

WESTERN SYDNEY UNIVERSITY

DOCTORAL THESIS

**A framework for understanding
and predicting the take up and
use of social networking tools in a
collaborative environment**

Author

Rhys Shaun Aldridge TAGUE

Student Identifier: 16170676

*A thesis submitted in fulfilment of the requirements
for a Doctorate of Philosophy*

July 2016

Acknowledgements

Before it began the question of *why* conjured in the front of my mind. Why would you spend so many years focusing on one thing? Why would you bother writing a thesis if only a handful would read it? Everyone I asked couldn't quite answer it. Until one said, *you learn a lot about yourself*.

I believe a PhD is more than just a topic. It is an understanding. An understanding that whatever you do in life is a result of the people around you. The people who help, hinder, inspire and most importantly love you. Your existence is a result of others, therefore what you achieve through life will be also. Others can make it hard, as they don't see the same path as you, and believe you're going about it all wrong. Yet, there are others who make it easier, as they encourage you to *keep at it*. Those are whom I would like to acknowledge.

My primary supervisor Professor Anthony Maeder, I could dare say had a magical way of lifting stress. It was truly just a result of his ability to think clearly, however I like to think it was magical. There were times where I needed a clear perspective on my thinking and he was the one who provided it. I remember the first time Anthony said, "*I want you as my student*.". At first... no thanks, but now I am so grateful that I was a student of his. I couldn't see myself with any other supervisor.

My co-supervisor Jim Basilakis thank you for your constructive criticisms and insights when in our general meetings. Without those I do not believe I would have produced the quality of work that I have.

To my friends, those moments of laughter, support and your ability to simply understand was second to none. I believe I have learnt a lot about what a true friend is. Also, the coffee! The simple act of grabbing a coffee was euphoria.

My Family. The most important, yet the most hard to say thank you to. How can I thank you for your love and support. What do you say? I'm very lucky to exist in a family where there is more than parents and siblings. We are friends. Best of friends. The only thing I can say is, I love you all and always will.

There were times I travelled back to the original question of *why?*. I think I have the answer now: *You learn a lot about yourself, and others*.

Declaration of Authorship

I, Rhys Shaun Aldridge TAGUE, declare that this thesis titled, 'A framework for understanding and predicting the take up and use of social networking tools in a collaborative environment' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

Contents

Acknowledgements	i
Declaration of Authorship	ii
Contents	iii
List of Tables	vii
List of Figures	viii
Abbreviations	x
Abstract	xii
1 Introduction	1
1.1 Publications	5
1.2 Research Question	6
1.3 Research Aim and Objectives	6
1.4 Methodology	7
1.4.1 Problem space	8
1.4.2 Background and Related Work	8
1.4.3 A Framework Model - Construct	8
1.4.4 Construct Use	9
1.4.4.1 Framework Implementation	9
1.4.4.2 Case Studies	9
1.4.4.3 Participant Use and Data Collection	9
1.4.5 Research Contribution	10
1.4.5.1 Framework Output	10
1.4.5.2 Reflect empirical findings back to literature	10
1.4.6 Explore Scope of Application	10
1.5 Thesis Road Map	10
1.5.1 Introduction	10
1.5.2 Background and Related Work	11
1.5.3 Framework Model - Construct	11

1.5.4	Case Studies	11
1.5.5	Participant Use and Data Collection	11
1.5.6	Social Group Role Signatures	11
1.5.7	Discussion	12
1.5.8	Conclusion	12
1.5.9	Reference List	12
1.5.10	Appendices	12
2	Background and Related Work	13
2.1	Introduction	13
2.2	Collaboration	14
2.2.1	Online Collaboration	19
2.3	Group Dynamics	21
2.3.1	Social Groups and Their Dynamic Structure	22
2.3.2	Social Group Roles	24
2.4	Online Collaborative Environments	27
2.5	User Modelling	31
2.5.1	Understanding and Predicting User Behaviour	34
2.6	Identification and Analysis of Roles Online	36
2.7	Summary of Findings	39
3	Framework Model - Construct	41
3.1	Introduction	41
3.2	Framework Life Cycle	42
3.2.1	Evaluation	44
3.2.2	Commitment	44
3.2.3	Change	45
3.2.4	Transition	45
3.3	Online Collaborative Environment	46
3.4	Functional and Characteristic Behaviours	46
3.4.1	Publisher	47
3.4.2	Annotator	48
3.4.3	Lurker	48
3.5	Implemented Framework	49
3.5.1	Visual Representation of Implemented Framework	49
3.5.2	Widget/Tools Metadata	50
3.5.3	Interaction Packages	53
3.5.4	Framework Domain	55
3.5.4.1	Framework Proxy	55
3.5.4.2	Framework Processing	58
4	Case Studies	74
4.1	Introduction	74
4.2	Espressobility	75
4.2.1	Description	75

4.2.2	Collaborative Activities	76
4.2.3	Explicit User Roles	85
4.3	Squashies	85
4.3.1	Description	85
4.3.2	Collaborative Activities	86
4.3.3	Explicit User Roles	92
4.4	Walk 2.0	93
4.4.1	Description	93
4.4.2	Collaborative Activities	94
4.4.3	Explicit User Roles	99
5	Participant Use and Data Collection	103
5.1	Introduction	103
5.2	Espressobility	104
5.2.1	Recruitment Method	104
5.2.2	User Registrations	104
5.2.3	Participant Use	105
5.2.4	Data Collected	109
5.3	Squashies	110
5.3.1	Recruitment Method	110
5.3.2	Participant Use	111
5.3.3	Data Collected	114
5.4	Walk 2.0	114
5.4.1	Recruitment Method	114
5.4.2	Participant Use	115
5.4.3	Data Collected	119
6	Social Group Role Signatures - Framework Output	121
6.1	Introduction	121
6.2	General Signatures	122
6.3	Context Based Signatures	128
6.4	Accommodating Characteristic Behaviours	136
6.5	Handling A Dynamic User Social Network Structure	139
7	Discussion	143
7.1	Introduction	143
7.2	Confirmation of Research Aims	143
7.3	Framework Comparison	147
7.4	General Discussion	149
8	Conclusion	154
	Reference List	156

Appendices	169
A <i>Assessing User Engagement in a Health Promotion Website Using Social Networking</i>	169
B <i>Adaptive Web Framework for Online Collaborative Environments</i> .	179
C <i>Classifying Collaborative Behavior in the Form of Behavioral Stereotypes in Collaborative Mobile Applications</i>	185
D <i>Espressobility Recruitment Brochure</i>	191
E <i>Squashies Recruitment Brochure</i>	193
F <i>Ethics Consent with Terms and Conditions of Participant Use</i> . . .	195

List of Tables

2.1	Types of Awareness	18
3.1	Data Dictionary for an Interaction Package	54
3.2	Data Dictionary for Interaction History Database Table	57
3.3	Data Dictionary for Domain Table	63
3.4	Data Dictionary for Neuron Table	64
3.5	Data Dictionary for Synapse Table	64
3.6	Data Dictionary for Synapse Weight Table	64
3.7	Fuzzy Set Terminology	69
3.8	Ranges for behaviour strength set terms	70
3.9	Ranges for behaviour strength set terms	71
4.1	Type of activities a user could create within the Expressibility application	77
4.2	Type of activities a user could create within the application	87
6.1	Top 10 most active users and their general social group role signatures within the Expressibility case study	124
6.2	Top 10 most active users and their general social group role signatures within the Squashies case study	124
6.3	Top 10 most active users and their general social group role signatures within the Walk 2.0 website case study	125
6.4	The context signatures for use 3199 and their use of the different activity types	135
6.5	Initial accommodation of behaviours for the Expressibility case study by showing general signatures of first user	138
6.6	Initial accommodation of behaviours for the Squashies case study by showing general signatures of first user	138
6.7	Initial accommodation of behaviours for the Walk 2.0 website case study by showing general signatures of first user	138
6.8	Signature of user 8 before and after the creation of a social-tie with user 6 within the Expressibility case study	142
6.9	Signature of user 8 with context of the review post tool before and after the creation of a social-tie with user 6 within the Expressibility case study	142

List of Figures

2.1	Three domains of collaboration [1]. The intersection highlighting the point of collaboration	15
2.2	A graphical representation of Moreland’s Model of Group Socialisation combined with Biddle’s Characteristic Behaviours	26
2.3	An interface of the Social Coding Platform GitHub	28
2.4	A screen capture of the Facebook News Feed	30
3.1	Framework Life Cycle represented as the social group role life cycle	43
3.2	Proposed framework emulating the social group role life cycle	43
3.3	Visual Representation of Framework	49
3.4	Widget/Tools metadata segment of framework	51
3.5	Framework Domain showing framework components	55
3.6	Visual Representation of Artificial Neural Network (ANN)	60
3.7	Flow Diagram of Dynamic Creation of the ANN Structure	62
3.8	Database schema for the ANN	65
3.9	Hyperbolic Tangent	66
3.10	Visual representation of linear membership (fuzziness)	68
3.11	Visual representation of Behaviour Strength	70
3.12	Visual representation of Behaviour Age	71
3.13	Visual Representation of Fuzzy Logic (FL) system	73
4.1	Screen shot of global Activity Stream within the Espressoability App	78
4.2	Screen shot of the Coffee location Activity Stream within the Espressoability App	79
4.3	A screen shot of posting tools for coffee locations	80
4.4	A screen shot of a user profile within the application	82
4.5	A screen shot of the location information dashboard	83
4.6	Screen shot of activity commenting	84
4.7	Screen shot of activity tagging	85
4.8	Screen shot of the activity stream of Squashies the mobile application	88
4.9	Screen shot of a class list view in the Squashies App	89
4.10	Screen shot of a class session view in the Squashies App	90
4.11	Screen shot of a user profile within the Squashies application	91
4.12	Screen shot of a class view within the Squashies Application	92
4.13	Screen shot of application settings within the Squashies Application	93
4.14	Screen shot of the site dashboard for the Walk 2.0 website	97

4.15	Screen shot of a profile page for the Walk 2.0 website	100
4.16	Screen shot of a group page within the Walk 2.0 website	101
4.17	Screen shot of forums within the Walk 2.0 website	102
5.1	Bar graph showing coffee locations created by users of Espressobility	106
5.2	Bar graph showing activities created by users for locations within Espressobility	107
5.3	The percentage of the different activity types created within the Espressobility application	108
5.4	Distribution of comments made by each user within Espressobility .	109
5.5	Number of class sessions attended by each user within Squashies . .	112
5.6	Distribution of activity types within Squashies	113
5.7	Number of activity actions per activity type within Squashies . . .	114
5.8	Distribution of activity item types within the Walk 2.0 website . . .	116
5.9	Segmentation of visibility for activity type Status Post within the Walk 2.0 website	116
5.10	Total activity items posted by the top 20 users within the Walk 2.0 website	117
5.11	Top 20 users with the most comments for activity items within the Walk 2.0 website	118
5.12	Top 20 users with the most thumbs up for activity items within the Walk 2.0 website	119
6.1	The general signatures for user 8 within the Espressobility case study over the user's life	126
6.2	The general signatures for user 3203 within Squashies over the user's life	127
6.3	The top 10 users and their actions within the Walk 2.0 website . . .	128
6.4	Location context based signatures for user 3 of Espressobility	130
6.5	User 1 and their determined general signature with their user profile context signature within Espressobility	131
6.6	Role signatures for user 1 with refined context based signatures with user 8 and tools/widgets within the Espressobility case study. . . .	132
6.7	Location social-tie with location tool signature comparison for user 26 within Espressobility	133
6.8	Comparison of context signatures for instructor user 3203 with class members and non member within the Squashies case study.	134
6.9	Distribution of total actions for user 3199 within the Squashies case study.	136
6.10	Visual signature of user 6 before and after the creation of a social-tie with user 8 within the Espressobility case study	140
6.11	Visual signature of user 6 with context of the review post tool before and after the creation of a social-tie with user 8 within the Espressobility case study	141

Abbreviations

ANN Artificial Neural Network

API Application Programming Interface

BPNN Back-Propagation Neural Network

COG Centre of Gravity

CRUD Create, Read, Update and Delete

CSCW Computer Supported Cooperative Work

FL Fuzzy Logic

HREC Human Research Ethics Committee

HTTP Hypertext Transfer Protocol

IPUSM Interaction Package and User Session Manager

JPS Joint Problem Space

JSON Javascript Object Notation

NEAF National Ethics Application Form

NFS Neuro-Fuzzy System

NHMRC National Health and Medical Research Council

NN Neural Network

OCE Online Collaborative Environment

OSN Online Social Networks

PHP Hypertext Preprocessor

RDMS Relational Database Management System

RESTful Representational State Transfer

SKI Shared Knowledge Indicator

SNA Social Network Analysis

SNS Social Networking Sites

UI User Interface

WSU Western Sydney University

WWW World Wide Web

XML EXtensible Markup Language

Abstract

Doctorate of Philosophy

A framework for understanding and predicting the take up and use of social networking tools in a collaborative environment

by Rhys Shaun Aldridge TAGUE

Online collaborative environments, such as social networking environments, enable users to work together to create, modify, and share media collaboratively. However, as users can be autonomous in their actions the ability to create and form a shared understanding of the *people*, *purpose*, and *process* of the collaborative effort can be complex. This complexity is compounded by the the natural implicit social and collaborative structure of these environments, a structure that can be modified by users dynamically and asynchronously.

Some have tried to make this implicitness explicit through data mining, and allocation of user roles. However such methods can fail to adapt to the changing nature of an environment's structure relating to habits of users and their social-connectedness. As a result, existing methods generally provide only a snapshot of the environment at a point in time. In addition, existing methods focus on whole user bases and the underlying social context of the environment. This makes them unsuitable for situations where the context of collaboration can change rapidly, for example the tools and widgets available for collaborative action and the users available for collaborative interactions.

There is a pre-existing model for understanding the dynamic structure of these environments called the "*Group Socialisation Model*". This model has been used to understand how social group roles form and change over time as they go through a life cycle. This model also contains a concept of characteristic behaviours or descriptors of behaviour that an individual can use to make judgement about another individual and to create an understanding of a role or social norm that may or may not be explicit. Although studies have used components of this model to provide a means of role identification or role composition within online collaborative environments, they have not managed to provide a higher level method or framework that can replicate the entire life cycle continuously over time within these environments.

Using the constructive research methodology this thesis presents a research construct in the form of a framework for replicating the social group role life cycle within online collaborative environments. The framework uses an artificial neural network with a unique capability of taking snapshots of its network structure. In conjunction with fuzzy logic inference, collaborative role signatures composed of characteristic behaviours can then be determined. In this work, three characteristic behaviours were identified from the literature for characterisation of stereotypical online behaviour to be used within a role signature: these were *publisher*, *annotator*, and *lurker*.

The use of the framework was demonstrated on three case studies. Two of the case studies were custom built mobile applications specifically for this study, and one was the Walk 2.0 website from a National Health and Medical Research Council project. All three case studies allowed for collaborative actions where users could interact with each other to create an dynamic and diverse environment. For the use of these case studies, ethics was approved by the Western Sydney University Human Research Ethic Committee and consistent strategies for recruitment were carried out.

The framework was thereby demonstrated to be capable of successfully determining role signatures composed of the above characteristic behaviours, for a range of contexts and individual users. Also, comparison of participant usage of case studies was carried out and it was established that the role signatures determined by the framework matched usage. In addition, the top contributors within the case studies were analysed to demonstrate the framework's capability of handling the dynamic and continual changing structure of an online collaborative environment.

The major contribution of this thesis is a framework construct developed to propose and demonstrate a new framework approach to successfully automate and carry out the social group role model life cycle within online collaborative environments. This is a significant component of foundational work towards providing designers of online collaborative environments with the capacity of understanding the various implicit roles and their characteristic behaviours for individual users. Such a capability could enable more specific individual personalisation or resource allocation, which could in turn improve the suitability of environments developed for collaboration online.

Chapter 1

Introduction

The vast range of dynamic behaviour seen in our world is a result of our relationships with others. Each relationship, personal, professional, or casual determines our behaviours towards others. For instance, a project team working towards the next milestone has an explicit structure of relationships composed of roles and behaviours identified for each member's interaction and contribution. Yet, a social outing with friends results in an implicit structure, and therefore so are the individual's roles and behaviours.

Relationships tend not to be complex, they either exist or they don't, yet the attributes of the relationship, such as behaviour, are. Online relationships have brought about their own challenges, yet the online world enables tracking and capturing user interactions to identify these attributes which can then be processed and analysed to be classified as behaviour. In terms of collaborating, understanding or even predicting these attributes, such as behaviour, to understand a user's role or intention within an [Online Collaborative Environment \(OCE\)](#) is complex and troublesome because of their dynamic and asynchronous nature.

The rise of the internet has amplified the ability of individuals to create relationships. Yet, there is a fundamental difference between relationships online and the offline. It is the ability of a user to be asynchronous with their activities and interactions. This results in a user being able to interact with historical media or action that exists in the present. This unique characteristic of the internet has moved it from a stale landscape of information to rich social media and collaborative environments where users are able share, modify, and create media together to achieve a task or objective.

Today, these environments are used by 88% of individuals in the 15-17 year old age group and 86% of the 18-24 year old age group [2]. It is the dominate form of internet usage, and has resulted in some social media environments having billions of user bases. For instance, Facebook, currently the largest [Social Networking Sites \(SNS\)](#) in the world, has 1.15 billion monthly active users [3]. The size of these environments are a result of their ability to foster the creation of relationships and provide various interactions past and present.

The types of the relationships created within these environments are similar to offline relationships where explicit and implicit relationships can be established. The two types of relationships can be thought of in the terms of awareness. Where explicit relationships are relationships users are clearly aware of, and implicit ones are those which are not. An example is GitHub, a highly focused [Computer Supported Cooperative Work \(CSCW\)](#) environment for code development, allows users to explicitly follow projects and other users, however users are also able to clone repositories of code allowing users to create implicit ties to these repositories and users [4]. Both type of relationships are constructive to collaboration as both types enable a user to contribute to collaborative efforts even if it is consumption - clone code repository.

Users can become aware of implicit relationships through awareness tools, such as activity streams, notifications, and interaction logs [5]. These tools present past interactions users may have had with other users or nodes, and therefore allow a user to infer relationships, which may or may not be part of a user's *following* or *friends* list. For example, a user clones a code repository and an activity status is published to an activity stream stating the user's interaction. Another user could infer based on the type of project cloned what the user is related to.

Even though these tools can take the implicit to the explicit for relationships they still cannot show a representation of a user's behaviour and its strength within a [OCE](#), and as a result a user's interaction can only be represented not their behaviour. This is a result of the Internet and its asynchronous nature. Therefore, identifying the behaviours and its strength from a singular interaction in a collaborative setting is futile. Instead, a user would have to track a user's history by collating and filtering past interactions with members and other nodes within a network. An effort which would be hindered by the implicit relationships and behaviours not visible to users. As a result, the user can only achieve a fragmented picture of a user's behaviour within a collaborative setting.

Complexity is added as social media and CSCW result in a user's environment being dynamic. A result of skeletal functionality where tools and widgets allow users to interact and create relationships, resulting in *their* environment. This dynamic capability creates a complexity around understanding behaviour, as a user may show varying behaviour between child environments created through possible user groupings. For instance, GitHub allows a user to follow and star other users or repositories and in each grouping their behaviour may be different [4]. The result of this is a user's behaviour is dependent on context of the relationships and interactions they perform and if these relationships are implicit it is impossible for an user to determine a user's behaviour.

Although behaviours can be dynamic, user's still present patterns of behaviour over time. Patterns of behaviour derived from their habitual interactions with tools and widgets. For example, an user uses a tool to annotate photos from a circle of users in a photo sharing environment, yet they may never use tools to upload new photos. This type of pattern could then change based on the circle of friends, where they do upload photos, yet never annotate. The variation of patterns could be infinite based on the tools and widgets available. Although infinite, the friends who interact with a particular user will only need to know about their behaviour within the context of their circle of friends.

These patterns of user behaviour can result in a implicit role assignment by other user's within an environment. In general the term role is used for classification of relations and behaviours to manage analytic complexity that can be a result of social systems [6]. As a result the term can be used as an umbrella to cover a range of attributes and metrics associated with users. This is especially so with varying environments and task oriented objectives or achievements. Although the term may result in an umbrella effect, there are varying degrees of abstraction from analysis of user psychological attributes with content published to the most granular level of simplistic interactions resulting in behavioural and relational patterns [6]. At the granular level roles are less dependent on context compared to abstracted roles such as *Social Networkers* where their purpose is to build and establish networks in a social networking environment [7]. Using a bottom up approach for the formation of such a role, at a granular level the user would be carrying out one of or many [Create, Read, Update and Delete \(CRUD\)](#) commands towards another user profile through the interface of tools and widgets needed to achieve such an interaction. Their continual use of such tools would lead to a

stereotypical behaviour and when clustered together could present a role signature. For instance, a user may publish to another user's profile continually, in essence networking, yet their behaviour continually displayed is publication to a relation, a **Create** or **Update** command from **CRUD**.

To provide this characterisation in real time in an collaborative environment would be comparable to social group role formation [8]. In social groups every individual within a group (an environment) contributes to the group's success towards its purpose. Each individual earns an *implicit* role to help achieve the group's purpose of existence. The role is determined through the Group Socialisation Model [8]. The model presents a life cycle for group socialisation that results in an *implicit* role for individuals. Within this life cycle there is an evaluation phase where other individuals of a group evaluate a user's characteristic behaviours and then determine an understanding of a user's role. The characteristic behaviours to determine a role is usually referred to an individuals *social norm* [9]. A social norm being a generalised view of what is acceptable behaviour amongst a group. Offline the variation of social norms is infinite, however online social norms are restricted to the interactions an individual can carry out. Their interactions deriving from the tools and widgets available to them or the semantic meaning of an action carried out to achieve a task. For instance, a user is able to share content amongst their network, and therefore it is an acceptable behaviour.

Offline the social group role life cycle is easily achieved by evaluating a individual's behaviour over time. Online however, as **OCE** are dynamic and asynchronous the ability to carry out the social group role life cycle is hindered and complex as a user cannot effectively start the evaluation of characteristic behaviours which is the first phase of the life cycle. Therefore, the users fail to form implicit roles altogether unless more they have a way to identify these characteristic behaviours.

This thesis presents a construct in the form of a research framework for automating the social group role life cycle within **OCE**. Believed to be the first of its kind, such a framework could determine role signatures composed of characteristic behaviours that could be used in a range of scenarios from helping users form *implicit* roles to personalisation of collaborative tools and interfaces supporting collaboration.

1.1 Publications

This research has resulted in various publications:

- Rhys Tague, Anthony J Maeder, Corneel Vandelanotte, Gregory S Kolt, Cristina M Caperchione, Richard R Rosenkranz, Trevor N Savage, and Anetta Hinchliffe. **Assessing User Engagement in a Health Promotion Website Using Social Networking.** In *Global Telehealth*, pages 84–92. IOS Press, 2014

This paper presents an early analysis of user participation within the Walk 2.0 case study. In addition to this, content analysis of content published was analysed to determine the possible reasons for use.

[Appendix A](#)

- Rhys Tague, Anthony Maeder, and Jim Basilakis. **Adaptive Web Framework for Online Collaborative Environments.** *Australian Journal of Intelligent Information Processing Systems*, 13(4):11–15, 2014

This paper proposes an early design of the research framework for the purpose of personalisation of tools based within an Online Collaborative Environment.

[Appendix B](#)

- Rhys Tague, Anthony Maeder, and Jim Basilakis. **Classifying collaborative behavior in the form of behavioral stereotypes in collaborative mobile applications.** In Aly A. Farag, editor, *Multimedia Technology IV*, chapter 165, pages 165–169. CRC Press, 2015

Initial output from the Espressobility cases study is presented in this paper. The components of the framework are also discussed. The classification of behavioural stereotypes are also presented.

[Appendix C](#)

1.2 Research Question

How does one characterise and automate the creation of social group role signatures through user collaborative behaviour within a dynamic and flexible online collaborative environment?

1.3 Research Aim and Objectives

The aim of this research is to provide a construct in the form of a research framework for automating the social group role life cycle for the creation of social group role signatures within online collaborative environments.

This research aim can be broken down into the following objectives:

- Propose a research framework construct for automating the social group role life cycle for the creation of social group role signatures with the use of:
 - An asynchronous, yet life cycle design
 - A dynamic and automated machine learning method for characterisation
- Investigate and identify stereotypical behaviour in online collaborative environments.
- Provide an implementation of the framework construct to be demonstrated in a real world online collaborative environments
- Demonstrate the implemented framework with various online collaborative environments
- Compare and analyse framework output against actual longitudinal user usage data.
- Explore the dynamic capability of the research construct with the implemented framework output.

To achieve these objectives the following outputs are required:

- An implementation of the research framework that can
 - Formulate general social group role signature composed of characteristic behaviours that can be compared over time
 - determine context based social group role signatures composed of characteristic behaviours
 - adapt to new characteristic behaviours over time
 - adapt to handle changing user social network structures
 - be generic and accommodating of different online collaborative environments
- Development of case studies to demonstrate real world application of the framework
- Identification of simplistic stereotypical behaviour to represent general characteristic behaviours for social group role signatures within multiple case studies

1.4 Methodology

This study uses a constructive research methodology. This methodology enables the creation of knowledge on how a problem can be solved, modelled, or understood in principle for a domain specific problem [13]. The creation of a construct is the core to this approach. A construct being an artifact, in the nature of either a model, framework, algorithm, software, or theory to establish practical functioning or theoretical contribution to a problem. This methodology is applicable to many disciplines, such as computer science, management, accounting, and medicine. The ability to identify a problem space, provide a solution in the form of a construct, and then demonstrate the construct enables one to provide a contribution through practical function or theoretical connection, thus providing a proof of concept.

The methodology can be characterised into 6 distinct phases [14]:

- Discover a practical problem space in the relevant discipline
- Obtain understanding of the problem space

- Provide an innovative construct solution
- Demonstrate that the use of the construct works
- Provide theoretical connection and research contribution
- Explore the scope of application

The following outlines the various steps for each phase undertaken for this study:

1.4.1 Problem space

The ability to characterise and automate the creation of social group role signatures within online collaborative environments is the problem space. The problem space has been outlined as a research question in section 1.2 and with the research aim in section 1.3.

1.4.2 Background and Related Work

An extensive literature review has been undertaken to understand and identify related work and problem space. This review involved using various keywords related to the fields of collaboration, CSCW, user modelling, SNS, social group dynamics, personalisation and recommendation systems. The findings of this review steered the development of the research framework. This is outlined in section 2.7.

1.4.3 A Framework Model - Construct

A framework model was proposed incorporating the learnings and gaps identified from the literature review. The framework model employs machine learning, meta data, social ties and stereotypical behaviour classifications. It was also designed to be dynamic and asynchronous, yet also capable of a sequential life cycle. The model is described in chapter 3.

1.4.4 Construct Use

1.4.4.1 Framework Implementation

Once the model was proposed an implementation of the framework was developed. This implementation was interfaced with various online collaborative environments. The implementation consisted of a client and server system where the client side was installed within the [User Interface \(UI\)](#) of the online collaborative environment, and the server side was independent to the online collaborative environment server system. The implementation is presented in section [3.5](#)

1.4.4.2 Case Studies

The resulting framework was interfaced with three different [OCE](#) to demonstrate the practicality of the construct in real world scenarios. There were two mobile applications with a local community cohort of participants and one [National Health and Medical Research Council \(NHMRC\)](#) funded project (reference number 589903) called the Walk 2.0 project. The mobile application case studies allowed participation registration during their operation and use. This allowed the framework to be demonstrated in an asynchronous environment with the dynamics of a continually changing collaborative environment. The Walk 2.0 project however, provided a historical trial arm of nation wide participant use allowing for the demonstration of the framework to operate on historical participant use. With these three case studies a real world application and practicality of the construct could be demonstrated. Without such a demonstration, the construct use and capability to carry out an asynchronous life cycle while also being able handle the dynamics of a continually changing collaborative environment with real time and historical participant use could not be shown. Each case study is discussed at length in Chapter [4](#).

1.4.4.3 Participant Use and Data Collection

Participant use within the mobile applications resulted in data collection for the framework. For the Walk 2.0 website the recruitment of participants was outside the scope of this study, however data collected from the Walk 2.0 website was used with the framework under the approved ethics of the Walk 2.0 study.

1.4.5 Research Contribution

1.4.5.1 Framework Output

The output of the framework was analysed against actual participant use of the case studies. In addition, the output for certain users were explored to test the frameworks capability to handle various aspects of an [OCE](#) in regards to uncertainty and dynamic capability of a user's social network. This is describe in chapter [6](#).

1.4.5.2 Reflect empirical findings back to literature

Once output was determined by the framework, confirmation on research objectives and comparison of relevant literature against the frameworks capability to determine social group roles signatures was discussed. This is presented in section [7.2](#) and [7.3](#)

1.4.6 Explore Scope of Application

A general discussion and future work was presented to explore future applications a this research. This is presented in section [7.4](#)

1.5 Thesis Road Map

The following is an outline of this research thesis and a short explanation of each chapter.

1.5.1 Introduction

The introduction gives an opening to the research and an brief outline of the research contained within this thesis.

1.5.2 Background and Related Work

An extensive literature review was carried out to give insight into the background of the research, and also related work. The literature review consisted searching for related keywords and domains of the research. The domains covered are collaboration, [CSCW](#), user modelling, [SNS](#), social group dynamics, personalisation and recommendation systems.

1.5.3 Framework Model - Construct

In this chapter the model of the framework is presented. This includes inspiration for the framework, the life cycle of the framework, and each component of the framework. A implementation of the framework is also presented in this chapter.

1.5.4 Case Studies

This chapter will explain each online collaborative environment where the framework has been applied, and the collaborative purpose and activities of each environment. How the collaborative activities were represented in the framework will also be discussed.

1.5.5 Participant Use and Data Collection

The method for recruitment of participants and their use of the case studies is presented in this chapter.

1.5.6 Social Group Role Signatures

This chapter presents the output of the framework and its capabilities. Certain users and their role signatures will be presented, and the ability of the framework being capable of handling uncertainty with social-ties will also be presented. The output of the framework will be compared with participant use to compare the accuracy of the frameworks output.

1.5.7 Discussion

A discussion of the empirical findings is presented in this chapter. These findings are compared back to the literature. Limitations of the framework is also discussed along with future work for automation of the social group role life cycle and the creation of role signatures. In addition this, various applications of the framework is also discussed.

1.5.8 Conclusion

An outline of the research and its contribution will be summarised within this chapter.

1.5.9 Reference List

Listing of literature used within the thesis will be presented within this chapter.

1.5.10 Appendices

Appendix items, such as papers published and participant recruitment information needed for this research.

Chapter 2

Background and Related Work

2.1 Introduction

As this research uses constructive research as its methodology a thorough and in-depth literature review of background and related work was conducted. This literature review will identify and explain the background of the research domains this research pertains to, theories of those domains and related work. The findings within the review have also provided inspiration and reasoning for the resulting framework construct. This review is also used for the discussion of this thesis to provide understanding and insight towards the empirical findings from the research construct.

For this review the following approach was carried out: Based on the research question and aims various article and journal databases were used; Google Scholar, Computers and Applied Sciences Complete (EBSCO), ProQuest Computer Science Collection (ProQuest), ScienceDirect (Elsevier) and Social Theory (Alexander Street Press). These databases were then queried with various search terms relating to each major domain of the research question and aims. These search terms formed the inclusion and exclusion criteria. Literature which did not contain or directly relate to these search terms were excluded. There were also various *implicit* questions which steered the literature review which were:

- What does it mean to collaborate, and is there a definitive model for collaboration?

- What models exist in research to model online collaborative behaviour?
- What approaches are used in the real world to model online collaborative behaviour?
- How can asynchronous behaviour be acquired in online collaborative environments?

The bibliographies of articles and journals identified as relevant were also used as a way to navigate the literature. The resulting keywords are as follows: *user model, behaviour model, interaction patterns, recommendation algorithm, computer supportive cooperative work, online social network, social networking site, online collaborative environments, online user behaviour, collaborative behaviour, group dynamics, social group roles, collaborative roles, role formation, characteristic behaviours, role identification, collaborative interface*. With these keywords modifying operators were also used, *AND* (\mathcal{E}), *OR* ($-$), *NOT* ($!$), to filter results.

Once all literature was collated, a synthesis was carried out to form an outline and structure for the literature review. This entailed grouping relevant literature together to form bodies of knowledge for each domain relevant to this research. For instance, literature that related more to a generic model of collaboration was not grouped with online collaborative environments although there were connections between the two. Once this was conducted this led to the qualitative review of the following background and related work.

2.2 Collaboration

What does one mean by the term collaboration? The Oxford Dictionary gives the definition as a simple statement, “*the action of working with someone to produce something*”, however this doesn’t encapsulate the many facets involved for *successful* collaboration. The study of collaboration is extensive and there are many theories relating to the different aspects needed for it to occur, however many of these theories are given within a particular context. Croker *et al* [1] demonstrate this through their work by attempting to define what the term collaboration means and how it differs from team work. However, they conclude it is a simple definition and tends to only be a component of collaboration. As a result, they then ask

the question is collaboration so general of a term that one cannot define it? They do however, present a simplistic three domain model where all three domains are persistent throughout the literature: **people, purpose, and processing**. Figure 2.1 gives a visual representation of their three domain model. The intersection of these three domains is where the phenomenon of collaboration occurs.

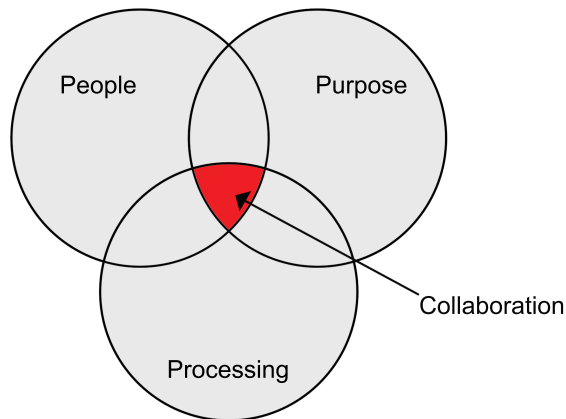


FIGURE 2.1: Three domains of collaboration [1]. The intersection highlighting the point of collaboration

Although Croker *et al* provides a generalised model and understanding of collaboration without context, the intricate behaviours of individuals involved in collaboration are not highlighted. Instead, one could suggest as long as these three domains exist collaboration will occur. Gray [15] further explores and provides a detailed foundation of the intricate details of collaboration stating collaboration as “*through which parties who see different aspects of a problem can constructively explore their difference and search for solutions that go beyond their own limited vision of what is possible.*”. When compared to Croker *et al* model, shows the occurrence of collaboration is more dependent on the individuals, or ‘parties’, and their perspective of the purpose of collaboration. This early work by Gray provided a strong foundation for these intricate details, yet more recent work by Thomson *et al* [16, 17] provides a more succinct definition:

Collaboration is the process in which autonomous or semi-autonomous actors interact through formal and information negotiation, jointly creating rules and structure governing their relationships and ways to act or decide on the issues that brought them together; it is a process involving shared norms and mutually beneficial interactions [17]

Thomson *et al* show collaboration is more than the simple statement of ‘*the action of working with someone to produce something*’. Collaboration is a result of actors and their dynamic communications and interactions with each other. This dynamic capability leads to the structure and rules of the collaboration being carried out.

Collaborative learning environments can highlight these interactions as participants who enter into these environments bring their unique knowledge and understandings [18]. When participants within the environment initially gather they go through a knowledge building phase by asking questions to establish a ‘*shared understanding*’. This understanding is used to gauge others for the purpose of determining gaps or capability of those who collaborate. In addition, the understanding allows collaborators to construct a state of mind for those who are collaborating [19].

Without this shared understanding, coordination and actions taken within a collaborative scenario will be more cooperative than collaborative [20]. As cooperative is more a division of labour, where collaborative is a mutual respect for a problem and engagement to achieve a coordinated effort - shared understanding. Once a shared understanding is achieved the movement towards a shared knowledge starts to form. Roschelle *et al* [20] argue to achieve shared knowledge for collaboration a **Joint Problem Space (JPS)** is needed to create structure within such a fluid and flexible phenomenon that is collaboration. Their suggested **JPS** structure integrates:

- goals
- description of the current problem state
- awareness of available problem solving actions
- associations that relate goals and features of the current problem state, and available actions

Although such a structure encapsulates a way to establish shared knowledge, like the term collaboration, one can see it cannot work for all types of collaboration. However, one aspect of Roschelle’s *et al* argument tries to be all inclusive with that collaborative problem solving is a result of a shared conceptual problem space,

languages used, situation and activities carried out. However these elements of a [JPS](#) is continually changing within some contexts of collaboration, such as online collaboration, and therefore the ability to create a clear structure for a [JPS](#) become problematic.

The work by Callazos *et al* [21] attempt to removed this complexity by presenting an indicator called a [Shared Knowledge Indicator \(SKI\)](#). This indicator identifies the level of shared knowledge a collaborative learning group may have. As shared knowledge is hard to map to every context, the [SKI](#) is used to show shared knowledge awareness. An awareness that allows an individual to understand the collaborative structure of collaborative groups [22]. Without this awareness individuals become isolated and disconnected to other individuals collaborating.

The method of questioning to discover a trail of interaction is one of the more simplistic methods for forming a shared understanding. Ogatta [22] *et al* provides a table, as presented in table 2.1, outlining the types of questions needed to achieve a better level of knowledge awareness. For collaboration each type of awareness is needed and as a result the questioning an individual will ask will be a combination of the various awareness domains.

Awareness	Example Questions
Social	What should I expect from other members of this group? How will I interact with this group? What role will I take in this group? What roles will other members of the group assume?
Task	What do I know about this topic and the structure of the task? What do others know about this topic and task? What tools are needed to complete task? How much time is required? How much time is available?
Concept	How does this task fit into what I already know about the concept? What else do I need to find out about this topic? Do I need to revise any of my current ideas in light of this new information? Can I create a hypothesis from my current knowledge to predict the task outcome?
Workspace	What are the other members of the group doing to complete the task? Where are they? What are they doing? What have they already done?
Knowledge	Who is discussing looking at the same knowledge that I am looking at now? Who has changed the knowledge since I have last looked at it? What knowledge are they discussing now? What knowledge of my input did they change?

TABLE 2.1: Types of Awareness

The facets of collaboration are vast and broad like the term. The literature highlights this and also the depth upon which one can go to understanding the complexity of collaborative behaviour. Behaviour that is derived from individuals appearing to act autonomously based on their dynamic communication and behaviour - a result of distributed understanding and *interfaces* to interact with

other collaborators. For this research, the definition Thomson *et al* [17] present in addition with a shared understanding [18] that is needed amongst individuals will be the guide and understanding when the term collaboration is used.

2.2.1 Online Collaboration

Is there a difference between collaborating online and offline? A valid question, as online so much is taken from an individual yet at the same time given. *Traditional collaboration* as discussed in the previous section is sometimes presented as a phenomenon as every individual brings an underlying potential structure to a collaborative process, however online, even though flexible, it is limited by the tools and interfaces available. Therefore there is predefined interactive behaviours to collaborate online. This restriction validates the question to distinguish collaboration online and offline.

