CARDIFF UNIVERSITY SCHOOL OF COMPUTER SCIENCE AND INFORMATICS

***40 Credits*:** CM3202 One Semester Project
***Author*:** Jamie Ide
***Supervisor*:** Dr. Alia Abdelmoty
Moderator: Dr. Padraig Corcoran

Initial Plan

"Uncovering Personal information in the Internet of Things"

Project Description

The concept of the “Internet of Things” (IoT) is more prevalent now than ever before and its adoption is showing absolutely no signs of slowing down. In fact, the IoT is anticipated to have unprecedented growth through to 2020 with the market expected to surpass $1.7 trillion with over 20.8 billion devices (Nathan Eddy, 2015). The concept surprisingly isn’t new and by nature is simple; The IoT can be described as a network of physical object’s (devices, gadgets or “things”) containing embedded technology that are connected through wired and wireless connections (and unique addressing schemes) which essentially creates a pervasive environment enabling interactions between physical and digital worlds (see Figure 1).

FIGURE 1: IOT ABSTRACT (PANDA SECURITY, 2015)

Support from stakeholders and market forces alike have enabled the establishment of the concept and consequently the features devices provide have progressed tremendously in response. Popular wearable devices now include location tracks, activity records and various health records and indicators. The increase in functionality and sheer number of “things” means an absurd amount of data is being generated by these devices (Cisco estimate by 2018 there to be 403ZB worth). Worryingly, a large proportion of this is personal data or data derived from people (e.g. activity records) which allows us to distinguish a person from other people in a group and therefore profile them (Github, 2014). Thus, the IoT concept provides justifiable exciting opportunities for consumers to make use of these devices and their growing functionality as they willingly provide their data. However, it also provides the need to investigate potential drawbacks from the data collected as a result of the interactions between these devices and humans in regards to compromising user’s privacy, hence the motivation for this project.

A major contribution of this project is to generate a literature review outlining what data IoT devices like the popular “Fitbit” collect, how they go about collecting it and how the data can be used to profile specific users. Then, in order to guide an analytical study of some representative data sets from a device, the dimensions of the problem will be modelled. The literature review, problem model and dataset analysis will demonstrate the extent of potential personal information and privacy threats that may be derived from the gadgets and this information partnered with conducting further research into current threats will allow me to provide users with a consolidated view of threats regarding IoT devices (not just a specific device). In addition, an initial survey and interview designed to examine whether

users have privacy or other concerns about the data being collected will be carried out, as will a subsequent survey and interview to compare any changes in behaviour or attitude once presented with the derived knowledge of the possible privacy vulnerabilities this project may uncover.

This project therefore presents a study of privacy implications from the data being collected by IoT devices in relation to user awareness and behaviour. Put simply, before consumer's consent to wearing a device, they should have the ability to know what data it collects, how to see the data it collects and ultimately what privacy vulnerabilities are associated with the data collected by the device. This project aims to Identify if this is the current position. If proven not to be the case, this project will attempt to provide recommendations (as deemed adequate from a focus group) for protecting personal privacy when using wearable devices and to increase awareness of potential privacy threats.

Aims and Objectives

Aim : Identify what data can be collected, how the data is collected, how the user can access this data and how the data can be used on wearable devices e.g. Fitbit
Objective: Examine terms and conditions provided by relevant distributors of these devices alongside actively scrutinizing functionality of a device (how and what data it collects). Research into possible applications and uses of data collected (both from a user and third party perspective) and research methods of attaining the data from the device.

Aim: Model the dimensions of the problem to identify potential changes which can drive recommendations for improving user awareness on privacy and an analytical study of data
Objective: Employ a Soft Systems Methodology modelling approach e.g. Identify situation considered problematic, express the problem situation, formulate root definitions of relevant system of purposeful activity, build conceptual model of the system from the root definition, compare model with real world situation, define possible changes and suggest actions to improve the problem situation (recommendations).

Aim: Investigate current attitudes and behaviour to privacy implications of users of wearable IoT devices
Objective: Extract beliefs, views and opinions through quantitative methods as to what current users think they know about what data is being collected from them when using wearable devices like Fitbit and whether they have any privacy or other concerns about the data being collected.

