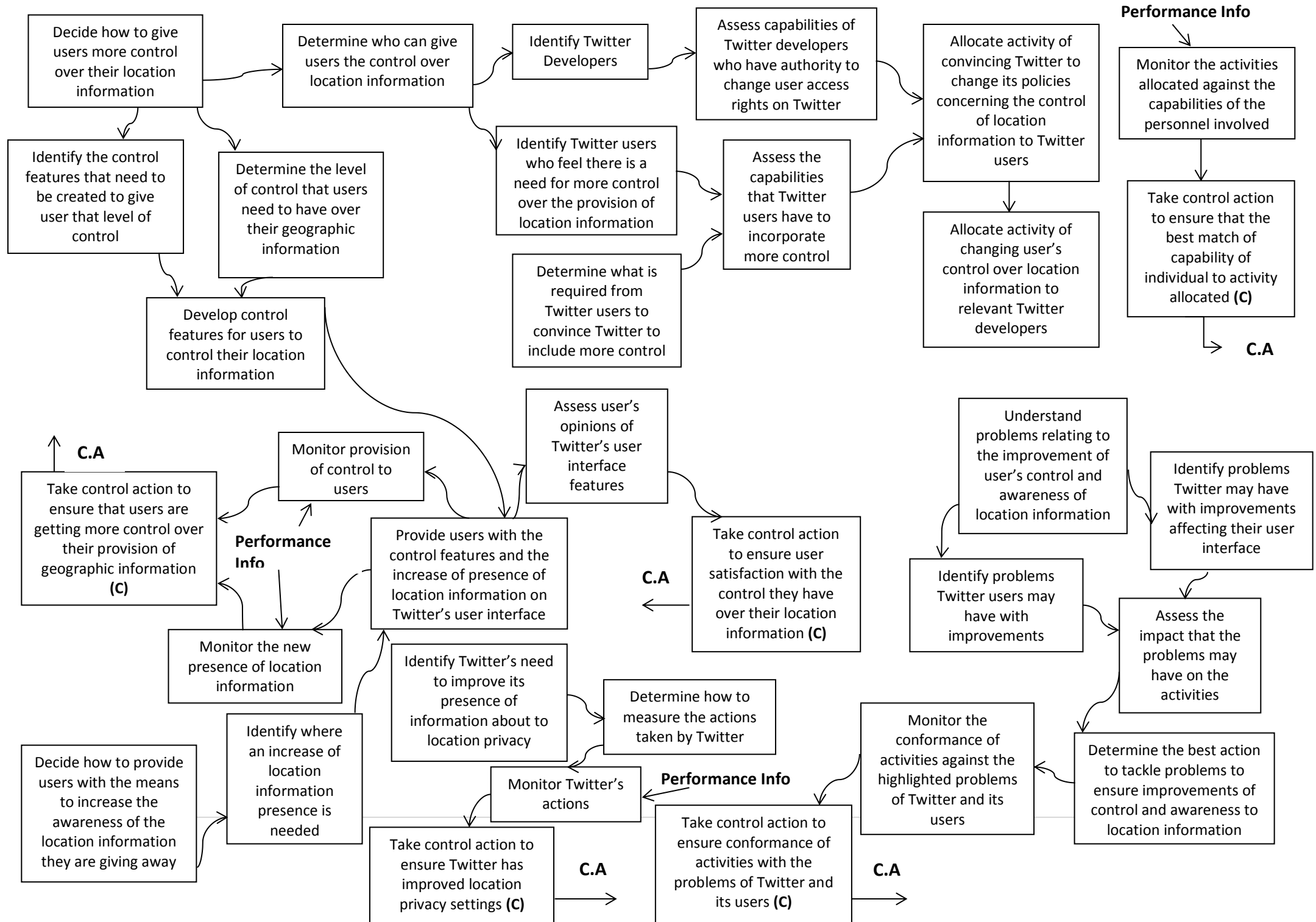
The next chain of activities is related to the actors of the system. In order for the system to be able to function, it has been identified that a set of actors are required. As stated previously, these have been seen to be the Twitter development team and also by Twitter users themselves. Firstly, the capabilities of both sets of actors need to be identified to see if they are capable of carrying out some of the activities needed for the system to function as it was meant too. From the discovery of the actor's capabilities, the activities highlighted in the previous paragraph can then be allocated accordingly so that those who have the abilities and capabilities to do so can carry them out. For example, it will be of importance to the system that Twitter developers that are capable of changing the interface, making decisions on changes and those who are in charge of privacy issues are identified. Once these people have been identified, they can then be allocated activities from the previous two chains of activities. The twitter developer's tasks are more obvious to highlight from the system as they are the people carrying out the changes. However, Twitter users are also actors, and they have activities of their own that need to be completed, in relation to the conceptual model and the system's overall functionality. The activities that are mostly related to the user are the activities concerning the actual determination of the problems. Since they are the ones using Twitter, it is them who are possibly being affected by the privacy concerns. Therefore, it will be up to the users to declare that they think that the current privacy policies are a concern and they have the activities of proving this through different mediums. There has been work carried out which is necessary for the task of disseminating a questionnaire to Twitter users to gauge the opinions' of users so that they can see if these concerns are universal. This would be one of the activities that could be carried out in order to determine the concerns of Twitter users. Unless these concerns are raised from within, then Twitter would more than likely continue on with its business as usual.

Another chain of activity that needs to be carried out by the Twitter development team is the activities relating to the possibly disruptions the changes of their interface and policies may have on things like its image, its interface's usability and other possible issues that may arise from the changes. By making changes, Twitter faces the possibility of making changes to things which may affect the way it is seen by the public and its users. For example, making a simple change such as the way the way a user sends out a tweet, to accommodate for the changes made to improve user's control of their location information may make the process longer or more

complex than needed. This may turn users away from using the site since they may not like the changes and feel it disrupts their experience on the site. Therefore, the development team need to first identify the possible problems that come with the changes and subsequently see how these changes may affect the public's opinion of them. Once they have been identified, there needs to be a "risk assessment" type of process carried out by the team to assess the impact that these changes may have to Twitter. After summing all of these issues up and carrying out all the task before-hand, the team can then determine the best possible way of addressing the problem of user's awareness of the privacy concerns when providing location information while having the least amount of impact on the interface, the image of the brand and on its users.

The final chain of activities is again relevant to the Twitter development team and also relevant to the Twitter users too. The activities relate to the actual functionality of the system and how Twitter goes about making the necessary changes for the users. The development team are obviously required to make the changes, whereas the users and possibly the team themselves, are required to monitor whether Twitter is actually taking the necessary steps to making things better. If users cannot see that work is not being carried out to make the necessary changes, then they can take the necessary actions to ensure that Twitter keeps on track of all the activities within the system to ensure that users are continually satisfied with the policies concerning the location information.