Although online interfaces restrict the ability of total freedom, today these are capable of achieving a fundamental element of collaboration. They allow users of interfaces to be *autonomous* with their interactions and communications instead of just cooperating - very much like the offline world. The way this is achieved is through various interfaces so individuals can create the shared understanding and awareness of the activities carried out within a collaborative scenario. *Activity theory* provides a theoretical model towards understanding awareness of actions, or *activities* taken place within a collaborative scenario. Traditionally the theory was used to provide an understanding of general human activity, yet it is also used to help understand *activities* and their use online and in software development [23].

Understanding activities online can help mediate collaboration as the theory can provide context of motive, goal, and condition of activities being carried out [24]. With this context, tools and interfaces can be used to help users create an understanding of activity and action performed over time. The context is a result of the theory using the principles of coordination, cooperation and co-construction [25]. Coordination being result of the act or acts a user carries out towards a common goal, cooperation is how it is achieved, and finally, co-construction relates to the ability of evaluation of tools and practices to achieve coordination and cooperation. Hemetsberger *et al* [25] use an [OCE](#) for open source software development to observe the principles of the theory to try and demonstrate this by what users *do* within the environment over a 4 month period. With a focus on core activities

users were able to, through the use of interfaces and tools, achieve a *coat tailing* effect where users replicated activity sequences from other successful collaborative projects within the environment. They were able to transfer context and procedure from one project to the next. This highlights the ability to create a shared understanding of the activities and tools needed to achieve a successful collaboration. Awareness is not explicitly defined in their work, however it is central to this shared understanding so users can be autonomous.

The concepts discussed so far are more for small group, or focused collaboration online. However, with the rise of social media, online collaboration can be with hundreds, if not thousands of individuals. As users having 100% autonomy and without the need to relate to other users they are able to operate solely with the tools and interfaces given to them within an [OCE](#) to create, modify and delete media for their own purpose, yet still play a role within a large collaborative effort. They have their own motivators, such as goals, tasks or activities they wish to perform. An example of this phenomenon is with Wikipedia and its Wikipedians [26]. Wikipedia, an online encyclopaedia, is a result of thousands of users coming together to create an open and editable encyclopaedia online. Today, the collaborative effort of the broad spectrum of users of Wikipedia has resulted in the replacement of the traditional hard copy encyclopaedia, such Britannica. Although the users of Wikipedia are not experts in areas of knowledge needed for an encyclopaedia article, their actions and motivators to create such articles has resulted in an encyclopaedia reputable to those in existence for over 244 years [27]. Rafaeli *et al* [26] present possible motivators to help explain why users carry out activities in such an environment. Motivators ranging from protecting one's ego to just having fun. Such work demonstrates how an individual online can be solely driven by their own goals or outcomes rather than that of a collaborative context.

Although Wikipedians may be driven by an individualistic goal or outcome, their ability to achieve such a goal or outcome comes back to their awareness of the interactions and actions performed within the environment. GitHub, an [OCE](#), allows individuals to work together to build complex software projects by creating transparency through its interfaces too [4]. Each user within GitHub is independent and their behaviour and interactions are made available for other users to discover their collaborative effort. This *Social Coding* of projects is comparable to Wikipedia with articles, as users do not have to commit to one overarching goal

or outcome of a small group. Dabbish *et al* [4] carried out a study to test the transparency of behaviour and interactions in relation to a user's ability to infer another user's collaborative effort. Through visible cues provided by interfaces and tools, their study showed individual users were able to infer a rich set of user attributes. This inference not only allowed users to understand another user's collaborative effort, but also the level of quality of code or project the user worked on.

In comparison to offline collaboration, online collaboration and its ability to provide interfaces and tools for collaboration, awareness through transparency, and 100% autonomous actions, enable a user to attain the same level of collaboration online although users are restricted by the provided interfaces and tools. Although a restriction, interfaces and tools as shown are also the enabler for a user to achieve 100% autonomous contribution and as a result differentiate the two types of collaboration contexts.

2.3 Group Dynamics

Individuals form groups as a group tends to be of interest to the individual [28]. This interest might be a hobby or a way to achieve an objective, such as raising profits for a company. Another could solely because of homophily [29], where individuals come together because of a similar interest or characteristic others possess. We see this in our society as individuals come together for hobbies, such as public speaking, because they have a genuine interest or it fulfils a need for the individual.

The formation of a group, or people coming together to form a group, is an incremental process [30]. That is, there may only be two or more individuals who come together to form a group at formation, and over time more individuals join. This incremental growth can lead to the groups objective, if defined, to be changed or refined based on the type of individuals who join the group. We see in our day-to-day life this incremental growth, it can be fast, such as individuals rallying together because of a natural disaster, or others slow, such as a rallying for a charitable cause.

The size of the group and the variation of individuals in the group can determine its power and capability. The term "*Two Minds are Better Than One*"

lightly describes the power of a group. Every new individual who joins the group results in the groups growth in power and capability in regards to its intellect and strength. Woolley *et al* [31] demonstrate this growth in intellect through a study they conducted on individuals who were randomly assigned to groups. Each group was given tasks to carry out and the group's performance was documented. It was found there were signs of improved intellect capability of each group as more people were added. They also showed a group would perform poorly if it was dominated by certain characteristically chosen individuals as the group was not collectively working together. Even though a group's power and capability is increased with size, a discovery by Ringelmann [32] many years ago showed the larger a group got the less performance a single individual would contribute to the goal or objective of the group. As a result, the groups actual output of power and capability is hindered by the size of the group and demonstrating that individuals in the group do not work as one. Therefore, the larger the group the more "*free riders*" or people who enjoy the benefits of the group without contribution, exist.

Depending on the type of group, such as a social group, will determine the type of actions, interactions and dynamics that go on within a group. Outlined so far are the general aspects of group dynamics for all group types which are: reason for group membership, group formation, and the power of groups. For this work only social groups and the collaboration process within such a group will be reviewed. This review will also include the group's structure and the formation of roles within such groups.

2.3.1 Social Groups and Their Dynamic Structure

Social groups can exist within society as a group of family members, friends, or work colleagues. Homophily [29] tends to be the cause of such groups, with each member having something in common with every member of the group. The dynamics of such groups are complex and tend to be implicit. They differentiate from other group types because they do not need explicit structure [8, 33]. The purpose of a social group to an individual is still about fulfilling a need the individual has [28], but the need is not explicitly expressed by the group.

A social group's structure is made up by the individual's interactions with other members of the group [8, 33]. Moreland [8] provides a model on how members of a

social group do this. They go through three stages with every member they interact with: the first is to evaluate other members, this evaluation process allows the member to understand where other members are in regards to their contribution to the group and their purpose within the group. Second, after a member has evaluated their group members they then commit to their evaluation. By committing they can understand what they can expect from the other members and the benefit the members will bring to the social interactions and goals or objectives of the group. The last phase of Moreland's model is the transition phase. As social groups are very dynamic, members of the group re-evaluate members constantly to monitor changes within the group. This re-evaluation is set off by their initial evaluation and deviation of members or a member has from this evaluation.

Palla *et al* [33] also shows how the success of a social group is dependent on this dynamic behaviour. Without this process the members carry out, the ability of the group to change or *Transition* to accommodate the group's members or its goal, the group will disintegrate. This work however highlights a significant point when it comes to the size of social groups. If a social group is relatively small, the group is based off the strong relationships and therefore the group will only stay together if these relationships continue to be present. However, if the size of a social group is considerable these strong relationships are not the stabilizer of the group, instead the stabilizer is the dynamic behaviour changing and the interactions conducted by the members. An example is a student body in a school. Students travel up through the years until they graduate and leave, but the students in the school as a whole is still represented as a social group, even though all the students can be replaced with new ones.

This dynamic behaviour assignment or dynamic structure removes boundaries, such as size, which would exist if the group had a detailed and pre-thought out structure. It allows the group to adapt to new challenges or objectives as they arrive. A common challenge being the entry of a new member. When a new member joins other members of the group would evaluate their position and determine their role or potential contribution to the social group. As a result, the group can grow indefinitely without problem.

The unlimited boundaries of social groups, because of its dynamic structure, allow for collective intelligence [34]. Tadeusz [34] describes the phenomenon of collective intelligence as a result of group behaviour and he also gives examples of such behaviour, one being seen in army ants [35]. Franks [35] show that Army ants on

their own or in small groups are unintelligent. To demonstrate this, if you place one hundred ants in one location each ant will follow the ant in front continuously around in a circle until the ants die of exhaustion. However, if you have a group with hundreds of thousands of ants the group becomes intelligent and can regulate a ants nest within one degree Celsius, even though there seems to be no structure and at sometimes may seem chaotic. A result of a single ant being able to evaluate the simplistic behaviour of thousands of ants around it to understand its behaviour forward.

We are now seeing this collective intelligence behaviour as a prevalent characteristic found within online social networks [36, 37]. Users are able to understand their placement within a group by other's actions. A type of behaviour not too dissimilar to what was discussed in section 2.2.1 where users are creating a shared understanding and awareness to achieve collaboration.

Even though there is a lot of literature on social theory [30], there are a lot of questions left unanswered for social groups, which leaves understanding of these groups desired. With the rise of online social networks and the easier access to the online social network's data the dynamic behaviour of social groups is becoming a focal point for research in social theory.

2.3.2 Social Group Roles

Every individual in a group contributes to a group's success. This contribution varies from simple existence within the group to leading or managing it. As a result, each member is assigned either directly or indirectly a role or an area which they look after to help achieve the group's success. For direct assignment of roles the resulting group tends to be called a team [38].

The roles within a team do not emerge, but instead are given out to individuals based on their behaviour according to Belbin [38]. Belbin, who created team role theory, provides a test for individuals within a team to complete which in turn determines their possible role within a team. From this test one of nine possible roles can be given to an individual: Plant, Monitor Evaluator, Co-Ordinator, Resource Investigators, Implementers, Completer Finishers, Team workers, Shapers and Specialists. For a team to be successful, in terms of completing its goal, all these roles must exist. He also suggests, all these roles be evenly distributed for

team success. However, with the rise of online collaboration, where a team can have thousands of members, these roles cannot be evenly distributed. Also the ability of individuals being able to assess and assign one of Belbin's roles to every member is not possible, as not all members know Belbin's team roles.

In social groups however, we don't tend to see direct assignment or testing of individuals to determine their role. Instead indirect assignment occurs as a result of their dynamic structure. This dynamic structure is a result of the group member's ability to evaluate other members. For example, each member inside a social group can carry out their own cognitive, communicative, and interactive tasks in a group, and as a result, these actions or behaviours determine the role the user is undertaking in the group. Biddle [9, 39] calls this *characteristic behaviour*, where the behaviour follows a certain type of norm which has formed by the social group. These social norms form from widespread characteristic behaviours throughout the social group [40]. The social norms are implicit so they are not explicitly given to the group. They're passed on from member to member by members watching other members characteristic behaviours.

Figure 2.2 gives a graphical representation of Moreland's Group Socialisation Model combined with Biddle's Characteristic Behaviours. The initial characteristic behaviours, which are usually from social norms, are evaluated by other members. After another member has finished their evaluation they commit to that person's role within the group until their characteristic behaviours change. After which, a transition phase starts and the process starts again. This model shows that a social group is continually changing as there is no end point which makes a role terminal.

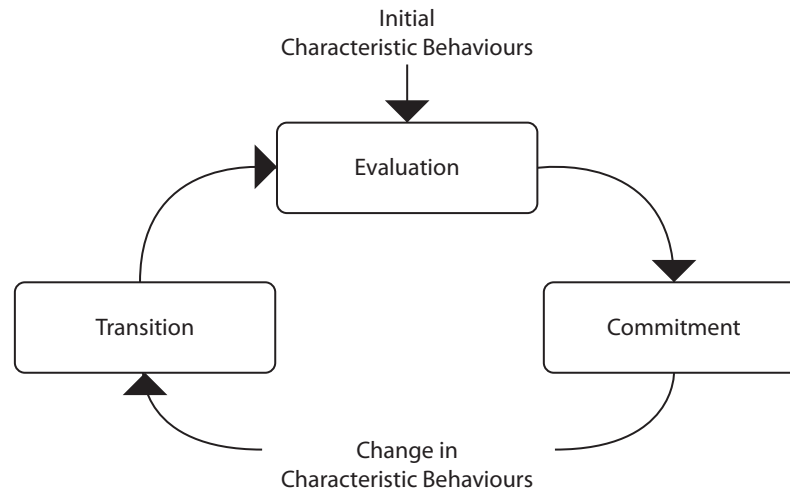


FIGURE 2.2: A graphical representation of Moreland's Model of Group Socialisation combined with Biddle's Characteristic Behaviours

This combined model could be compared to the establishment of shared understanding, however on an individual level. A method for an individual to establish autonomy within a social group without having to understand an entire group's overall goal or outcome, yet being able to contribute because of their ability to evaluate other members to establish their own understanding and awareness of their group members behaviours and therefore their roles. One can see this kind of behaviour in open source software development [41]. When an open source project initially starts there is no real hierarchy present, however, overtime the contributions and behaviours shown by the project members form the hierarchy, the core project leaders and roles. As a result, the characteristic behaviours present in certain members become acceptable social norms, and in this case of open source development the social norms of leadership or project management.

There are many theories on why roles form, which Biddle has reviewed extensively [9], yet this is outside the scope of this study. For this work the focus will be on how characteristic behaviours in a social group determines an individuals role or status. This focus will be explored more in section 2.6 methods of identification and analysis of user roles online.

2.4 Online Collaborative Environments

Initially the [World Wide Web \(WWW\)](#) was stale when it came interactions, only allowing users to consume media not create it. Those who could create were developers in tune with the technologies to achieve the production of a web-page or media. Although primitive in interaction, it still allowed collective intelligence. If a user happened to take interest in a web-page and they had the capability as a developer, they could link to that web-page of interest. This simple capability of linking enabled the construction of a vast network of knowledge and media all connected and accessible. Not until “*Web 2.0*” and its new design patterns were end-users empowered to interact with a web-page and each other. Web 2.0 brought focus towards the end user producing media instead of developers. Instead of the consumption of a web-page a user could use a web service to participate in the creation of media. It enabled a service to continually update itself based on its user base and not its developer base, resulting in an environment’s architecture based on user participation and not production of media [42]. Such an environment enabled users to start interacting with each other and ultimately collaborating with each other within environments.

This focus of an architecture of participation resulted in the creation of [OCE](#) limited in their media, yet complex in their functionality. Developers approached the development of such environments with the understanding that the resulting media will be from the actions shown by its users and their use of the functionality provided. This approach can be problematic to such environments as users may not collaborate since the perceived value of the environment cannot be seen initially due to the lack of media [43, 44]. To circumvent this, seeding of these environments is a common practice [45]. Seeding is where users who are early adopters, or influencers, are encouraged to start participation for some type of reward. Their efforts create awareness for other users about the potential of such an environment. In effect, creating a shared understanding.

Although participation is the focus of [OCE](#), the result of this participation can be very explicit or implicit based on the type of the environment.

Formal environments, such as groupware, where goals and objectives are explicitly set out to the user base, tend to focus on participation based on a user’s role [46, 47]. For instance, in a collaborative learning environment the role of instructor

and student are natural to the collaborative context. Therefore, the tools available to each role are explicit. For instance, an instructor is capable of producing asynchronous media for students to consume specifically for learning [48]. This consumption by users can be carried out in a collaborative or individual way to help the students learn.

This type of explicitness is also found in purpose driven environments where an environment's existence is for one particular purpose. It does not relate to the user base, but the service offered by the environment. Open source software development environments are an example of purpose driven environments. GitHub, a social coding online collaborative environment, is an example [4]. No roles are given to users within such an environment, yet the tools and widgets provided by the environment are explicit and therefore enable a user to adopt varying styles of participation. Fig 2.3 demonstrates the different interfaces available which enable users to take on different modes of interaction. Users are able to contribute to the code base, monitor and maintain issues, or even create a Wiki for the project's documentation. A user is then able to take on many roles to fulfil the purpose of the environment. Another example of a purpose driven environment is a *Wiki*. There are no roles which people can explicitly take on, however, they are able to interact in explicit ways resulting in one user being able to carry out specific roles for an explicit purpose and for a Wiki it is the creation of an article [7].

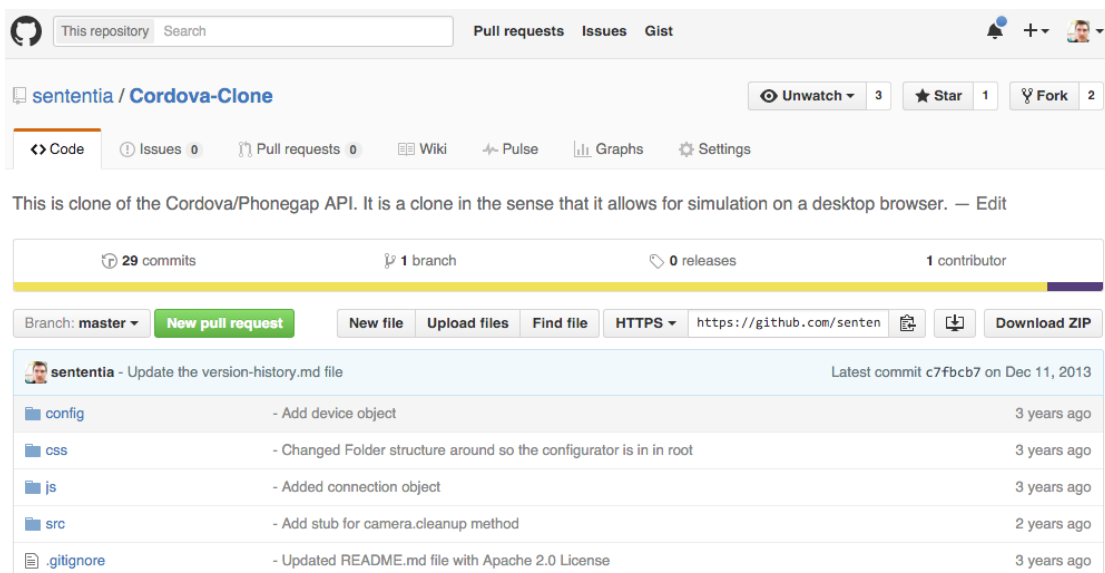


FIGURE 2.3: An interface of the Social Coding Platform GitHub

Online Social Networks (OSN) do not tend to possess this explicitness of purpose or formal collaborative environments. Instead the purpose of a user's participation is a result of their own goals or objectives which may or may not relate to the other users [49]. For instance, a user may join a group and participate with that group on the basis of homophily [50]. Where homophily when two or more users have an interest in a common, such as a culture, topic or goal. This establishes a shared understanding for the purpose of collaboration. The purpose is more inferred by users rather than made explicitly by them [51]. It is this implicit attribute of **OSN** that have allowed them to achieve the dominate form of **OCE**. User's are empowered to be completely autonomous with their participation. A user can create their own formal or purpose built environment for their own goals and objectives and then participate according to *their* environment. An example of this implicitness of an **OSN** is within Flickr [52]. Users tag photos and the purpose or objective of tagging those photos is entirely dependent on the individual. One would think tagging photos would be for organisation, or descriptors of what the image is composed of, however users may tag for the reason of sharing by tagging other users in the photo. This type of user behaviour is not explicitly defined by the tools and widgets available, and as a result the behaviour is implicit and so has to be inferred by other users so they can create a shared understanding around the behaviour of the other user.

Although a user can be autonomous with their actions within an **OSN**, like all **OCE** they have a context for collaboration even if the user does not adhere to it. This is even so with the largest **OSN** Facebook. The over arching context of this **OCE** is to create *collective social action* [44]. Although a user may form their own collaborative goals or objectives within the environment, the environment's collaborative context stays the same. This is because the tools and widgets available within the environment form the foundations of the collaborative context through mediating collaboration. An environments context is important and needs to be taken into consideration when designing an environment, however one must know and expect the user could have another collaborative context in mind [53, 54].

OCE have now become pervasive in terms of access. Users are able to access the same environment through different interfaces to continue their participation. One particular growth area for these environments is personal devices [55]. They accompany the traditional web-application **OCE**, and allow the user to participate in an *always connected* fashion [56]. However, a user's participation through these

devices can be restrictive as these personal devices tend to have a limited view port. One such example of this is GitHub’s mobile phone application [57]. Compared to GitHub’s traditional desktop web client, a user has to adopt a general behaviour of consume and browse media instead of actively contribute to a code base they may belong to as the collaborative process of maintaining code would be difficult through such a small view port.

The tools and widgets which are available to users within OCE are crucial to the collaborative process. They promote awareness or creation of media by users for users to connect, share, modify or interact with. An example of this is an *activity timeline* tool that allows a user to navigate a chronological list of different actions or media [58, 59]. These actions and media tend to be a result of the user’s interaction with generic tools used within an environment, such as content authoring [60]. Facebook, the largest social network in the world, provides such a tool which they refer to as a user’s *Facebook News Feed* represented by Fig 2.4. This feed allows users to engage with other users and their media, updates, and activity to discover the user’s participation within the environment over a period of time [61].

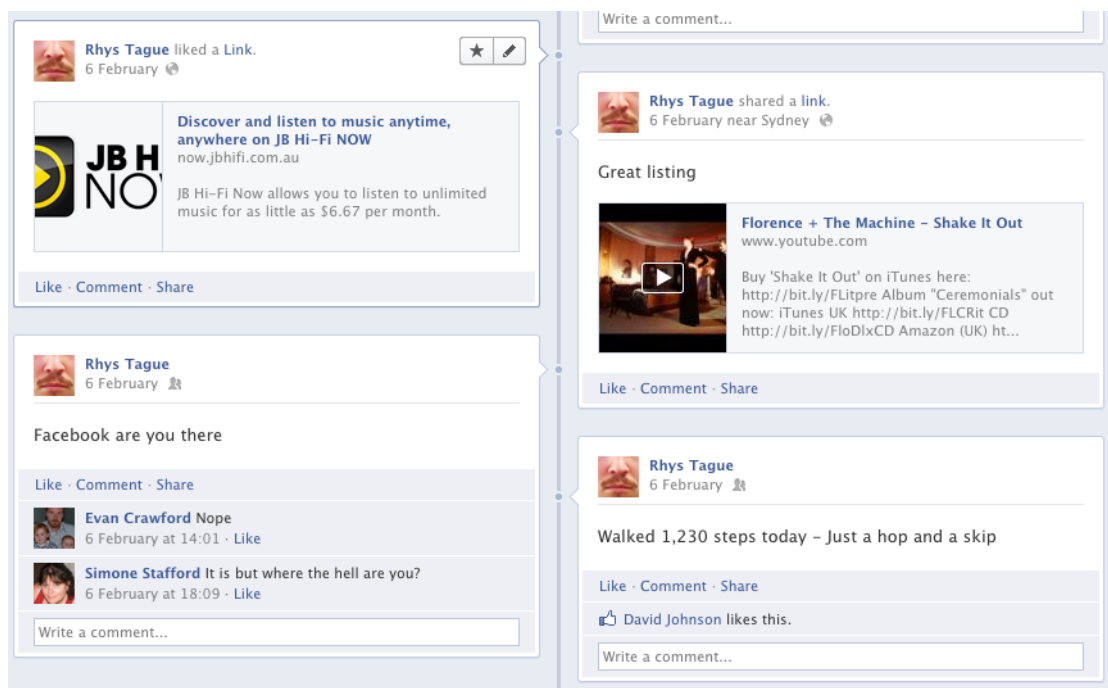


FIGURE 2.4: A screen capture of the Facebook News Feed

OCE can be broad and narrow with in regards to the collaborative process. As explained, there are those that have roles and purpose explicitly defined, and others providing a skeletal functionality for users to define their own roles and

goals or objectives for collaboration, resulting in an implicit structure. For this study, the latter form of environments will be used. Environments which do not have explicit collaborative structure, yet provide tools for users to create one.

2.5 User Modelling

Users can now exist in [OCE](#) that are implicit in nature - ones where users define the goals, purpose and participation structure. As a result the one size fits all when it comes to interaction is limited and ineffective [62]. Therefore, knowing a user's behaviour allows a developer, designer, or researcher to understand and predict the take up and use of an environment at an individual, group, or population level. With such an understanding better decisions or interaction design for user participation, experience and personalisation can be achieved [63]. The collection and storage of a user's behaviour is called user modelling.

Today, the type of the information that can be collected about a user can be diverse, complex or simple. However, this information is useless if some sort of understanding does not exist for the purpose of collection. For instance, if an environment collects a user's age and their browsing history, yet no analysis is carried out after collection what was the purpose of collection? Regardless of this, the more information one collects about a user the more defined a user model can be, and therefore the analysis can be second to collection. According to Kobsa though [64], a user model should consist of the following:

- Assumptions about one or more types of user characteristics, such as their knowledge, misconceptions, goals, plans, preference, tasks, and abilities.
- Representation of relevant common characteristics of users pertaining to specific user subgroups of the application system (so-called stereotypes).
- Classification of users belonging to one or more of these subgroups, and the integration of the typical characteristics of these sub groups into the current individual user model.
- Recording of users' behaviour, particularly their past interaction with the system.
- Formation of assumptions about the user based on the interaction history.

- Generalisation of the interaction histories of many users into stereotypes.
- Drawing of additional assumptions about the current user based on initial ones.
- Consistency maintenance in the user model
- Provision of the current assumptions about the user, as well as justifications for these assumptions

Although Kobsa states a user model should consist of the above items, it is also known that not all of them are present in user models today. Even if only a few items listed are present in a user model it can still result in a successful model of user behaviour as the items collected tend to be used for specific functions.

As the structure and characteristics of a user model is based on its purpose, the type of information to collect should be thought out before implementation, however this is not always performed in a such a way. For instance, the function of tagging media within an environment is a task a user can carry out as an activity. Only after the tagging by users has established does one try and build a user model to understand a user's behaviour or interests from the tags used [65]. This is usually a result of not knowing what the user will do, then waiting for data to be collected before trying to understand what the user did to make informed decisions or personalisation of content through recommender systems [66]. This dilemma is so common that the term *cold start* is given to the scenario of a user model with no data or structure.

The *cold start* dilemma is sometimes stemmed by using a technique called *collaborative filtering* [67]. A method to harness existing user models with data and structure to make a prediction of user behaviour based on the current state of a user without a user model. This method is used heavily when it comes to content consumption and recommendation. Amazon, one of the largest eCommerce sites in the world, uses collaborative filtering for recommending products to currently browsing users, based on historical purchases and browsing habits of other users. [68, 69].

Cold starting is not the only dilemma with user modelling. With the capability of capturing vast quantities of user behaviour the size of the model becomes a performance penalty for a system. With such large amounts of data to process a

system can appear sluggish and problematic as the system is trying to understand and predict the user's behaviour. Yet, without this information about a user the probability of successfully understanding and predicting behaviour of a user lowers. Therefore, one may forgo a user model if it becomes a bottleneck to a system. Others may focus more on the ability to exclude historical user behaviour from the processing and focus more on recent behaviour. Luz *et al* [70] present a method for creating a balance between performance and successful user understanding and predicting, creating a lightweight user model for recommendation and feedback to the user. By focusing on an asynchronous method of processing through steps. This allows the system to comply with the user's actions concurrently with the processing of the user model.

The word behaviour is given to almost all user model elements as interests, beliefs, goals, and interactions can be categorised as a result of a user's behaviour [62, 64]. What a user model collects depends on its purpose which may change over its life. Therefore, the developer of the user model needs to maintain it so it is fit for purpose. For this reason the user model tends to be an extension to the system using it instead of it being part of the system operation. Some work suggest a user model should be a service to a system instead of a module or component [62, 71]. With such a service system can upload user behaviour to a remote system, and then be provided with results or recommendations. Such a system enables roaming user profiles that can be used for users who belong to multiple environments.

Although the term user behaviour suggests all behaviour is a result of the users actions this can be misleading. For instance, within OCE, in particular OSN, users work together and establish behaviours towards one another and as a result users establish a group behaviour within an environment that is not individualistic [72]. For instance, if a user is wanting to collaborate they may use a particular tool provided within an environment as a form of communication, even though the chosen tool is not the explicit tool for communication defined by developers. A group of users may establish a shared understanding that this tool is the main form of communication within an environment. As a result a user model may be distorted towards a particular action because it is seen as an individualistic characteristic to the user yet is an implicit collaborative behaviour. This type of behaviour is commonly known as *intentional social action* or social norm, where users establish *we-intention* [73] - a group based intention instead of an individualistic one. There has been work on providing theoretical models of social intentional action

within an online social network [74] and other collaborative tools [75], however this area is still in its infancy in terms of formal user modelling techniques to discover this *we-intention* behaviour. Instead, to accommodate this group based behaviour there has been some recent work with combining user models together to create a group model [76, 77]. By doing this one can create an understanding and even recommendation based on group preferences. However, this is very recent work, and is yet to mature.

2.5.1 Understanding and Predicting User Behaviour

How a user model is processed is dependent on the purpose of the user model, much like the reason for behaviour collection. Simplistic statistical methods are popular to understand volumes of certain behaviours, yet understanding and predicting why a user has a particular behaviour requires more complex methods.

Simplistic methods tend to result in analytical data, such as number of posts or tags towards media [78, 79]. This type of data is good for understanding growth of an environment in regards to the production of media and growth in a user base size, however limited in terms of why its user base is growing. Analytical data is also good to track impressions, such as media creation or consumption towards media or user profiles. For instance, understanding social cascades within an environment can only be done through an analytical approach [80]. Having such an insight for an environment, in the sense of media production or consumption, allows one to determine the health of a user base and understand user attrition [79].

When one wants to understand and even predict user behaviour simplistic analytical methods are not sufficient. This is because the drive why a user does something is not always clear. To understand behaviour one needs to know why or how a user produces media. Once known, one can personalise, design, develop a users environment or even influence a user [81–84]. For example Hauser *et al* [83] provide a method for morphing a user’s interface to align with the user’s behaviour and cognitive style. They achieve this through Markov model decision making with click stream data the user has created by interacting and navigating the site. Knowing a user’s behaviour or cognitive style not only benefits the user, but also the developer/owners of the site. By aligning a website with a user’s behaviour or cognitive style one can increase conversion rates which in turn leads

to larger profit margins. This is seen extensively with site content for eCommerce sites or commercial focused services [84]. Commonly called content filtering, content is presented to users possessing particular traits discovered by processing a user model for a specific purpose, and therefore unveil user behaviour, such as relational behaviour, amongst content. With simple analytical methods this type of behaviour is hard to find.

The methods to achieve this understanding and predicting is by using soft-computing, or more commonly known as machine learning, techniques [85]. There are many techniques and each have their advantages and disadvantages. Some of the more common techniques for user model processing are: Fuzzy Logic (FL) [86, 87], Neural Network (NN) [82, 88], Neuro-Fuzzy System (NFS) [89–91], Markov Chain Models [92], and Bayesian Methods [83, 93]. Traditional statistical methods are also used for data mining amongst user models, such as clustering algorithms [94]. As every method differs and has a extensive literature background, only the relevant techniques to this study will be presented.

Fuzzy Logic, even though not a formal machine learning technique, is used often to derive a level of certainty from uncertainty [85]. Instead of a binary result of 0 or 1, traditional of logical statements, Fuzzy Logic can create a fuzzy set of values between 0 and 1 and then infer a result from this set. This results in *degrees of truth* very much like how our mind concludes decisions. For instance, the concept of tall is ambiguous and one can not state someone is tall as a fact unless including all degrees of truths resulting in absolute truth, or 1 in logical terms. With Fuzzy Logic one could say someone is 0.62 of tallness giving the degree of truth and it would be absolute truth [95]. This degree of truth is useful to user modelling as one can provide a threshold towards the degree of truth to classify users. One limitation of Fuzzy Logic is that it can not learn from data, as a result it is often used for user modelling for recent actions within an environment. Such an example is shown in the work by Henry *et al* [96] by using Fuzzy Logic to adapt product recommendations within an environment to align with recent user behaviour towards various products. The aim is to adapt an environments interface with little predictable user behaviour.

Sharing traits of how the brain processes information, a Neural Network is a common form for processing user models. Their ability to be trained and retrained to create a form of experience allows a Neural Network to identify patterns within data. This pattern recognition makes them popular in user modelling as users can

be classified and or grouped together. There are many types of Neural Network and some use supervised and unsupervised learning methods. Chou *et al* [88] demonstrate the use of a supervised learning Neural Network by using a back-propagation Neural Network to classify navigation behaviour from click-stream data. The use of the Neural Network allowed for better understanding, recommendation and prediction of site content for classified users. Although they are popular with user modelling, one of the dilemmas with discerning results from a Neural Network is how the results are interpreted [85]. To overcome this other methods are used alongside Neural Network, one popular method is Fuzzy Inference giving the method the name Neuro-Fuzzy System.

A Neuro-Fuzzy System uses the power of a Neural Network, and its ability to learn from experience (historical data), in combination of Fuzzy Logic to infer a definitive result from the Neural Network output [97]. This method overcomes the common limitation of black-box processing which tend to be found with Neural Networks, and also the limitation of finding suitable classification values for Fuzzy Logic [85]. Because of these characteristics it provides a powerful means of understanding and predicting user behaviour from a user model. Stathacopoulou *et al* [98] present this behaviour with their use of a Neural-Fuzzy System for diagnostic capabilities towards student learning style within an online collaborative learning environment. The Neural Network provided the generalisation capability, while the fuzzy logic used the generalised results from the neural network to infer the student's learning style.

The choice of method for processing a user model is a result of the user model used while also considering the type of output desired. As a Neuro-Fuzzy System has the ability learn from user data, while also inferring a definitive classification, this method has been chosen for this study.

2.6 Identification and Analysis of Roles Online

The theory of social roles has already been discussed in 2.3.2, yet the identification and analysis of such roles online is yet to be explored. The reason to identify and analyse roles offline is the same online as shown in the work by Wu *et al* [99]. By knowing a user's role one can interact within context with another user based on their known role. However, a user's role is not always explicit and therefore the

context of interaction towards a user or interface is not always known. This is one of the reasons why it is important to understand a user's role. For instance, in a collaborative context understanding a user's role explicitly or implicitly could help provide resources and/or tasks to users who carry out that action needed in the collaborative process. Work done by Yang *et al* [100] suggests this with their work on identifying an 'editor' role within Wikipedia. An editor is not known to the user base as each user can operate in an autonomous nature, as a result Yang *et al* create a specific *role signature* and use this signature to identify users who show traits of an editor. Knowing this they hope to provide an environment where tools or tasks can be routed to the correct users instead of being discovered by those users. With such an ability, it could possibly make those users more productive.

Identifying and analysing roles is not always for the purpose of interaction personalisation. It is quite often used to simply understand the structure of a user base or an environment [101] especially in the social context [51]. Knowing the structure of an environment allows a researcher or a developer to understand why there may be certain attrition rates within groups or an environment for instance, or why certain users interact with particular users [102]. This desire for understanding is the fundamental drive of sociology and other social science fields.

The type of roles identified online depend on the environment and the reason to identify the role. Roles may be functional [103, 104], while others are purely for social role structure understanding [105]. Functional roles can be thought of as roles that relate to the functional activities of an online community. This is similar to functional departments within an organisation, such as finance and marketing [103]. These roles have activity signatures that allows a user to constitute a role based on the purpose of those activities. For instance, a user may have a functional role of an editor or contributor. An editor would have activities relating to reviewing and updating content, while a contributor would have activities relating to publishing and media creation. When comparing traditional organisations with functional departments, online functional roles are not always known and sometimes are a result of analysis after continual use instead of planning and design. As a result identification methods are used to discover or create a role signature [100].

Approaches to identification relate to user modelling where information about a user's behaviour is collected and analysed for patterns as discussed in section 2.5. However, to analyse one user and then use one user's behaviour as a role would be

of no use as isolated behaviour present in one user may not be present in another. Therefore, repeated behaviour for one user, or behaviour shared amongst many users, tends to be used to identify roles [106]. The context of identification is also needed to understand and help classify the repeated behaviour. This can be simple behaviour such as posting or commenting to a certain Blogosphere repeatedly, enabling one to establish a role of *Standard Bloggers* for instance [101]. The complexity of context can extend beyond user behaviour and also be a result of social structure and psychological attributes [6]. Psychological attributes, although not used often and outside the scope of this study, tend to be found through content analysis, where social structure is found through [Social Network Analysis \(SNA\)](#) methods. Regardless of the behaviour the complexity of the role online is created from it's context.

Creating, discovering, and analysing social or network structure is the most common form used to establish a context for a role signature [6, 7, 100, 102]. The term *structural signature* is also used throughout the literature [107], however for consistency in this study the term role signature will be used to represent both unique behaviour and social or network structure. As roles exist in interconnected environments, such as [OCE](#), [SNA](#) is typically the main tool used to establish structure within an environment [108, 109]. [SNA](#) examines both content and semantics of a network-based structure. There are many approaches one can take, yet the approach is dependent on the purpose of analysis or understanding [109]. For instance, if one wanted to use a simple analytic approach they could establish structure within an environment based on information diffusion [110]. Through analysis of this diffusion patterns could be established and then used to form the basis of a role signature. However if a role signature is established in regards to behaviour, and social or network structure, it must be unique yet present for one user or many over time.

Once roles signatures or patterns of behaviour are identified within an environment the role naming convention used are either behaviour related, original or literature cited names decided by the researcher or the identifier of the signatures or patterns. For instance, work by Buntain [111] *et al* use a social network called *reddit* and its network structure to identify the *answer-person* role. This is a role which is displayed by users who continually answer questions, or responds to media rather than create or publish new media. As a result the *answer-person* is a role represented by behaviour, however this particular role is also present in

literature from where the term was a result of data visualisation of interactive behaviour of users [107]. Forming original role names is common and are the result of purpose or environment specific environments [104]. Role names tend to relate to a cohort of users who display some relation in terms of behaviour within an online environment.

2.7 Summary of Findings

Although there are many methods and approaches to identifying roles online none could be found which adopt the *Group Socialisation Model*. This might be a result of the model being a life cycle, which is sequential, and the ability to map this life cycle to an OCE being too complex. OCE are dynamic, flexible and asynchronous in nature and therefore not sequential like a life cycle. It could also be because the definition of collaboration is broad and the ability of a user to be individualistic and autonomous with their action, yet still be collaborative, hard to evaluate and derive a characteristic behaviour in a real world collaborative scenario. Another possible reason could be because the social group role life cycle does not identify roles. It uses the understanding that social roles are *implicit* and that they are established by role signatures formed by characteristic behaviours agreed upon by other users. Therefore, the dynamic and fluid nature of role signatures being able to grow and form based on user action too hard to foresee or identify with out explicit structures or approaches.

The rigidity of approaches for forming roles or role signatures was strong throughout the literature. This goes against the nature of collaboration, as collaboration can sometimes be referred to as a *phenomenon* within the literature because it is so fluid and dynamic. It also highlights these approaches and their incompatibilities with real world applications for OCE. This can be seen with many approaches that employ a structural or analysis method to forming roles. Approaches that predetermine a characteristic behaviour with rules, activities or features of an OCE. After which data mining is used to determine the strength of a behaviour. This goes against the literature as it presents characteristic behaviour as an output of collaboration based on ones collaborative effort. Therefore, how can one predetermine the rules, activities or features a user will align with if it is not known until collaboration occurs. This could be the reason why many approaches harvest data to analyse with their approach instead of real world applications.