Aim: Identify what information/patterns (high level data mining) and subsequent privacy threats can be derived from analysing data of wearable devices
Objective: Analyse datasets from wearable devices, comparing what can be achieved through the collection of location tracks, activity records and other health indicators over a long period of time (e.g. 1 year) and compare that against instantaneous snapshots that

the user normally see on their devices (e.g. daily/weekly updates) to identify potential privacy threats.

Aim: Identify current threats to users "Privacy" using wearable devices
Objective: Use the analysis conducted in the previous step to identify possible threats to personal privacy that are implicit in the data and are not directly presented to the user through the normal interfaces. Conduct further research to identify any potential existing threats which have already been uncovered.

Aim: Identify any changes in behaviour or attitudes to privacy threats when providing previous participants with results of analytical study and research (potential privacy threats)
Objective: Re-Investigation using same quantitative method employed previously to original participants in order to compare any changes in behaviour or attitudes once confronted with identified potential threats.

Aim: Identify a set of recommendations for protecting privacy when using wearable IoT devices like Fitbit
Objective: Describe a suggested course of action to be taken to solve a particular privacy threat.

Ethics

Reviewing the ethical guidelines on the Cardiff University website (Spasic, 2015) has enabled me to identify that in order to adhere to Cardiff University's policies for carrying out ethical research and to safeguard professionalism and integrity, I will need to fill in and submit an ethical approval form. This is due to collection of potentially private data from wearable devices and extraction of knowledge through quantitative methods (interviews, surveys).

I will also ensure:

1. Anyone assisting me in my project will receive a clear statement of what the project is about, what it involves, and what their part in it will be. If they are willing to take part they must confirm this in writing.
    a. They will have the right to remain **anonymous**
    b. Anyone who subsequently withdraws their consent will have that decision respected such that it will render any data they provided unusable.
2. Questionnaires/Surveys will require approval from my supervisor.
3. Data will be used solely for its intended purpose in relation to this project
    a. In accordance with the Data Protection Act, I will not divulge any personal details relating to anyone without their expressed permission.
4. I will only collect and use the **volume** of data deemed necessary to complete this project.

Work Plan

Red = Critical Green = Risk

On-going Work throughout the Project:

Task 1	Research work in relation to this project which has been carried out in order to broaden and deepen knowledge in this area
Task 2	Progress meetings with Supervisor (Weekly, and Critical Quarterly Supervisor Review Meetings (CQSRM))
Permanent Tasks	Write up Final Report.

Week 1 – Commencing 25th January

Task 3	Develop Initial Plan

Deliverables **-** Initial Plan

Week 2 – Commencing 1st February

Task 5	Research all information described in initial Aim and Objective.
Task 6	Produce literature for research completed into studies carried out in this field of work with respect to the privacy threats on wearable devices: • What are the user's current attitudes and levels of awareness towards privacy threats?
Task 7	Begin to model the dimensions of the problem e.g. SSM
Task 8	Initiate the collection of a dataset **(Risk – I cannot get data from a willing participant who has been using a relevant device – supervisor has used Fitbit and I have used Apple Watch so can model this data as mitigation)**

Deliverables – Written report of the findings from task 5 and completion of literature review into current attitudes

Week 3 – Commencing 8th February

Task 7	Complete the SSM modelling.
Task 9	Develop Questionnaire for Initial survey of user's attitudes based of literature produced in task 6 • Obtain approval for questionnaire • Pilot questionnaire to ensure it is fundamentally correct in its aim to meet deliverables outlined and coherent by nature • Distribute first Questionnaire
Task 8	Continue collecting relevant data sets (explore programs for analysis, and begin to design set of questions to implement on data elements collected)

Deliverables – SSM Model, First Questionnaire.