The conceptual model is drawn out in full within the appendices, giving the previous paragraphs in a presentable, easy-to-read form.

Requirements Section

In order to be able to develop a user interface to incorporate the ideas which have been backed from the questionnaire, there will need to be a set of functional and non-functional requirements developed. To allow for a clear developmental process of the features, it was decided to split the requirements into sections, each section specifying requirements for a particular feature. Along with each requirement, there is also a set of acceptance criteria which links to the requirements. These are included so that once the features have been implemented within the interface that they can prove that all of the features meet the requirements that had been outlined. The requirements specification has also adapted the MoSCoW notation to all of the requirements, which has been outlined by DSDM Atern (No Date). This was included to prioritise the importance of some requirements to the feature, over others. Nonetheless, it should be targeted that most, if not all requirements are met so that the needs of the users are satisfied with the inclusion of the features into the interface. In terms of the MoSCoW notation, some of the requirements may seem to be higher up on the priority list for that feature. However, it was highlighted within the conceptual model that Twitter must try to remain as closely associated to itself as possible without deviating too far since it does not want to turn any of its users away from using the site because of changes which may affect the usability or the presentation of the site, for example. Therefore, highest priority is put on the maintenance of Twitters image within the eye of the public.

Requirements and the Reasons

A feature to raise user's awareness of what information they are giving away with each tweet

1.1.   The interface must have an increased presence of location information.

1.2.   The system should include an increased presence of support for users regarding the provision of location information (linking to requirements 4).

1.3.   The interface could clearly display all relevant location information to be given away in the pending tweet, within the tweet submission box.

Acceptance Criteria: 1.   There is an increased presence on the new interface as opposed to what it was on the existing interface.

2.   All relevant location information that is to be provided is shown to a user before sending a tweet.

3.   The location information is displayed in the format from which it is collected.   (Example, if the user location is collected in co-ordinate form then it should be show to the user in this form, not any other that maybe misleading.)

The first feature that was to be included was to raise user's awareness of what information they are giving away with each tweet. The requirements behind this feature were fundamentally to increase the amount of information that would be helpful to users as opposed to what Twitter had already in place on their interface. From the questionnaire results it was determined that users felt that Twitter did not do a good enough job of explaining things which related to location information with a tweet. Therefore, the idea was an increased presence throughout the site. In relation, one of the requirements highlighted the need to include location information within the tweet submission box, which could be included within the interface since it would clearly outline to the user what was being sent out before they actually sent it out. Nevertheless, if the logistics of incorporating this type of information into such a small box was not conceivable without manipulating Twitter's interface in a way in which it could adversely affect it, then it would not be a feasible requirement.

On top of the increased presence, it was also a requirement to feature an increased presence of support that users can use in order to help raise their awareness of the

implications. The increased presence of support could take the form of many different outputs, such as video, text and by other means deemed necessary. Although this does not increase the presence of location information directly, the inclusion of such help features should help the users to become more aware if they do use the additional features open to them. Even though the analysis from the questionnaire showed that many users do not find a document such as a "terms of service" useful when trying to understand things, the different outputs must be made available to cater to as many user as possible so that the message and the information can be shared as wide as possible.

A feature to hide personal location information from your tweet

2.1.   The user must be allowed to change their privacy settings so that no location information is included within a tweet which is to be sent out.

2.2.   The user should be allowed to select a button within the tweet submission box, which prevents location information being included within a tweet.

2.3.   The system would not store any information about the user that was sent with this feature.

2.4.   The user must have access to change their privacy settings from all devices and platforms that are used by Twitter users.

Acceptance Criteria: 1. All of the existing privacy settings Twitter has in place in relation to location remains unchanged.

2. The different levels of granularity of which the location information can be collected in are modifiable to the user within their privacy settings. (Example, specific place names, specific place co-ordinates, general place names etc.)

3. The "safe tweet" concept must send out a tweet containing no sensitive information relating to the user.

4. The "safe tweet" must only contain the user twitter handle and the tweet content.

This idea behind this feature was for users to be able to send out a tweet without any sensitive information, such as their location, being present within the tweet. In terms of a feature that could be included that would incorporate this idea, it would be necessary that no information from the user is collected once they use this feature to send out a tweet as opposed to sending out a regular tweet. For example, sending a regular tweet from a WIFI spot will contain a number of location

information units including the place name and sporadically the information may be even more in detail. However, by selecting to send out a tweet with the new feature, this will hide all of this information from everyone else since this information will not be collected and stored by Twitter. The requirements for this feature leave the development open to a number of different ideas. Some of the ideas which have been identified as a possible implementation feature could be a clickable option that a user selects before sending out a tweet that hides all information. Another possible solution could be a separate tweet submission box that once used, hides all location information but the existing tweet submission box remains present for users whom wish to use it.

A feature allowing you to control what information you give away and who can view this information when you tweet.

3.1.   The user must be allowed to control their default settings for what information and to whom it is sent to from within their privacy settings.

3.2.   The user could be allowed to control the degree to what people see their location information and what type of location information they see.

3.3.   The user should be allowed to view and change their settings on any given tweet from an option within the tweet submission box.

3.3.1.   The user should be allowed to change their default settings from a list of checkboxes for that specified tweet only.

3.3.2.   The system must change the user's privacy settings back to default after a tweet has been sent with any modifications.

3.3.3.   The system must not display or contain within a tweet any location information regarding a user who has specified that that information should not be sent out with a tweet, be it a set of co-ordinates or a place name.

3.4.   The system must not allow any location information to be viewed from Twitter's interface, from collecting tweets from Twitter's API or through any other method if a user has specified this in their privacy settings.

Acceptance Criteria: 1.   The different levels of granularity of which the location information can be collected in are modifiable to the user within their privacy settings and the one off tweet in the form of checkboxes. (Example, specific place names, specific place co-ordinates, general place names etc.)

2.   Only relevant information that has been selected to be collected from the user is collected and stored within a tweet.

3. Any information that has not been selected to be collected from the user is not stored by Twitter.

The thinking behind this feature was to offer users more control over their own location information. This is reflected within the requirements that state that the new feature must offer the user the chance to allow whatever preference they want, in terms of both who sees this information and what type of information they see. Twitter currently does offer users this type of control already. However, with requirement 3.2 stating that users could control the degree of their location information, meaning that they can control to what level this information is presented and to whom it is presented to, then it could offer users more control than is already been offered to them. This requirement is coined with the "could" phrase as it would mean having to change location information from one form into another for the purpose of the functionality of the feature, example changing a set of co-ordinates to a place name to cover or hide the co-ordinates. This may not be what Twitter wants to offer its users, since this degree of control over their information could be perceived as unnecessary.