Although many approaches are rigid, the approaches that use generalisation methods, such as collaborative filtering, were able to provide a level of dynamic characterisation. Although, their focus was more on characterisation of the user base instead of the collaborative context of individual users. This is limiting as different [OCE](#) allow individuals to have their own purpose for and process to collaborate. Therefore, a user base wide classification may not be beneficial for more focused collaboration and meet the three domains of collaboration: *people*, *purpose* and *processing*.

This literature review has identified a gap for automating the formation of *implicit* roles within dynamic, flexible and asynchronous environments such as [OCE](#). More specifically, there is no framework that uses characteristic behaviours in conjunction with the social group role life cycle defined by the Group Socialisation Model to provide an automated way to produce role signatures within dynamic and asynchronous [OCE](#).

Chapter 3

Framework Model - Construct

3.1 Introduction

For this study a framework has been designed and created to form a construct for the purpose of knowledge creation. The framework uses a combination of common and innovative practices while also using domain specific theories to solve the problem of modelling social group role signatures for [OCE](#). By creating a framework that is able to produce dynamic and adaptive social group role based signatures, one can identify, characterise and possibly predict a user's collaborative behaviour within an [OCE](#).

This chapter presents a visual guide to the framework model through its life cycle, its components and its theoretical simulation of social group role formation. The implementation of the framework is also discussed within this chapter, which will lead into the following chapter discussing the case study environments where the framework was implemented.

This chapter is sectioned as follows:

- Framework Life Cycle
- Online Collaborative Environment
- Functional and Characteristic Behaviours
- Visual Representation of Implemented Framework

- Widget/Tools Metadata
- Interaction Packages
- Framework Domain
- Framework Implementation Model

3.2 Framework Life Cycle

At the heart of online collaboration is autonomy towards establishing a shared understanding. Users are able to operate independently and in an autonomous nature of each other, yet still share a common understanding towards a collaborative effort. As a result, any framework that exists in this environment also needs to be autonomous. It needs to be capable of changing and accommodating new behaviour, user to user collaboration, and new interactions that may be implemented within an [OCE](#). In addition, it also needs to identify different collaborative contexts that may exist. In the offline world, social group role theory already provides a theoretical model for this type of framework as discussed in [2.3.2](#). However, after the extensive literature review no framework for [OCE](#) have been identified where a user's characteristic behaviours, provided by Biddle [39], have been combined with Morelands Group Socialisation [8]. This results in the following framework being a first of its kind and will set foundational work for providing a dynamic and autonomous framework for characterising social group roles through user characteristic behaviours.

The theoretical model provided by Moreland and Biddle has already been discussed at length in [2.3.2](#). In that section figure [2.2](#) shows there is a life cycle to social group roles. Where an initial characteristic behaviour demonstrated by an individual starts the formation of a social group role, and then others use this initial behaviour to commit to person's social group role until a change is seen after evaluation. This can all be seen in Figure [3.1](#) where the the framework replicates the phases of the social group role life cycle. For detail on how this is implemented figure [3.2](#) presents the components needed to achieve the life cycle in an automated and asynchronous manner to work within an [OCE](#).

The life cycle is not an overarching life cycle for the entire user base. Instead, the framework will carry out this life cycle for each individual user. This allows the

framework to accommodate new users as they join an [OCE](#). It is also in line with how humans naturally carry out the social group role life cycle by carrying out this process for each individual within an environment or a group.

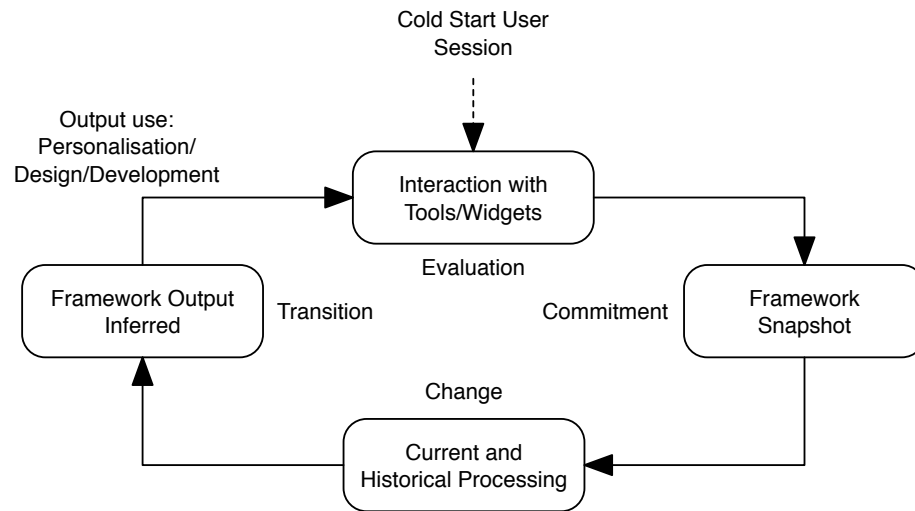


FIGURE 3.1: Framework Life Cycle represented as the social group role life cycle

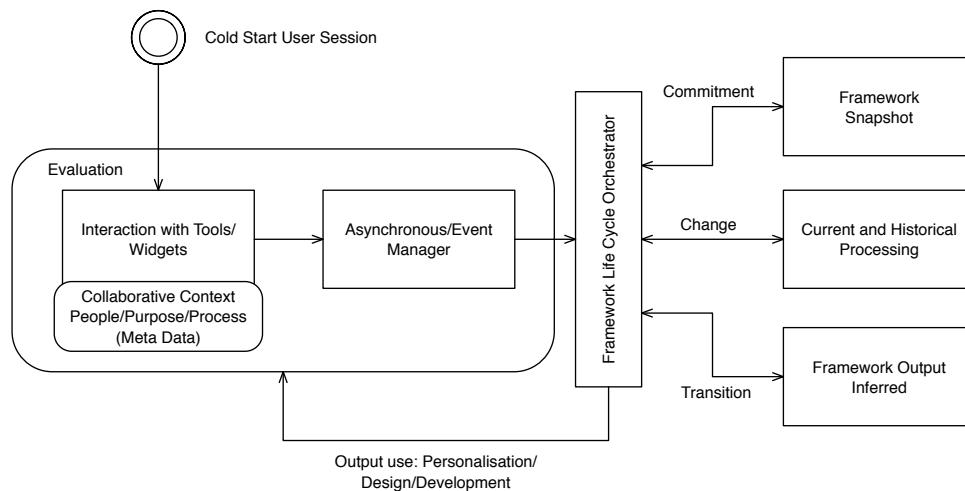


FIGURE 3.2: Proposed framework emulating the social group role life cycle

To show the framework's emulation of the social group role model, the life cycle will be explained through the different phases found within this model. To fully understand this life cycle one should adopt the mindset of the framework taking the role of an individual within a group setting amongst one or more individuals.

3.2.1 Evaluation

When comparing to the social group role model, the evaluation phase is a process of an individual monitoring or evaluating one or more individual's behaviour. The framework achieves this by monitoring and evaluating the interaction of the user within an [OCE](#). For the life cycle to start an initial behaviour needs to be established. In the environment where the framework uses this is through a cold start user session. A cold start user session is a user with no previous record of existence within the framework. Like all user sessions, the user goes through the evaluation phase where they interact with the environment through tools and widgets that have semantic data stored in metadata describing the collaborative context. The metadata possess the social tie, the tool or widget identification, and the characteristic behaviours that will be a result of interacting with the tool or widget. For every interaction an interaction package is created, cached within the environment and passed to the Asynchronous/Event Manager. Once at the Asynchronous/Event Manager the resulting interaction package is passed to the framework life cycle orchestrator to be moved to the next phase of the life cycle.

3.2.2 Commitment

The commitment phase is where, based on evaluation, an individual commits towards the evaluated behaviour. This commitment phase offline is in the form of an implicit understanding of an individual's actions towards a task, a group, or an individual. This understanding is of course subjective for the individual carrying out this process. The framework does not have the capability to be subjective, instead it uses machine learning. During this phase the framework takes the user session and commits it to its experience database. This commitment involves storing original behaviour in the form of original interaction packages, and training the machine learning system, a back-propagation neural network, with new behaviour. During this phase the framework also timestamps and creates a

snapshot of its current structure. This snapshot is then used during the change phase.

3.2.3 Change

Once committed, an individual will hold that commitment until there is a change shown in an individual's characteristic behaviours. This change is discovered by an individual comparing their commitment towards another individual until they believe there is a substantial change in characteristic behaviours to warrant a transition. As the change is implicit this is highly subjective and dependent on the individual. One could argue it is another form of evaluation, however this part of the model is more aligned with comparison than evaluation, comparing current behaviour with historical behaviour. The framework achieves this phase by processing snapshots taken of the experience during the commitment stage. The difference between these snapshots can vary depending on the last user session. For instance one user may log in every day, and another may only log in every week or longer. Because of this the difference between the two snapshots can be substantial or negligible. Once multiple snapshots have been processed the result can then be passed onto the transition phase of the framework.

3.2.4 Transition

The transition phase within the social group role model is where an individual will use the change, or difference between historical and current characteristic behaviours, to decide to either transition their own behaviours towards an individual or to go through the evaluation phase again to commit to a new social group role for another user. Within the framework this phase is the use of the result sets processed in the change phase. Fuzzy rules are used to compare the result sets to infer an output of characteristic behaviours for a particular user over a period of time, keeping the transition phase more stable instead of erratic which can be the result when using unique user sessions.

How an individual handles the transition phase offline is subjective, however within the framework this can be objective. One can use the output for personalisation,

design and development considerations, resource allocation or simply understanding social structure make up within an environment. The output is in the form of a role signature based on the characteristic behaviours found within a [OCE](#).

Once the transition phase is over, the life cycle starts again with a new user session.

3.3 Online Collaborative Environment

Although an [OCE](#) is not part of the framework it is however what the framework interfaces with and is used to demonstrate real world application of the implement research construct. For this reason, the [OCE](#) needs to possess certain attributes for the framework to execute its life cycle. These attributes must relate to the three domains of collaboration - people, purpose and processing. One can assume if an environment is an [OCE](#) then it will have these attributes for the necessity of successful collaboration. How the construct receives these three attributes is not a consideration for the framework construct but more an implementation issue. In section [3.5.2](#), the implementation of the framework with the use of tools and widgets, this is achieved through metadata where the three domains of collaboration form the metadata of the tools and widgets within the environment.

3.4 Functional and Characteristic Behaviours

Functional and characteristic behaviours are used to determine the role of an individual as they form a major part of a role signature established by a group [[9](#), [39](#)]. The possibilities for the different types of behaviours an individual can possess is endless, yet online these behaviours are restricted by the functions and interfaces a user has access to. These functions and interfaces are a result of [CRUD](#) and processes within an environment. Although the interfaces may be limited based on function and process, the semantic representation of behaviour is not.

The semantic representation of functional or characteristic behaviours is a result of the [OCE](#). For instance, for a social coding environment the behaviour *commit* is an update command, yet the term commit is a semantic meaning for the updating of a code repository. This semantic representation of an [OCE](#) function and interaction

creates environment dependent characteristic behaviours. For a framework to be able to accommodate this semantic representation of behaviour it must be able to dynamically adopt new behaviours to truly represent a role signature. However, as semantics is a result of the environment's function or purpose the ability to identify a semantic behaviour relevant for a possible role signature is not precise. An identified behaviour may be unique to one particular function or interaction that occurs at irregular intervals and therefore not significant enough to form part of a role signature.

The framework created has the capability to dynamically adopt new functional or characteristic behaviours over time, and this will be explained in further detail in how the framework processes the commitment phase (section 3.2.2). For this study, stereotypical behaviour was chosen to represent generic characteristic behaviours. By using generic behaviour one can demonstrate a generic output for various OCE. The stereotypical behaviours used for this study are inspired by literature for user behaviour, and social network usage patterns [60]. Although this study uses stereotypical behaviour, if an environment requires a more complex and diverse range of characteristic behaviours this can be achieved in the framework by providing those behaviours in the tool/widget metadata. When a user demonstrates this behaviour the framework will accommodate such new behaviour and will present it within its output.

The stereotypical behaviours that will represent characteristic behaviours of a role signature are: *publisher*, *annotator*, and *lurker*. Each of these are described below.

3.4.1 Publisher

Within CRUD the first function is Create. Although creation represents the creation of *any* new data, many tools within an environment allows users to publish, *create*, new media for consumption by other users. For instance micro blogging allows an individual to publish small broadcasts of media for other users to interact with and annotate [112]. Therefore, any interaction a user carries out that publishes *new* media for other users within the OCE will be identified as the characteristic behaviour ***publisher***.

3.4.2 Annotator

Where the publisher behaviour creates new media for others, an annotator **Updates** or publishes in relation to pre-existing media. For instance, those who reply to published media via comments is publishing new media yet in relation to pre-existing media [46, 101]. Comments are not the only form of annotation. Users are able to *Like* [113, 114], edit [7, 18], or tag [52] to change or add meaning to published media. Therefore, interaction where a user publishes in relation to pre-existing media will be identified as the characteristic behaviour ***annotator***.

3.4.3 Lurker

The term Lurker is commonly represented as a role, a set of behaviours [115, 116]. If one analyses these behaviours the action of a lurker is consuming media or **Reading** media. Consumption of media is the most common form of behaviour shown online [117], and therefore the term Lurker is already a stereotypical behaviour.

With the common use of the term and how it is the result of any interaction where a user doesn't not publish or annotate, yet consumes media through browsing or interacting, the final characteristic behaviour identified as ***lurker***.

3.5 Implemented Framework

3.5.1 Visual Representation of Implemented Framework

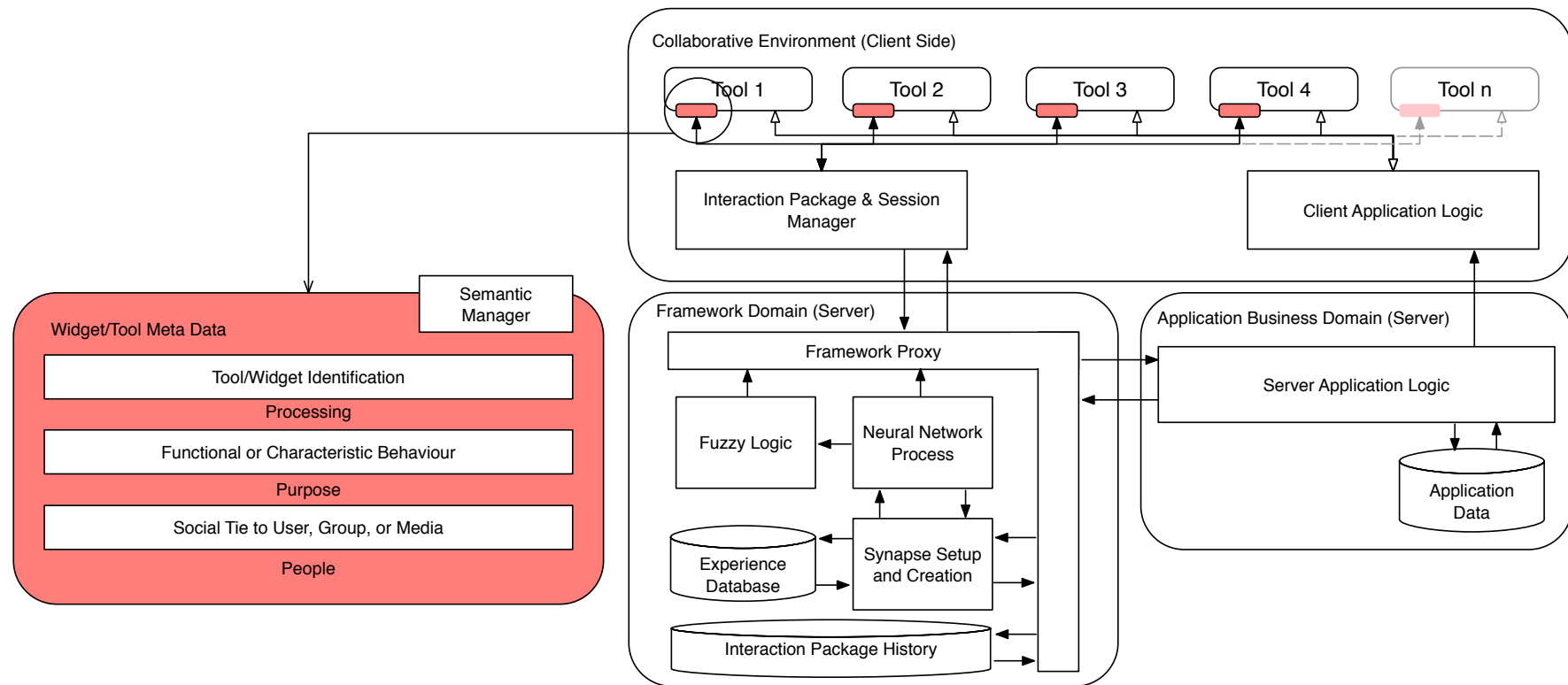


FIGURE 3.3: Visual Representation of Framework

Figure 3.3 provides an overall representation of the implemented research framework. From this figure one can see the framework is modular in design with a proxy, present in the Framework Domain, allowing for interaction between modules. Because of this modular design the framework is loosely coupled to the business logic of an OCE. This enables the framework to serve many environments through a service based architecture. The framework proxy handles the requests sent to the framework to orchestrate the framework life cycle so the OCE does not need to be aware of its implementation.

Although the Framework Domain has a proxy to allow OCE to interface with it, there still needs to be an Interaction Package and User Session Manager (IPUSM) present on the client side of the OCE. This provides the medium for maintaining and monitoring semantic meta data about a user's behaviour. Even though this manager may be on the client side, it is also loosely coupled by design allowing it to be installed without modifying the widget/tools provided by the OCE. This loose coupling is achieved through the use of an event based architecture allowing the manager to listen and not interrupt the normal execution of the environment.

The following is a detailed breakdown of the client and server modules to provide how the implemented framework achieves the social group role life cycle asynchronously.

3.5.2 Widget/Tools Metadata

An OCE will have many tools and widgets, and each one of these need semantic data. This semantic data is needed by the framework to understand the context of interaction performed with the tools or widget. The semantic manager is responsible for this data by creating metadata that is assigned to a tool. This metadata stores descriptors of context for the collaborative action undertaken within an environment. Referring to figure 3.4 one can see the descriptors held by this metadata comprise of the three domains of collaboration, people, purpose and processing. These three domains are made up of three elements of the environment: the tool/widget unique identifier, characteristic behaviours, and social tie.

The tool/widget unique identifier is assigned by the developer during development. This identifier is not something the developer would need to specifically assign for

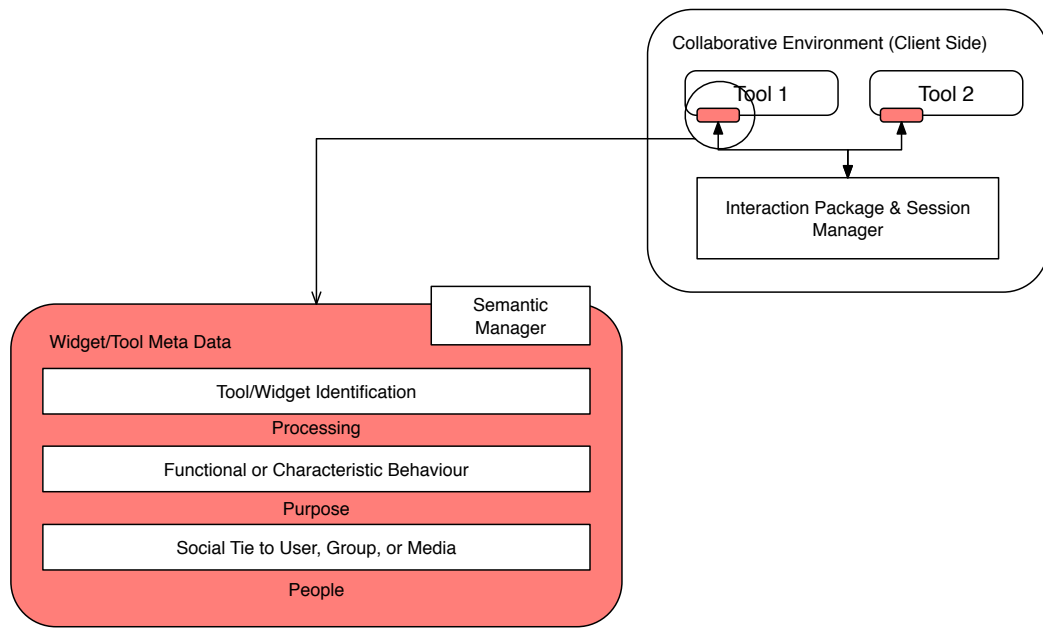


FIGURE 3.4: Widget/Tools metadata segment of framework

use with the framework, as it would already exist in some capacity for the needs of the environment's business logic. If one was wanting to traverse interaction packages cached by the framework thought should be given to this identifier so that it is descriptive in some capacity and not just for business logic. A string representation, such as "*social-post-tool*" would be a good identifier. In addition to it being descriptive, the identifier needs to be unique to prevent the package manager creating interaction packages for multiple tools yet represented as one.

Although particular characteristic behaviours were identified in the last section, this can be any semantic behaviour the tool or widget is capable of. As some tools and widgets are able to carry out multiple behaviours they must be listed within this metadata. If there are multiple characteristic behaviours for a tool or widget the tool or widget must be able to provide some form of identification, by firing events, when an action has been completed. As a result, this may lead to complex rules to establish which behaviour has been carried out, however this is at the discretion of the framework implementer and knowing when events are fired to determine the behaviour. The rules used for the case studies will be presented in the case studies chapter.

The final metadata item is the social tie. As the environment is an [OCE](#) any interaction within the environment *should* be resulting in a collaborative effort. If a user uses a tool in isolation and it is not contributing to a collaborative action,

it is deemed a personal tool or widget and therefore does not have a social tie. A social tie is any connection that can form within an environment between two nodes. A node is a user profile to user profile, or user profile to media item. As a social tie can be a result of interaction, such as publication of new media, it is the tools responsibility to fulfil this attribute at some point during interaction. This is common for almost all tools as the framework needs to know who or what the user is interacting with.

The representation of this metadata can be in many forms. A structured representation is preferable, such as [EXtensible Markup Language \(XML\)](#), as it provides a markup to identify each attribute when processing. The representation is created and managed by the semantic manager. Once the semantic manager creates the metadata for the tool, it will provide an event upon which the interaction package manager will listen to so it can access and prepare the semantic data for the interaction package that will be sent to the framework.

Listing 3.1 presents an example metadata representation in [XML](#). From this one can see the characteristic behaviours the tool can perform are in the behaviours tag, the tool identifier within the tool tag, and the social-tie in the tie tag. The social-tie is established by using the user's profile identifier, and the node identifier within the environment concatenated together. For instance, a user identifier of 289 and node identifier of a media item 3243 would create the social-tie 'u289-m3243'. The character in front of each identifier is chosen based on the node type. For this example, as the first identifier is a user profile the character 'u' was chosen. For the second, 'm' was chosen for media. The choice of character is a design choice, however a character should be used to ensure identifiers are unique.

The semantic manager is not implemented into the tools or widgets, it resides on the client side of the [OCE](#). It is an event based manager that listens to when a user or the system completes an action within the environment. These actions, if not explicitly defined by the developer for the framework, are part of the normal operation of the tool or widget. Listening to these actions enables the semantic manager to be decoupled from the [OCE](#), yet able to work within the framework.

The semantic manager is the first event based component of the framework. The use of event based modules within the framework removes the burden of the framework processing from the business logic. Processing can be carried out by different processes or remotely.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <data>
3   <behaviours>
4     <behaviour>publisher</behaviour>
5     <behaviour>annotator</behaviour>
6     <behaviour>lurker</behaviour>
7   </behaviours>
8   <tool>wall-post</tool>
9   <tie>u289-m3243</tie>
10 </data>
```

LISTING 3.1: Metadata Representation

3.5.3 Interaction Packages

The framework does not use traditional interaction outputs, such as clicks, or navigation logs. Instead, the framework uses the metadata provided by the semantic manager to create interaction packages. By using interaction packages instead of crude interaction outputs one can analyse an interaction within context in real time. The creation of interaction packages is carried out by the [IPUSM](#) which establishes a collaboration context with the user's purpose, processing, and people (social tie).

The [IPUSM](#) is like the semantic manager, event based manager. [IPUSM](#) role is to source and process interactions when an event occurs. It resides on the client side of the [OCE](#) and is decoupled from the business logic of the environment. It is the interface between the client and server side components of the framework. It achieves this by using an [Application Programming Interface \(API\)](#) for passing user sessions and interaction packages to the server for instantiating the framework's life cycle.

The interaction package created by the [IPUSM](#) consists of available characteristic behaviours, social tie, tool or widget identification, characteristic behaviour demonstrated, current user, and a timestamp when the interaction was carried out. The package is represented in [Javascript Object Notation \(JSON\)](#) which allows a string based representation of an object instance to be easily accessible and transcoded on the server.

Listing 3.2 presents an example of an JSON interaction package. Key/value pairs are used to represent different variables within the interaction package. Table 3.1 is a data dictionary for the various keys and values found in Listing 3.2.

```

1 // JSON
2 {
3     "behaviours": ["publisher", "annotator", "lurker"],
4     "tool": "wall-post",
5     "key": "publisher",
6     "user": 3243,
7     "tie": "u289-m3243",
8     "timestamp": 1444605840
9 }

```

LISTING 3.2: Interaction Package Example

As the client and server side of the framework are independent of each other, the IPUSM caches (stores on the client locally) until successful interfacing with the server is established. This is needed as many OCE rely on an Internet connection to provide functionality.

Key	Description
<i>behaviours</i>	Behaviours represent the characteristic behaviours available by the tool or widget where the user has carried out an interaction
<i>tool</i>	Is the unique identifier of the tool or widget
<i>key</i>	The characteristic behaviour demonstrated by the user during interaction according to framework implementer.
<i>user</i>	The unique user identifier
<i>tie</i>	The concatenating of a node identifier (user, group, or media) and the user identifier forming a unique social tie identifier.
<i>timestamp</i>	The numbers of seconds since the epoch of the system. This number can be converted to a date and time representation.

TABLE 3.1: Data Dictionary for an Interaction Package

3.5.4 Framework Domain

The framework exists on its own domain as the framework does not need to be part of the OCE domain. The framework is capable of being remote to the OCE and therefore act as a service if needed. This is useful when wanting to process a pre-existing OCE to map the role signatures of the user base at a later point in time or to use multiple OCE with one framework implementation.

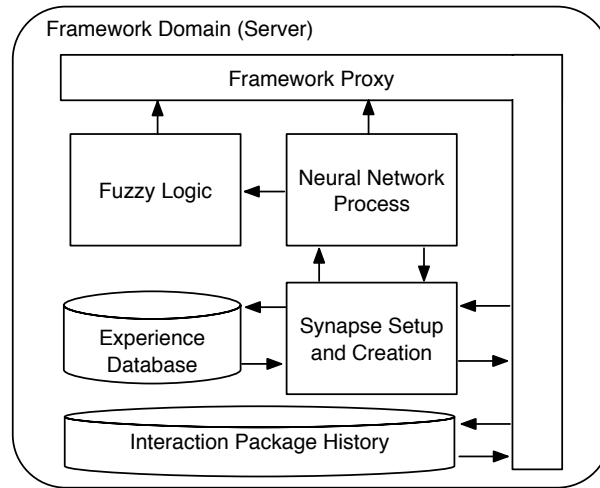


FIGURE 3.5: Framework Domain showing framework components

The framework domain is made up of various components as seen in figure 3.5. These components work together to establish and infer a collaborative role signature for a user. Each component is highly coupled to one another, meaning the components are dependent on each other to achieve the desired output.

3.5.4.1 Framework Proxy

Without orchestration the framework would not be able to carry out its life cycle, and would be limited in terms of its ability to handle interaction with other domains. The framework proxy provides this orchestration for its internal components, while also allowing a simplistic interface for outside domains and input. The proxy also manages the acceptance of interaction packages while also storing them so no interaction is lost.

Through the use of a proxy the framework achieves a *black box approach* to external domains. A user of the framework does not need to know how the framework

achieves its output they only need to know of the input required for the system [118]. Again, this decouples the framework from a [OCE](#) so it can be applied to many environments without dependency. How it achieves this can be seen in the Visual Representation of the Framework, figure 3.3, where the interaction from the [IPUSM](#) and the Server Application Logic is present. These external domains do not have access to the frameworks components only the proxy. External domains provide requests of function upon which the framework responds with a [Hypertext Transfer Protocol \(HTTP\)](#) response. This is achieved through a [Representational State Transfer \(RESTful\)](#) service provided by the framework proxy.

As requests can be from any external domain, the originating domain has to be able provide an identifier to have the proxy respond to it. This is needed as a client application of an environment may provide a direct request of the framework and as the client is an application and not a domain the framework would reject the request. This feature also allows segmentation within the framework so it can handle multiple environments without mixing previous experiences. Although not implemented in case studies presented in the next chapter, the framework will also allow roaming user profiles so a user's interaction history can be segmented by identified domains.

The proxy has simple requests that can be made. Through these requests the different phases of the life cycle are carried out. The framework does not perform the life cycle on every request. Instead, each request completes one or more phases of the life cycle. This approach was chosen as the framework works in an asynchronous environment and therefore the requests for each phase of life cycle will be also. For instance, an environment may want to commit without the need to process an output. In addition, the processing needed for different phases, such as commitment, have heavier processing loads than others, such as transitioning. Therefore, to have a framework perform the life cycle on every request would create a bottleneck for any environment wanting it for personalisation or customisation where real time output is needed.

The following is an outline of the requests available for the framework proxy. These requests can be made by any domain that is able to identify itself with a unique identifier available within the framework proxy, set by the implementer of the framework.

Capture

The framework proxy is able to store interaction packages without processing. This ensures all interactions are captured in raw format so the history of interaction is not lost after processing. The raw format is stored within a database table that can be called upon in the future to process. The data dictionary used for the database is found within table 3.2.

The capture request is the most common request the interaction [IPUSM](#) makes.

Column	Data Type	Description
id	Integer	Tuple identifier
domain	String	Unique identifier for an OCE
user	Integer	Identifier for user within an OCE
package	String	The JSON string representing the interaction package
creation	DateTime	The date and time the tuple was created within the database
processed	Boolean	A boolean flag representing if interaction package has already been processed within the framework

TABLE 3.2: Data Dictionary for Interaction History Database Table

Commit

The commit request carries out the processing of stored interaction packages. It achieves this by coordinating the retrieval of interaction packages in the interaction package history database, and passing it into the frameworks processing. This process is a computational expensive exercise as synapse creation and setup is carried out to perform back-propagation within the [Back-Propagation Neural Network \(BPNN\)](#). This type of [NN](#) uses **supervised learning** and therefore requires coordinating input and processing of historical data to update, process and add to the experience database. This historical data are packages stored in the interaction package history database that are flagged as not processed.

The commit request should be carried out either by a timed event or manually. The reason for this is to prevent the framework processing to often. An [OCE](#) can use a new user session as the event and this would ensure sufficient time lapse between committing and processing. This also aligns with the framework life cycle.

The process carried out within this request will be explained in the next section Framework Processing.

Signature

The signature request produces the output from the framework. The framework proxy uses the request, which has a [JSON](#) body, to feed into the framework to infer a role signature. Listing 3.3 represents the request body delivered to the framework proxy by the signature request. From this listing one can see it is similar to an interaction package, however there are two additional fields - *start_datetime* and *end_datetime*. These *start_datetime* and *end_datetime* date and time fields enable the proxy to segment the framework processing based on experience states in time. They are only required when a signature request is wanted for a particular time period.

Although in listing 3.3 the fields *tool*, *user*, and *tie* are present, only one or more is required. The framework only needs one or more input field to determine a role signature. However, with every input field given, the context of the signature is refined. Therefore, depending on the purpose and application of the role signature will determine the level of context needed for a signature request.

How the output is achieved from this request will be explained in the next section [3.5.4.2](#)