Week 4 – Commencing 15th February

Task 2	(CQSRM)
Task 8	Continue collecting dataset, begin to make use of chosen program and analyse the dataset by implementing designed questions against small dataset ready to be applied to larger dataset once gathered.
Task 9	**Whilst waiting for results of questionnaire**, interview a select number of people using same questions from questionnaire. (can then understand first hand user's opinion's on data being collected and potential privacy threats). **RISK – Potential delay in response from questionnaire, planned distribution early as mitigation (alongside researching methods for getting maximum exposure) and will conduct interviews to gather instant feedback, allocated other work to ensure continued progress.**

Deliverables – *Datasets, Interview Results*

Week 5 – Commencing 22nd February

Task 8	Finish collection and analysis on Dataset (larger) using chosen program
Task 9	Collect results from initial questionnaire and write up results

Deliverables – Collection and write up of first Questionnaire, Dataset Analysis

Week 6 – Commencing 29th February

Task 10	Research existing threats to privacy regarding IoT devices
Task 11	Design a second questionnaire including; Findings from dataset analysis, findings from researching existing threats and incorporating results from initial questionnaire and interview (follow same procedure as initial e.g. approval, pilot)

Deliverables - Report on existing privacy threats regarding IoT devices, Second Questionnaire.

Week 7 – Commencing 7th March

Task 2	(CQSRM)
Task 11	**Wait on results from second questionnaire**. Again, interview a select number of people using same questions from the questionnaire to allow myself the opportunity to understand first hand if there is a change in user attitude once prevented with privacy threats. **RISK – Delay in response, same mitigation applied as previous.**

Deliverables - Second Interview results written up.

Week 8 – Commencing 14th March

Task 11	Collect results from second questionnaire and write up results.
Task 12	Carry out comparison of both sets of questionnaire results.

Deliverables - Collection of results from second questionnaires, comparison and of the question questionnaires.

Easter Recess – Commencing 21st March - 8th April

During this period, I will continue to write up the final report whilst also comparing my results with any previous results from other related projects.

Week 9 – Commencing 11th April

Task 2	(CQSRM)
Task 13	Identify any recommendations which can be established from results, arrange focus group to gather opinions on these recommendations with the intention of increasing awareness of privacy threats.

Deliverables – Focus group results and any subsequent recommendations

Week 10 - Commencing 18th April

Permanent Task	Write up Final Report (Including conclusions, reflection on learning during project)

Week 11 – Commencing 25th April

Permanent Task	Final Report. First Draft.

Deliverables – First Draft.

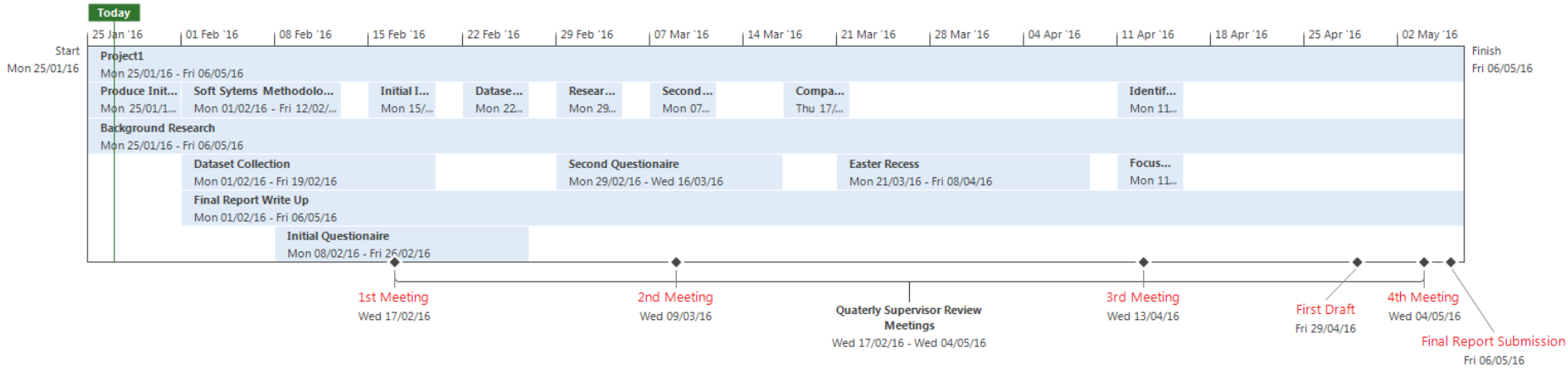
Week 12 – Finishing 6th May

Task 2	(Final CQSRM)
Permanent Task	Final Report. Proof read, make relevant alterations and submit project.