These requirements also highlight the possibility of having one-off location settings where once the tweet has been sent, the settings change back to the settings that are pre-determined within the user's settings. This gives users the control of changing up their settings on-the-fly which may be what they want to do to show off a place they may have been that they want to show with their friends. However, their main location settings state that all location information is removed from a tweet. With a new feature designed with these requirements, it offers the user this control to change their preference for that one tweet without having to change their settings and forgetting to change back, leaving the vulnerable to having their location information manipulated in future.

Videos within Twitter's "Help" menus explaining

4.1. The system must have help videos which are appropriate to the task of providing user with an opportunity to improve their awareness of the implications of providing location information and how they can improve their control over such information.
4.2. The videos must be readily available for users to access whenever they deem necessary.
4.3. The videos should be suited to all Twitter users.
4.4. The videos must be compatible for viewing on all media platforms available to Twitter users such as smart phone and tablet apps, internet browsers etc.

Acceptance Criteria: 1. All videos are always available to users.

2. All videos are kept up-to-date with providing user the knowledge that is relevant to the current policies behind location information.

3. All videos are accessible in an acceptably agreed number of languages.

The videos were meant to be included within the interface as an extra way of raising awareness of user's knowledge to the implications of providing location information. The requirements for the videos are simply to make sure that all the videos that are shown to users are appropriate to them. This would mean that the topics are all related by a common purpose, the content is diverse so that it can be viewed and understood by all users and so that the videos can be viewed from all compatible Twitter applications and variations of Twitter's interface.

A feature to warn users before agreeing to use third-party application with your Twitter account

5.1. The system must have a page, a pop-up message or other form of providing information before allowing the user to agree to use their Twitter accounts to login into third party applications.

5.1.1. The system must explain what agreeing to this will mean in terms of the information that will then become available to the third party application.

5.1.2. The system must explain what the third party application will use their information for.

5.1.3. The user must be able to confirm to the information presented to them before agreeing to link their account to third party applications.

5.2. The system should make the user aware once a third party application makes any changes to their policies or terms of services.

5.2.1. The system should make the user re-agree to link their Twitter account to third party applications after reading the changes that have been made.

5.2.2. The system must make sure to remove the user's information from the third party application if they decide to not agree with the new changes that have been made.

5.2.3. The system must only allow the third party application to a user's information if they agree to the new changes.

Acceptance Criteria: 1. Once the user selects to link Twitter account with third party application, presentable information is shown to user at the appropriate time.

2. The information presented to the user must be relevant to third party applications using information from user's Twitter accounts.

3. The user must select to acknowledge the information before agreeing.

The idea behind the final set of requirements were focusing on the third party aspect of the research and the idea that users were concerned by the fact that third party applications have access to their location information from Twitter if they agree to link the accounts. The thinking was to make sure that the user has to go through an extra page of information before agreeing to link the accounts. This is different from the existing set-up of third party applications that use this method because once the user has logged into their Twitter accounts, the accounts are linked. Therefore, the requirements outline that the feature must have an additional "terms of service" like page after agreeing to link the accounts. This would ensure that all the implications that are linked to connecting to a third party application are presented to the user in a presentable and timely manner. Although the user research carried out found that users are highly unlikely to read through such documents, it seems to be the best way of getting this information across to the user, alongside the use of videos within the menus. With these two methods of raising awareness in place, there must be an effort made on the part of the user to take the time to understand what is going on if they do wish to take part.

Design of Prototype Features

From the set of requirements that were specified, there was a need to develop a set of prototypes for some of the features that had been highlighted in the questionnaire. This need was highlighted from the fact that users of Twitter themselves felt that these features were necessary. Therefore, to develop the best features for the interface, there needs to be some level of user involvement with the design process since they will be the ones using them eventually. Subsequently, it was decided to develop a set of prototype designs for which a select number of Twitter users could see and visualise within their intended environment (in the places within the Twitter interface) and gauge their opinions of those features.

The designs of the prototypes were very basic with no functionality involved in the process. The prototypes are simply a design that shows users the concept of the feature and how they would fit within the interface. They will then be presented with a set of questions which will attempt to gauge an understanding of their feelings of each feature and from this; the features can be taken through for further development, given they have a good backing from the user testers.

Prototype One – Increased User Awareness

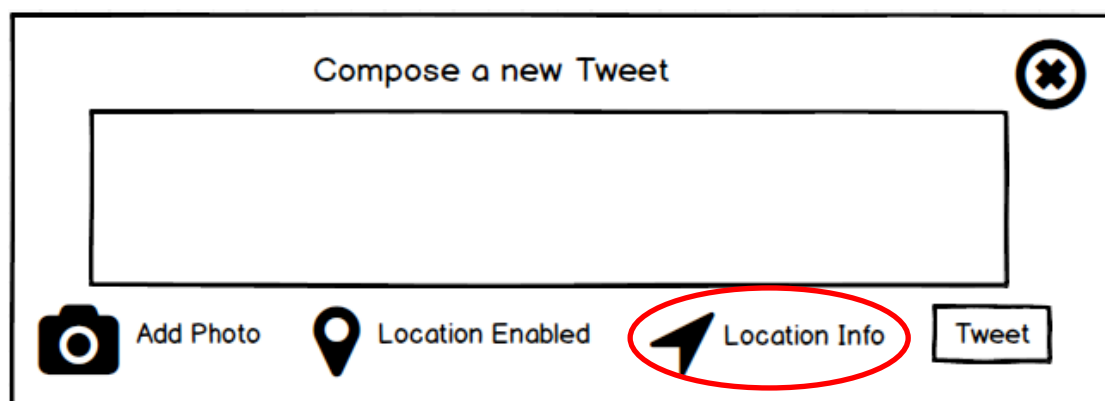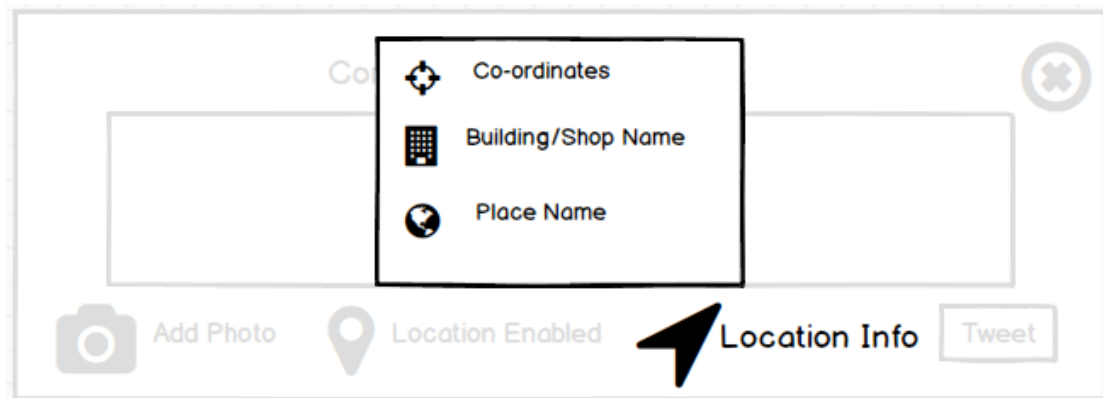**Figure 8: New feature "Location Info" added to Tweet Submission box**

Figure 9: Hover over with the cursor shows user location information for Twitter's website, clickable icon for Twitter's application for mobile and tablet devices

The first prototype was designed with the idea of raising user awareness of their own location information they will be giving away with each tweet. The design concept is that once the user hovers over the newly created icon, as shown in *figures 8 and 9,* and the location information will pop up on the screen, fading out all other features on the screen to emphasis the presence of the location information box. For mobile and tablet applications, the design is similar except there will need to be a touch on the icon or lettering from the user to bring about the pop up box. The design, although simple, is an effective way of communicating such information to the user.