```
1 // JSON
2 {
3     "tool": "wall-post",
4     "user": 3243,
5     "tie": "u289-m3243"
6     "start_datetime": 1444605840,
7     "end_datetime": 1444505320,
8 }
```

LISTING 3.3: Signature Request Body

3.5.4.2 Framework Processing

A [NFS](#) is used for processing a role signature. This [NFS](#) is decoupled into segments that are used through out the processing to achieve an output. Decoupling these segments also enables independent operation of the stages of processing via the framework proxy.

After extensive research on various methods the choice of a [NFS](#) was chosen as it is a common approach to user modelling, and it also gave the greatest flexibility and control over the supervised learning method [85]. This is because of its ability to set up the [ANN](#) structure before processing based on the request made. Also the [FL](#) allowed an explicit method to determine strength based on age of behaviour.

Regardless of the machine learning method used, for it to work within the framework it must be able to take snapshots of its dynamic structure and state, compare historical and current role signatures composed of characteristic behaviours, so it can infer a present role signature.

Artificial Neural Network

The [ANN](#) used by the framework processing is a [BPNN](#). This type of neural network uses back-propagation for learning from experience. It uses sigmoid neurons instead of perceptron neurons. Where perceptron neurons output a 0 or 1, a sigmoid neuron can achieve degree of 1, such as 0.23. Sigmoid neurons also have the characteristic of: if there is a small change in neuron synapse weight, there is a small change in the neurons output. This is beneficial as one can train incrementally allow the output to change incrementally also. This characteristic of sigmoid neurons is a result of a sigmoid transfer function - a function that has a *S* shaped curve [119].

The back-propagation algorithm is a common algorithm for [ANN](#) [120]. It is considered a **supervised learning algorithm** as the resulting [ANN](#) needs to know the desired output for each input value. By knowing the desired output and input one can calculate the gradient of a loss function for all weights within the [ANN](#). Therefore, one can modify this gradient based on the experience provided, or more commonly know as, train the [ANN](#), and update the [ANN](#) structure.

The theory of [ANN](#) is outside the scope of this research. A generic [ANN](#) approach was used based on the [ANN](#) found in *Programming Collective Intelligence: Building Smart Web 2.0 Applications* [121]. This particular approach to an [ANN](#) was chosen because of its ability to create a dynamic structure at the time of processing, therefore being able to accommodate the dynamics of an [OCE](#). Although this approach for an [ANN](#) has been well documented [121], the resulting implementation and characteristics of the developed [ANN](#) for the framework differ. Therefore, the implementation and characteristics of the [ANN](#) need to be explained and outlined. For instance, the [ANN](#) chosen has been extended for this research with the

novel ability to allow for the creation of snapshots of experience used by the ANN. This is a characteristic that is unique for this implementation of the framework.

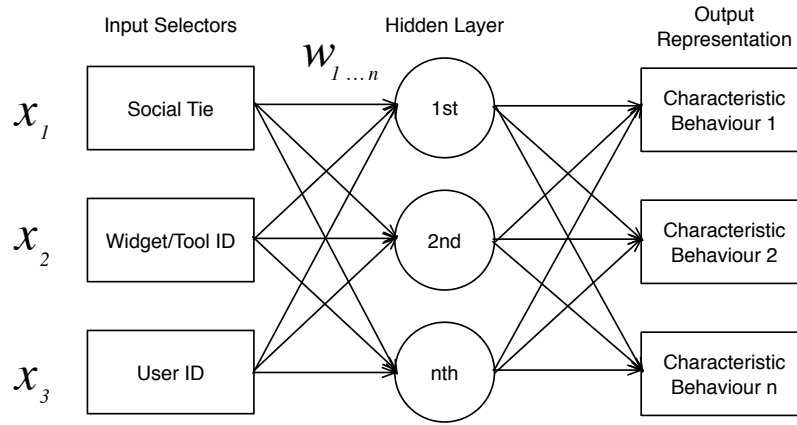


FIGURE 3.6: Visual Representation of ANN

Dynamic Creation - Synapse Setup and Creation

The structure of the ANN is not hard coded within the framework. Instead a structure is created every time a request is made. This is crucial to the design of the framework as new users, behaviour, tools/widgets, and social-ties will be created continually throughout the life of the OCE. This means the input and related output for the ANN will change over time.

Figure 3.6 provides a visual representation of the ANN structure. Based on the inputs and desired outputs this structure is created by the framework when requested. The structure has one hidden layer. This hidden layer is made up of sigmoid neurons related to each input selector received by the neural network. These neurons are created every time a unique 'set' of inputs is presented to the framework. Unique being the combination of inputs, not necessarily a new input. This is to ensure each input has a weighted synapse towards a relevant neuron.

The dynamic creation is achieved by storing the synapses, synapse weights, and neuron identifiers in a database called the Experience Database. When the framework proxy initiates a request, the inputs passed to the framework will be used as selectors to select related neurons, synapses and synapse weights in the database.

This process is also carried out for all outputs. However, all characteristic behaviours ‘experienced’ for an [OCE](#) will be used, not a selection. This is done because behaviour displayed throughout a user’s profile life can be used to establish role signatures.

This dynamic structure creation allows the structure to represent all related experiences before processing. As a result, the framework can estimate a role signature even if all input provided has not yet been experienced. This is a capability of an [ANN](#). For instance, a social-tie presented as input has not yet been committed. The [ANN](#), based on the other presented input selectors, will make an estimation to approximate the behaviours for the unknown social-tie.

Figure 3.7 presents a flow chart of the [ANN](#) dynamic creation. This figure shows the process for inputs. It is identical for outputs except for the single process of selecting all experienced characteristic behaviours instead of receiving them from the framework proxy.

During this dynamic creation the process of selecting snapshots is carried out. When a request is made with *start_datetime* and/or *end_datetime* within the body of the request, as shown in listing 3.3, these fields are used for conditional selection of neurons and synapses within the database. This leaves the resulting structure at a point in time, and as a result the processing can be carried out on this structure at a later point in time. If there are no date and time fields, the most current snapshot is processed.

As a snapshot exists at a particular time this can result in neurons and synapses not existing or outside the date range specified. When this occurs the neurons are excluded or the synapses are set to zero so they are effectively insignificant for processing. This ability ensures the resulting [ANN](#) structure is still valid for processing, yet the structure used for processing is at a earlier point in time.

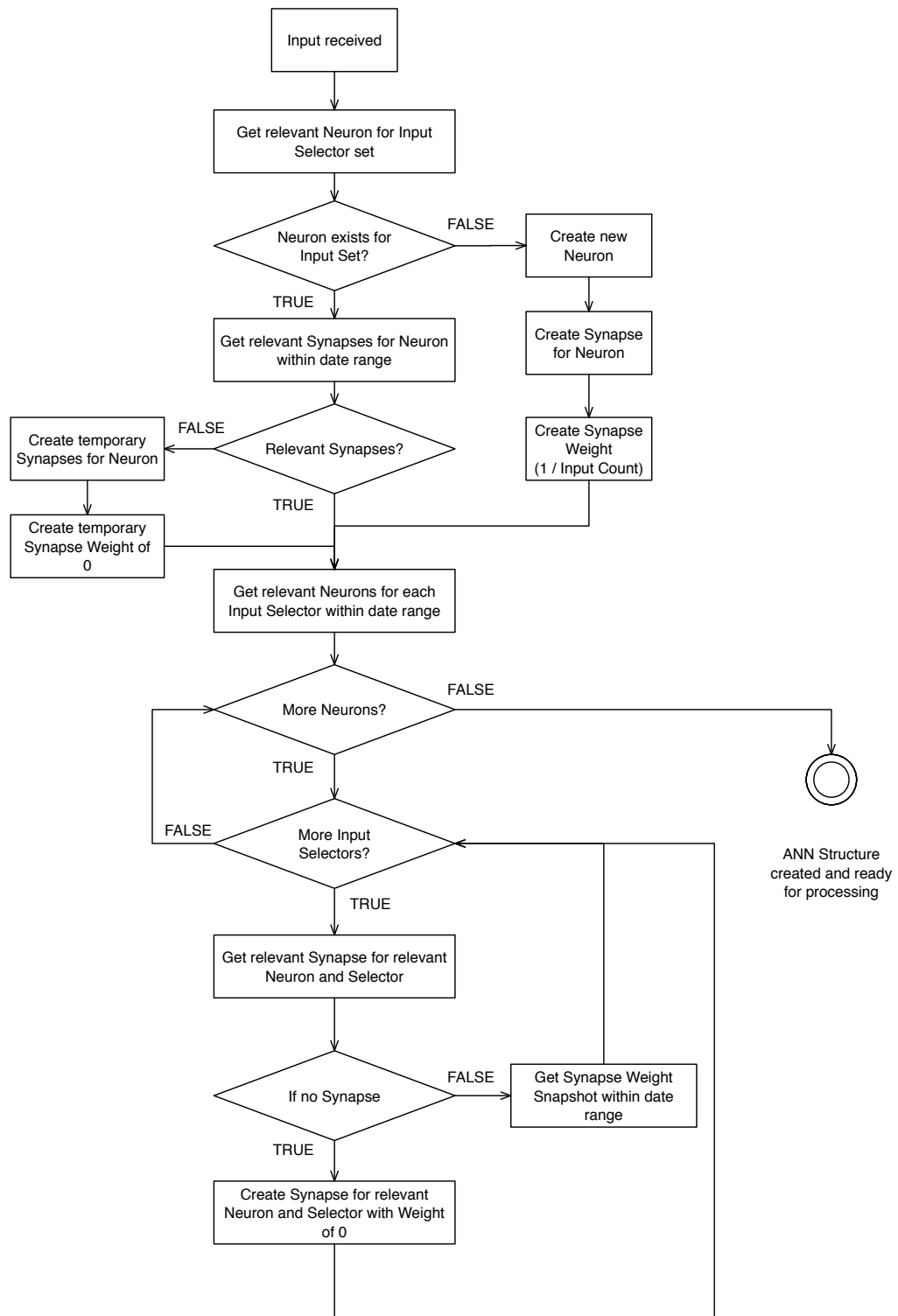


FIGURE 3.7: Flow Diagram of Dynamic Creation of the ANN Structure

Experience Database

The experience database holds not only the synapse weights, but also the structural elements and snapshots of these structural elements for the ANN.

The tables within the database reflect the structural elements within in the ANN. A Relational Database Management System (RDMS) should be used to ensure relational integrity between these tables as it will also ensure a valid structure for the ANN when created.

Figure 3.8 presents a database schema for the experience database. One can see the domain table is related to the neuron table. This ensures when the structure of the ANN is created only the neurons for the related domain are used. This capability enables the framework to be used across multiple domains holding various OCE.

Tables 3.3, 3.4, 3.5, and 3.6 present the data dictionary of columns within the tables of the experience database.

Domain Table

Column	Data Type	Description
id	INTEGER	Tuple identifier
name	VARCHAR	A unique name of a valid domain that is capable of interfacing with the framework
valid	BOOL	Boolean flag representing if the domain is still valid for framework interfacing

TABLE 3.3: Data Dictionary for Domain Table

Neuron Table

Column	Data Type	Description
id	INTEGER	Tuple identifier
name	VARCHAR	A unique name comprising of a concatenated selector set
domain	INTEGER	Foreign key representing a related domain table primary key
datetime	DATETIME	Date and time the neuron was created

TABLE 3.4: Data Dictionary for Neuron Table

Synapse Table

Column	Data Type	Description
id	INTEGER	Tuple identifier
selector	VARCHAR	A selector representing an input passed to the framework
neuron	INTEGER	Foreign key representing a related neuron table primary key
type	INTEGER	Foreign key representing a related type table primary key

TABLE 3.5: Data Dictionary for Synapse Table

Synapse Weight Table

Column	Data Type	Description
id	INTEGER	Tuple identifier
synapse	INTEGER	Foreign key representing a related neuron table primary key
weight	DOUBLE	A double representing the weight of the synapse
datetime	DATETIME	Date and time the synapse weight was created

TABLE 3.6: Data Dictionary for Synapse Weight Table

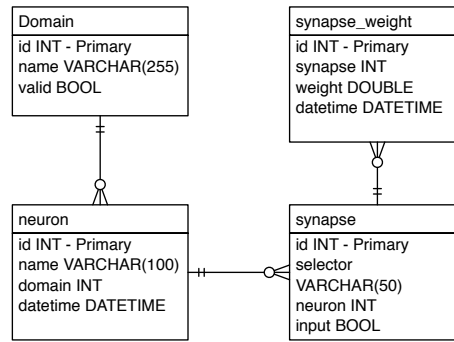


FIGURE 3.8: Database schema for the ANN

Neural Network Process

Once the dynamic structure of the ANN has been established the processing is comparable to any other ANN implementation using a matrix-based approach. The explanation of matrix-based approaches for neural networks is outside the scope of this research. However, the sigmoid transfer function and calculation of change present in the chosen ANN approach will be outlined for understanding.

There are two types of processing a BPNN can perform, feed forward and back-propagation. Feed forward is the processes of going through each layer of the ANN and working out the value of each neuron based on the synapse weights, then using those values to determine the output. By using a sigmoid neuron you can choose, by the type of sigmoid transfer function used, how the ANN reacts to the input presented. The sigmoid transfer function used was the *hyperbolic tangent*. This is a common approach to ANN as many programming languages already provide the *tanh* method. The *tanh* method is:

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1} = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (3.1)$$

Figure 3.9 presents a graphed hyperbolic tangent showing the *S* curve the function produces. From this one can see the rate of change for *y* in relation to *x*. This relation is an important attribute to the framework as the output can be represented as a degree of 1, and therefore, so can a characteristic behaviour. This aligns with behaviour as it is not on or off, but present to an extent. Also, the *S* curve ensures the behaviour shown in a user session does not become their dominate behaviour

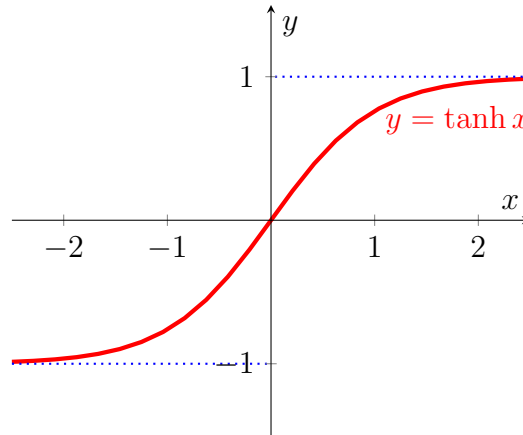


FIGURE 3.9: Hyperbolic Tangent

in a single user session. Only by a user demonstrating a behaviour multiple times over a period of time does that behaviour move closer to 1.

Back-propagation is the process of updating the synapse weights in relation to actual output experienced. This is achieved by providing a training key, an output experienced that is weighted, so when the synapse weights are updated the synapse weight representing the output experienced has a change moving towards 1, and all other weights moving away from 1. Change occurs to all synapse weights within the structure created from the input and desired output presented. This change is determined by calculating deltas for each synapse weight through the derivative of *tanh*:

$$f'(z) = 1 - (f(z))^2 \quad (3.2)$$

Once a set of deltas are established a learning rate is applied to each synapse weight and then multiplied by each established delta. Once back-propagation has occurred and has resulted in an update to the [ANN](#) structure, a snapshot is taken. This consists of creating new tuples within the *synapse_weight* for all synapses that have had their weights updated. By creating new tuples within the *synapse_weight* table it ensures the previous state of the [ANN](#) structure still exists.

Fuzzy Logic

The [FL](#) system is the last phase of the social group role model, which is the transition phase. It is used to create a *crisp* value for each characteristic behaviour over time using output from the [ANN](#). Although the output from the [ANN](#) can be used to represent a role signature, those outputs are in a state of fuzziness. This is because time, strength and age of behaviour over time is not represented. For instance, this could be why would a characteristic behaviour from the [ANN](#) be valid when it is a month old - a result of no new user sessions for a month or more. Through [FL](#) a characteristic behaviour will be inferred from historical behaviour and current behaviour based on strength and age.

The [FL](#) process is also important as it creates stability in role signatures. A user may demonstrate a strong behaviour within a particular user session, however historically the behaviour is weak, therefore the behaviour displayed within that particular user session is a possible anomaly.

The input for the fuzzy logic is the output of two snapshots provided by the [ANN](#). These two snapshots are determined by the signature request with the fields *start_datetime* and *end_datetime* provided in the request body - as in listing [3.3](#). As these fields may not be actual snapshot dates but a range, the framework will choose the youngest and oldest snapshot between these two dates. If these fields are not provided in the request, the framework will use the latest snapshot and the snapshot directly before the latest.

The structure of fuzzy logic used is the Mamdani-style fuzzy inference [\[122\]](#). It was chosen for its simplicity and common approach to fuzzy logic. To determine the degree upon which a value belongs to a fuzzy set the membership functions used are *linear fit functions* [\[123\]](#). These membership functions are not curved or a mathematical equation. Instead, they result in a linear membership between sets. Figure [3.10](#) visually represents this linear membership. There is a gradual linear inclusion, therefore a value can exist within a set to a degree. Listing [3.4](#) provides an implementation of the membership function in the Python programming language. From this listing, one can see the fuzziness passed into the function is used to determine membership, and then the degree of membership.

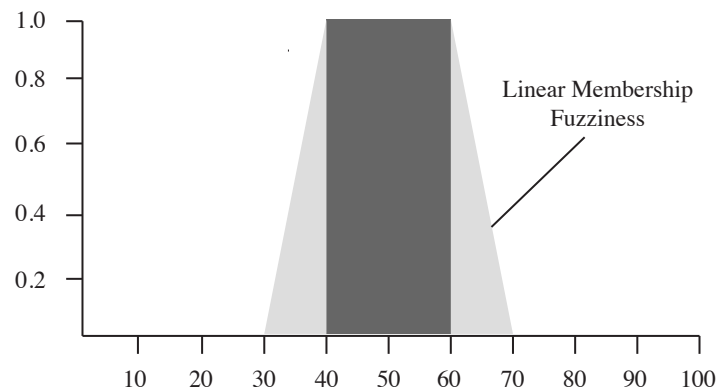


FIGURE 3.10: Visual representation of linear membership (fuzziness)

```

1
2 def fuzzify(startLimit, endLimit, fuzziness, value, noEdge=None):
3     """Fuzzify value according to start and end limit and the amount
4     of fuzziness
5
6     Keyword Arguments:
7     startLimit — The set starting value
8     endLimit — The set ending value
9     fuzziness — The amount of linear fuzziness
10    value — Value to be fuzzified
11    noEdge — If one sided set
12    """
13    value = float(value)
14    startLimit = float(startLimit)
15    endLimit = int(endLimit)
16    if value >= (startLimit+fuzziness) and value <= (endLimit-
17    fuzziness):
18        return 1
19    if noEdge == 'left' and value < (startLimit + fuzziness):
20        return 1
21    elif value < (startLimit + fuzziness) and value > startLimit:
22        if startLimit < fuzziness:
23            fuzzy = value/float(fuzziness)
24        else:
25            fuzzy = (value%startLimit)/float(fuzziness)
26        return fuzzy
27    if noEdge == 'right' and value > (endLimit - fuzziness):
28        return 1
29    elif value > (endLimit-fuzziness) and value < endLimit:

```