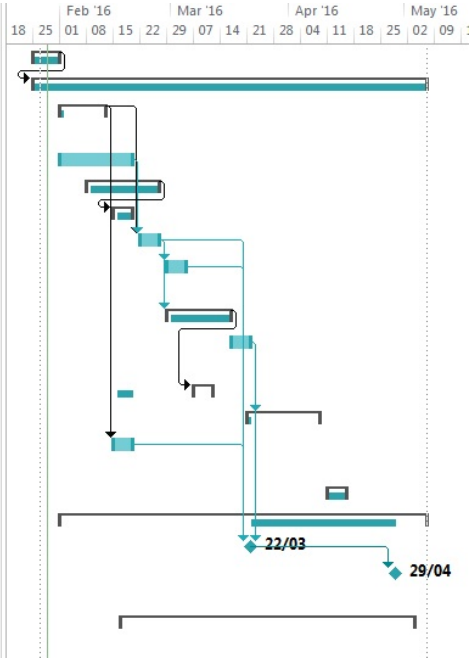
Deliverables – Final Report.

Risk – **Conflicting workload as module run in parallel to this project has coursework due, mitigated as have lowered allocated work load on lead up to submission to ensure adequate time can be spent completing that work.**

Gantt chart



	Task Name	Duration	Start	Finish	Predecessors	Successors	Add New Column
1	**Produce Initial Plan**	**6 days**	**Mon 25/01/16**	**Sun 31/01/16**		**9**	
9	**Background Research**	**75 days**	**Mon 25/01/16**	**Fri 06/05/16**	**1**		
13	**Soft Sytems Methodology Modelling**	**10 days**	**Mon 01/02/16**	**Fri 12/02/16**		**36,24**	
17	Dataset Collection	15 days	Mon 01/02/16	Fri 19/02/16		24	
18	**Initial Questionaire**	**15 days**	**Mon 08/02/16**	**Fri 26/02/16**		**22**	
22	**Initial Interview**	**5 days**	**Mon 15/02/16**	**Fri 19/02/16**	**18**		
24	Dataset Analysis	5 days	Mon 22/02/16	Fri 26/02/16	17,13	26,25,42	
25	Research into current privacy threats	5 days	Mon 29/02/16	Fri 04/03/16	24	42	
26	**Second Questionaire**	**13 days**	**Mon 29/02/16**	**Wed 16/03/16**	**24**	**31**	
30	Comparison of Questionaire results	3 days	Thu 17/03/16	Mon 21/03/16		33,42	
31	**Second Interview**	**5 days**	**Mon 07/03/16**	**Fri 11/03/16**	**26**		
33	**Easter Recess**	**15 days**	**Mon 21/03/16**	**Fri 08/04/16**	**30**		
36	Identify and Establish Recommendations	5 days	Mon 15/02/16	Fri 19/02/16	13	42	
37	**Focus Group**	**5 days**	**Mon 11/04/16**	**Fri 15/04/16**			
41	**Final Report Write Up**	**70 days**	**Mon 01/02/16**	**Fri 06/05/16**			
42	First Draft	0 days	Tue 22/03/16	Tue 22/03/16	36,30,25,24	43	
43	Final Report Submission	0 days	Fri 29/04/16	Fri 29/04/16	42		
44	**Quarterly Supervisor Review Meetings**	**55 days**	**Wed 17/02/16**	**Wed 04/05/16**			



References

CISCO. (2015). *The internet of things and big data: Unlocking the power.* Available: http://www.zdnet.com/article/the-internet-of-things-and-big-data-unlocking-the-power/. Last accessed 26th January 2016.

Github. (2014). *A Colloquial Definition of Big, Open, and Personal Data.* Available: https://github.com/theodi/data-definitions. Last accessed 26th Jan 2016.

Nathan Eddy. (2015). *Gartner: 21 Billion IoT Devices To Invade By 2020.* Available: http://www.informationweek.com/mobile/mobile-devices/gartner-21-billion-iot-devices-to-invade-by-2020/d/d-id/1323081. Last accessed 25th January 2016.

PANDA SECURITY. (2015). *Thousands of errors found in multiple Internet of Things devices.* Available: http://www.pandasecurity.com/mediacenter/family-safety/internet-of-things-security-devices/. Last accessed 31st Jan 2016.

Spasic, I. (2015). Research ethics - Cardiff school of computer science and informatics, Available at: http://users.cs.cf.ac.uk/I.Spasic/ethics/. Last accessed 25th January 2016.