Reasons why the hover method is effective for users who use computers is that there are not any other unnecessary page changes that need to take place, it's a commonly used concept throughout other websites and it's quick. Also, the feature was designed so that the icon was positioned alongside the "tweet" button, so that in a majority of cases users will hover over this icon anyway so their location information will be presented to them. Nevertheless, the hover feature may seem like an irritating feature that continuously keeps popping up when a user may not want it too. The reason why the hover method was selected over the clicking method was because of the positioning of the icon within its environment. Having the icon very close to the "tweet" icon could mean that some users accidentally click one button rather than the other so by having the hover method eliminates the possibility of this happening.

Prototype Two – Improving User Control of Settings

Figure 10: User has the "Security and privacy" section of the settings menu open, where they select to "Remove all location information from a tweet" in the process, removing all proceeding options from being selectable

Figure 11: Location information options available to the user, where they can alter them to their own preference

The second prototype was designed with the idea of improving user's control over their own location information, in relation to what control is already provided to them. The design concept is one that has been adopted by Twitter within its settings menus, after researching into it. The user is faced with a set of setting for which they select with their own preference. The order of settings is as such, so that if the user does select the first option to remove all location information from a tweet, then the other options become defunct since their selections will have no effect. Therefore, if the first setting is selected then all other information regarding this is faded out, as shown in *figure 10*. If the first setting is unselected, it frees up the user to have more control over their location information, as shown in *figure 11*.

The decision was to provide users with a higher level of control over their information than they already had, with more options available to them. Whereas previously a user was only able to remove location tagging, now they are able to remove all information that may be linked to where they had sent the tweet from instead. On top of that, before they could only opt to show only select users this information or everyone but now they have the added option of allowing nobody to view this information if they deem it necessary. Furthermore, the user has more control over what location information is given out and in what form with this design. They can choose to specify how their location information is viewed within a tweet be a specific location such as a set of co-ordinates or a vague location which would then take those co-ordinates and then specify a vague place from where those co-ordinates are positioned, for example making the place as vague as Cardiff, rather than a specific area. This was decided as the optimal decision for design since it gives the user a lot of control which will make them feel at most comfort when using the application.

Prototype Three – Improving User Control

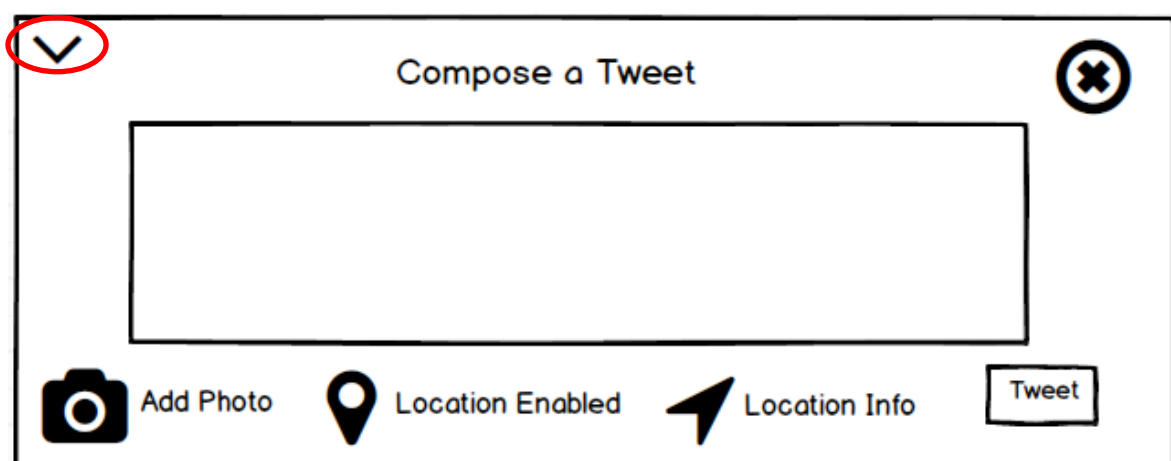**Figure 12: New dropdown feature added to tweet submission box**

Figure 13: The dropdown feature provides users with privacy settings from the help settings, which allows them to change these for a one off tweet

The third prototype is very similar to what the second prototype offers users, except that it gives users the chance to change these settings on a one off tweet. This was set out within the requirements section as a possibility for development to give users more control over a tweet. The simple concept again offers users a functionality that would fit in well with Twitter's simple-to-use interface, as a chance of offering this type of control to users. This way, Twitter is not deviating too far away from what its users are costumed to seeing when they use their site or applications, so like the other new feature; it should not affect their image.

The feature is a simple pull down menu when the user selects the little down arrow, shown in *figure 12*, which then brings about the privacy settings from the menus for which the user can change for a one off tweet, as shown in *figure 13*. The difference between both features is that prototype 2 is a permanent setting, whereas prototype 3 only affects that tweet the user sends, and then the settings revert back to how the user had them.

User Testing the Prototypes

Because of the novelty of the ideas behind the prototypes, it was best practise to gauge possible user's opinions of what they think about the features and what they expect to get out of them. Therefore, there was a need to identify what answer would be of most benefit to the design process that could come from the users.

The first main priority was to gauge what the user thought of the feature, a first impression of what they thought about it and if it was a good idea to include or not. It was decided to split this into a few different questions because simply asking for the users first impressions would give very basic answers that would not be very beneficial to the improvement of the features. Therefore, it was decided that the user was asked about whether they felt they would use the feature a lot once it had been deployed into public use, do they see the feature as being useful to users with its inclusion, do they think the feature would be easily understandable from the way it is presently designed, on top of gauging their initial impressions on things such as the idea behind the feature, the design and any other comments they wish to make about it. The second main priority was to identify anything that the user would change about the feature, so that these ideas could be taken on board and carried through to include in within the second set of requirements, which would be a more refined version of the requirements that were already stated. This question would be asked towards the end of the interview, after giving them enough time to take in the feature and get their thoughts about so that they make any suggestions that they feel would be necessary for the improvement of the feature.