```

29     fuzzy = (fuzziness - (value - (endLimit - fuzziness))) / float(
        fuzziness)
30     return fuzzy
31     return 0

```

LISTING 3.4: Fuzzy Membership Function implemented in the Python programming language

Strength and age of behaviour are the categories of fuzziness used by the framework. The amount of linear fuzziness chosen for this study is between 15-20% of the range of all possible sets. For instance, If all possible sets are within the range of 100, then the fuzziness of a set is 15 or 20. The input to the [FL](#) is the output from the [ANN](#), where the [ANN](#) output is between -1 and 1. These values cannot be used with the [FL](#). Therefore, they are normalised so they result in a value between 0 and 100. For instance, an output from the [ANN](#) of -0.23 becomes 38.5 .

Table 3.7 presents the set terms for each category of fuzziness. The strength category will be used for current and historical [ANN](#) output from the [ANN](#). It will also be used to determine the final characteristic value based on [FL](#) rules. The age category will be used to represent the difference in time between current and historical snapshots.

Fuzzy Set Terminology

Strength	Age
Weak	Young
Neutral	Old
Strong	

TABLE 3.7: Fuzzy Set Terminology

Although the linear membership of each set is configurable within the framework, the set areas for this study is found within figure 3.11. The value ranges of each behaviour strength set are presented in the table 3.8. With a linear fuzziness of 15-20% between sets, one can see how each behaviour strength overlaps. This configuration of behaviour strength sets have been chosen so the characteristic behaviours are not so volatile because of the sigmoid behaviour of the [ANN](#). At 50 it resembles an output of 0 from the [ANN](#). The value of 0 is where the rate of change

within the ANN is most significant. Therefore characteristic behaviours that led to common and regular interaction were offset from this point of 0 to ensure when a behaviour is strong, common and regular for a user it is not in a position that can change rapidly. This is why *Weak* is the largest set of behaviour, as it is capturing behaviour that is not common, and *Neutral* is capturing behaviour that could go towards *Strong* or *Weak*.

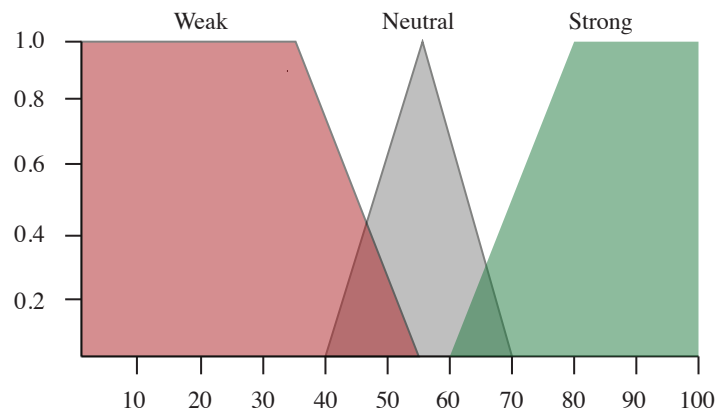


FIGURE 3.11: Visual representation of Behaviour Strength

Behaviour Strength Set Ranges

Set Term	Start	End
Weak	0	55
Neutral	40	70
Strong	60	100

TABLE 3.8: Ranges for behaviour strength set terms

There are two terms within the age set category, *Young* and *Old*. For this study 30 days has been chosen as the total range for time. This is configurable and will be dependent on the OCE it is implemented in. The chosen set ranges, as presented within table 3.9, are also dependent on the OCE. For instance an environment may see daily use, so a *Young* set range would be smaller than 15 days and therefore 15 days would be considered *Old*. Figure 3.12 presents a visual representation of Behaviour Age sets, demonstrating a fuzziness between the two sets of 10 days.

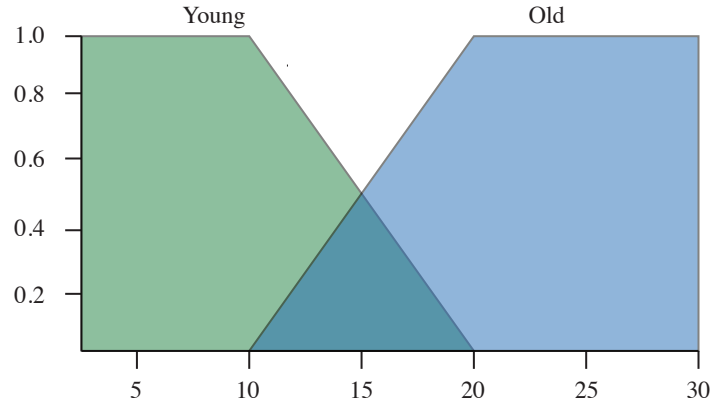


FIGURE 3.12: Visual representation of Behaviour Age

Behaviour Age Set Ranges

Set Term	Start	End
Young	0	10
Old	10	30

TABLE 3.9: Ranges for behaviour strength set terms

Offline individuals do not transition to a new social group role instantaneously on every engagement, and so these [FL](#) rules try to simulate this. The rules favour a *Weak* characteristic behaviour over a *Neutral* behaviour so there is a limited chance of a behaviour being *Strong* when it is an anomaly at a point in time. This is achieved by using the strength of a current characteristic behaviour and either an old or young historical behaviour.

The [FL](#) rules are as follows:

Let current characteristic behaviour be denoted as set C , historical characteristic behaviour as set H , and days between two snapshots as set A . The output of the rule, the concluding behaviour, is denoted as set B .

1. **IF** C is *Weak* **AND** H is *Weak* **THEN** B is *Weak*
2. **IF** C is *Weak* **AND** H is *Neutral* **THEN** B is *Weak*
3. **IF** C is *Weak* **AND** H is *Strong* **AND** A is Young **THEN** B is *Neutral*

4. **IF** C is *Weak* **AND** H is *Strong* **AND** A is *Old* **THEN** B is *Neutral*
5. **IF** C is *Neutral* **AND** H is *Weak* **AND** A is *Young* **THEN** B is *Neutral*
6. **IF** C is *Neutral* **AND** H is *Weak* **AND** A is *Old* **THEN** B is *Weak*
7. **IF** C is *Neutral* **AND** H is *Neutral* **THEN** B is *Neutral*
8. **IF** C is *Neutral* **AND** H is *Strong* **AND** A is *Young* **THEN** B is *Strong*
9. **IF** C is *Neutral* **AND** H is *Strong* **AND** A is *Old* **THEN** B is *Neutral*
10. **IF** C is *Strong* **AND** H is *Weak* **THEN** B is *Weak*
11. **IF** C is *Strong* **AND** H is *Neutral* **AND** A is *Young* **THEN** B is *Strong*
12. **IF** C is *Strong* **AND** H is *Neutral* **AND** A is *Old* **THEN** B is *Neutral*
13. **IF** C is *Strong* **AND** H is *Strong* **THEN** B is *Strong*

All rules are evaluated on the inputs to create a fuzzy set for defuzzification. All rules consist of the **AND** operator. The method used for this operator is the traditional *min(minimum)* method. This means for all conditions of a rule, the smallest value achieved for a rule will be passed to the fuzzy set for defuzzification.

Once all rules are evaluated aggregation occurs using the common *max(maximum)* method to create a fuzzy set. This fuzzy set is then defuzzified using the common defuzzification method *centroid*. This method mathematically finds the [Centre of Gravity \(COG\)](#) and is expressed as:

$$\text{COG} = \frac{\sum_{x=a}^b \mu_A(x)x}{\sum_{x=a}^b \mu_A(x)} \quad (3.3)$$

The output of this defuzzification method is a crisp value for each characteristic behaviour between 0 and 100.

The whole [FL](#) process is presented in Figure [3.13](#). One can see that 3 input values, fuzzy rules, and fuzzy inference results in a crisp value.

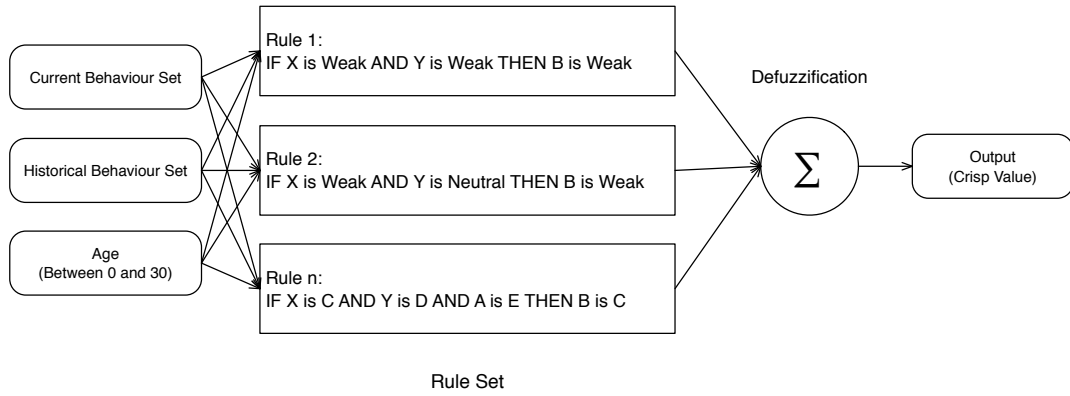


FIGURE 3.13: Visual Representation of FL system

To demonstrate this, the crisp value from the following example output from the ANN is as follows:

Current Behaviour Set = (publisher:0.09, annotator:-0.36, lurker:-0.01)

Historical Behaviour Set = (publisher:0.57, annotator:0.10, lurker:-0.20)

Age difference between sets (days) = 16

FL Output:

Role Signature = (**publisher:** 72.14%, **annotator:** 25.0%, **lurker:** 32.5%)

This example output demonstrates crisp values representing characteristic behaviours over time. This example also provides representation of the final output of the framework - a social group role signature.

Chapter 4

Case Studies

4.1 Introduction

After the framework was designed and implemented it was interfaced with three different case studies. Two of these case studies were developed specifically for this research. The third was a pre-existing [NHMRC](#) approved research project. The choice for these case studies were a result of being opportunistic and having generic online collaborative environments to demonstrate the framework with. For the mobile applications, Espressobility was chosen and developed for its simplistic collaborative goal of reviewing coffee locations, Squashies was chosen and developed because of the opportunity to use a group based environment where users might know each other online and offline. The Walk 2.0 Project was chosen because of the opportunity interface the framework with a historical dataset to demonstrate the frameworks capability to handle historical instead of real time data found within an [OCE](#).

Although the interfacing of the framework for these case studies were the same, the framework was ported to two different programming languages to accommodate the different development environments used for the case studies.

This chapter will present the cases studies with the following structure:

- Case study description
- Discussion of each collaborative activity provided by the case study along with the tools and widgets used within the environment

- Any explicit user roles present within an environment

Although each case study has a different collaborative context, as discussed in the literature review, and is only mediated by the tools and functionality provided by environment, it will not be discussed at length in this chapter. Only the collaborative functionality, tools and widgets that could lead to collaborative behaviour will be discussed. All other details of the case studies, that is, the recruitment time lines or phases, ethics, and participant use of these case studies will be presented in the next chapter, Chapter 5 Participant Use and Data Collection.

4.2 Espressoability

4.2.1 Description

Espressoability is a mobile phone application specifically built as a case study for the framework. The purpose of the application was to allow users to collaborate around coffee locations identified and published by users, such as Cafés. It achieved this by enabling users to be sole creators of content through their interactions with locations and others. Users were provided with tools and widgets for rating, tagging, commenting, and reviewing coffee and Cafés. Through their collaborative effort an Espressoability rank between 1 and 10 was given to each coffee establishment. The resulting effort provided a resource for users to find the best *coffee scene* within their local area. Users were also able to socialise with other users by using social networking tools for following, and posting messages to each other.

The application was built using hybrid mobile technologies so it could be used on both iOS and Android mobile phones. This hybrid technology was *PhoneGap*. At the time of development PhoneGap was a new technology and the features available were still in their infancy. This led to a significant increase in development time, and also unpredictable behaviour within the application for select users. As a result, the application was only released for mobile phones supporting the Android operating system.

The development of this case study was in the Python programming language. This resulted in the framework domain also being built in the Python programming language. Although, this is not necessary for the framework, it was chosen to host the framework on the same server hosting the mobile phones application [API](#) to save on resources. Although the framework and application were on the same server they were two virtual domains operating on one server. This conforms with the framework residing on its own domain and other domains interfacing with it. When requests were made to the server they were made to <http://framework.espressobility.com> (offline), a sub-domain to the application domain - there is no dependency even though it was as sub-domain.

Although the framework domain was programmed in Python on the server, the Semantic Manager and the [IPUSM](#), which resides on the client, were written in Javascript. Again this was chosen as the *PhoneGap* client used Javascript and HTML5.

4.2.2 Collaborative Activities

Espressobility enables a user to explore and interact with coffee locations within their community. The application promoted autonomous interaction so users were able to collaborate in their own time and in their own way. Through the use of collaborative activities a user could contribute to the knowledge domain of coffee locations around a user's geo-location.

Upon first joining the environment users were presented with a home page showing an **activity stream**. The purpose of the activity stream was to promote awareness for the user-base so they could see the activities carried out by other users. The activities were presented in chronological order as seen in [figure 4.1](#). The type of activities the activity stream presents within Espressobility are described in [table 4.1](#).

Activity	Description
Coffee Review	Any user could provide a review 140 characters for a coffee beverage at a coffee establishment
Food Review	Any user could provide a review 140 characters for food served at a coffee establishment
Customer Service Review	Any user could provide a review of 140 characters for customer services experienced at a coffee establishment
Price Review	Any user could provide a review of 140 characters in relation to the pricing at a coffee establishment
Location Review	Any user could provide a review of 140 characters in relation to the geographical location and atmosphere of a coffee establishment
Photo Review	Any user could take a photo and add a 140 character review at a coffee establishment
Create Establishment	Any user could create a new coffee establishment if not found within the application
Update Establishment	Any user could update a coffee establishments details
Establishment favoured	Any user could indicate a coffee establishment was one of their favourite establishments

TABLE 4.1: Type of activities a user could create within the Espressobility application

The type of activity was represented by small indicators in the top right corner of each activity. In figure 4.1 one can see this with **Point 1** and the small knife and fork icon representing a food review next to the Espressobility Rank given to the coffee establishment.

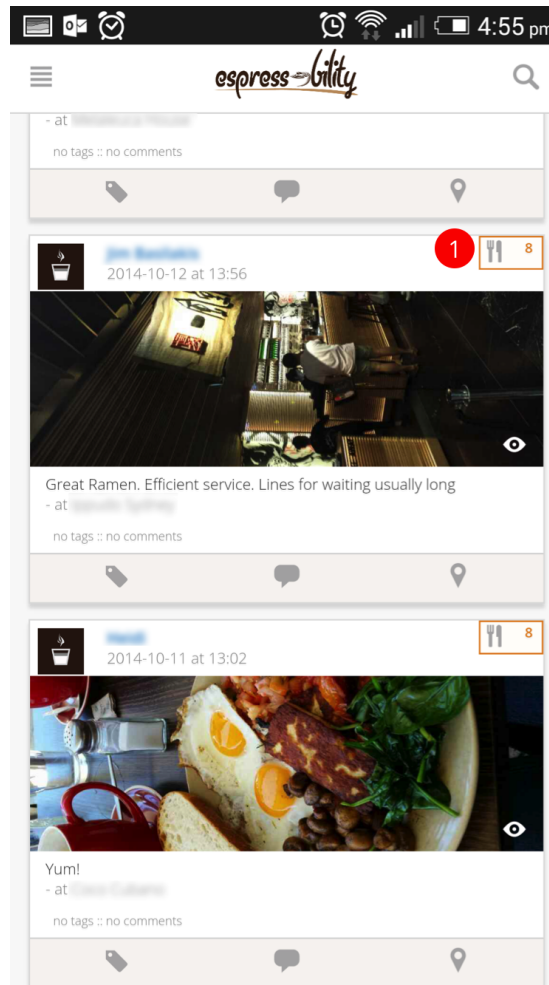


FIGURE 4.1: Screen shot of global Activity Stream within the EspressoBility App

Activity streams were not limited to a global space. Every coffee location and user had their own personal activity stream showing related activity. This is seen in figure 4.2. At the top of the activity stream is an establishment users could interact with. Activity streams had the capability of **infinity scrolling**. This feature allowed users to continually scroll through streams without having wait for new items to be loaded. This was achieved by the activity stream signalling the end of a set list of activity items and requesting more items from the server. Every time the application requested more items the **IPUSM** would listen to the event and create an interaction package with *lurker* as the characteristic behaviour demonstrated. Depending on the type of activity stream, such as global, personal or location, the social tie used for the interaction would be the relating user profile, establishment or a general tie to the application.

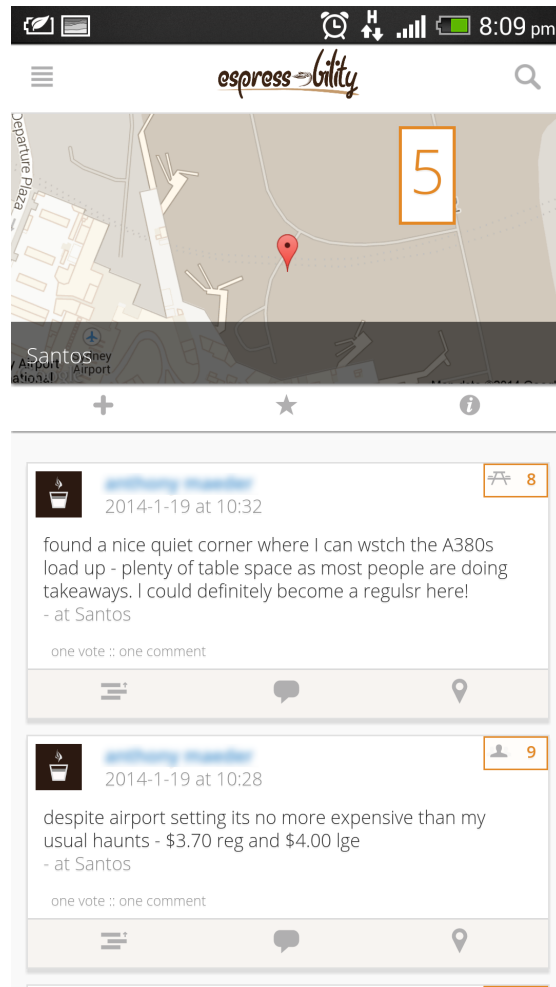


FIGURE 4.2: Screen shot of the Coffee location Activity Stream within the EspressoBility App

Users were able to create activities by interacting with different locations and users. Figure 4.3 is a screen shot of the various tools available for users to use in relation to particular location. Point 1 opens the **review post** tool that allows a user to add a new review towards the location. Horizontally from Point 5 the different types of reviews are presented. The first, which is selected, is coffee, followed by food, price, service and location. The interface for each type is the same, only the type of review changes when selected. Point 6 is a range value from 0 to 10 and is the value of a particular review. Point 7 allows a user to change the type of review to a photo review, with the user being able to take a photo which will be used with the posted review. Once a review has been completed by the user, the event fired by the application for posting the review is captured by the IPUSM. The tool for posting a review is the publication tool and therefore the

characteristic behaviour of *publisher* was used to represent this behaviour. The location identifier was used as the social tie.

Point 2 within figure 4.3 allows a user to add a location as their **favourite location**. This was a simple toggle. When a user favoured a location a post to the activity stream was made. Once favoured a user could then quickly find their favourite locations through the favourites menu item found within the application navigation. As the favoured action resulted in a publication, the *publisher* behaviour was used to represent the interaction.

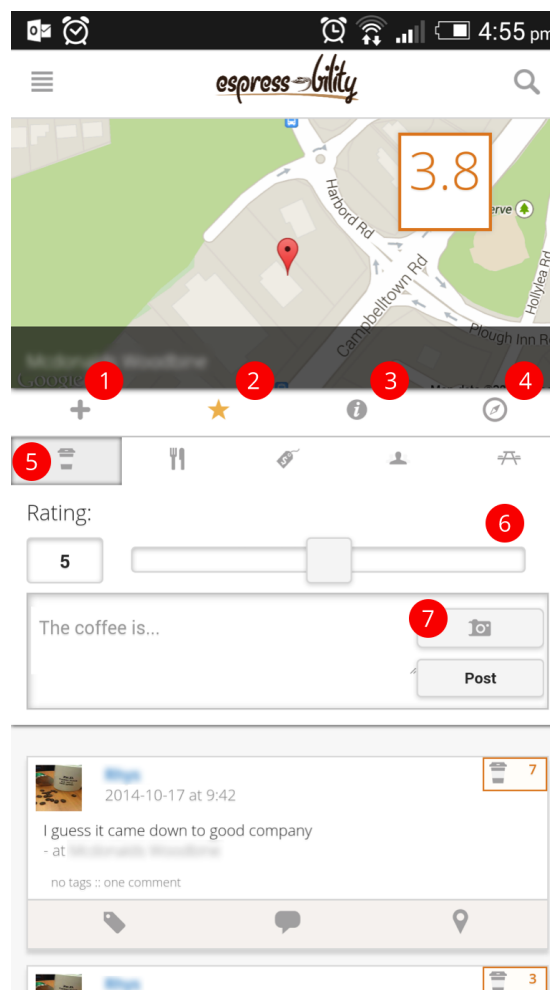


FIGURE 4.3: A screen shot of posting tools for coffee locations

Point 3 within figure 4.3 allows a user to open an **information dashboard** for the location. This dashboard is shown in figure 4.5. It shows a total of EspressoBility ranks and also a brief description about the location. When a user opens this information dashboard the IPUSM listens for the event and an interaction package with the *lurker* behaviour, location identifier as the social tie, and the information

dashboard is created. The brief description, Point 1 within figure 4.5, is created when a new location is created. However, this description can be updated by any user after the locations creation. When this is carried out the IPUSM listens for the change and when it occurs gives the behaviour of *annotator* as the user is not creating new content, they are simply altering or updating old content. This information dashboard had two possible characteristic behaviours that a user was capable of showing while interacting with it. These two behaviours were present in the metadata for the tool and the IPUSM was configured to listen for the events for when these behaviours were carried out. This configuration was based on when certain buttons were pressed within the application.

When a user visited another user's profile page they were also able to create an activity by **posting a message** directly to that particular user. Figure 4.4 shows this with a user profile page and at Point 1 the plus icon allows another user to post to the user's activity stream. The characteristic behaviour given to this action was *publisher* and the social tie the user identifier for the user profile where the post was made.

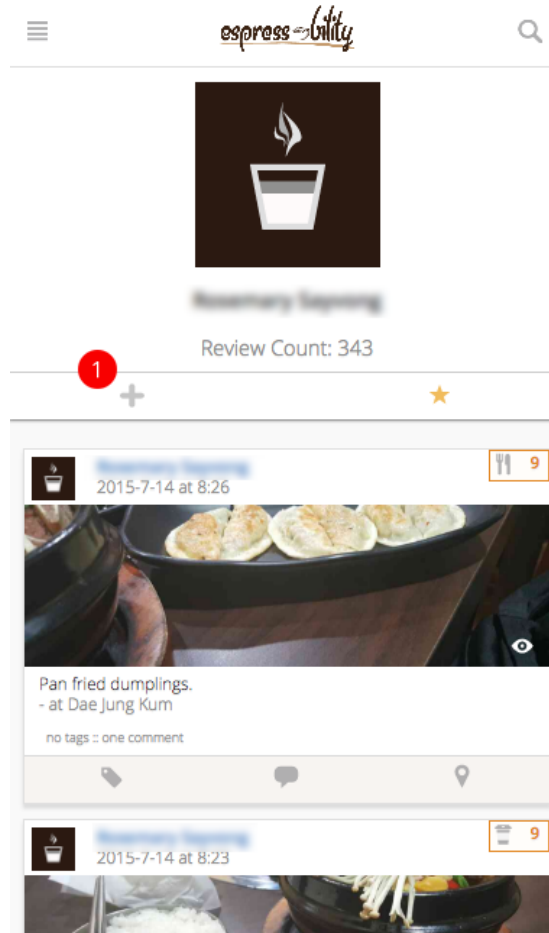


FIGURE 4.4: A screen shot of a user profile within the application

For each activity published, a user can interact with that activity through **tagging** and **commenting**. The capability of tagging can be seen within figure 4.6. Tagging allowed users to create their own taxonomy for different types of posts. As seen in figure 4.6 the terms, ‘*big serving*’ and ‘*breakfast*’ were created by users to classify the post. Once a term was created a link was formed for other users to click and find other activities tagged with the same term. As an activity could have two social ties, the location the review was for, and the user who posted the review, the IPUSM creates two interaction packages for each social tie. As the user is annotating the activity with a term the *annotator* characteristic behaviour is used.

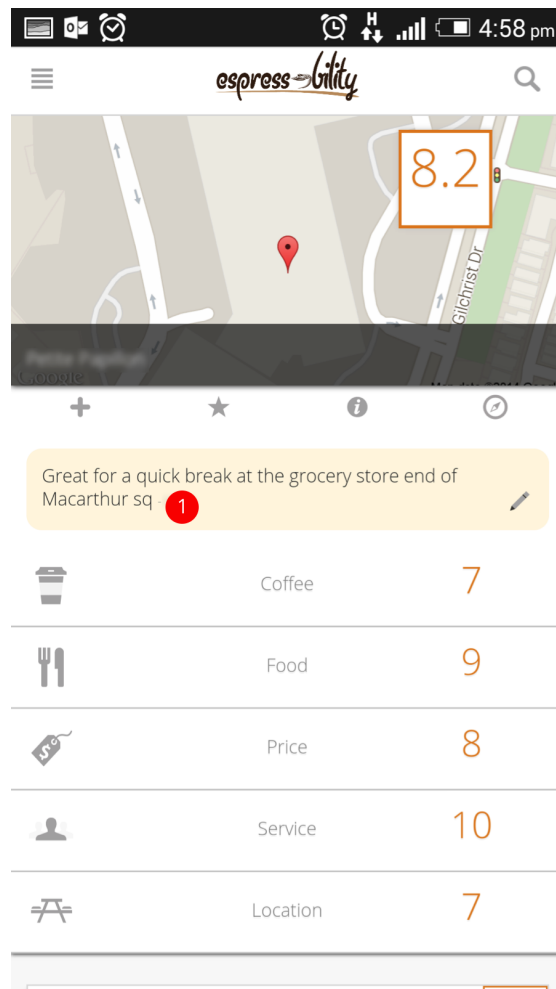


FIGURE 4.5: A screen shot of the location information dashboard

Activity commenting allows a user to add a comment to an activity to create a dialogue around the activity. Figure 4.7 presents the comments previously made for the activity and also the comment input box. As a user may open up comments, yet not make comment there is a possibility of two characteristic behaviours shown by the user. One is *lurker*, the other *annotator*. If a user opens up comments for an activity, yet does not make comment the *lurker* behaviour is used within the interaction package. However, if the user makes comment the original *lurker* behaviour is dropped in favour for *annotator*.

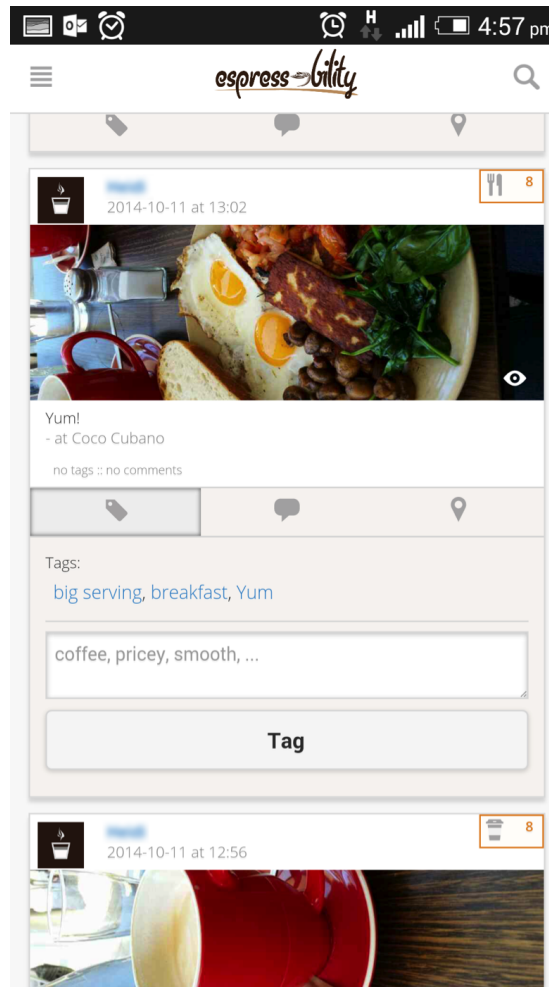


FIGURE 4.6: Screen shot of activity commenting

Coffee locations could be created by users based on their geo-location. They would search for a location and if it did not exist they could use **location creation**. By creating a new location a user created another node within the application that other users could interact with and establish social-ties with. Once a location was created an activity would be posted to the global activity stream showing the creation of the new location. This allowed users to interact with it like all other activities while also being aware of the new location. By creating a new location a user was creating new content, therefore the *publisher* characteristic behaviour represented the action. Although this action was collaborative, there was no explicit social-ties to relate with the action of creating a new location. As a result, the application identifier was used as the social-tie within the tool metadata.

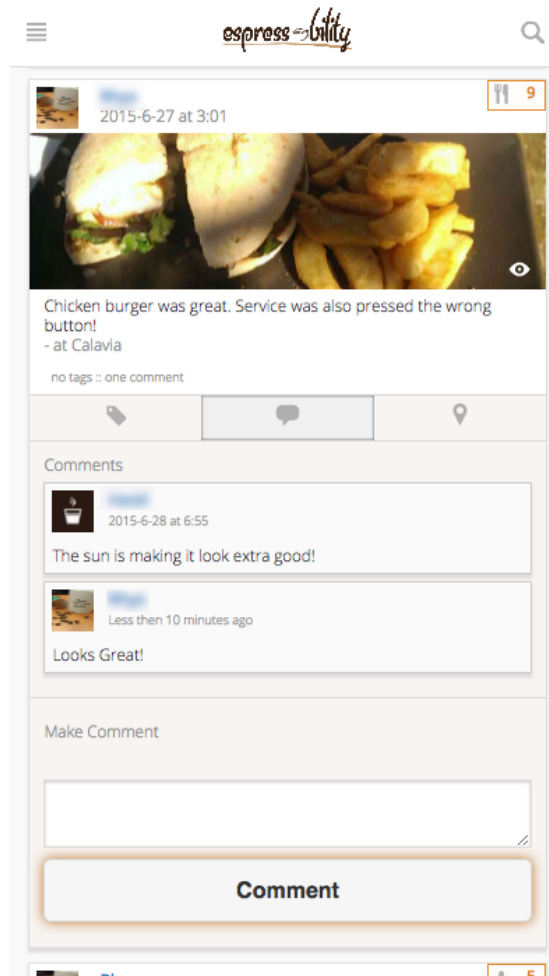


FIGURE 4.7: Screen shot of activity tagging

4.2.3 Explicit User Roles

There were no explicit user roles within this case study. Every user who joined had the same functionality as every other user.

4.3 Squashies

4.3.1 Description

Squashies is a mobile application developed as a case study so the framework could interface with with a cohort of users within an environment where explicit roles exist in a collaborative context. In cooperation with a local Gym, Squashlands

Gym and Fitness, within the Greater Western Sydney region of NSW, Australia, the application was provided to the members, instructors, and staff of the gym. The purpose of the application is to encourage use of the gym through exercise classes, tracking workouts and socialising with other members. Both members and instructors of the gym were able to download and install the application to their personal mobile phone.

After the length of development with *PhoneGap* for *Espressobility* it was chosen to develop a native application for the iOS operating system. The business logic layer for the application was built with the [Hypertext Preprocessor \(PHP\)](#) programming language. Again, to save on resources the framework was ported to [PHP](#) and resided on the same server as the business logic. The framework was on a sub-domain to the case study - <http://framework.squashies.com.au> - and the mobile application and business logic layer would make request to the framework domain. As the mobile application was built using iOS technologies, the semantic manager and the [IPUSM](#) were built in it also.

4.3.2 Collaborative Activities

Squashies was a custom mobile application for the members of Squashlands Gym and Fitness in the Greater Western Sydney Region of NSW, Australia. Members and gym instructors were able to collaborate together using various tools and widgets to socialise, create and manage gym classes, and interact with the gym outside of the confines of the gym.

Once a member was registered they would be presented with an activity stream. This activity stream held different types of activities happening within the gym or with members. Figure 4.8 provides a screen shot of the global activity stream for all activities carried out within the application. An activity stream enabled **infinity scrolling** so as the user scrolled the list of items continued to grow. As this action is carried out the [IPUSM](#) creates an interaction package with the behaviour of *lurker*. The social-tie provided by the semantic manager was dependent on the relation the activity stream had. The global activity stream used a general application identifier, if it was a personal activity stream for a user profile the user identifier was used.

The type of activities available for a user to create are outlined in table 4.2. Each one of these activities when created would appear in the global activity screen.

Activity	Description
Class attendance	Any user could indicate if they were attending a gym class or not (Point A in figure 4.8)
New class available	A gym instructor could create a new class and have that class published
Gym class session	A user could post about a class they attended
New cardio session	A user could post about a cardiovascular session they did within the gym
New Strength Training Session	A user could post about a strength training they did within the gym
Check in	A user could check in to the gym to indicate they have been to the gym (Point E figure 4.8)
Gym notice	Instructors and staff of the gym could produce a notice about the gym
Progress Post	A member could post about their progress towards their exercise goals
Status Post	A user could post to the activity stream with a general status (Point B figure 4.8)

TABLE 4.2: Type of activities a user could create within the application

It was easy for a user to *post an activity*. Point F in figure 4.8 shows a plus symbol which represents ‘add activity’. By pressing add activity it would open another screen and based on the user’s role and if they were part of a class they would be presented with different options to post. When a post was created, if the post was in relation to a class, the gym, a session, or the user, the social-tie would be established by the semantic manager. If an activity was related to a class, the class identifier would be the social-tie. If it was a gym post, it would be the user identifier of the instructor or the staff who posted it. If it was a personal session, the user posting the activity would be the social-tie. This would be achieved by creating a social-tie with the user identifier repeated twice. For instance if a user had an identifier of *363* then the social-tie would be ‘*u363-u363*’. The semantic

manager can achieve this as a social-tie is not the user, it is a *unique relationship* with a user, in this instance a user has a relationship with themselves.

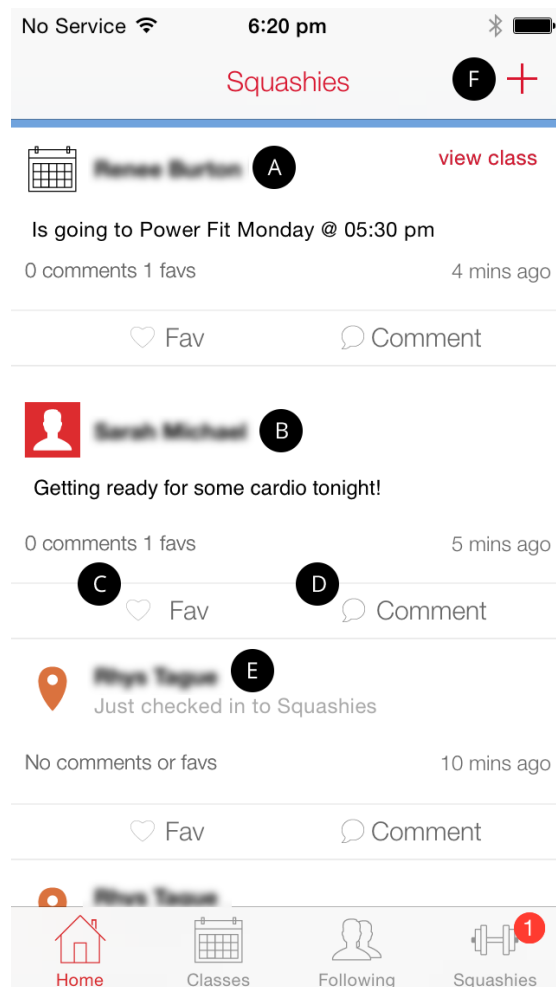


FIGURE 4.8: Screen shot of the activity stream of Squashies the mobile application

Although the act of posting different activities is similar in nature, each activity type posting had a different tool identifier. The reason for this was because not only was the context different for the activity, some activities, such as personal cardio session, had more details to enter instead of just a message. For the cardio session the amount of minutes of cardio could be entered in addition to a message, therefore it is a different tool to a personal status post and other activity types. All activity postings were represented as the *publisher* characteristic behaviour.

Instructors were able to **create classes** for gym members to join and interact with. Figure 4.9 presents a screen shot of the different class sessions available for members that were created by instructors. When a class session was selected by a

user they were presented with the class session view as seen in figure 4.10. Within a class session view the user could view others attending and also indicate if they intended to attend the session. When a user indicated they were attending, an activity was posted to the activity stream and their behaviour was represented as *publisher*. After this action two interaction packages were created by the IPUSM. One for the type of class and the other for the classes instructor. This was chosen as a user may choose to attend a class based on an instructor and not a class, and therefore the behaviour may be a result of the instructor not the class.

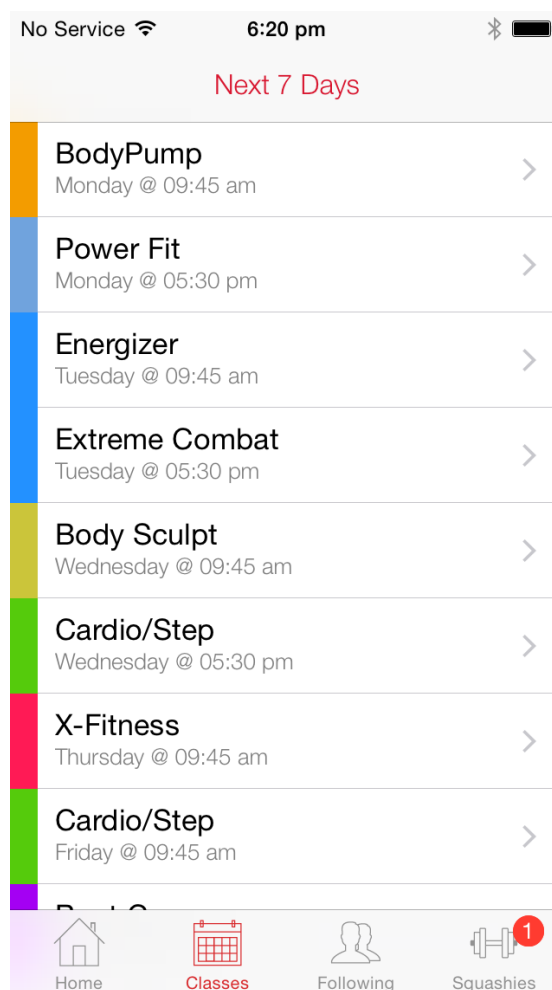


FIGURE 4.9: Screen shot of a class list view in the Squashies App

Once an activity was created, a user could interact with the activity through commenting and marking it as one of their favourites. Point C and D in figure 4.8 present the favourite and comment buttons. When a user selects the favourite button the button toggles from a white heart to a red heart indicating the user has favoured it. When a user favourites an item the item is listed within a user's

favourite activities under their profile page - Point A in figure 4.11. The favourite action was represented as an *annotator* characteristic behaviour.

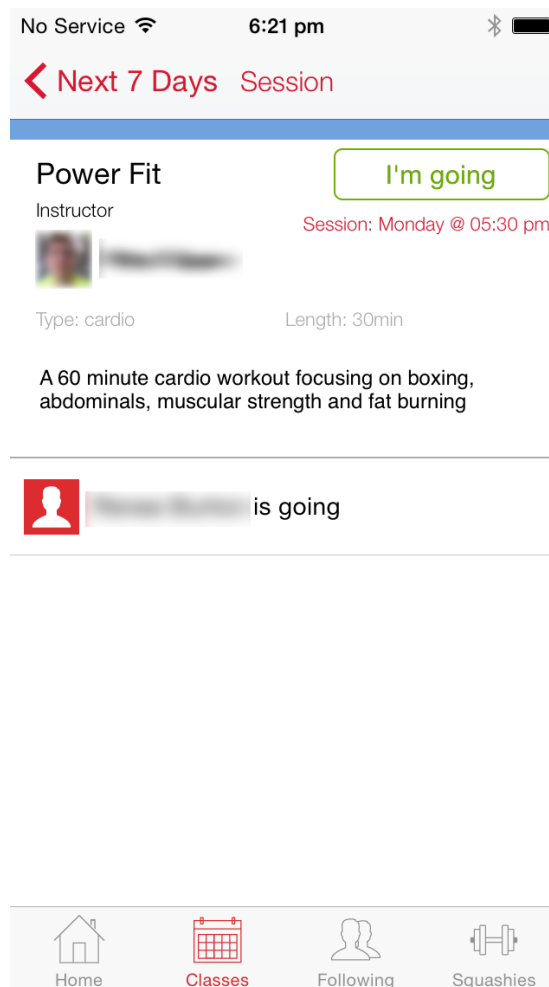


FIGURE 4.10: Screen shot of a class session view in the Squashies App

When a user pressed comment on an activity a new window would open and present the various comments for the activity, as well as the date and time the activity was favoured by other users. At the bottom of this window a user could enter in a message to leave a comment. The commenting window allowed the user to demonstrate two characteristic behaviours, *annotator* and *lurker*. If the user did not post a comment towards the activity after they leave the commenting window, the *lurker* behaviour was used, however if they did comment, the *annotator* behaviour would represent the posting of a comment.