Below is an outline of what would have been expected from the results of the interviews, against the actual feelings of the features that were given by the users that undertook the interviews, which is outlined within the "Results and Evaluation" section of the report. The expectations were compiled from the underlying assumptions that were created from the results of questionnaire and also some of the facts that were revealed from the background research. The actual results were broken down and compiled to give the general consensus from all answers from each question. The questions which were asked in the interviews are also shown below.

1. What are your initial thoughts about the feature? Regarding things such as the design, the idea of the feature, etc.

2. Do you think that it would be a feature that you would use often when you send out a tweet?

3. Do you think that the user would be able to understand the feature from how it has been designed?

4. Do you think that the feature would be useful or not? Would the new feature help to improve user awareness and control?

5. Would there be anything you would change about the feature? Regarding the actual design, the idea of the feature, etc.

Expectation

- In terms of the usefulness of the features, there is a general expectation that a lot, if not all respondents would find the features to be useful because of the fact that the users being interviewed are going to be from a similar pool of people who would have taken the questionnaire in the past. Because the responses from the questionnaire were very favourable to the inclusion of these three features, it can be assumed that the feedback on the three features usefulness will be positive.

- It is expected that the feedback on the initial impressions of the features to be mixed, because each person will have their own ideas in mind of how they would like the feature to be presented to them on the interface. For example, the hover feature in prototype one may be favourable to some and not to others, whereas the clickable pull down menu from prototype three may bring about the same mixed response. Because of the novelty of the features, there is also an expectation that the features may take some explaining before the user understands the feature clearly and they know exactly what each is trying to achieve.

- Again similar to what is stated before; there is an expectation of possible changes that users will voice about because each person has their own way of how they see the feature playing out in their mind. Although this is expected, it should be mentioned that taking all comments into consideration is not advised because having too many different changes being made may result in a feature that is not recognisable to the end user, being contrived. The expression "too many cooks" is applicable to this situation, and making sure that only the most constructive comments that make the most sense need to be considered.

- Another point that may be expressed from the users is that they cannot see the difference between the old interface and the ones being presented them. This would be an opinion which is most welcomed from the point of this project, since the features have been designed in such a way that on the surface, they do not affect the user's experience while using the interface. But until the user actually uses each of the new feature that is when their experience will change and hopefully then the usability and performance is affected greatly so that it becomes a problem to users.

Results and Evaluation

User Test Evaluation

Having carried out a number of interviews to gauge an idea of users initial thoughts towards the prototype designs which had been developed, the results have been summarised a produce in the sections below. The reason for user testing was to be able to understand what the user liked and did not like about the designs and to take on board in changes which users felt strongly about and thought that making the change would benefit the feature positively.

Prototype One

- In terms of the respondents initial impressions, the overriding feeling was that generally the design of the prototype was good, in particular the way the section of the location information was compiled with the different categories of information being presented to the user and also the design of fading everything out to give priority to the feature once it had been selected by the user. Nonetheless, there was a less positive response to the hover idea which had been proposed for Twitter's website. The feeling was to keep things consistent along all platforms so that the clickable feature could be included, rather than the hover.
- There was no real consistency with the respondents answer to whether they felt they would use the feature, but one common denominator was that all said that they would probably use it at one point, be it for the novelty of it or because they were interested in the information in the pop-up box.
- In terms of whether they felt the feature was useful, all found it to be useful for one reason or another but mainly because it served its purpose for achieving the task of raising user awareness.
- There was a good feeling about the design as was mentioned previously, and the respondents felt that the user would eventually get used to the feature over time and that they would not have a problem with using it. Nevertheless, there were comments made about making additions to the feature so that the user could fully understand its purpose for being added to the interface.

Prototype Two

- The initial thoughts were varied for this feature, but most were positive. Some of the common thoughts among the respondents were that the feature seemed simple to use with a very self-explanatory design with the use of the text to explain each option. There was a possible flaw with the text that was picked up by one respondent, whom suggested that the positioning of the text on the page should be reconsidered because it could be misleading as to which piece of text relates to which set of options.

- The feeling from the respondents were sort of contradictory, because they all thought that the feature would be a useful addition but at the same time, none of them felt that they would use it themselves. So what was made from these comments is that they felt the feature would be useful to a lot of other people just not themselves, which should be taken as a positive.
- The main positives that were highlighted from the respondents were the idea of removing all location information from a tweet and not having to worry thereafter. The other positive aspect they felt was the way in which most of the options were labelled with a clear definition of what control you would make having selected that option.
- There were negative comments made about this feature. Some of the options were not labelled appropriately according to the respondents, in particular, the options "as collected" and "approved users only". They mentioned a possible reconsideration of the wording of these two options. Also one respondent even commented on the necessity of giving users the control over the level of information they can provide. They did not feel it was something that users would like to have, even though this was not what came to be from the questionnaire results.

Prototype Three

- The first impressions about the drop-down list option and the one-off aspect of this feature gained the most positive responses from users. Another comment made was that the drop-down list keeps the original interface decluttered from having unnecessary options on it. Other responses which were made which were positive were towards the quick change aspect of changing the user's settings using this feature rather than having to go through all those settings pages. However this comment was contradicted by another response who felt having both this feature and prototype two included in the interface was unnecessary and one should be included over the other.
- The feature also received mixed responses in terms of the options being recognisable from the settings menu. Firstly, one user felt that unless the user actually took the time to look into the settings to find these options, they would be unfamiliar with their purpose once they click on the drop-down item. Because of this, and the fact that this feature includes no help text like the feature from the second prototype, the user may become lost as to the purpose of this in the first place. Contradicting this argument were those who felt that users would be able to recognise this feature from being within the menus and this would mean that users would be at ease with its purpose because of it.
- One of the recommended changes that were common among a couple of the responders was to make the options within the drop-down list more self-explanatory just like they are within the feature from prototype two. But this goes against the positive point made about keeping the page decluttered

which is the overall goal as this keeps the interface along the same lines as how it is being presented now by Twitter.

Evaluating the User Feedback

Having summarised the interview responses, it was clear that the overriding feeling towards the three features were positive. The general feedback towards all three were that the users felt they would be a useful inclusion in the interface and having them there would take steps towards achieving what had been outlined at the beginning of the project. Nonetheless, there was some negative feedback towards some of the design, in particular the first and second feature. There was some negative feedback towards the hover feature which should be looked at and potentially changed if the general feeling towards it stays negative. Also the feeling towards giving the user control over the level of information was not as keen as what was predicted. There was not a lot of feedback about this, and when it was mentioned it was a general feeling of why is it necessary to be included.