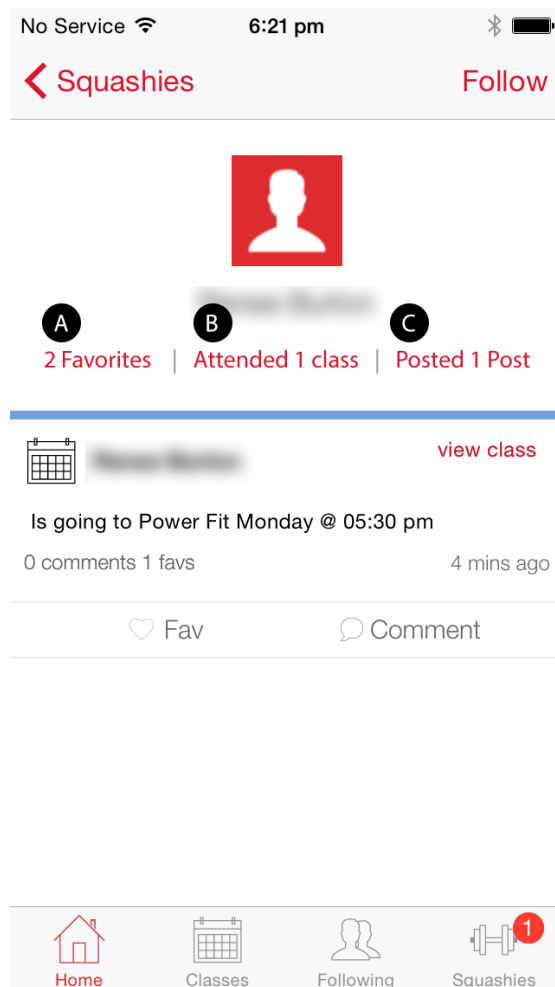


FIGURE 4.11: Screen shot of a user profile within the Squashies application

A user was able to navigate to another users profile page to view their personal activity stream. Such a user profile page is presented in figure 4.11. Once a user was at a user profile page the different activity types a user created could be filtered. Point A, B and C of figure 4.11 highlight the links that enabled the filtering of a user's personal activity stream. The filtering was based on activity type so a user could quickly find an activity they were looking for. When a user carried out this action the behaviour of *lurker* would be associated with it. The social-tie would be the user identifier of the user profile page. This type of functionality was also available for classes, as seen in figure 4.12

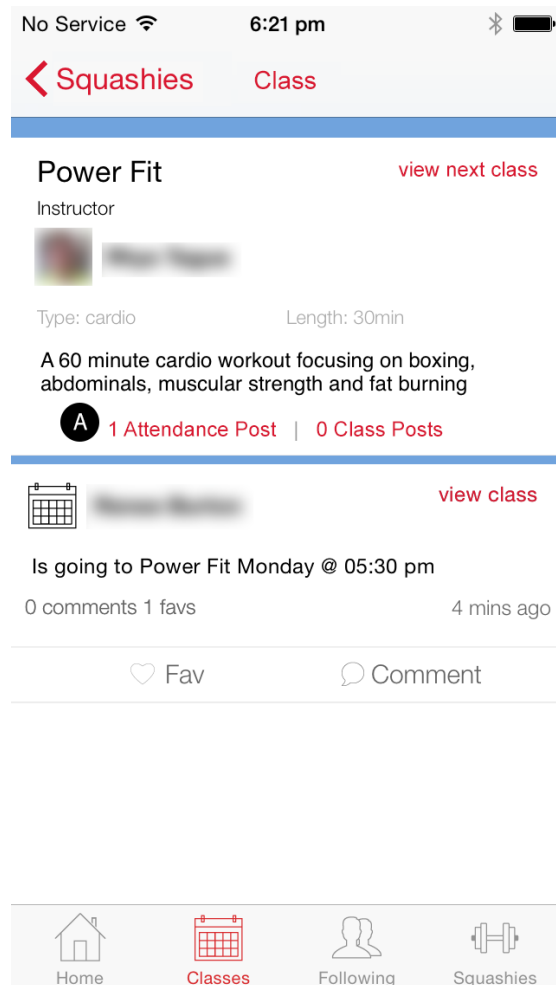


FIGURE 4.12: Screen shot of a class view within the Squashies Application

4.3.3 Explicit User Roles

There were two distinct roles within the Squashies Application: Members of the Gym and Instructors of the Gym. The member role was automatically given to users who signed up for the application. They could be considered a generic user within the application. The instructor role however, had additional functional capabilities to the member role. If a user was an instructor they had the ability to create classes and class sessions. They also had the ability to send a class status to all members who were attending a class session. For the instructor role there was no changes in the form of how the framework was interfaced with when comparing with the member role. However, the IPUSM did capture the actions of an instructor when they created a class and posted to all those attending a class session. The behaviour given to these actions was *publisher*.

For a member to become an instructor they had to lodge a request to become an instructor in the settings page of the application - Point A in figure 4.13. Upon submission of this request, gym staff would have to log in to an administration interface and approve or deny the request.

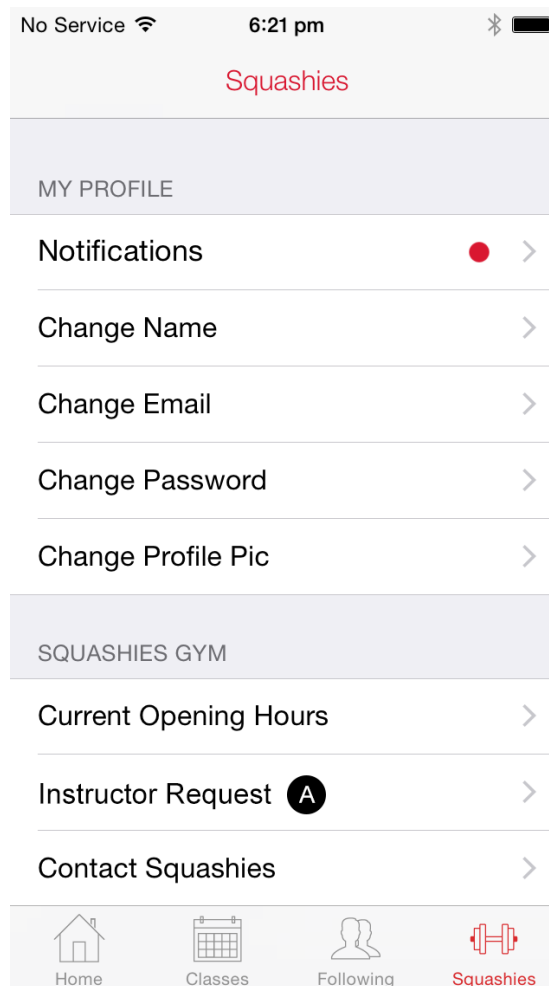


FIGURE 4.13: Screen shot of application settings within the Squashies Application

4.4 Walk 2.0

4.4.1 Description

The Walk 2.0 website was not a case study purposely built for this research. Instead it was a trial arm of a [NHMRC](#) project with the reference number 589903 investigating the use of a *Web 2.0* website promoting health-related physical activity [124]. As a result, the creation of the Walk 2.0 website was outside the

scope of this study. Although the creation was out of research scope, the Walk 2.0 website was used to provide a way to demonstrate the framework's capability to interface with a pre-existing environment while also highlighting its general application and ability to meet the research aims with such an environment.

The interfacing was done post-trial to the Walk 2.0 project. This meant the environment wasn't real time like the other case studies. This resulted in mock interaction events based on the data produced by users of the Walk 2.0 website were used. The mocking of interaction data was through the use of tuple creation timestamps within the Walk 2.0 database. A time stamp would represent when an event occurred and this would be used by the [IPUSM](#) to create an interaction package. One limitation for this approach was the ability to use the characteristic behaviour of *lurker*. This behaviour could no longer be used as the tools and widgets used by the framework to listen for an event to capture the related data were not present. As a result, for this case study a new characteristic behaviour was introduced to align with the objectives of the Walk 2.0 website. This behaviour was *logger*. The *logger* behaviour was chosen to represent any action of logging passive activity.

Although the Walk 2.0 website was not a case study built specifically for this research, the development of the site was carried out by the researcher of this study. The development environment for the website was in [PHP](#). For this reason the framework implementation used for the Squashies case study was used to interface with the Walk 2.0 environment. At the time of interfacing the Walk 2.0 website was no longer in operation publicly and therefore interfacing was done locally on a computer simulating both client and server environment for Walk 2.0. This led to a modified semantic manager and [IPUSM](#) to inspect the tuples within the Walk 2.0 database to establish metadata, and then fire an event for the [IPUSM](#) to capture. Although these changes to Semantic Manager, the use of the framework with the Walk 2.0 website illustrates the capability of the framework to interface with environments not purposely built for the framework.

4.4.2 Collaborative Activities

The purpose of the Walk 2.0 website was to use Web 2.0 tools and widgets for logging and engaging with passive-physical activity such as walking. Users were encouraged to login everyday and log their 'step count' for the previous day. When

they logged their step count they were able to engage with other users as their step count was also published publicly within the environment. Setting goals, both individually and group based, was also a major focus of the Walk 2.0 study and this saw users compete in leader boards and ask questions within a forum for best approaches to achieve activity goals.

Once a user finished registration they were presented with their Walk 2.0 dashboard. The dashboard presented an overview of the interactions within the site. It was also a springboard for users to find and create new information. Figure 4.14 presents a user dashboard and the various areas of interaction a user could undertake. Point A within the figure highlights the core objective of the Walk 2.0 website which was a **passive activity logger**. Users could log steps taken, or the amount of minutes of moderate or vigorous activity completed which then was converted to steps. As the framework could not listen to lurking events because of its use post-trial to the website's availability, the record of logged steps or minutes of activity within the database was used by the semantic manager and this action was given the characteristic behaviour of *logger*. The social-tie used for logging action was the general application identifier and the passive activity logger would be the tool/widget identifier.

In addition to logging steps, the user had the option to add a comment to the log saved. This comment would be published to the 'Stepper Stream', which was a **global activity stream**. Point B in figure 4.14 presents this activity stream upon which the comment would be published. If a user did publish a comment with their log, in addition to the *logger* behaviour, another interaction package was created to capture the log comment. The characteristic behaviour used for the comment was *publisher*.

The Stepper Stream enabled users to directly post general status about their progress to the user base. This was achieved through the input box above the items in the Stepper Stream - as seen at Point D within figure 4.14. If a user published a general, post the action would be similar to a user posting a comment with their logged steps. The resulting interaction package would contain the characteristic behaviour of *publisher* and the application identifier as the social-tie. However, the tool identifier was the Stepper Stream not the activity logger.

There were more than one activity streams available to users. There were also **personal activity streams** for each user and group - as seen at Point A in figure

4.15. These activity streams held items relating to the user or group and provided the same functionality as the global activity stream, Stepper Stream, however the tool identifier was different.

Users were able to interact with the items within an activity stream. They could give an activity a ‘**Thumbs Up**’ if they liked a particular item. The Thumbs Up button can be seen at Point C in figure 4.14. For every Thumbs Up the IPUSM created an interaction package using the behaviour *annotator* and the social-tie of the user who was the creator of the item. If the item was part of a group, two interactions packages were created, one for the user identifier and one for the group identifier.

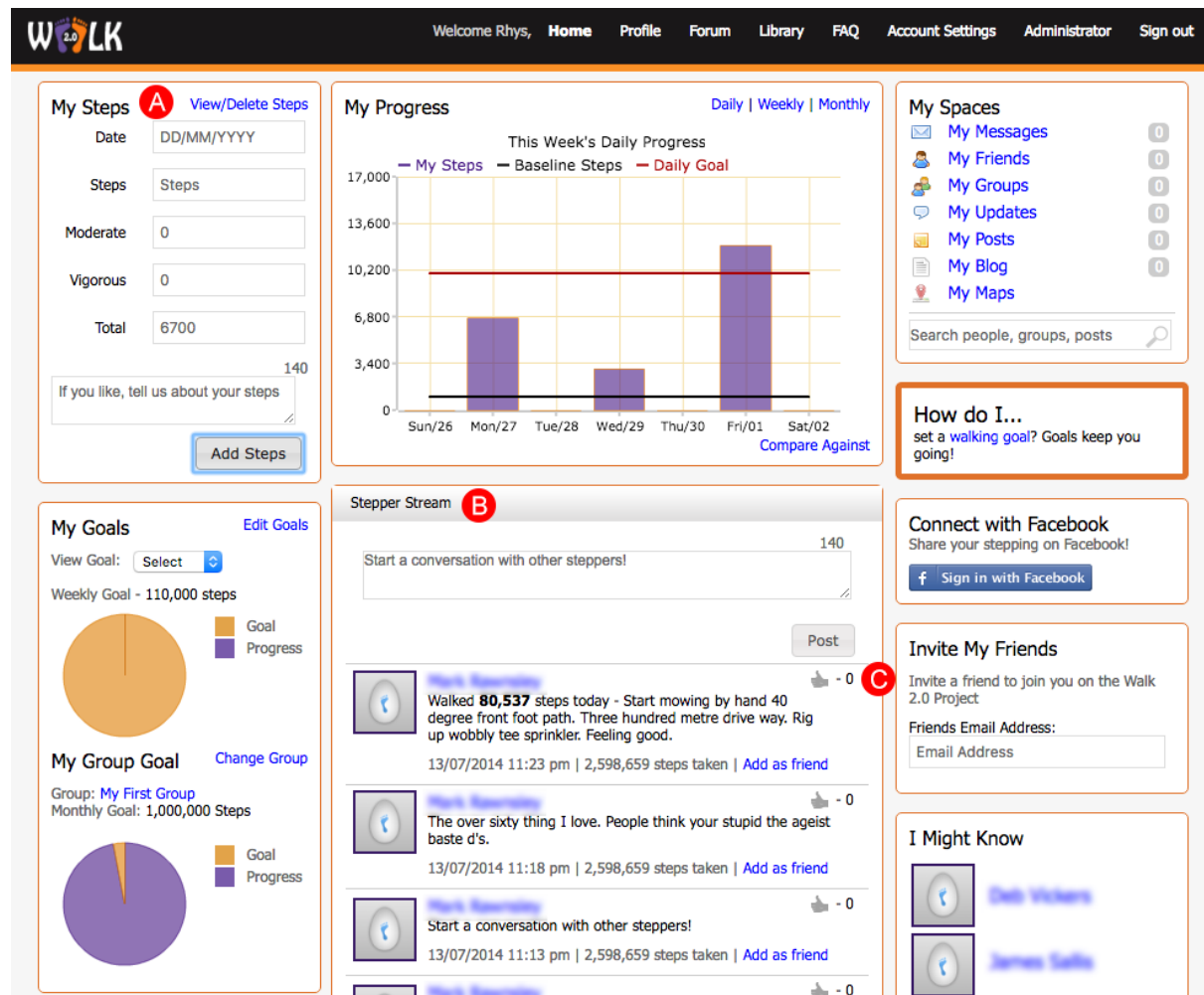


FIGURE 4.14: Screen shot of the site dashboard for the Walk 2.0 website

In addition to ‘Thumbs Up’, users were able to **make comment** on an item. Commenting was only permitted with items where a user was ‘friends’ with the user who created it. Commenting resulted in a similar interaction package to ‘Thumbs Up’. The characteristic behaviour of *annotator* was used, the social-tie was the creator of the activity item, and the tool was **stream item**.

Users were able to join groups to collectively achieve large goals, such as 1 million steps a month. The only collaborative activities a user could perform within a group was post items to the **group activity stream** and interact with those items as previously mentioned - as seen in figure 4.16. As these actions were identical to other activity streams, the *IPUSM* also performed the same behaviour with the other activity streams. However, the semantic manager included the group identifier for any action that occurred within the group. This ensured up to two interactions packages, one for the user identifier and one for the group identifier, were created if a user performed an activity in the group activity stream.

Posting forums were available for users to post and reply to topics of interests. Figure 4.17 presents a screen shot of the various forums available for users to post topics to. If a user posted a topic to a forum the *IPUSM* created an interaction package consisting of the characteristic behaviour of *publisher*, the forum identifier as the social-tie, and the posting tool identifier for the tool/widget. The forum identifier was chosen as the social-tie for a particular forum, as this could be a personal interest of a user. Therefore, to choose a generic identifier, such as the application identifier, the difference of behaviour between various forums would be lost. Users were also able to reply to topic posts posted by themselves or other users. If a user replied to a topic the *IPUSM* would create a package with the creator of the topic as the social-tie, the characteristic behaviour of *annotator*, and the tool identifier for the reply tool.

Users were able to maintain a simplistic **personal blog** about their life. This tool worked very similar to a general status post. However, the user was not limited in character length. They were able to write a post of considerable length and publish to their profile page. Other users were then able to provide comment. If a user published a personal blog post the social-tie used was their own user identifier with the characteristic behaviour of *publisher*. Similar to a general status post, if another user made a comment the social-tie was the user who published the personal blog post and the characteristic behaviour was *comment*.

There were many other collaborative activities available within the Walk 2.0 website, however for this study these were not considered for use with the framework. This was due to the lack of the Walk 2.0 website's user base using these extra activities.

4.4.3 Explicit User Roles

There were no explicit user roles within this case study. Every user who joined had the same functionality as every other user.

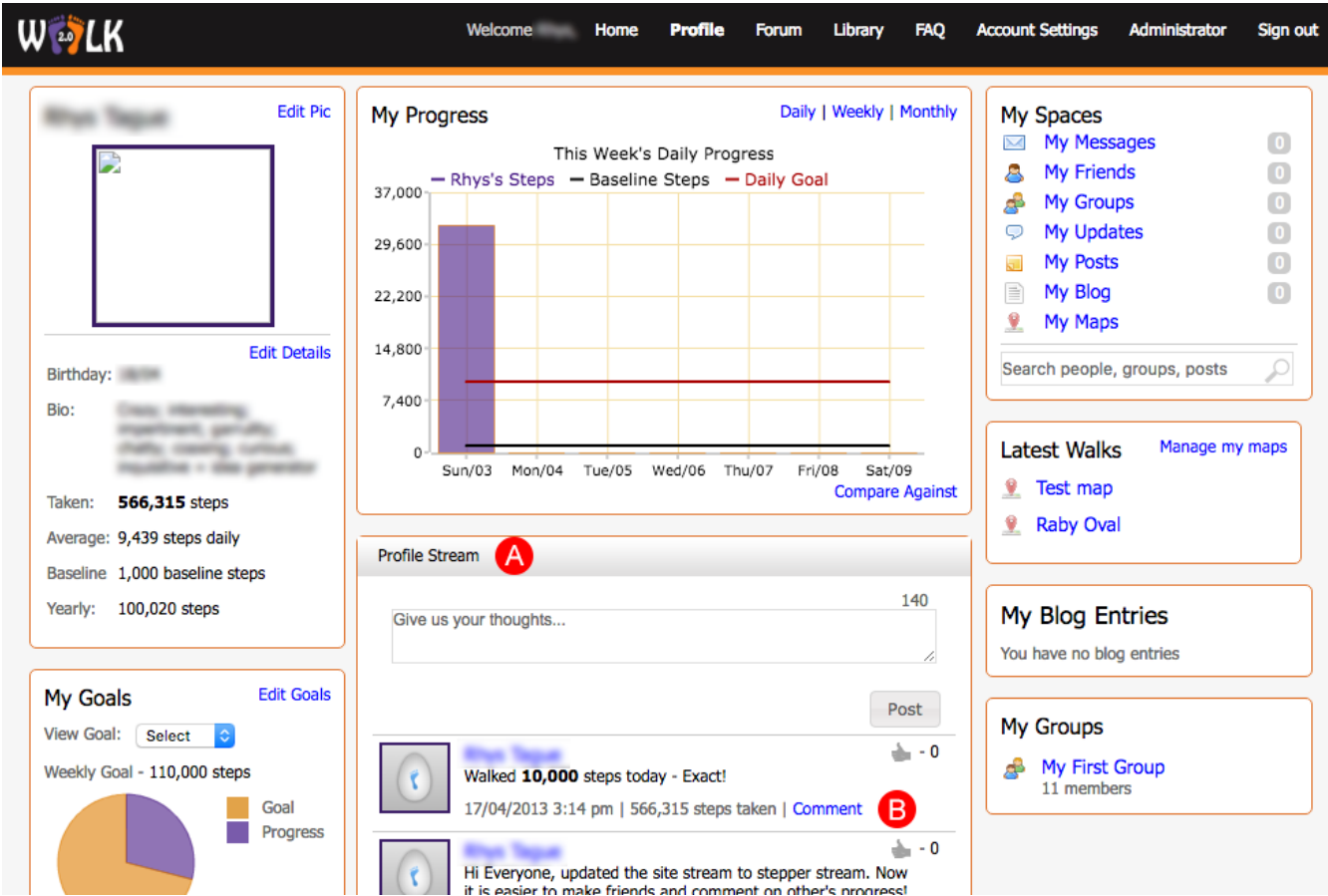


FIGURE 4.15: Screen shot of a profile page for the Walk 2.0 website

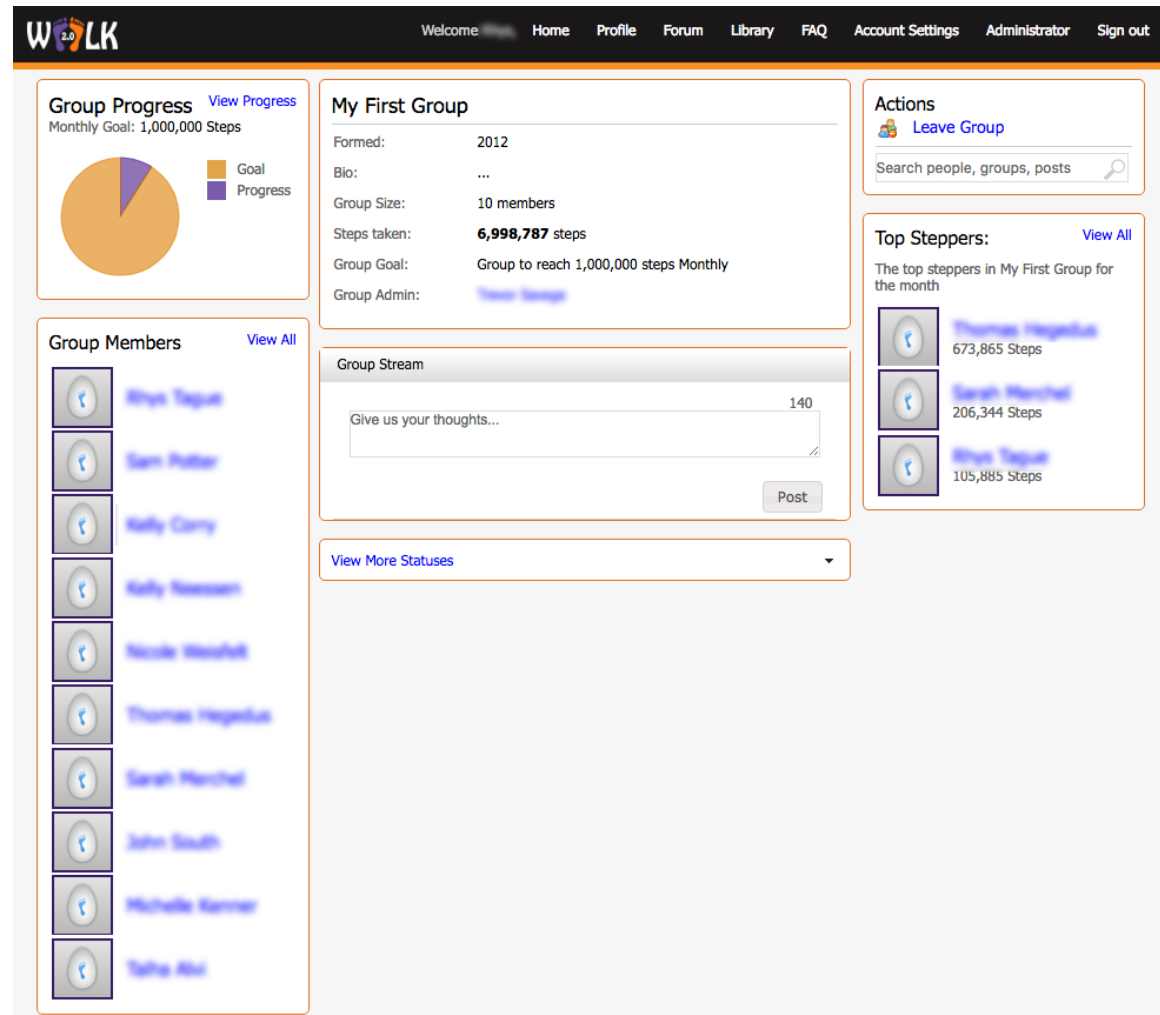


FIGURE 4.16: Screen shot of a group page within the Walk 2.0 website

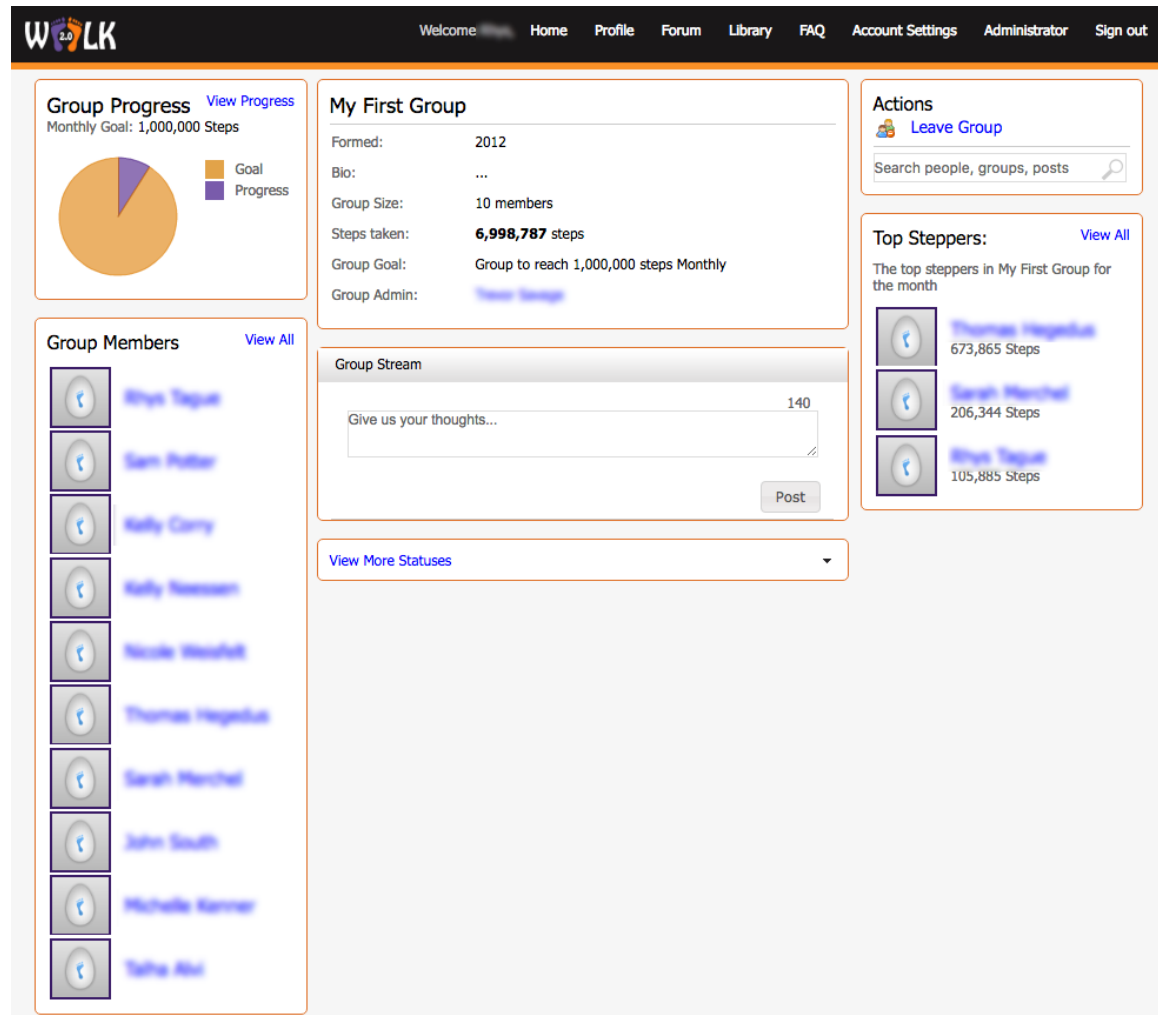


FIGURE 4.17: Screen shot of forums within the Walk 2.0 website

Chapter 5

Participant Use and Data Collection

5.1 Introduction

As each case study involved human participants, ethics approval was needed for recruitment and participation. This chapter briefly introduces the recruitment methods used for each case study and the outcome of those methods. This chapter also presents crude participant use and data created by the participants within the case studies in the form of simple statistics of media produced or interactions shown by all users. This participant use presented will be aligned with the collaborative activities presented in Chapter 4. By presenting the participant use it provides context of how users have used each case study. It will also help provide context and understanding for the framework's output. For instance, the Espressobility case study promoted information creation around geographical locations that served coffee, knowing participant use for this collaborative activity will identify users that aligned with this activity and so allow comparison to the framework's output and the actual participant use.

Participant use for all three case studies will be discussed as each case study had various activities or implicit goals for collaboration that did not exist in the other two. Therefore, presenting this will demonstrate in the next chapter, Chapter 6, how the framework can still create role signatures even though the participant use with an [OCE](#) may be fundamentally different.

The data collected from participant use is in the form of interaction packages and the dynamic structure of the ANN. Therefore, an outline of the amount of interaction packages created by the framework and the resulting storage of the dynamic structure of the ANN will also be presented.

5.2 Espressoability

5.2.1 Recruitment Method

Espressoability was created specifically for this study. As a result a [National Ethics Application Form \(NEAF\)](#) was completed and ethics was approved by the [Western Sydney University \(WSU\) Human Research Ethics Committee \(HREC\)](#) with the approval number H10021. Participants were not engaged to register with the mobile application, instead recruitment was word of mouth, and simple promotional material posted around [WSU](#) campuses - [Appendix D](#). This method of recruitment was used to promote a natural growth in the user base over time. The application was released in the Android Play Store on the 18th of October 2013 and taken down on the 7th of October 2015. This resulted in an ‘up time’ of 1 year 11 months 2 weeks and 5 days. Over this period recruitment was open to all individuals with a mobile device with the Android operating system.

5.2.2 User Registrations

When the application was taken down from the Android Play Store there were 36 users. Of these 36 users, 6 users were highly active (users who were daily or weekly users). The low user count is believed to be a result of the technology, which was *PhoneGap*, that caused inconsistencies with the application’s user interface and incompatibilities across devices. These inconsistencies and incompatibilities were reduced through application updates over the life of the application. There were 7 updates handling bugs and device incompatibilities.

Of the registrations 85% occurred within the first 3 months of the application’s release to the Android Play Store. After this period the remaining registrations occurred erratically with the last registration on the 21st of March 2015.

5.2.3 Participant Use

Although there was a small user base for *Espressobility*, the use of the application was high. This resulted in dynamic and continually changing content for the users to collaborate with.

Upon launch of the application there were no coffee locations for review or interaction. When the application was taken down from the Android Play Store there were 179 unique coffee locations distributed over Australia in New South Wales, Queensland, and Melbourne. Figure 5.1 shows the distribution of locations created by individual users within the application. From this, one can see the 179 locations were created by 10 users, with a single user creating 86 of those locations. Using the user identifier as a representation of when a user registered, early users of the application contributed the most locations.

Every user had the ability to carry out an activity towards a location, for instance, a coffee review. From 17 users 1,126 activity items were created. Figure 5.2 presents the distribution of activity creation amongst these 17 users. One user, user 3, accounted for the creation of 47.7% of all activities. Although a high activity creation count, this particular user had an average of 6.11 and a standard deviation of 6.77 activities over 88 locations. When compared to other users this average was consistent with the following top 3 active users who had an average of 6.85 and a standard deviation of 6.48 activities per location. This means although one user contributed 47.7% of activities, their activity creation per location was consistent with the most active users of the application. The only difference being user 3 visited more locations, which is consistent with their location creation count.

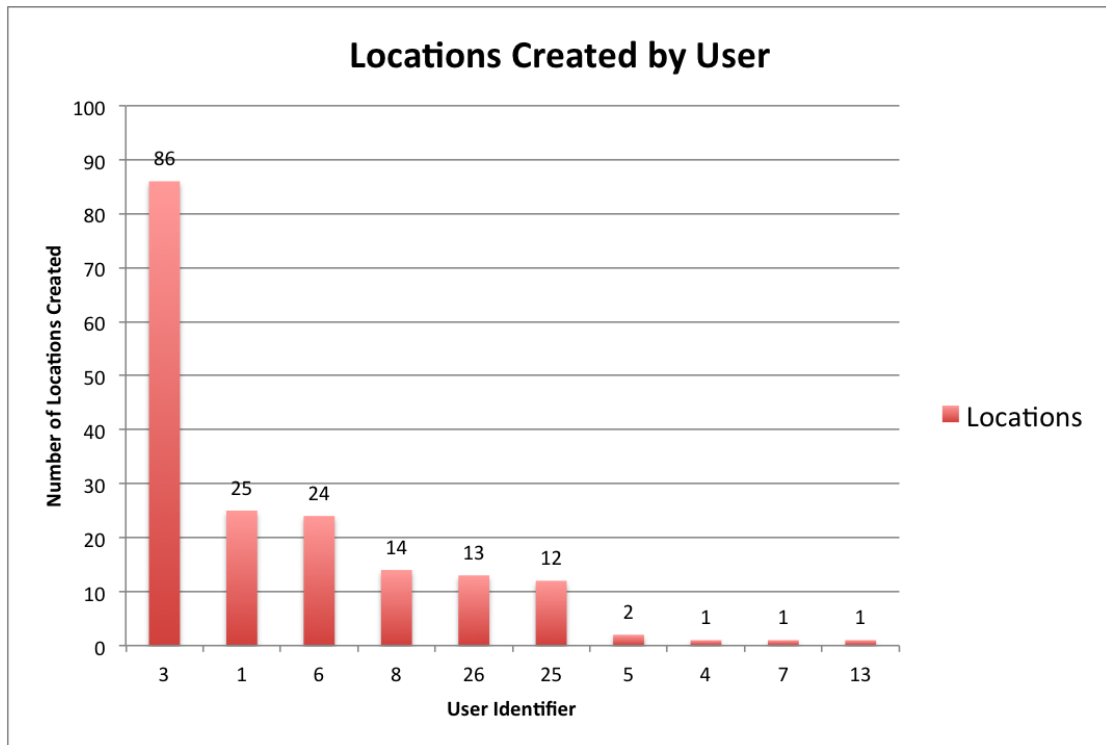


FIGURE 5.1: Bar graph showing coffee locations created by users of Espresso-bility

The average amount of activities published for a particular location was 6.66, however the standard deviation was 14.44. This was a result of some locations having high activity counts compared to others. For instance, one location had 118 activities published. This location was a local Cafe near one of the campuses of [WSU](#) and therefore had high traffic from users. This location and the high traffic indicates many users may be related to [WSU](#). This could be a result of the recruitment posters posted around the [WSU](#) campuses.

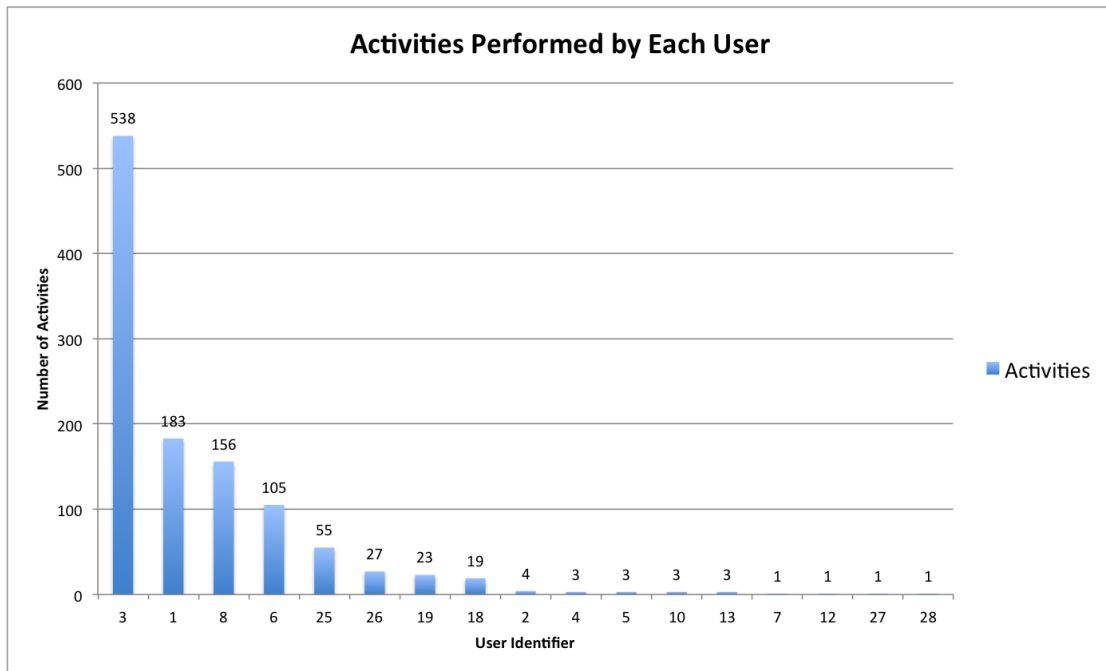


FIGURE 5.2: Bar graph showing activities created by users for locations within Espressoability

Figure 5.3 presents the distribution of the different activity types published within the application. From this one can see that of all the activity types available, the text review type was favoured by the users with 53% of all activities produced being a location text review.

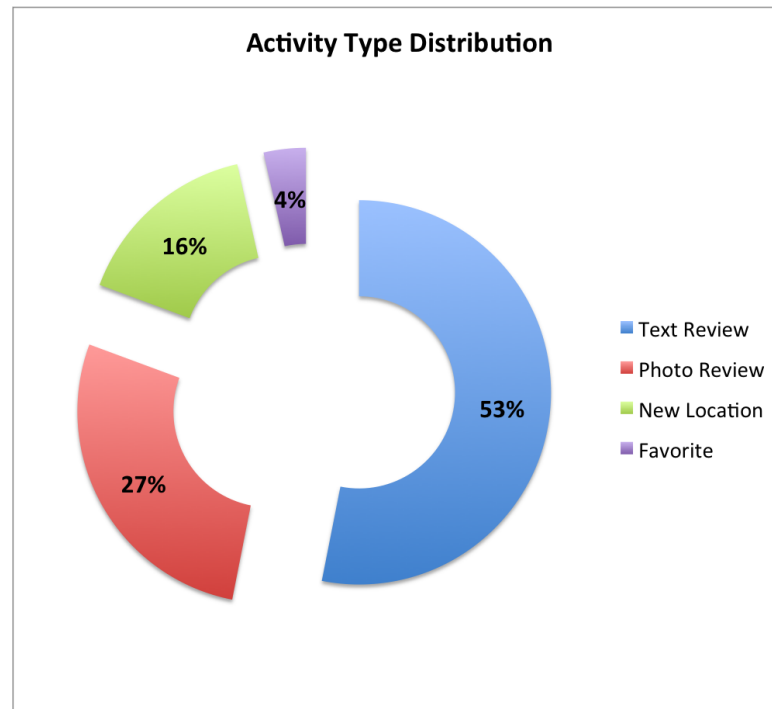


FIGURE 5.3: The percentage of the different activity types created within the Espressobility application

Of all activities created there were 276 activities which received one or more comments. In total there were 648 comments and an average of 2.34 comments made for each activity with a comment. 11 users were responsible for these comments, with one user creating 226 comments. Figure 5.4 presents the distribution of comments made by users. It is interesting to note with user 3, the user who had created the most locations and activities, was not the highest in terms of comments produced. Instead, they produced only 90 comments.

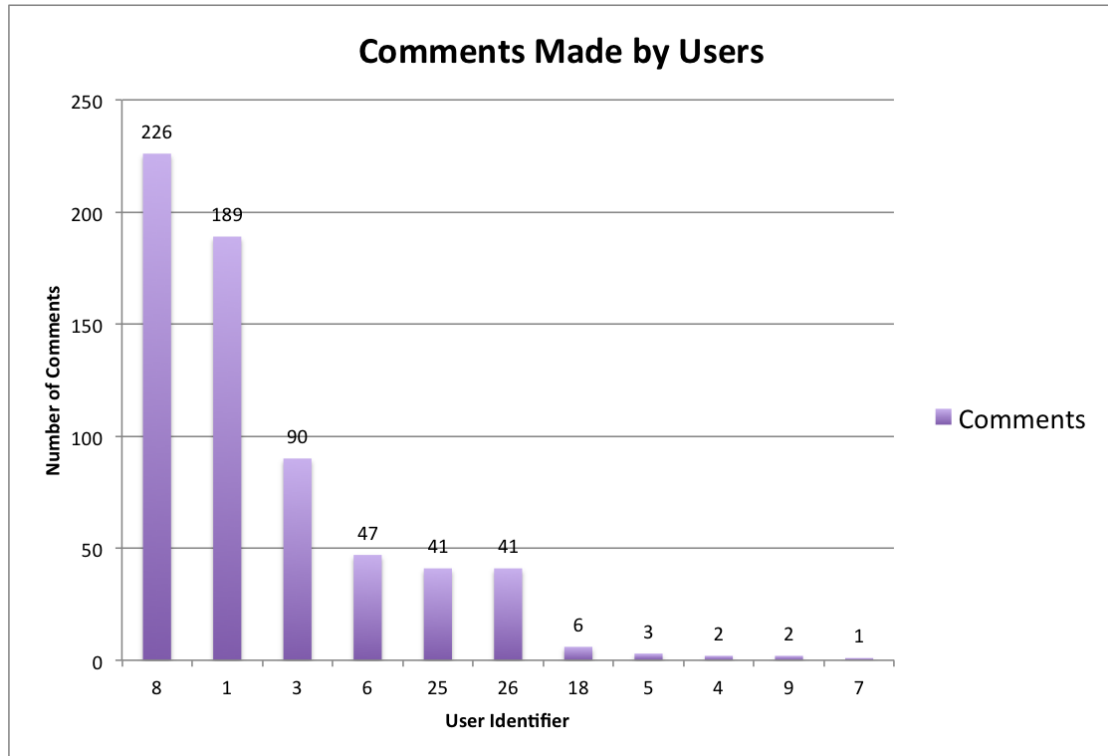


FIGURE 5.4: Distribution of comments made by each user within Expressobility

Tagging of activities was not common amongst users. Only 44 tags were produced by 6 users across 26 activities. Of the 44 tags, one user was responsible for 32 tags.

As shown, the use of the application can be considered high for a small user base. With this brief outline of participant use within Expressobility a context can be established for the output determined by the framework.

5.2.4 Data Collected

The data collected for the framework was a result of participant use. As outline in Chapter 4, the semantic manager and the [IPUSM](#) listened for events to occur within the application upon which an interaction package was created. These interaction packages were then sent to the framework domain via a commit request. From the above usage, as well as the addition of events within the application which did not create data, such as consuming media, 4,892 interaction packages were created and committed to the framework domain. This resulted in 259 nodes and 2,038 synapses for the [ANN](#). Although there is a large set of

nodes and synapses, as explained in Chapter 3, the dynamic structure of the ANN only uses relevant nodes related to the passed input selectors and therefore not all nodes will be used for each signature request.

The capability of the framework to take snapshots of the dynamic structure led to 307,912 synapse weight records. This data enables the ability to build the dynamic structure of the ANN through date ranges.

The resulting characteristic behaviours for role signatures within the framework, based on the behaviours presented by the users of the application were *publisher*, *annotator*, and *lurker*.

5.3 Squashies

5.3.1 Recruitment Method

The NEAF application created for Espressobility was amended to include the Squashies application and participants in the form of gym instructors and gym members. Amendments were approved and the recruitment method of participants was similar to Espressobility. Instead of recruiting at WSU campuses however, the gym had promotional material posted around high traffic zones within the gym - Appendix E. This promotional material highlighted the benefits of using the application with a gym membership.

The application was posted to the Apple App Store on the 16th of February 2015. On the 24th of February 2016 the Apple App Store released the application to the general public within the Australian region. For a period of 12 months the application ran uninterrupted until it was taken down on the 24th of February 2016. During this period registration for the application was open to any member of the public, however registration was focused on active members of the Squashies gym. During this recruitment period 63 users registered for the application. Of these 63 users, 57 were general members, and 6 became instructors.

The application was taken down to remove the capabilities of interfacing with the framework and allow Squashlands Gym and Fitness to take over management of the application to further develop it for their needs.

5.3.2 Participant Use

As gym classes and announcements were available through the Squashies mobile application the use of the application was high. This remained high throughout the life of the application until it was taken down from the Apple App Store.

Staff initially created a timetable of different class types for members to join and interact with. Eventually, those users who were granted the role of a gym instructor took ownership of the creation of class types and class sessions. At the end of the application's life there were 26 unique class types. Although there were a small amount of classes there were 699 class sessions. Given a class session could be once a week this is a substantial amount of sessions users could interact with. For the time the application was available, 12 months, this resulted in an average of 13.44 class sessions per week. Each instructor had an average of 2.24 class sessions per week.

Users were able to interact with class types created by instructors by indicating they were attending a session or posting a message in relation to a class session or class type. Not including gym instructors, 21 members indicated they were attending 271 class sessions with an average of 12.90 class sessions each. This attendance count is considered high as the user base of Squashies was a small subset of the actual gym membership, an estimated 1,500 members, and not all members attend gym classes. Figure 5.5 presents this distribution of class sessions attended by users. The first 6 users within this figure were gym instructors. Their attendance was naturally high as they always indicated they were going to their own class session.

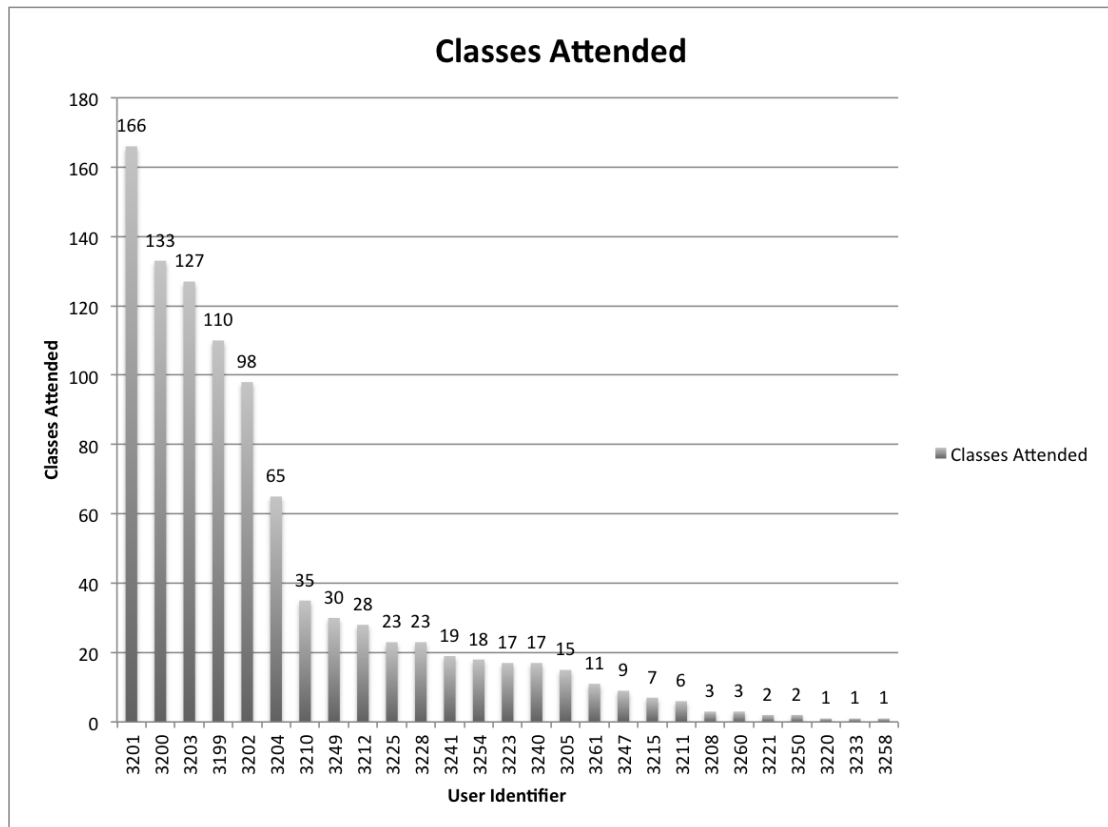


FIGURE 5.5: Number of class sessions attended by each user within Squashies

Indicating attendance to a class session was higher than posting a class message. Both types of users created a total of 114 messages towards 13 classes. 45.61% of the class messages were produced by the instructors relating to their class session or class type.

Including class session attendance and class messages, users created 3,141 items in total which were presented in various activity streams. Figure 5.6 presents the distribution of these items created by the users. The Check in activity was performed the most with 1,332 items, followed by class attendance with 970. Where users performed an activity relating to themselves, such as a status post, the number of items were low.

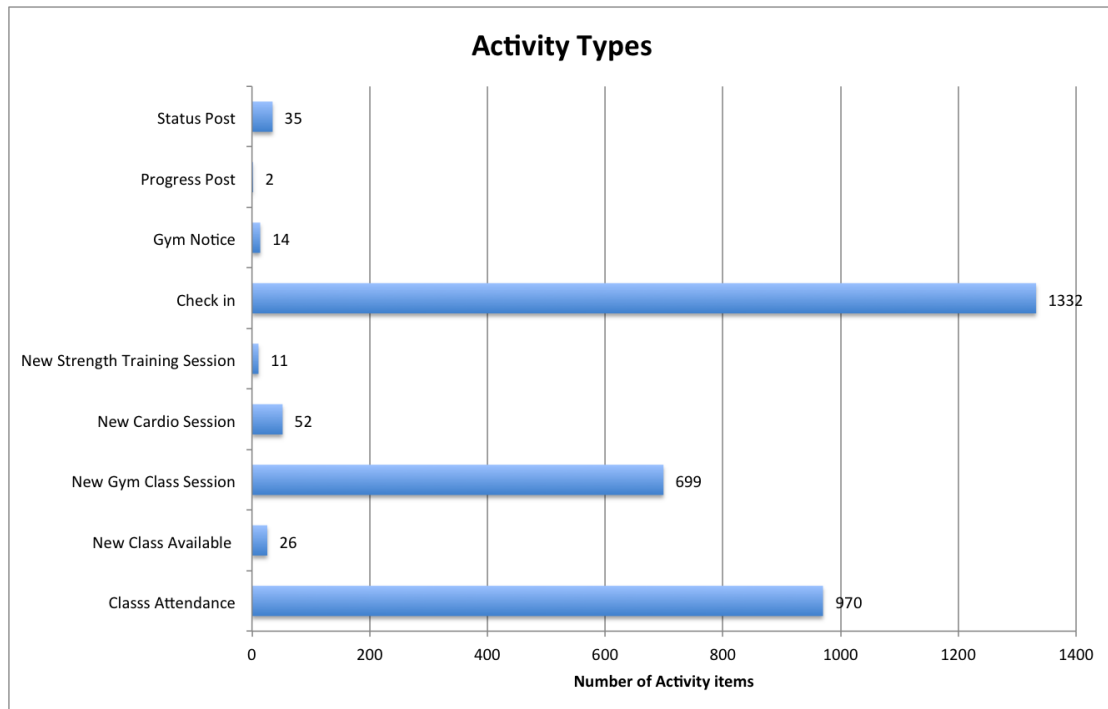


FIGURE 5.6: Distribution of activity types within Squashies

Each activity item created enabled other users to interact with it by commenting or making the item a favourite item. Figure 5.7 presents the number of actions per activity type within Squashies. The ability to favourite an item was used strongly across all item types. Commenting was also seen on every type, yet it was considerably less than a user making an item a favourite. An example of this was with the activity type Check in. Making an item a favourite was carried out 512 times, however only 128 comments were made.

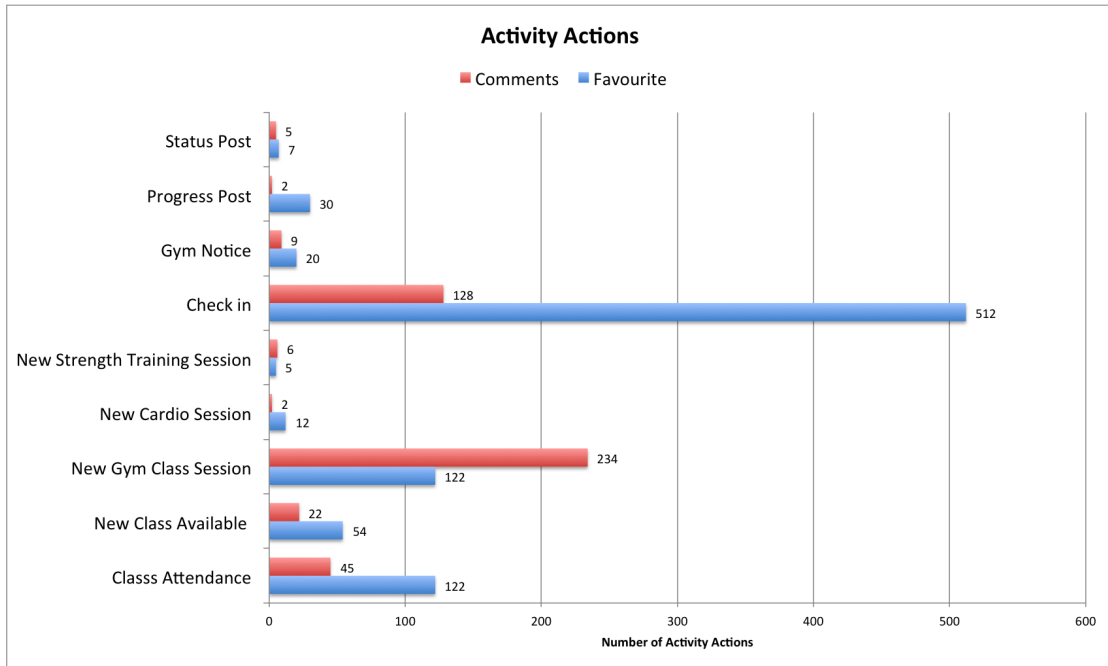


FIGURE 5.7: Number of activity actions per activity type within Squashies

5.3.3 Data Collected

Participant use with the collaborative activities outlined in Chapter 4 for Squashies the framework created 9,238 interaction packages. Through commitment requests these interaction packages resulted in 1,043 nodes and 7,203 synapses for the ANN. The snapshots of the ANN created 608,239 synapse weight records.

From the available characteristic behaviours within the meta data that could be demonstrated within Squashies the framework could use three characteristic behaviours for creating role signatures: *publisher*, *annotator*, and *lurker*.

5.4 Walk 2.0

5.4.1 Recruitment Method

The recruitment method used by the Walk 2.0 study is outside the scope of this project. However, ethics was received by the WSU HREC with the ethics approval number H8767. It was a registered trial with the Australian New Zealand Clinical

Trials Registry Number: ACTRN12611000253909, and World Health Organization Universal Trial Number: U111-1119-1755.

The recruitment resulted in 1,205 users with the first user registering on the 22nd of February 2011 and the last user registering on 30th of June 2014. The recruitment was not consistent because of the stages and different trials conducted for the Walk 2.0 study, however as the framework was interfaced with the final dataset of the study the entire user base will be processed and analysed between the first and last user registration.

5.4.2 Participant Use

The purpose of the Walk 2.0 website was to log passive physical activity, such as walking. With this purpose users managed to log 62,607 sessions of passive physical activity. These logs resulted in 791,969,995 steps taken by all users. Of the 1,205 users who registered for the Walk 2.0 website only 544 logged passive physical activity. Of these 544 users, the average number of logs per user was 112.05 and a standard deviation of 81.43. The highest log count by a single user was 924 logs.

Logging passive physical activity within the Walk 2.0 website was a personal activity, yet a user had the option to publish their log to an activity stream with a comment. There were 4,350 of the 62,607 logs that were published to the public activity stream. Only 174 users of the 544 users who logged their activity published their log with a comment. The logs published made up 83.66% of all items that were published. This can be seen in figure 5.8 where the distribution of the activity types are presented.

For the activity type Status Post 727 items were published, however this activity type had the ability to be published to different activity streams. Of the 727 items published 57 were published directly to a user's profile activity stream and 126 were published within a group. Figure 5.9 presents this segmentation of Status Posts based on their visibility.

For all types of activities published, figure 5.10 presents the top 20 users and their count of activity items posted to the community. From this figure the top 20 users achieved a publication count with an average of 141.5 posts per user.

This demonstrates the top 20 users were highly active users when it came to publications of various types of activity items.

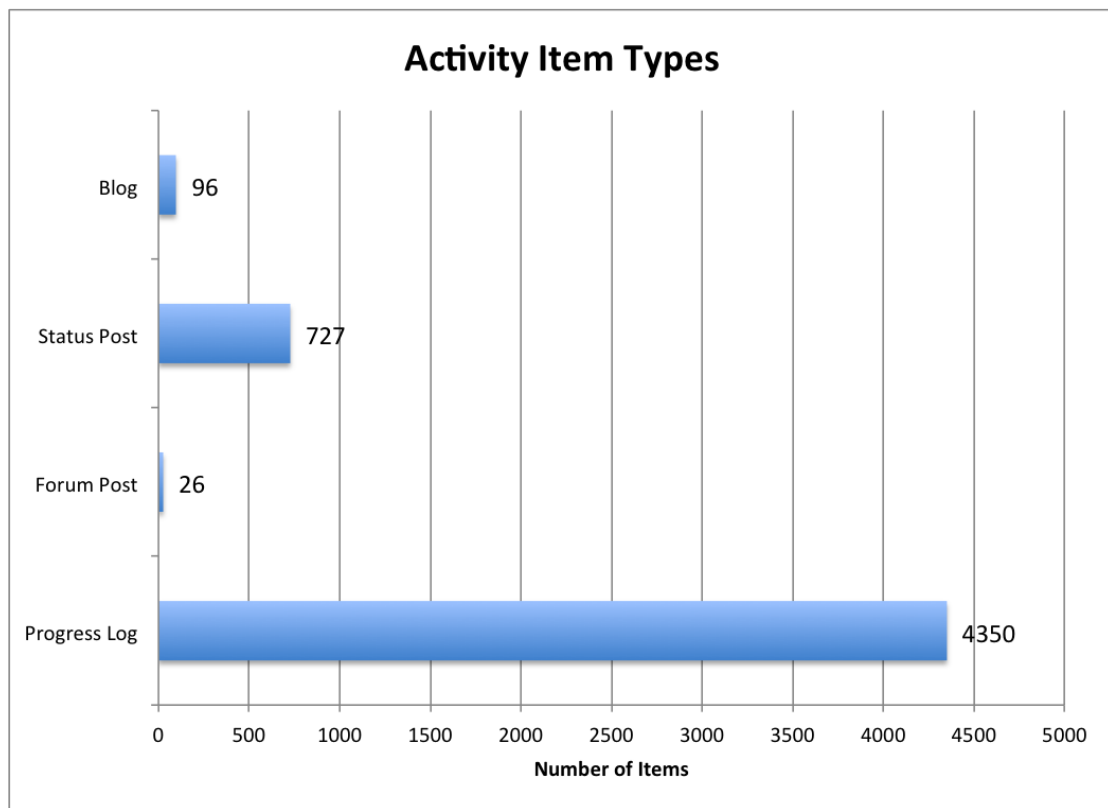


FIGURE 5.8: Distribution of activity item types within the Walk 2.0 website

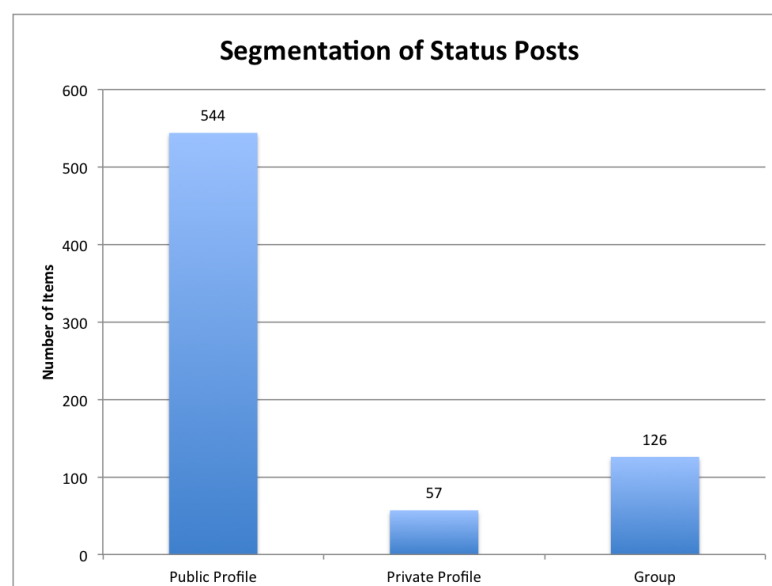


FIGURE 5.9: Segmentation of visibility for activity type Status Post within the Walk 2.0 website

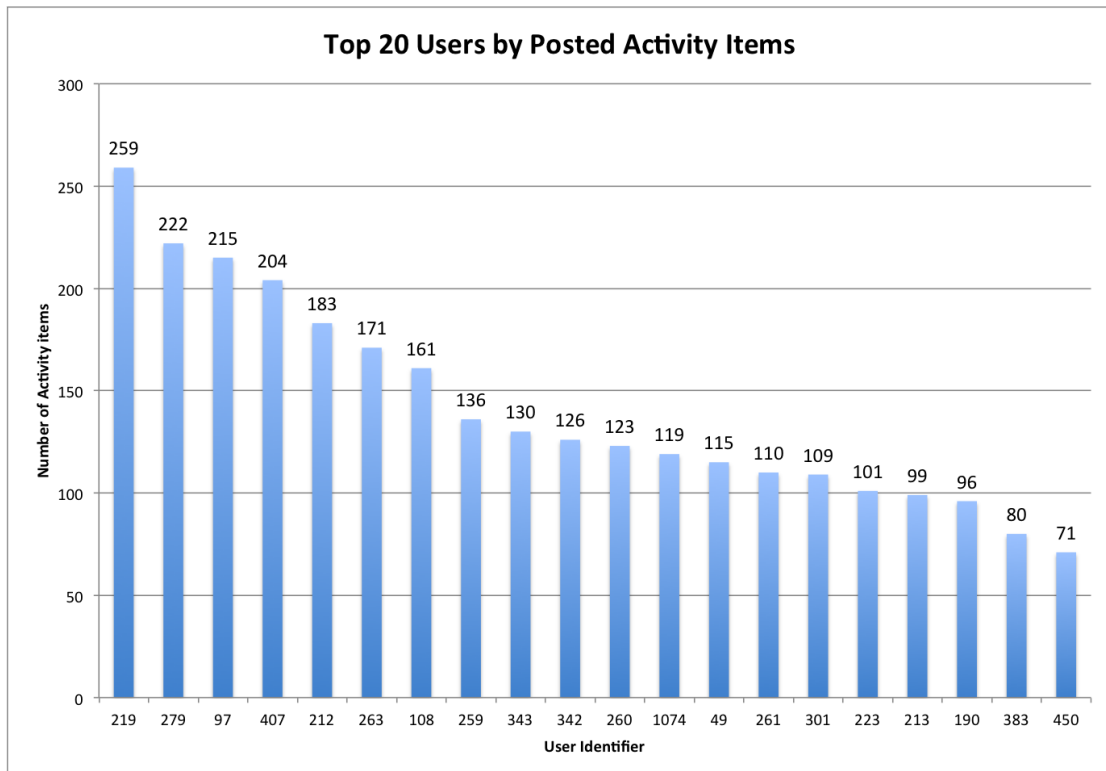


FIGURE 5.10: Total activity items posted by the top 20 users within the Walk 2.0 website

Users were able to interact by providing comment and ‘Thumbs up’ for all activity types. There were 1,606 comments made by 110 users towards activity items. An average of 14.60 comments were made by each user, with a standard deviation of 35.84. The figure 5.11 presents the top 20 users with the highest count of comments made. One can see the first 4 users accounted for 44.58% of all comments made, followed by a consistent decrease in the amount of comments made by each user.

The ability to provide ‘Thumbs Up’ to an activity item allowed 109 users to tag 2,501 activity items and give 4,265 ‘Thumbs up’. When compared to comments, the use of the ‘Thumbs Up’ was substantially more with an average of 39.12 ‘Thumbs Up’ per user. As a ‘Thumbs Up’ could only be done once for each activity item for each user it meant an average of 39.12 activity items had interaction by a single user. Figure 5.12 presents the top 20 users with the highest amount of ‘Thumbs Up’.

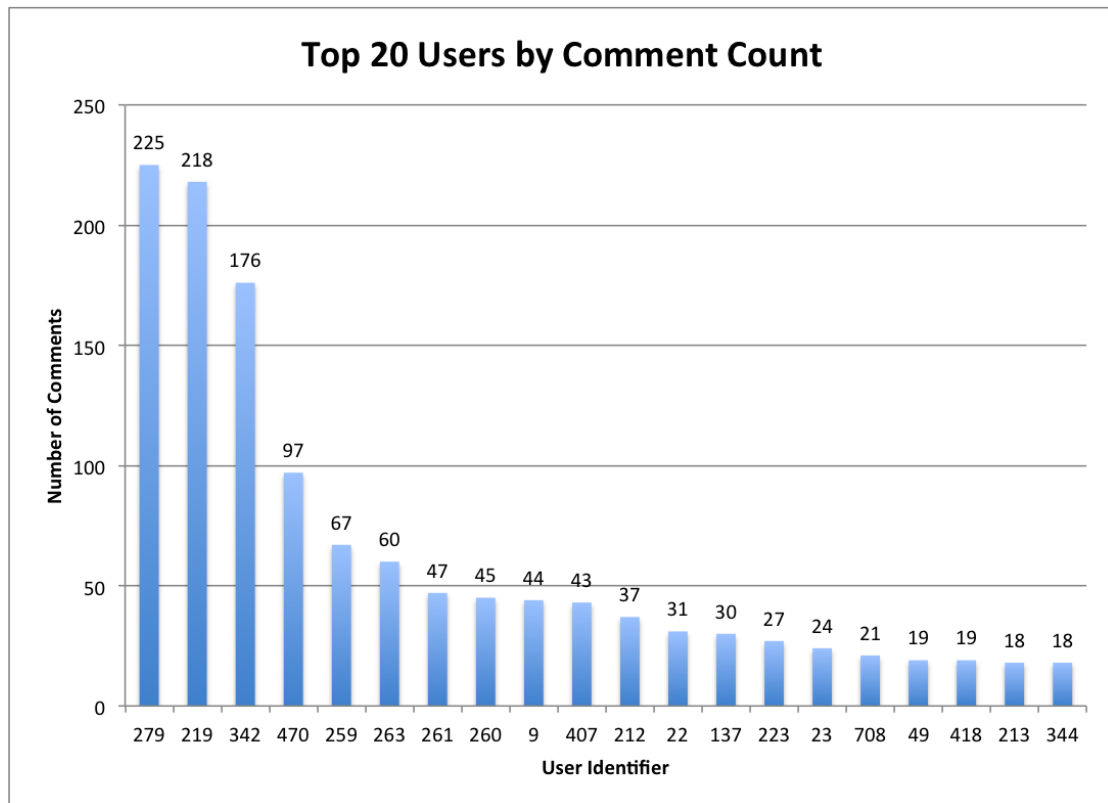


FIGURE 5.11: Top 20 users with the most comments for activity items within the Walk 2.0 website

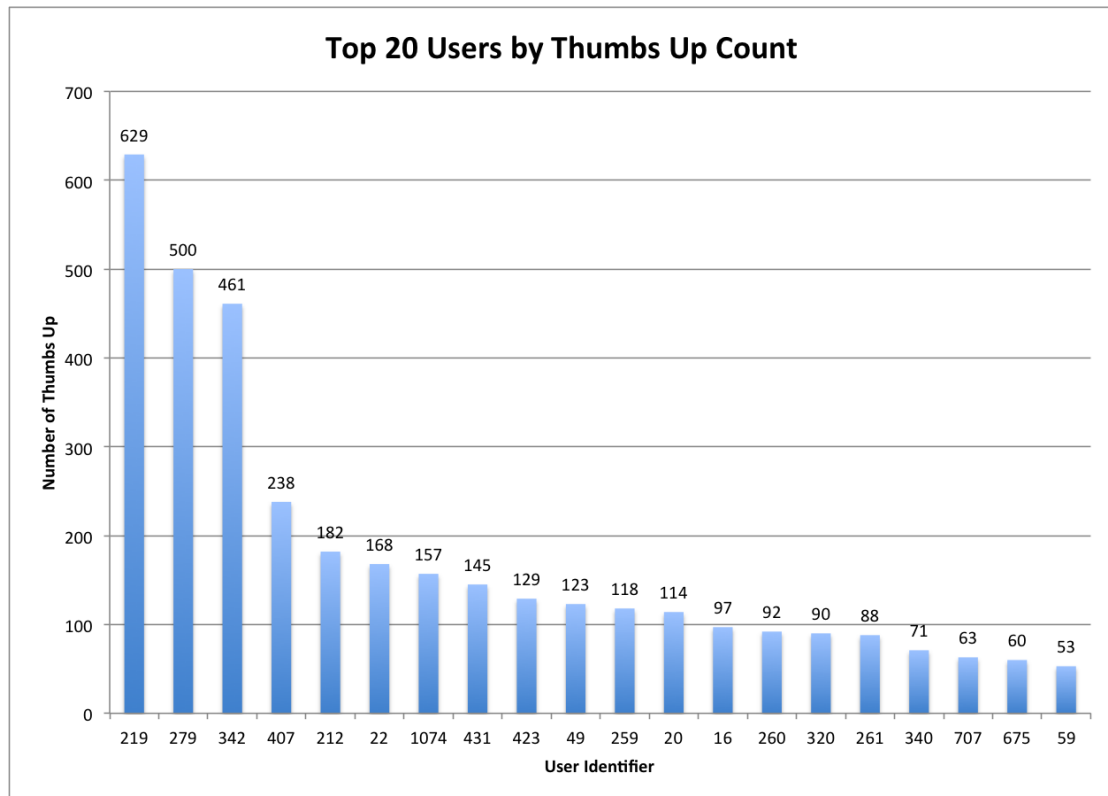


FIGURE 5.12: Top 20 users with the most thumbs up for activity items within the Walk 2.0 website

5.4.3 Data Collected

As the framework was interfaced post-trial to the Walk 2.0 study, this led to the framework not being able to capture all implicit interaction (interaction that did not result in saved data). It also led to the framework being modified to handle simulated events based on the data within the Walk 2.0 database. The modification was a script that traversed the records of the relevant tables within the Walk 2.0 database which would manually fire events so the semantic manager could create metadata for each record. As the tool identifier was not associated with the records, these were manually assigned within the script when processing each table. Once all records were processed for all tables the temporary metadata representations were organised into chronological order. The script would then fire an event manually for each metadata representation so the [IPUSM](#) could capture and handle it. Once the [IPUSM](#) handled an event the normal life cycle of the framework was carried out.

From the data within the Walk 2.0 database 79,748 interaction packages were created. After commitment requests were made to the framework 6,023 nodes with 32,537 synapses were created. As a new user session could not be identified because of post-trial analysis, a snapshot of the ANN would only occur if a previous snapshot with the same input was older than 24 hours. Snapshots of the ANN led to 947,927 synapse weight records for the 32,537 synapses.

The final characteristic behaviours available to the semantic manager for the tools and widgets were: *publisher*, *annotator*, and *logger*. Through the actions of the users all three characteristic behaviours were demonstrated by users allowing the framework to acquire all three behaviours for use in determining role signatures.

Chapter 6

Social Group Role Signatures - Framework Output

6.1 Introduction

The data collected from participant use was used by the framework to create an output in terms of a social group role signature composed of characteristic behaviours for individual users. This chapter presents an analysis of this by demonstrating the framework's ability to:

- determine a general social group role signature for a user
- determine a social group role signature for a user with context
- accommodate new characteristic behaviours over time that may exist within an [OCE](#)
- handle the dynamic and changing structure of an [OCE](#) in terms of user social-ties and tools/widgets

These capabilities will be achieved by analysing interesting users or groups of users with the determined role signatures from the framework for each user and then comparing a user's actual use within a case study. Interesting are those users who demonstrate unique or common patterns of interaction towards other users or within a case study that can demonstrate the frameworks capability.

Not all signatures for all users will be presented. This is because the framework can only produce one signature for one set of input per role signature request. Also, because of the large number of social-tie combinations a user may have, it would not be feasible to present all possible signatures.

6.2 General Signatures

The framework is capable of determining a social group role signature for each input combination experienced by the framework. This experience is achieved from the semantic data and the generated interaction packages. Although the semantic data and interaction packages create the input combinations, which consist of the user identifier, widget/tool identifier, and the social-tie, not all three inputs are needed to determine a signature. This is a result of the dynamic capability of the framework and its ability to build a signature based one or more inputs.

As a user can have many social-ties and use many tools there is no single unique role signature for a user. This is because role signatures change based on the context provided for collaboration. This context is create with the widget/tool identifier and social-tie. Although there is no single unique role signature, a general signature still can be achieved. A general signature is defined as a signature without context, only the user identifier is used within a signature request to determine a social group role signature. When this type of signature request is made the framework acquires only the relevant experience towards the user without context.

Tables 6.1, 6.2, and 6.3 present the resulting general role signatures of the top 10 most active users within each case study. The tables also present the ANN result for a historical and current snapshot. The days between snapshots are based on a user's latest user session and the one preceding. As a result, each row in this table has a unique time period as new user sessions were created by users at different time periods. These results were achieved by the framework by a signature request **with only the user identifier** present as input.

When comparing the resulting role signature for user 3 within table 6.1 and their general use of the environment, there is a match between a user's general participant use and determined characteristic behaviours. For instance, with 538 posts of activities and 86 locations created, user 3 had the most publications of all users

within the Espressoability case study. Their *publisher* strength was determined to strong at 85%. When comparing their *annotator* strength it was neutral at 55%. This also matches their overall annotations as it was considerably less to their publication count with 90 comments made. Although the *lurker* characteristic behaviour did not result in any data to compare, the weakness of this behaviour at 32.50% does suggest the user did not take time browsing and instead focused on their ability to publish content. This suggestion will be explored further when context based signatures are determined for this user in section 6.3.

Another interesting user within table 6.1 is user 8. The framework has shown the user was weak with the *publisher* characteristic behaviour at 40% yet strong with both *annotator* and *lurker* at 85%. When comparing their use of the environment this is also reflected with their comment count. They had the highest comment count with 256 comments. They also had the third highest post count of activities within the case study, however the framework has determined their *publisher* behaviour to be weak. This is because this general signature is the final signature determined by the framework. If the framework is given a date range within the signature request the framework is able to use historical snapshots to present past general signatures. This ability can be seen in figure 6.1 where general signatures are presented for user 8 over the user's life within the case study. Within this figure one can also see the variability of the user's general signature. Their *publisher* characteristic behaviour was a consistent neutral to weak strength, yet their *annotator* behaviour after 12 months was consistently strong. This matches their participant use also. 64.32% of all comments made by this user was within the last 12 months while their creation of locations and posting of activities was small with 42 over a period of 6 months.

	<i>Historical ANN Results</i>			<i>Latest ANN Result</i>				<i>Inferred Signature</i>		
User	Publisher	Annotator	Lurker	Publisher	Annotator	Lurker	Days	Publisher	Annotator	Lurker
3	0.78	0.1	-0.25	0.82	0.1	-0.1	4	85.00%	55.00%	32.50%
1	0.4	0.35	0.1	0.2	0.5	0.15	10	85.00%	79.54%	55.00%
6	0.23	-0.1	0.3	0.36	-0.15	0.28	18	70.00%	29.28%	71.36%
8	0.15	0.5	0.7	-0.02	0.6	0.73	5	40.00%	85.00%	85.00%
26	0.1	0.05	0.25	0.02	0.14	0.28	21	40.00%	48.07%	65.00%
25	0.06	-0.22	0.56	-0.17	0.16	0.31	28	27.06%	25.00%	74.41%
5	-0.33	-0.28	0.18	-0.27	-0.43	0.22	30+	25.00%	25.00%	55.00%
4	-0.68	0.78	0.53	-0.53	-0.63	0.42	14	25.00%	25.00%	85.00%
7	0.72	0.38	0.13	-0.32	0.1	0.27	17	55.00%	72.14%	68.40%
13	0.68	-0.79	0.52	-0.05	-0.81	0.61	22	55.00%	25.00%	85.00%

TABLE 6.1: Top 10 most active users and their general social group role signatures within the Espressoability case study

	<i>Historical ANN Results</i>			<i>Latest ANN Result</i>				<i>Inferred Signature</i>		
User	Publisher	Annotator	Lurker	Publisher	Annotator	Lurker	Days	Publisher	Annotator	Lurker
3201	0.76	-0.13	-0.28	0.78	0.12	-0.01	3	76.81%	25.00%	25.00%
3200	0.36	0.23	0.14	0.41	0.14	-0.05	2	85.00%	61.27%	37.00%
3203	0.15	0.43	0.55	0.28	0.24	0.29	4	70.00%	85.00%	85.00%
3199	0.30	0.13	-0.21	0.32	0.05	0.12	6	75.76%	48.07%	55.00%
3202	-0.28	0.42	0.23	-0.05	0.23	0.33	7	37.00%	85.00%	77.07%
3204	0.22	-0.23	-0.15	-0.15	0.05	0.01	13	32.49%	41.15%	40.30%
3210	0.32	0.13	0.71	0.38	0.09	0.53	3	82.00%	53.42%	85.00%
3249	0.21	-0.73	0.32	0.24	-0.74	0.43	9	63.18%	25.00%	85.00%
3213	0.01	0.14	0.61	0.03	0.22	0.59	16	43.26%	59.28%	85.00%
3225	-0.12	0.28	0.56	0.06	0.22	0.67	2	47.00%	70.00%	85.00%

TABLE 6.2: Top 10 most active users and their general social group role signatures within the Squashies case study

	<i>Historical ANN Results</i>			<i>Latest ANN Result</i>				<i>Inferred Signature</i>		
User	Publisher	Annotator	Lurker	Publisher	Annotator	Logger	Days	Publisher	Annotator	Logger
219	0.78	0.65	0.86	0.80	0.62	0.87	3	85.00%	85.00%	85.00%
279	0.52	0.76	0.82	0.48	0.72	0.84	5	85.00%	85.00%	85.00%
97	0.34	-0.43	0.72	0.28	-0.49	0.74	9	78.33%	25.00%	85.00%
407	0.58	0.13	0.72	0.51	0.07	0.73	12	85.00%	50.71%	85.00%
212	0.45	0.13	0.59	0.35	0.04	0.64	4	85.00%	46.81%	85.00%
263	0.48	-0.23	0.71	0.53	-0.28	0.73	14	85.00%	25.00%	85.00%
108	0.12	-0.72	0.71	-0.01	-0.74	0.72	18	41.05%	25.00%	85.00%
259	0.45	0.27	0.74	0.42	0.21	0.75	13	79.19%	30.05%	85.00%
343	0.16	-0.81	0.69	0.09	-0.82	0.73	19	53.23%	25.00%	85.00%
342	0.35	-0.81	0.69	0.43	-0.82	0.73	12	85.00%	30.06%	85.00%

TABLE 6.3: Top 10 most active users and their general social group role signatures within the Walk 2.0 website case study

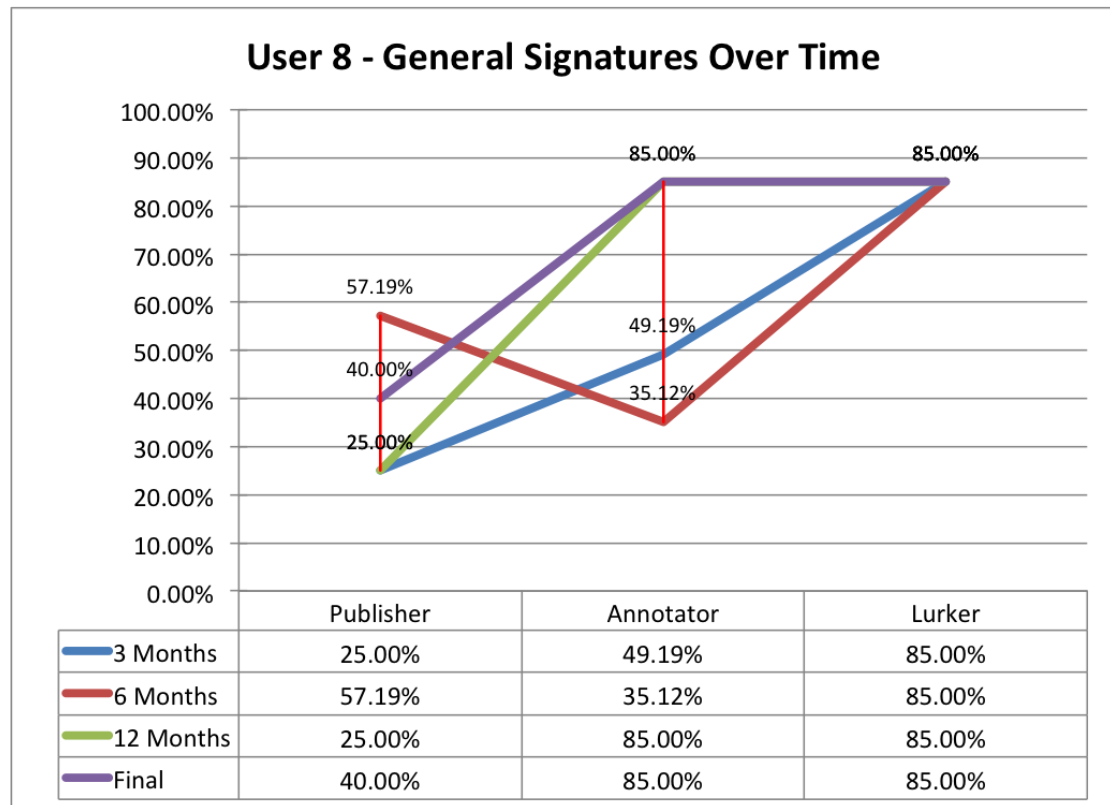


FIGURE 6.1: The general signatures for user 8 within the Espressoability case study over the user's life

The Squashies case study had two explicit roles for its users. Users who were general members of the gym and instructors of classes within the gym. Although not indicated in table 6.2, the first 4 users within this table are instructors. Users: 3201, 3200, 3203, 3199. All four of these users had a strong *publisher* characteristic behaviour. This aligns with their activity publication, as all instructors had to publish class sessions to other members and also state they were attending the class sessions they were conducting. Of these instructors, user 3203 had strong behaviour for all characteristic behaviours available within the case study. However, initially this user did not have such a general signature. Figure 6.2 presents the general signature for user 3203 at 3 months, 6 months and 12 months. This figure illustrates the user was determined to be weak in both *lurker* and *annotator* for the first 6 months of use, however it slowly increased to the point they demonstrated strong behaviour in all characteristic behaviours. When comparing their comments and favourite actions this matches the time line of behaviour change. Initially the user did not provide any other publication other than their class types, class sessions and their attendance. However, as time went on they

were commenting and favoured activity items created by users who were attending their classes.

When comparing gym members with instructors there is a noticeable difference between the two types of users. The *publisher* strength is weak or neutral for gym members, yet strong for instructors. This could be a result of a gym instructor needing to publish their class session for attendance, although when comparing the gym members use they published one check in on average for every new user session. During this user session a check in tended to be their only publication. They would however, favourite other gym member check ins or activities in the same user session.

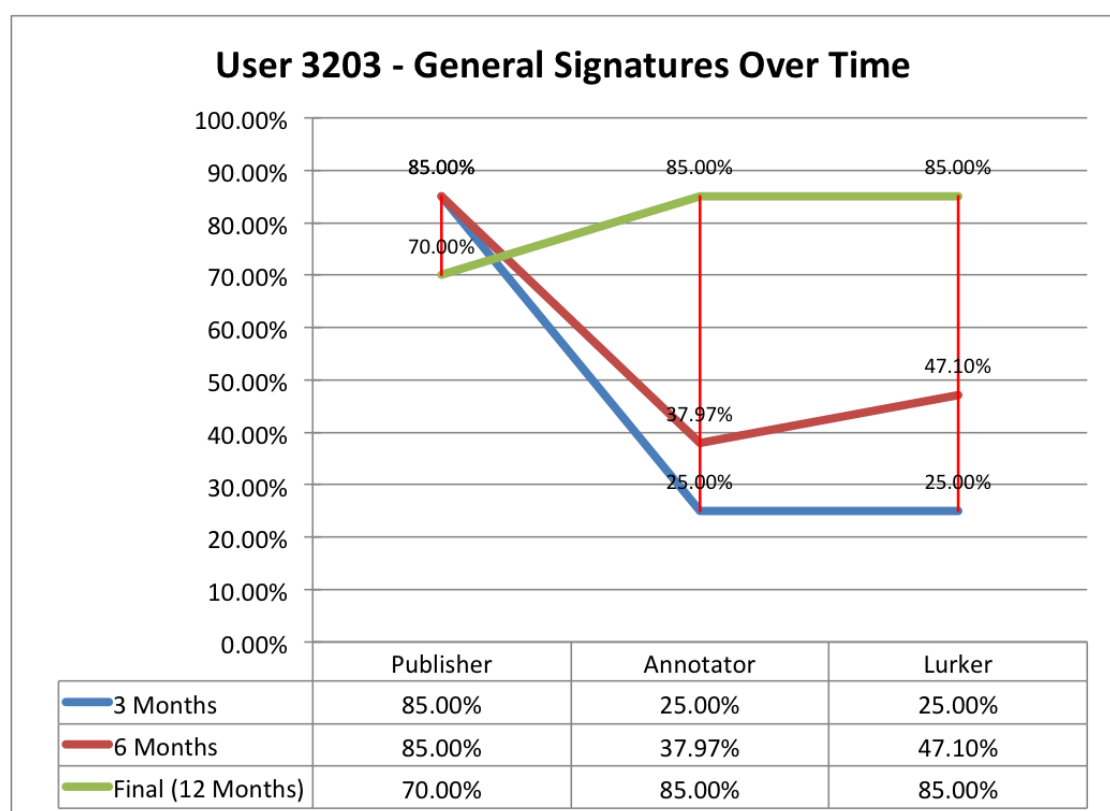


FIGURE 6.2: The general signatures for user 3203 within Squashies over the user's life

The top 10 users of the Walk 2.0 website all have one characteristic in common, they all have been determined to be strong with the *logger* characteristic behaviour. This result aligns with the objective of the Walk 2.0 website which was for users to log passive physical activity. The strong behaviour also matches the log count for each user. The log count was on average the most common activity users performed. This can be seen with figure 6.3 which presents the count of actions carried out by the top 10 users of the Walk 2.0 website.

Both user 219 and 279 were determined to be strong across all characteristic behaviours. When comparing to their actions within figure 6.3 this is reflected. They were the only users who demonstrated a constant use over each available activity. Although user 279 had a much larger ‘Thumbs Up’ count, the framework still determined their *annotator* behaviour to be strong at 85%. This is a result of the FL layer. This layer determines a strong behaviour as 85% and therefore will not go higher because of the fuzziness created with age and historical snapshots.

The general signatures determined by the framework for the case studies match the amount of actions undertaken by each user. Further discussion on these general signatures and how they can be used will be in chapter 7

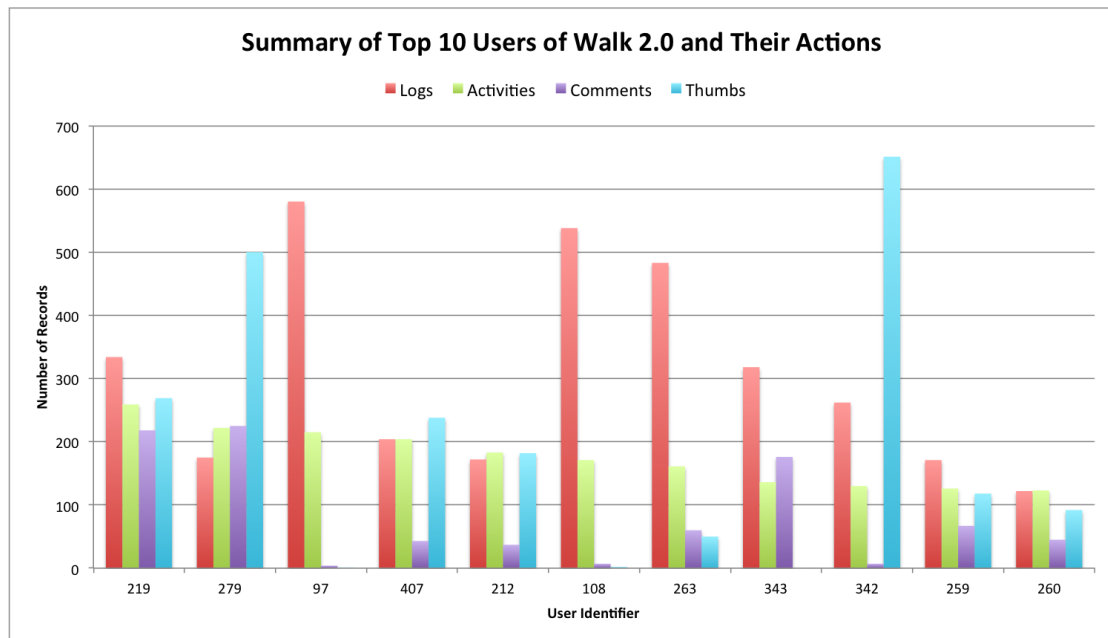


FIGURE 6.3: The top 10 users and their actions within the Walk 2.0 website

6.3 Context Based Signatures

The previous section presented general signatures for users within the case studies. Although the framework can determine a general signature for a user, the uniqueness of the framework comes from its ability to determine a role signature within collaborative context. This context is established through the social-ties created and also the tools used by a user. This section presents use of this context to for the framework to determine context based signatures for select users. These

signatures will be compared to a user's general signature and also their actual actions taken within context inside a case study.

Social-ties created by a user indicate a collaborative action where a user has carried out some action towards another user that resulted in a relational tie. By using a user identifier in combination with a social-tie a context is created for a collaborative action. Using a social-tie within a request enables the framework to determine role signature that is unique for the user and the context provided.

Within the Espressobility case study user 3 was determined to have a general signature where *publisher* was strong with 85%, but *annotator* and *lurker* were neutral and weak. This general signature was suggested to be a result of the user only focusing on their ability to publish and did not take time to engage or browse content created by other users. By providing the **user identifier and social-tie** within a signature request a context can be established to explore this suggestion. Three context based signatures for user 3 are presented in Figure 6.4 to illustrate this. Within this figure three social-ties have been used with three of the locations the user created and interacted with the most. For all context signatures it was determined by the framework that the *publisher* behaviour was strong while *annotator* and *lurker* was either neutral or weak. This is in line with the general signature also presented for this user. This reiterates the suggestion the user did not demonstrate a strong desire to know what was happening at their favourite locations or who was interacting with it.

User 1 within the Espressobility case study had a general signature of 85% *publisher*, 79.54% *annotator* and 55.00% *lurker*. When comparing their actual use of the case study this is reflected with the amount of activities carried out by the user. However, when requesting context based signatures for this user in relation to a social-tie and tools/widgets used their signature changes. Figure 6.5 presents this with the general signature of user 1, and the signature of a different context. A context where the social-tie in this figure is a tie with the user's own user identifier, which is *u1*, resulting in the social-tie *u1-u1*. This illustrates the *annotator* behaviour is strong while both *publisher* and *lurker* are weak. This makes it different to their general signature. When comparing with actual use of the environment, this user would tag their own activities published. They were also the user who produced 32 of all the 44 tags present within Espressobility.

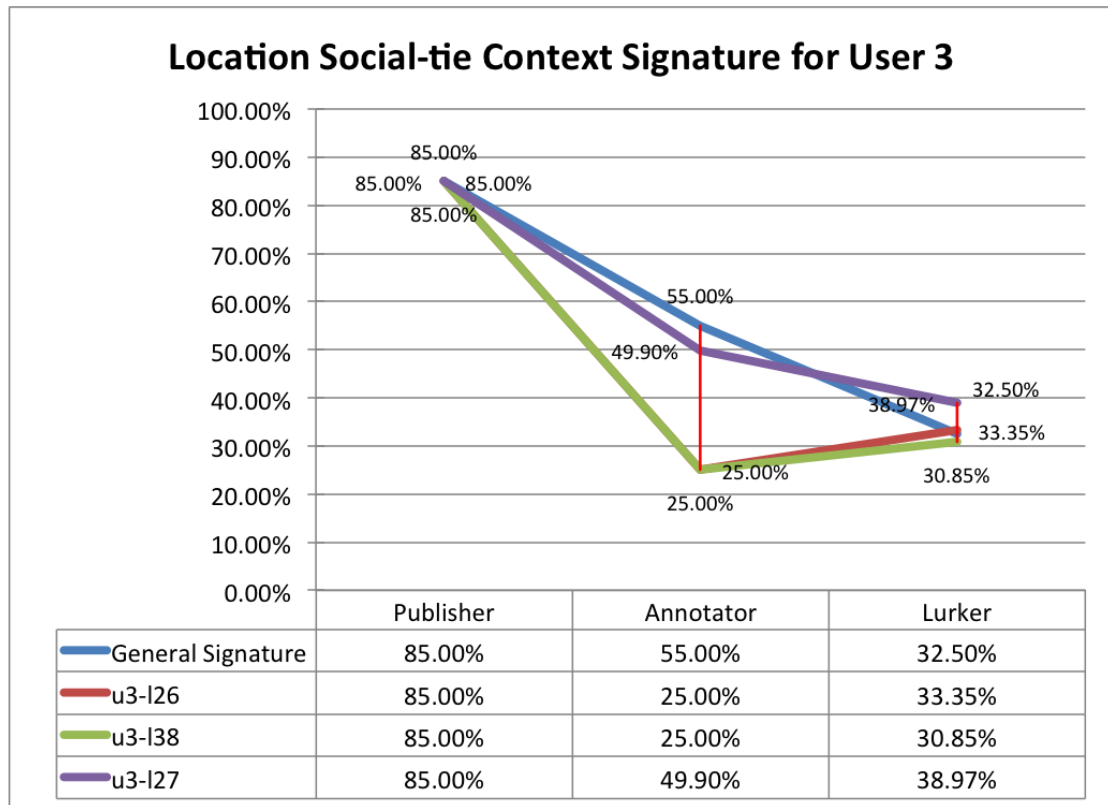


FIGURE 6.4: Location context based signatures for user 3 of Espressobility

When providing a more detailed context by providing all three input fields, the framework can determine slight changes in behaviour even though social-tie remains the same. This can be seen in figure 6.6. Within this figure User 1's general signature is present, the signature for the social tie with user 8 present, and also two more signatures in with the tools the user used while interacting with user 8. From this one can see there is a slight change between the *publisher* behaviour when a tool identifier is used in conjunction with a social-tie. The *publisher* behaviour decreases from neutral to weak providing a more refined context for each signature.

This refinement is not beneficial when a tool only offers one action for a user to perform one characteristic behaviour. However, it is beneficial when a tool is able to do two or more actions resulting in multiple behaviours. An example of this is the location tool within Espressobility. A user is able to publish a review, explore details about the location and edit the location. Figure 6.7 presents a comparison of signatures in regards to this tool with user 26 within Espressobility. The role signature with the context of the tool shows a dramatic difference with *annotator* suggesting the user did not favour this ability of the tool. However, the shape is

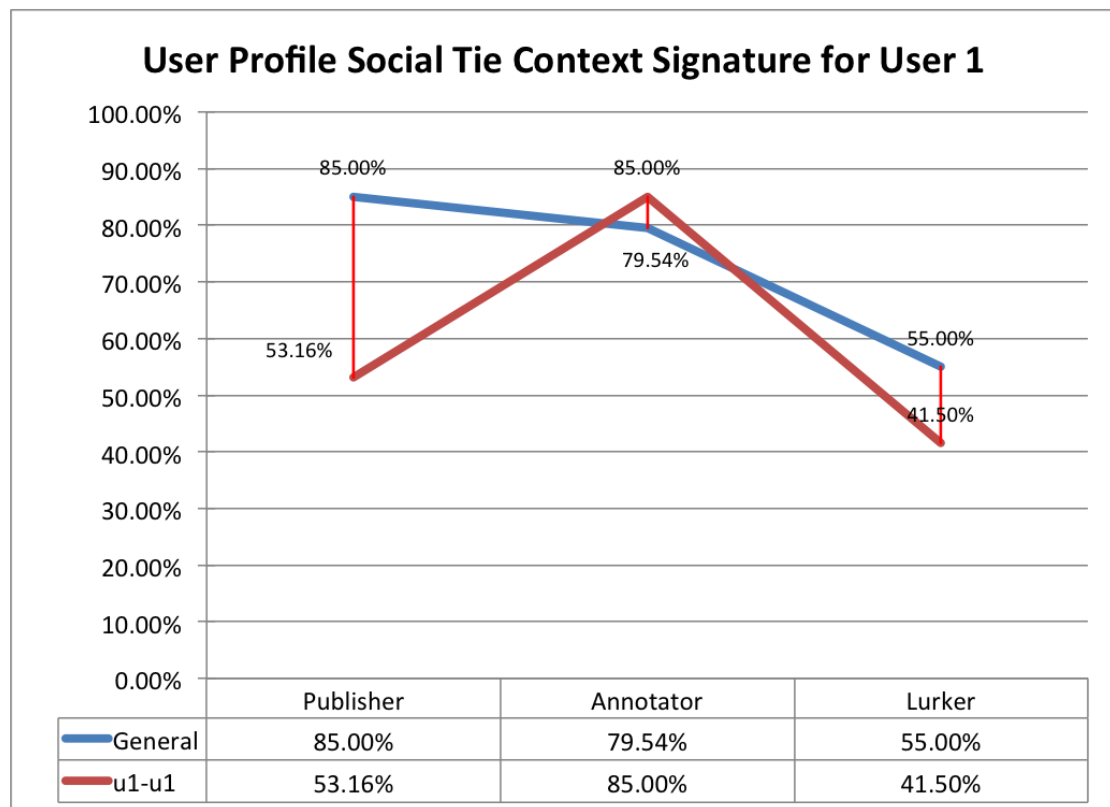


FIGURE 6.5: User 1 and their determined general signature with their user profile context signature within Espressobility

similar to the signature when only the location social-tie is present, which also suggests the context of the tool identifier refined the signature. When comparing to actual use this matches the amount of edits user 26 carried out. They only did two edits, while their published activity count was 7.

Figure 6.8 presents a general signature for an instructor, user 3203, and also three context based signatures within Squashies. The three context based signatures are the result of two users an instructor has a social-tie with and a class they have created. One of the users, user 3243, has attended the instructors class while the other user, user 3228, did not. Although these may be isolated signatures, looking at the third signature, which has the class context, the signature for the user who attended the instructors class has been determined to be similar. When comparing with actual use of the case study, instructors favoured on average 3.92 activity items posted by users who attended their class, while only favoured on average 0.27 activity items of all users who did not attend their class.

Squashies enabled users to create various activity types for publication to activity streams. Each activity type was regarded as a tool with its own unique identifier.

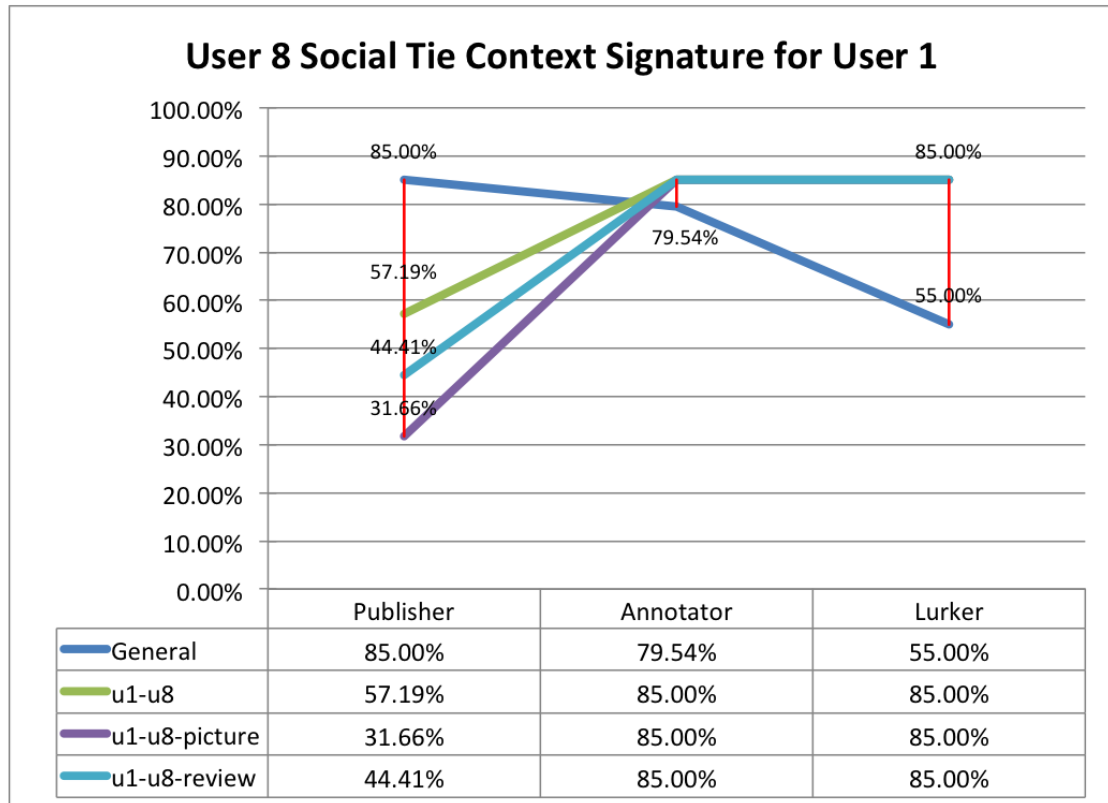


FIGURE 6.6: Role signatures for user 1 with refined context based signatures with user 8 and tools/widgets within the Epsressobility case study.

By providing a user identifier and tool identifier the framework can determine a context based role signature towards a tool for a particular user. Table 6.4 presents the determined context signatures for user 3199 of the Squashies case study with the context of each activity type. From this table one can see the framework has provided different signatures for each tool used. For some tools the user could not carry out the *publisher* behaviour, such as the gym notice, as they did not have the authority. For these tools the framework determines the *publisher* behaviour on what experience exists for the user in regards to the *publisher* behaviour. It will then build the dynamic structure of the ANN and determine a signature if it has experienced the user using the tool in the past or not. It also does this for all tools where a user has had little use or no use. An example of this is shown in Figure 6.9 where it presents the actions carried out by user 3199. It illustrates there has been little use of various tools, however the framework has still determined a signature based on the experience collected for the user.

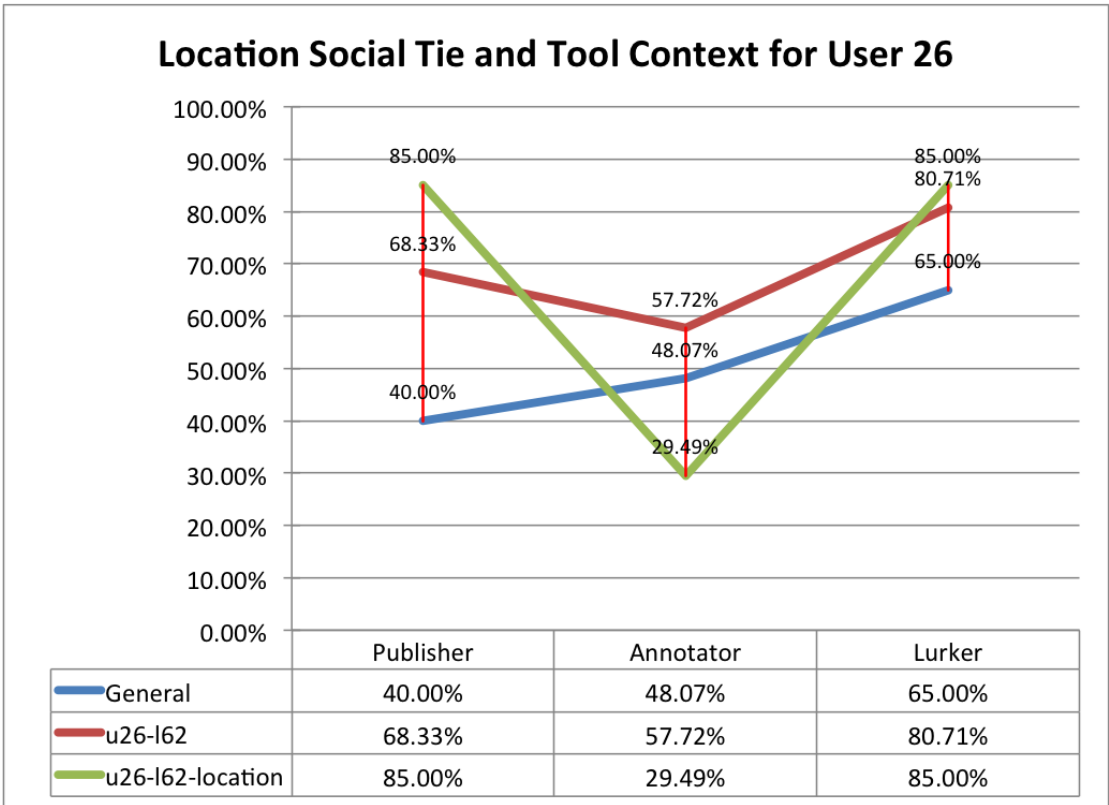


FIGURE 6.7: Location social-tie with location tool signature comparison for user 26 within Epsressobility

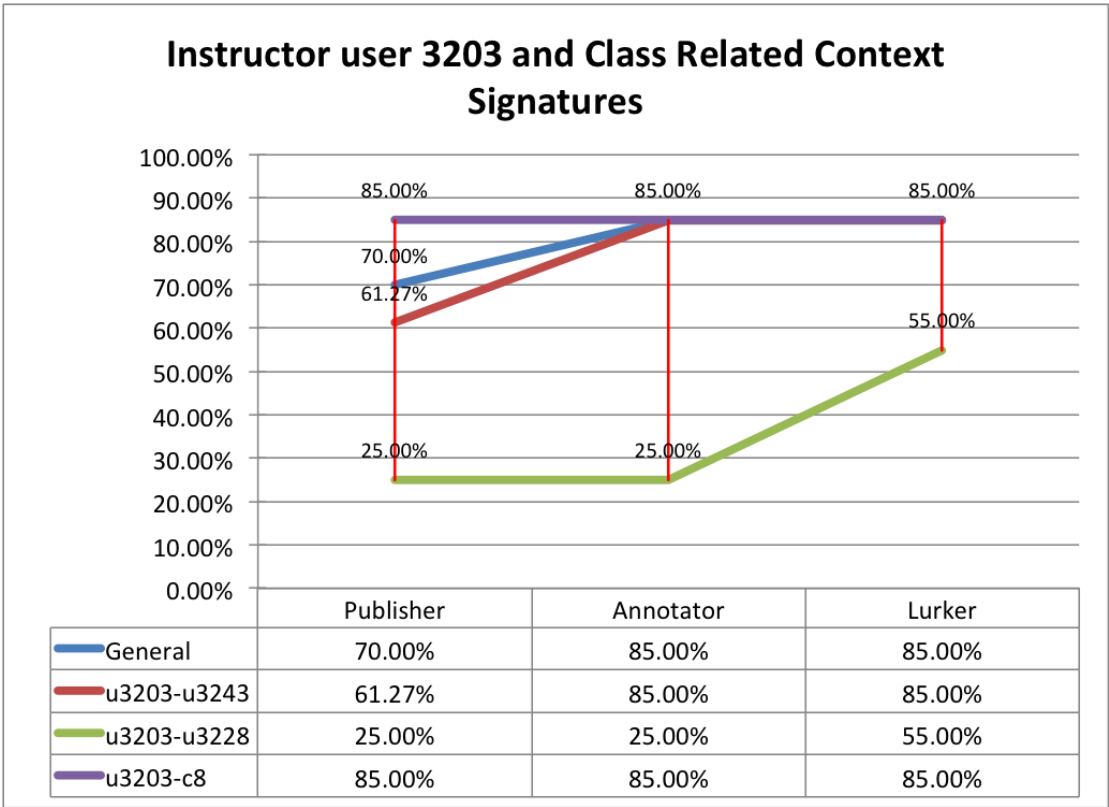


FIGURE 6.8: Comparison of context signatures for instructor user 3203 with class members and non member within the Squahies case study.

	<i>Historical ANN Results</i>			<i>Latest ANN Result</i>				<i>Inferred Signature</i>		
Sig. For	Publisher	Annotator	Lurker	Publisher	Annotator	Lurker	Days	Publisher	Annotator	Lurker
General	0.30	0.13	-0.21	0.32	0.05	0.12	6	75.76%	48.07%	55.00%
Status Post	0.21	0.02	-0.12	0.19	-0.06	0.03	6	57.19%	41.15%	30.85%
Progress Post	0.08	0.15	0.05	0.17	0.22	0.19	6	51.53%	59.28%	47.10%
Gym Notice	0.15	0.18	-0.09	0.11	0.14	0.1	6	55.00%	55.00%	33.35%
Check in	0.45	0.38	0.26	0.51	0.32	0.28	6	85.00%	85.00%	70.00%
Strength Session	0.12	0.08	-0.29	0.10	0.07	-0.20	6	55.00%	52.00%	25.00%
Cardio Session	0.14	0.09	-0.25	0.07	0.03	-0.28	6	55.00%	53.16%	25.00%
Class Session	0.79	0.81	0.68	0.82	0.79	0.7	6	85.00%	85.00%	85.00%
New Class	0.38	0.21	0.12	0.32	0.14	0.09	6	82.00%	57.19%	55.00%
Class Attendance	0.73	0.59	0.69	0.78	0.53	0.72	6	85.00%	85.00%	85.00%

TABLE 6.4: The context signatures for use 3199 and their use of the different activity types

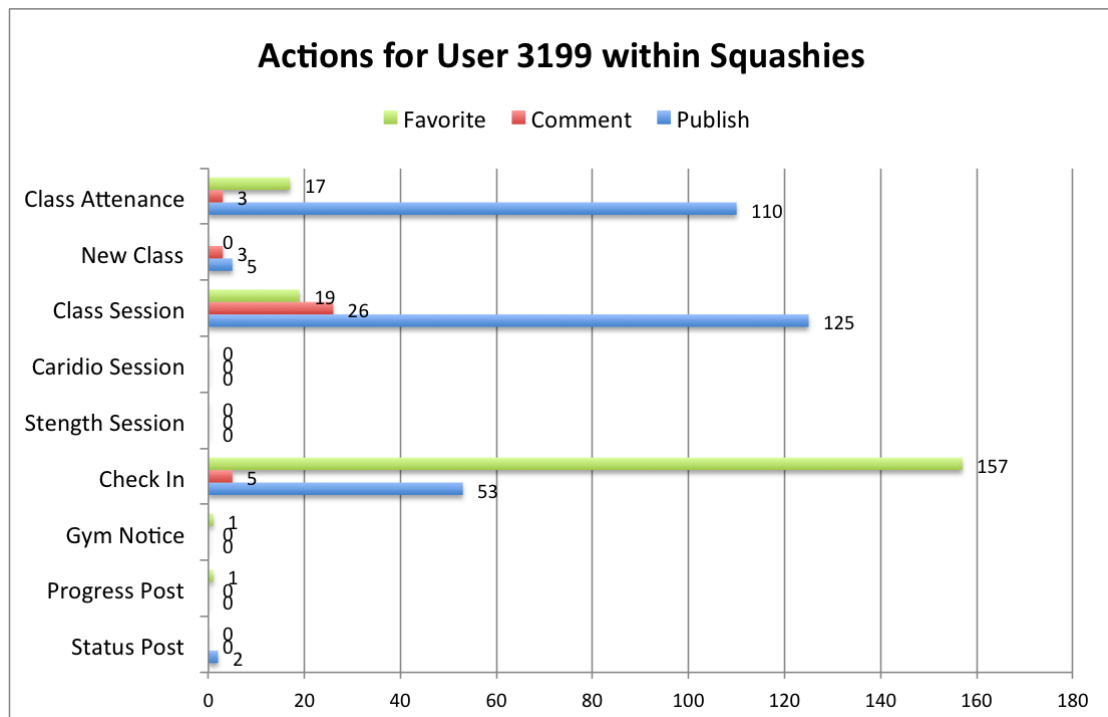


FIGURE 6.9: Distribution of total actions for user 3199 within the Squashies case study.

6.4 Accommodating Characteristic Behaviours

The characteristic behaviours within a signature are based on the commitment requests to the framework. This is why in chapter 5 the resulting characteristics were presented at the end of data collection. This is because a behaviour should not exist unless that behaviour has been demonstrated by one or more users. The three stereotypical behaviours chosen to represent characteristic behaviours, with exception of walk, were all present at the end of data collection. This meant at some point in time all behaviours were displayed by the user base.

As there were only three characteristic behaviours this resulted in users from both Espressobility and Squashies presenting the behaviours within a user's first user session. However, as the interaction packages were cached simulated user sessions could be created to segment interaction packages to demonstrate when behaviour was first introduced to the framework from a case study. The results of the simulated user sessions for Espressobility and Squashies can be seen in tables 6.5 and 6.6. These signatures are general signatures determined for the first users of the case studies. From these tables one can see behaviour is strong upon initial

commit. This is a result of how the framework commits new behaviours. The strength of a new behaviour is dependent on how many behaviours are present within the framework when the behaviour is committed. This is calculated by the division of one by the present count of behaviours in the framework. For instance, if there are 3 behaviours the new behaviour will be given an output synapse weight of 0.33. As the behaviours grow new behaviours will not be as strong initially. Using the simulated user sessions for Espressobility and Squashies it took 6 snapshots for user 1 to demonstrate all characteristic behaviours. However, a role signature could be determined with only one behaviour after 2. This is the same for the Squashies cases study.

The Walk 2.0 website did not require simulation. The first user presented new behaviour within each new user session after registration. Table 6.7 presents the first user and their determined general signatures. Although the *publisher* behaviour was committed the FL removed the behaviour from the resulting signature. This was also the case for the fourth user session where the user presented the *annotator* behaviour. This is because a behaviour strength could not be inferred by the FL as it was not present in the historical ANN result. This example demonstrates the capability of the framework adapting quickly to new behaviour present within an OCE. Only after 6 user sessions was a general role signature determined for a user.

Although these examples are of the framework accommodating behaviour for initial users, the ability of accommodating behaviour of the framework can be carried through out the life of the OCE. If a new tool is implemented within an environment that presents a new behaviour the framework has not experienced, once a user uses that capability of the tool the framework will then commit the new behaviour. Once committed all future role signatures created will include the new behaviour and its strength for a user.

	<i>Historical ANN Results</i>			<i>Latest ANN Result</i>				<i>Inferred Signature</i>		
User	Publisher	Annotator	Lurker	Publisher	Annotator	Lurker	Days	Publisher	Annotator	Lurker
1			1			1	0			85.00%
1	0.48		0.81	0.39		0.61	3	85.00%		85.00%
1	0.32	0.28	0.47	0.33	0.29	0.36	2	76.60%	71.53%	85.00%

TABLE 6.5: Initial accommodation of behaviours for the EspressoBility case study by showing general signatures of first user

	<i>Historical ANN Results</i>			<i>Latest ANN Result</i>				<i>Inferred Signature</i>		