Changed Deliverables

It was stated within the initial plan that the intention was to be able to deliver two questionnaires, one relating to the underlying thoughts of the users and the other relating to their thoughts about the privacy implications after the findings that would have come from the first questionnaire and also to gauge what features they would expect from the interface to help them feel more secure with these implications. Due to a restriction with time, it was decided to amalgamate the two questionnaires into one, and split the questionnaire into two sections. This meant that in terms of the deliverables there has only been one questionnaire delivered and one set of interview results, rather than the interviews and focus groups results. Furthermore, the initial plan had stated that there would be tweet analysis carried out from tweets that were derived using a Twitter analysis tool. However, the author was unable to get the tool to function within their own working space to be able to generate a stream of tweet from which they could carry out analysis. Therefore the deliverable of analysing a set of sample tweets has altered to analysing existing work carried out in a similar field of work. Nevertheless, the analysis carried out in this work along with the background research completed on the different aspects that have already been studied by others gave the project a good base to work from.

Extra Deliverables

In terms of work completed that was not outlined or not clear from the initial plan, there has been work carried out to deliver a CATWOE analysis, root definition and conceptual model which was not initially specified. These three deliverables have been developed to clarify the underlying problem and they cover the core problems that exist with both Twitter Users and Twitter in terms of how they are affected by the implications of providing location information. Following on from this work, there has also been work delivered which has taken steps towards improving user

security on the social web, in the form of deriving a set of requirements and subsequent prototypes for designs of features that may be used. The requirements outlined within the report, although could have been more thought out in terms of the preparation completed before outlining them, do meet the core functionality of what the features should be meeting. They address what will be required from each feature both in terms of what would be expected from the user and also from what Twitter would expect from them in line with the user experience it is trying to maintain. It would be naïve to believe that all the requirements were suitable for the feature, or that all requirements that would have been necessary are included within the set that are provided until the features are actually fully developed and tested. These requirements coupled with the fact that prototype designs have been created, tested with current users of Twitter's feedback and they received a positive response from those users, all suggests that the report has taken positives steps towards delivering a set of functional features which will allow for the improvement of protecting users' privacy on the social webs, which was the intention from the initial plan.

Evaluation of Work Completed

There is generally a positive feeling to how the project was planned, carried out and delivered. Firstly, the belief is that the choice of methodology for the project was correct, in particular the use of soft system methodology tools to develop a root definition and a conceptual model allowed for a definitive description of the problem faced by those involved in the situation. Secondly, carrying out the questionnaire well, not to the extent of which was expected in terms of the number of responses, but overall the responses that were generated did prove the underlying assumptions of the project to be true. This choice methodology would have been a useful tool to use for the user testing of the prototype designs also, because having a larger number of responses than what was able to be achieved through interviews would have given the prototype design more legitimacy in terms of how they were perceived by the potential end user. It is the belief that the project was built on the strength of the underlying background research and analysis carried out during the beginning stages of the project, and these fully support the decisions which were taken when developing things such as the root definition, conceptual model and prototype designs. Nevertheless, it is clear that the weakness of the project was being able to prioritise the work which was of most importance to the completion of the project, which would have allowed more time to be carried out on other work that would have helped validate the success of meeting the aims that were set from the beginning of the project.

Future Work

The author has completed some of the necessary initial steps in trying to achieve the aims which have been detailed from the outset of the project. Due to the time constraints set on the project, some of the next steps in project were unable to be undertaken and also steps preceding this point were missed which would have supported the project further. These ideas have been outlined in this section of the report.

- Begin the design process of the two features; one of which was the inclusion of a page which had tutorial type videos which could be used to educate the audience of the implications of providing location information, the other being the feature involving the increased restrictions on linking a users' Twitter account to third party applications.
- Develop and analyse user personas that will help visualise what the end user may be and it will allow for the designs of the prototypes and further development to be user-focused, to make sure that the users can get the most from the features.
- Develop and carry out a questionnaire for further feedback on the prototype designs, which will allow for a more refined second set of requirements and acceptance criteria.
- Carry out further designs of the features to clarify to the implementation team what exactly needs to be implemented. HTA diagrams or activity diagrams could be developed to show how the user will use the feature in order to achieve the end results that they are expecting from it. Use case diagrams can be developed to show the interactions between the user and the interface and the Twitter developers and the interface.
- Articulate a risk assessment to determine how the new introduction of the new features may affect the interface.
- Bring the results and feedback of the project to those who have the authority to carry out the development of the prototype designs so that they can be implemented into Twitter's interface.
- Test the features against the acceptance criteria which had been outlined within the second set of requirement, these being the more refined requirements, so that the features are fully functional to how they were intended to be. This task entails that the features meet both the functional requirements that are necessary for the users to use the features, and meet the non-functional requirements so that things such as performance, usability, etc. are maintained to an acceptable level.
- Initiate a trial period for Twitter users to be introduced to the new features and develop an understanding of their initial thoughts about the features through the use of polls, surveys or other means of generating user feedback.
- Fully introduce the feature into the interface.
- Maintain the features so that the functionality remains to what is expected from the users.

- Evaluate the usefulness and the effectiveness of the features that have been added through means of user feedback to test whether the number of users concerned with the implications of providing location information has altered since the introduction. Further actions can then be taken to continue the development and improvement of the features and the integration of further features which will also take steps towards achieving the aims of the project, if the response from users is positive.

Conclusions

This project has seen the first necessary steps being taken towards promoting the beliefs that there is need for an increase in the awareness of location information and the increase of user control to satisfy their safety while using social networking sites. These beliefs have been promoted by the completion of the following activities:

- A thorough background research and analysis into the ongoing work and research carried out by others in this field of work.
- A purposeful questionnaire that was designed to gauge the underlying beliefs about the problem from the users themselves, and the relative analysis compiled from the results to determine these beliefs.
- The application of Soft System Methodology tools to be able to clearly outline the problems that the main parties involved, the users and the Twitter developers, in the process are faced with and to be able to clearly define the activities which need to be put in place in order for those parties to overcome those problems.
- The specification of an initial set of requirement which will bring about the necessary functionality for which the feature can be of most use to the end-user. This is provided that the requirements are refined and then implemented and tested to the acceptance criteria outlined.
- The development and initial user feedback of a set of prototype designs for which will closely align with how the final implementations of the features will look on the interface. These designs were received with a mostly positive review also.

This project has seen the development of underlying beliefs of many people that there is a problem with the provision of location information and there needs to be actions taken in order to provide users with the safety they require. The initial introduction of the three new features have been found to be perhaps some of the keys to achieving this provision of safety to users and given they are developed on in future, there will be less of an emphasis surrounding the implications surround the provision of location information than there currently is today.

Reflection on Learning

Having undertaken this project, I have learned to appreciate the level of work required to carry out and deliver a project worthy of its work. From the outset of this project, I had a very vague idea of what the big picture of the project would be and what I wanted to be able to deliver at the end of it because I had been viewing the project as one big piece of coursework, instead of breaking it down into smaller, more manageable sections. With the vague picture in my head, I found that it was difficult to break down project into logical steps that I needed to complete in order to achieve what I wanted to in the end. So my objective was to follow the tasks outlined within the initial plan as closely as possible but I found that even those tasks seemed a little too high level and I found that I was never entirely satisfied with what I had been completing in relation to what I had outlined I would be delivering within the initial plan. I would say that around the time of carrying out the user analysis with the questionnaire and carrying out the Soft System Methodologies to help me define the problem faced by the entities involved with the location information that started to get a clearer picture in my head of what I needed to deliver in order to achieve a feasible solution for all parties. This was because that with two pieces of analysis in front on me, I could clearly see what the actors of the system needed and the questionnaire allowed me to see how they felt these solutions would be best delivered to them from Twitter.

Along the project I hit a few stumbling blocks, none more so than at the beginning of the project were I had outlined I would use an existing method of collecting a stream of tweets and analysing the results. It meant that I had to configure the existing code that had been created by another person, in order for the tool to function efficiently on my laptop, which I was unable to achieve. In hindsight, I spent a lot longer on trying to get this tool working than I should have, I could have spent this time more efficiently elsewhere in the project. Nonetheless, I felt that by having this standard of analysis in my project from a stream of tweets would show to what degree the implications that were possible from collecting and manipulating tweets from users and felt that by presenting this in the form of a questionnaire would turn users' opinion on the matter. It meant that towards the latter stage of the project I had less time to spend on tasks involving the specification of the requirements and the design of the prototypes, which could have used that time. Having used tools such as creating different personas and HTA Diagrams in previous university projects, I believe that including these into my project would have helped give more validity to my requirements especially, and from that allowed me to create a more refined set of prototypes to show off to the pending recipients of those questionnaires. The creation of different personas would have allowed me to visualise the features being used by different people in different scenarios and this could have helped me outline what tasks I wanted to achieve from each of the features. This coupled with the use of HTA diagrams, which could have been used to outline these tasks more clearly from the project point of view, would have again, given more backing to my requirements.

One of the main phases of the project was undertaking the process of designing and carrying out the dissemination of the questionnaire. Having previously never carried out a questionnaire, let alone to this scale where the results were going to be very influential to my final project, I felt that I needed to take on a lot of background research to find the best solutions for this phase of the project. It allowed me to understand the best way of designing a questionnaire to be able to contrive results from the respondents in a form that would best help my analysis of the problem. However looking back at the outcome, I felt as though I missed a lot of opportunities with the questionnaire from which I could have gauged a lot more feedback from the respondents, especially in terms of the features that could be help to solve the problem relating to the implications of providing location information on social networking sites. To elaborate, I only asked the recipient whether they felt each feature would be necessary. Naturally, a lot of responses would show that they thought the feature would be necessary because in theory all of the feature sound like they could address the problems. Nonetheless, I felt that I did not get the full opinions from the respondents, especially a deep enough view where I could get an understanding what aspects of the feature were that they felt that in theory would be necessary or not. Another opportunity I felt that was missed out on was the number of responses I had received on the questionnaire. Although I did disseminate the questionnaire into many areas which I felt would get the best range of responses from a different range of users that I could reach out too, having only 42 responses was a little disappointing for what I was hoping to achieve. Another opportunity I believe that I missed out on was with the prototypes of the interface design which I came up with. Perhaps with more time which I did not have towards the end, I could have designed a questionnaire to gauge more opinions from my prototype designs than what I got with the interviews and therefore allowing me to get a better understanding for what users would have wanted from the designs. Overall I felt that choosing the combination of Google Forms and Microsoft Excel to produce the questionnaire and the resulting analysis was a good choice. Having spoken to a number of people who completed the questionnaire, they felt that the design was good and the questionnaire was simple to use. In terms of the analysis, having worked with Excel pretty much all through my education, I was more the comfortable deriving the relevant information I needed from the results.

Another phase of the project was the use of Soft System Methodology. I knew, from having undertaken a project previously which had entailed the use of SSM in the past, that the benefits from carrying out such analysis as a CATWOE analysis and a root definition analysis would help me see the problem at hand. I believe that by doing these two things helped and allowed me to look at the project from different points of view, not just the obvious point of the users. It allowed me to foresee that there may be reluctance from both actors involved and that when I came up with the set of requirements and designs which I did, I had to take these reluctances into consideration. This is why I choose to words some of the requirements with "could" rather than "must" since all parties must be seeing the same picture for the requirements that are a must, and in some of the cases, may be Twitter may feel that the requirements are not feasible with the image they were trying to keep. In

terms of the design of the conceptual model, I believe that I could have perhaps developed into some of the problem areas further. An example being that I stated within the model that an activity would be to assess the capabilities of the Twitter user and Twitter, but I did not take into account other parties or entities whom use the location information as a means of carrying out their business actions. An example being those who use location information to target advertisement, third party applications that use their customers that often link their Twitter accounts to that application, in order to access this source of information. By exploring this part of the problem what have added another dimension to the analysis.

The requirements and prototypes were the next phase in the process. Having previously stated earlier, I believe that requirements could have had more backing from a stand point of where they were derived and could have been achieved through a number of tools rather than a paragraph explaining this. Nonetheless I believe that the requirements clearly explain what each feature should be doing, both functionally and especially non-functionally. I think that outlining the non-functional requirements clearly was important since these are aspects of the feature that affect things such as the interfaces performance, usability, look and other things which can affect Twitter negatively which was identified as a problem from the SSM. Having used it in the past, I believe that the use of Balsamiq Mockup for the prototypes provided me with an adequate set of tools for which I was able to design a clear set of prototypes to show off what was trying to be achieved from each feature. I felt that it was a simple to use, quick package for designing and in any future project involved with design I will be looking to use this package again.

Overall I believe that the project has helped me to learn how to tackle projects of this scale more effectively in future and has helped me to become a better independent worker, in terms of having to be able to motivate myself to get on with work that was outstanding and make sure everything was completed within the timescale set. Looking back, perhaps I should have made better use of the design of and use of the Gantt chart which I created from the initial plan, which would have made the project smoother and would have spared me more time towards the end for other tasks. The big mistake that I made was managing my resources at the beginning which I have mentioned, but overall I am very pleased with the outcome of the report I have documented.

References

Alrayes, F and Abdelmoty, A. 2014. Privacy Concerns in Location-based Social Networks. Dissertation, Cardiff University.

Barkuus, L and Dey, A. 2003. *Location-Based Services for Mobile Telephony: a Study of Users' Privacy Concerns* [Online]. Citeseerx. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.527&rep=rep1&type=pdf [Accessed: 02 March 2015].

Bennett, S. 2014. *Minimum Age Requirements: Twitter, Facebook, Instagram, Snapchat, WhatsApp, Secret [Infographic]* [Online]. The Social Times. Available at: http://www.adweek.com/socialtimes/social-media-minimum-age/501920 [Accessed: 28 March 2015].

Crandell, D. 2010. *Inferring social ties from geographic coincidences* [Online]. Proceedings of the National Academy of Sciences of the United States of America. Available at: http://www.pnas.org/content/107/52/22436.full [Accessed: 01 March 2015].

DSDM Atern. 2008. *10. MoSCoW Prioritisation* [Online]. Driving Strategy Delivering More. Available at: http://www.dsdm.org/content/10-moscow-prioritisation [Accessed: 02 April 2015].

Etherington, D. 2014. *Twitter Acquires Longtime Partner and Social Data Provider Gnip* [Online]. TechCrunch. Available at: http://techcrunch.com/2014/04/15/twitter-acquires-longtime-partner-and-social-data-analytics-provider-gnip/ [Accessed: 19 February 2015].

Fiegerman, S. 2014. *Report: Twitter Will Have Nearly 400 Million Users by 2018* [Online]. Mashable UK. Available at: http://mashable.com/2014/05/27/twitter-user-growth-2018/ [Accessed: 12 February 2015].

Friedland, G and Sommer, R. *Cybercasing the Joint: On the Privacy Implications of Geo-Tagging* [Online]. Data Science Association. Available at: http://datascienceassn.org/sites/default/files/Cybercasing%20the%20Joint%20-%20On%20the%20Privacy%20Implications%20of%20Geo-Tagging.pdf [Accessed: 01 March 2015].

Giyga. 2012. *Privacy Concerns Keep Users Away From Social Sign-In* [Online]. eMarketer. Available at: http://www.emarketer.com/Article/Privacy-Concerns-Keep-Users-Away-Social-Sign-In/1009613 [Accessed: 21 February 2015].

Glancy, R. 2014. *Will you read this article about terms and conditions? You really should do* [Online]. The Guardian. Available at: http://www.theguardian.com/commentisfree/2014/apr/24/terms-and-conditions-online-small-print-information [Accessed: 26 February].

Gnip. *Plugged In Partners* [Online]. GNIP. Available at: https://gnip.com/ [Accessed: 19 February 2015].

Gruteser, M and Grunwald, D. 2003. *Annonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking* [Online]. Usenix. Available at: https://www.usenix.org/legacy/event/mobisys03/tech/full_papers/gruteser/gruteser.pdf [Accessed: 17 March 2015].

Humphries, D. 2014. *Only 8 Percent Read Service Contracts* [Online]. Software Advice. Available at: http://www.softwareadvice.com/security/industryview/data-privacy-report-2014/ [Accessed: 26 February 2015].

Lai, A. 2014. *The Data Digest: The Evolution of Consumer Attitudes on Privacy* [Online]. Forrester Research. Available at: http://blogs.forrester.com/anjali_lai/14-08-01-the_data_digest_the_evolution_of_consumer_attitudes_on_privacy [Accessed: 27 February 2014].

Lederer, S et al. 2003. *Who Wants to Know What When? Privacy Preference Determinants in Ubiquitous Computing* [Online]. Citeseerx. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.86.9506&rep=rep1&type=pdf [Accessed: 02 March 2015].

Madden, M. 2012. *Privacy Management on Social Media Sites* [Online]. Pew Research Center. Available at: http://www.pewinternet.org/2012/02/24/main-findings-12/ [Accessed: 25 February 2015].

Madden, M et al. 2013. *Teens, Social Media, Privacy* [Online]. Pew Research Center. Available at: http://www.lateledipenelope.it/public/52dff2e35b812.pdf [Accessed: 26 March 2015].

Moffitt, J. 2014. *Twitter Geographical Metadata* [Online]. GNIP Support. Available at: http://support.gnip.com/articles/geo-intro.html [Accessed: 20 February 2015].

Nikitha, K and Kumar, N. 2014. *Event Detection and Analysis System* [Online]. International Journal of Computational Science, Mathematics and Engineering. Available at: http://www.ijcsme.com/IJCSME-V1-I5-009.pdf [Accessed: 11 March 2015].

O'Brien, K. 2015. *Citizens 'need more control' over their data* [Online]. The Irish Examiner. Available at: http://www.irishexaminer.com/ireland/citizens-need-more-control-over-their-data-309532.html#.VMuwpGkpld4.twitter [Accessed: 18 February 2015].

Pandey, A. 2011. *What is CATWOE Analysis?* [Online]. BPM geek. Available at: http://bpmgeek.com/blog/what-catwoe-analysis [Accessed: 04 April 2015].

Pew Research Center. *Questionnaire Design* [Online]. Pew Research Center. Available at: http://www.pewresearch.org/methodology/u-s-survey-research/questionnaire-design/ [Accessed: 26 March 2015].

Raicu, I. 2012. *Are Attitudes about Privacy Changing?* [Online]. Santa Clara University. Available at: http://www.scu.edu/ethics-center/privacy/attitudes/ [Accessed: 25 February 2015].

Smith, L. 2014. *Analysing Social Web Data.* Dissertation, Cardiff University.

Snekkenes, E. 2001. *Concepts for Personal Location Privacy Policies* [Online]. Ansatt. Available at: https://www.ansatt.hig.no/einars/papers/ACM_EC01_13_09_2001.pdf [Accessed: 11 March 2015].

Statista. 2013. *Average age of active Twitter users in selected countries as of October 2013 (in years)* [Online]. Statista. Available at: http://www.statista.com/statistics/283120/average-age-of-twitter-users-in-selected-countries/ [Accessed: 28 March 2015].

Statista. 2014. *Number of monthly active Twitter users worldwide from 1st quarter 2010 to 4th quarter 2014 (in millions)* [Online]. Statista. Available at: http://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/ [Accessed: 12 February 2015].

Twitter Help Center. 2014. *FAQs about adding location to your Tweets* [Online]. Twitter Help Center. Available at: https://support.twitter.com/articles/78525-faqs-about-adding-location-to-your-tweets [Accessed: 20 February 2015].

Twitter Help Center. 2014. Using location services on mobile devices [Online]. Twitter Help Center. Available at: https://support.twitter.com/articles/118492-using-location-services-on-mobile-devices [Accessed: 20 March 2015].

Wilson, T. 2005. *How GPS Phones Work* [Online]. How Stuff Works. Available at: http://electronics.howstuffworks.com/gps-phone1.htm [Accessed: 20 March 2015].

Zickuhr, K. 2013. *Location-Based Services* [Online]. Pew Research Center. Available at: http://www.pewinternet.org/2013/09/12/location-based-services/ [Accessed: 19 February 2015].