User	Publisher	Annotator	Lurker	Publisher	Annotator	Lurker	Days	Publisher	Annotator	Lurker
3916	1			1			0	85.00%		
3916	0.84		0.42	0.78		0.49	7	85.00%		85.00%
3916	0.78	0.27	0.39	0.74	0.22	0.32	4	85.00%	68.40%	83.42%

TABLE 6.6: Initial accommodation of behaviours for the Squashies case study by showing general signatures of first user

	<i>Historical ANN Results</i>			<i>Latest ANN Result</i>				<i>Inferred Signature</i>		
User	Publisher	Annotator	Logger	Publisher	Annotator	Logger	Days	Publisher	Annotator	Logger
9			1	0.5		0.81	1			85.00%
9	0.34		0.82	0.29	0.24	0.76	8	85.00%		85.00%
9	0.39	0.18	0.78	0.25	0.1	0.80	4	83.98%	55.00%	85.00%

TABLE 6.7: Initial accommodation of behaviours for the Walk 2.0 website case study by showing general signatures of first user

6.5 Handling A Dynamic User Social Network Structure

The user social network structure of an [OCE](#) is dynamic and is capable of changing over time. This occurs as new users and tools enter an environment and increase the amount of social-ties towards other users or media generated from tools. Because of the dynamic structure of the [ANN](#) within the framework it is able to accomodate these new social-ties before a user interacts with it and after.

Table [6.8](#) presents the determined signatures of user 6 and the context of the social-tie *u6-u8* before and after its creation. Also within this table is the general signature for the user before the social-tie is established. The second row is the same framework snapshot for the general signature was determined, but with the context of the social-tie before it was established by the user. The difference between the general signature and the social tie can be seen within figure [6.10](#). Both signatures share the same characteristics, however with the context of the social-tie the framework has determined the behaviours to be slightly weaker. The final signature within the table is after the social tie was established. The *publisher* dropped to neutral strength while *annotator* increased from weak to neutral, but *lurker* remained the same as the other two signatures. When comparing to actual usage of the case study user 6 created a social-tie by commenting on a review post by user 8.

When using the social-tie for context there are only slight differences between strengths. Using the same general signature, yet adding the tool/widget identifier to the context in addition to the social-tie a different signature is produced. Table [6.9](#) presents the user 6 again with the social tie *u6-u8*, however the context of the tool used for the social tie creation has been provided. Using the figure [6.11](#) to represent the signatures visually one can see the role signature after the social-tie is more aligned with the role signature before a social-tie was established. In this instance, there was an increase in the *annotator* behaviour taking it to strong instead of neutral, where the *logger* behaviour decreased going from strong to neutral.

The above example demonstrates the ability of the framework being capable of determining a role signature where a context was yet to be experienced by the framework. The framework only using what it had learnt from other interaction

packages about the user, it inferred a signature for future actions for a user towards another user. Within a dynamic and changing collaborative environment this capability could provide a means for personalisation towards potential and existing social-ties.

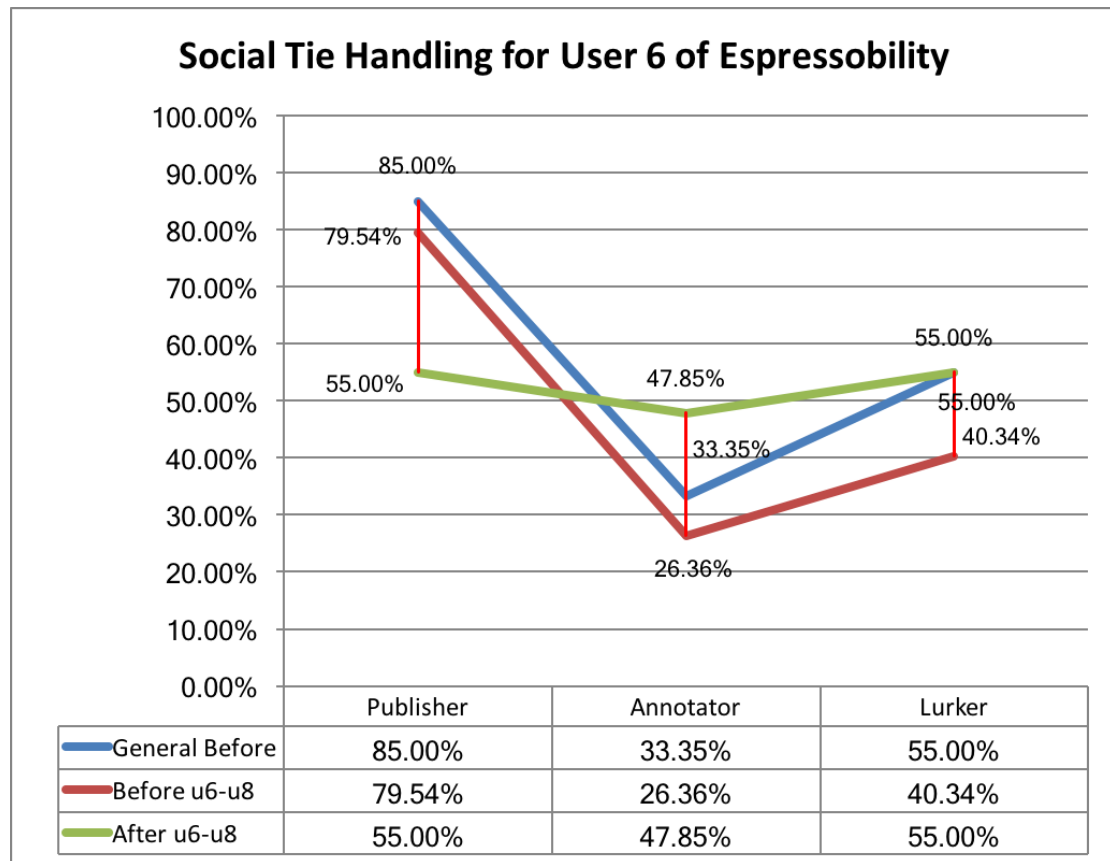


FIGURE 6.10: Visual signature of user 6 before and after the creation of a social-tie with user 8 within the Expressibility case study

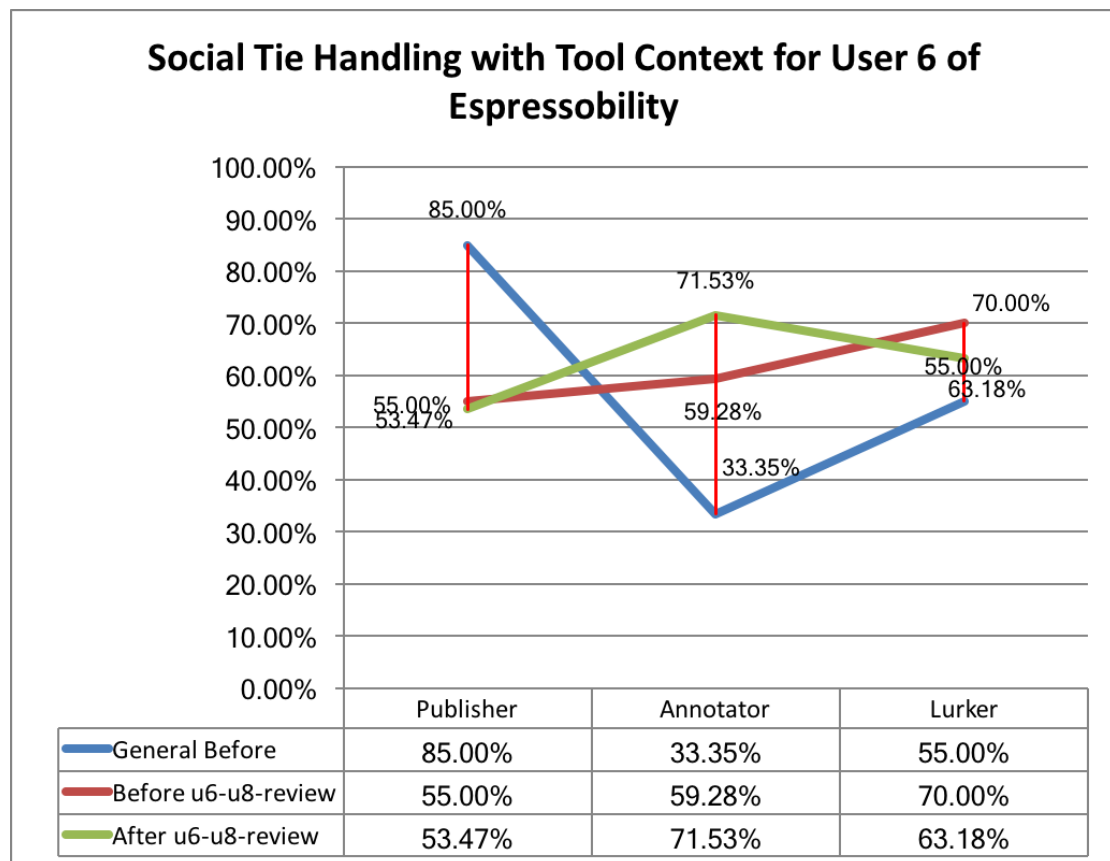


FIGURE 6.11: Visual signature of user 6 with context of the review post tool before and after the creation of a social-tie with user 8 within the Espressoability case study

	<i>Historical ANN Results</i>			<i>Latest ANN Result</i>				<i>Inferred Signature</i>		
Sig. For	Publisher	Annotator	Lurker	Publisher	Annotator	Lurker	Days	Publisher	Annotator	Lurker
General Before	0.48	-0.09	0.15	0.41	-0.04	0.19	5	85.00%	33.35%	55.00%
Before u6-u8	0.36	-0.18	-0.07	0.32	-0.27	0.23	5	79.54%	26.36%	40.34%
After u6-u8	0.19	0.05	0.11	0.13	0.04	0.19	7	55.00%	47.85%	55.00%

TABLE 6.8: Signature of user 8 before and after the creation of a social-tie with user 6 within the Espressoability case study

	<i>Historical ANN Results</i>			<i>Latest ANN Result</i>				<i>Inferred Signature</i>		
Sig. For	Publisher	Annotator	Lurker	Publisher	Annotator	Lurker	Days	Publisher	Annotator	Lurker
General Before	0.48	-0.09	0.15	0.41	-0.04	0.19	5	85.00%	33.35%	55.00%
Before u6-u8-review	0.12	0.26	0.28	0.1	0.24	0.24	5	55.00%	59.28%	70.00%
After u6-u8-review	0.09	0.29	0.24	0.12	0.24	0.14	7	53.47%	71.53%	63.18%

TABLE 6.9: Signature of user 8 with context of the review post tool before and after the creation of a social-tie with user 6 within the Espressoability case study

Chapter 7

Discussion

7.1 Introduction

This chapter presents a discussion of the research undertaken. This will include a response to the research aims in relation to the research undertaken by restating the research aims and discussing the results achieved. Also, a discussion on how the framework compares to other models and methods for establishing roles and role signatures will be presented. Finally, future applications and future work will be discussed to explore the possible directions of this research.

7.2 Confirmation of Research Aims

As stated in Chapter 1 the aim of this research was to provide a construct in the form of a research framework for automating the social group role model life cycle for the creation of social group role signatures within an [OCE](#). One of the complexities of such an aim is the dynamic nature of an [OCE](#). Within an [OCE](#) users can be autonomous in an asynchronous manner and therefore their actions are captured and are capable of existing throughout time. This disrupts the natural cycle of the social group role life cycle as all possible behaviours throughout time are present for a user to evaluate. Therefore the problem for a user to correctly understand and/or identify another user's role or social norm composition within a group at a moment in time becomes apparent. In addition to this, users are able to carry out implicit behaviour that is not public, such as the *lurker* characteristic

behaviour. This causes a user to form an awareness of all characteristic behaviours of another user. With an aim that held such complexities, the proposed framework needed to be able to overcome these to successfully carry out the life cycle continuously overtime.

Even though the ability to provide such a construct was the aim of the research, various objectives were set out to confirm it. The first was to provide the construct with the use of an asynchronous, yet life cycle design, and a dynamic and automated machine learning method for characterisation.

For the asynchronous, yet life cycle design, the proposed use of an event based asynchronous approach with requests to handle and organise the phases of the social group role life cycle has been demonstrated as a successful approach. It enabled the phases of the life cycle to operate independently of each other. It achieved this when a request was made and the framework could organise the life cycle sequentially via the framework proxy. This was demonstrated by the framework evaluating and committing user sessions before the life cycle phases *change* and *transition* were requested to determine a role signature. For example, if a signature was requested the framework would organise and check all phases of the life cycle were ready and capable of being carried out in a sequential manner. However, if only the *commitment* phase was requested only that phase was carried out.

The use of a [NFS](#) for the machine learning method also proved to be a valid approach. However, unique customisation of the [ANN](#) was needed to handle the changing and dynamic social structures and behaviours within an [OCE](#) over time. This led to the addition of snapshot capability for the [ANN](#) structure so the [ANN](#) could create historical output. This capability and the choice of approach with the [ANN](#) and its ability to dynamically build its structure on request played a significant role in the success of the framework's dynamic and automated ability. Without it a static [ANN](#) would present the same problem as other methods where historical processing is lost [85]. The [FL](#) layer of the machine learning method also played a significant role in stabilizing and inferring a role signature over time. Without it a role signature would not be taking in account the age of behaviour and therefore could misrepresent the strength of a characteristic behaviour leading to a stale user model [125].

The second objective of the aim was to have the framework determine general role signatures composed of characteristic behaviours. As demonstrated in Chapter 6 the framework achieved this for the top 10 users of each case study. After exploring various users, it was demonstrated that the framework could also determine multiple general role signatures overtime showing change in characteristic behaviours. This was shown when comparing a user's general signature to their total use of a case study at different points in time: 3 months, 6 months and 12 months. Their usage of the tools and related behaviours available within the case study matched the determined characteristic behaviours within a general role signature. This was also seen with context based role signatures.

The capability of the framework to produce context based roles signatures was the third objective. Similar to the frameworks capability to determine general role signatures it also demonstrated its ability to determine context role signatures. However, the framework also demonstrated its ability to refine the role signature as a more detailed context was given. This was illustrated in figure 6.6 when comparing a general role signature of a user with context based signatures. The strength of the characteristic behaviours changed to accommodate the new context of a social-tie, and then it was refined again with the tool/widget identifier was used in addition to the social-tie. When comparing the determined context based role signatures with the actual use of the case studies, this was the same as the general role signatures where the usage matched that of the signatures.

The framework being able to accommodate new behaviour over time was the fourth objective. This objective was achieved with the use of a back-propagation ANN. As a back-propagation ANN is a supervised learning method the framework can be trained on new behaviour when a user demonstrates it within an environment. This was demonstrated with the behaviour take up by initial users and the final characteristic behaviour composition of determined role signatures for each case study. When the framework was first applied to the case studies it contained no experience about the users and the characteristic behaviours that are available within a case study and only through a user performing a characteristic behaviour could the framework learn it. Once learned though, the dynamic structure of the ANN could then use all learned behaviour as part of a role signature.

The same method accommodating behaviour over time also allowed the framework to accommodate changing social structures. This capability was the fifth objective. Within the case studies, users changed their social structures by forming

new social-ties with different users and media. This led to the framework having to accommodate this change. This was demonstrated with user 6 within the *Espressobility* case study. The framework was able to determine a role signature without a known social-tie and then the framework determined a more refined role signature once the framework was provided with interaction packages towards the social-tie. The first role signature determined was similar to user 6's general role signature, however once the social-tie existed the role signature was refined to suit the updated social structure.

The chosen characteristic behaviours used for the role signatures determined by the framework were three stereotypical behaviours: *publisher*, *annotator*, and *lurker*. However, for the Walk 2.0 website *logger* was substituted for *lurker* as demonstration of the framework was post-trial for the Walk 2.0 project. The identification of these characteristic behaviours was the sixth objective of the research aim. It was needed to have generic behaviour which could be used with the framework to demonstrate its capability across multiple case studies. The identified stereotypical behaviours represented the general behaviours derived from [CRUD](#) actions. They were used as generic characteristic behaviours which could be used for a more generic role signature composition to demonstrate and to compare across multiple case studies.

The seventh objective of the research aim was to create a framework so that it was generic and accommodating of multiple [OCE](#). This was achieved through a request architecture within the framework domain. It enabled the framework to be external to an [OCE](#) and therefore not dependent on the [OCE](#) interacting with it. The framework was also capable of handling multiple [OCE](#) through one implementation of the framework. Although this was not demonstrated with the case studies as they operated at different time periods and resource availability was scarce, the framework was still demonstrated successfully across multiple case studies with the same code base. The only changes between case studies were the development environment and the implementation of the semantic manager and the [IPUSM](#). The semantic manager and the [IPUSM](#) were different as they needed to specify which events to listen to and the mapping of user sessions for the [IPUSM](#). This task would be needed for each environment as each environment will have different behaviour associated with tools. They will also have different methods of storing a user session. With only these changes, the use of the framework over multiple case studies demonstrates the last objective achieved.

The last objective was the development of case studies to demonstrate real world application of the framework. Two case studies were developed specifically for this study, and one was acquired from another research project. The two case studies developed were two mobile applications that were available to the public. They both required the approval of ethics for participant use and recruitment. These development and use of these case studies allowed real world application of the framework as it was applied before the a case studies release to the public. This is because the ability to apply the framework to a case study before release enabled real time data collection that enabled the ability to demonstrate the automation of the social group role life cycle in a real world scenario. The case study that was acquired from another research framework, although it was post-trial use, provided demonstration of the framework with an existing data set. These case studies demonstrated application of the framework and its ability to be generic to work across many [OCE](#) that are live and in use and those that are offline and are static.

7.3 Framework Comparison

The asynchronous and implicit nature of an [OCE](#) makes the continuous identification and analysis of collaborative and social group roles online complex [102]. This could be the reason why the range of methods for identification of roles within an [OCE](#) are diverse. Many are static and are carried out at a particular point in time. This could be a result of each [OCE](#) being different, in terms of tools and purpose, and as a result so is the data. It could also be an output of a known dataset being easier to analyse then a changing dataset. These methods are not incorrect, as they provide insight into the detailed structure of environments, however they fail to work for real world scenarios because of their focus on structure. The framework presented does not focus on structure but conforms to the dynamic and fluidity of [OCE](#). This allows scenarios where knowing a user's role or role signature at the point of interaction could help allocate resources or assist with collaborative efforts.

The dynamic and adaptability of the framework towards various [OCE](#) is the fundamental difference between other methods. Generally other methods commonly use the approach of setting structures or rules for signatures to create roles [100, 106]. However, a flaw in such an approach is the laborious task of mining the social

structures and identifying behaviour after there is enough data to do so. This task could be why many approaches use existing OCE to source their data and as a result always look in the past after the collaboration has already taken place [7, 111]. Another flaw in such an approach is having the structures and identified behaviour tied to a particular OCE limiting their use for other OCE. This was not the case with the framework. After two users sessions, a historical one and a current one, the framework could determine a role signature. There was no need to manually set structures and identify behaviour from data harvested from an OCE.

Identification or categorisation of users into roles within an OCE is a common objective of many studies [7, 51, 101, 102, 104]. However, identification and categorisation doesn't align with the social group role model. This is because if an OCE is dynamic and continually changing the identification or characterisation of roles would be also [102]. This is amplified with the collaboration model of three domains: *people*, *purpose*, and *processing*. Each user will have a different composition of attributes for each domain so they may not associate with one particular role all the time. With this understanding it was a conscious decision to not label or categorise a user to a single role. Instead, following the social group role model, the characteristic behaviours provide the descriptors for an "implicit role" for a particular context. This approach of not identifying or categorising users into roles yet creating role signatures enables the framework to be more dynamic as typecasting of users is prevented. Meaning, there is no explicit label specifying a set of characteristic behaviours a user must have. Instead, an OCE can have a range of behaviours the framework can use to determine a role signature instead of a predetermined role or classification a user belongs to.

The determined role signatures used by framework matched the user participation. However, as strength of characteristic behaviours will be a generalisation of behaviour, a stereotype, this framework will never output an absolute representation of behaviour as it will always be an 'approximation' like many user models [62]. Behaviour is subjective, therefore even if the ANN used another sigmoid function, or the FL had more detailed membership functions it still would result in an approximated strength representation of the characteristic behaviours shown by a user which may not be the actual user's preference or a behaviour the user associates with. Given this, the purpose of the research was not to provide a mathematically accurate framework, but a framework that could characterise

and automate the social group role life cycle in real time. This is why the results of the framework were compared with actual participant use. By matching the participation of users and their determined role signatures one could validate the automation of characterising common interactions carried out by a user within a dynamic and flexible environment.

Overall the framework achieved automation of the social group life cycle providing a universal **real world** framework for contrasting OCE. Although other methods of role identification that are static or coupled to a particular OCE might be more accurate in terms of classification or role identification, they will always be limited in application towards other OCE because of their dynamic and flexible nature. Therefore, it is believed this framework is the first generic/universal framework to provide automation of social group role identification within OCE.

7.4 General Discussion

The proposed framework has demonstrated its ability to characterise and automate the creation of social group role signatures within OCE. The success of this demonstration provides real world validation of foundational work that is believed to be a first of its kind for determining social group role signatures through the social group role life cycle for these type of environments. With such a framework for determining role signatures in a real world application, they can be used for the design, personalisation, and resource allocation for users that may help facilitate and support the collaboration amongst users.

The novel component of the framework is its ability to accommodate asynchronous behaviour within an asynchronous environment, while providing a life cycle to role signature formation base on social-ties. This allows the framework to work for real world applications. This was the reason for the decision of using case studies instead of harvesting data from pre-existing OCE, such as Twitter or LinkedIn. By using case studies one can carry out processing as behaviour occurs over time instead of a linear or sequential process as seen with the Walk 2.0 website. Harvesting also does not allow for the capture of implicit actions carried out by users. This was also seen with the Walk 2.0 website as the framework was interfaced post-trial and the data for the characteristic behaviour *lurker* was lost and so the *logger* behaviour was chosen as a substitute.

Each case study presented a different data set size. Because of the nature of the framework creating a role signature for an individual user the differences in data set sizes had no impact on the study. This was because the framework was capable of determining a role signature within two user sessions - although they could be incomplete role signatures. If there was more data for each user the NN would of had a more refined result, however the fuzzy logic using age to determine strength would cancel out this because it generalised the output from the ANN to prevent pigeon-holing a user. Therefore, the only difference of more data would be a slower change between strengths. This slower change could of been achieved by changing the learning rate of the NN if it was needed.

Although the size of the datasets did not have an impact on the study, a larger dataset would of allowed exploration into more users with the same general role signatures. This would of been desirable with the Espressobility case study. This case study had a small user base with only a few users demonstrating large amounts of behaviour, and with those users only two general role signatures were similar over time. To have many users with the same general role signature and then explore their behaviour amongst many users over time would of been desirable.

A limitation of this work was the limited functionality of the case studies. Although they enabled demonstration of the framework, not being able to have the framework create role signatures for all nodes within a network was a limitation. If there was more time, any node within an OCE could of been used. For instance, if there was group based activities present the group could of been used as a social-tie by the framework. Instead this study only used the user identifier and optional contexts of another social-tie and/or a tool/widget identifier. This could be future work where groups of users with a group identifier is used as an additional option to a user identifier to provide group based role signatures. This could provide a tiered approach to role signatures, where the framework creates more defined context role signatures. For instance, one could use the user identifier, a group identifier, a social-tie and a tool identifier to provide a refined context role signature based on behaviour to another user within the same group.

Presented in this work were three generic stereotypical behaviours used for characteristic behaviours. The use of three stereotypical behaviours, although provides a general mapping of behaviour for all OCE, limited the complexity of role signatures, and so limits the ability to make complex decisions about users because

it is a generalisation. In future work one could include more refined semantic related behaviours for differing OCE. The framework is already capable of adopting these refined behaviours because of its supervised learning approach - as seen with the substitution of *logger* with the Walk 2.0 case study - so there would be no modification to the current framework just the interfacing with a new OCE. An example of this could be an environment where social coding was the collaborative goal, and a semantic meaning for various actions within such an environment were used. For instance, *'pusher'* could be a semantic behaviour where the user pushes new code to a repository for other coders to collaborate on. This approach would lead to more refined, complex, and related role signatures instead of stereotypical behaviour as seen in this research. This approach would be more effective for user interface personalisation or in practice when providing resources to users as the role signature is less of a generalisation and more detailed and focused for the context of the collaborative environment.

Modification of the framework could also be carried out to explore the nuances of the different components of the framework. For instance, *linear fit functions* were used for membership functions of the FL layer of the framework domain. Changing these membership functions to other functions, such as *sigmoidal* or *bell*, could be carried out to see the effect it would have on the determined strength of the characteristic behaviours. Changes to the membership thresholds could also be undertaken for different OCE to experiment the rate of change in strength. However, this type of work would require a continual live case study so not to lose the implicit data that is shown by users. This type of work, however, would only refine the output so that it would be more granular for characteristic behaviour strength, it would not change the overall composition of behaviours present within a role signature.

OCE are becoming more pervasive in our daily lives. They are also introducing new ways to empower the user to be autonomous with their actions towards other users and media. As this continues the user's ability to create a shared understanding for collaboration will become more complex and diverse making it harder to collaborate effectively. This will be because a user will find it difficult to *evaluate* and *commit* the behaviours of other users and so will have trouble carrying out the social group role life cycle that all individuals do. By providing a role signature derived from the social group role life cycle this complexity could be removed. With the use of the framework users could use the determined

role signatures to understand the characteristic behaviours other users present for different collaborative contexts. It could make it easier for users to commit to implicit roles for other users within a collaborative context as they have a role signature to *evaluate*. By being able to evaluate it could allow their own social group role life cycle to occur. In addition to this, users could evaluate other users and their historical role signatures for similar collaborative contexts to find a user who was strong with a particular behaviour. As a result, a stronger and more diverse collaborative context could be established. However, this relies on the user's understanding of the role signatures determined. This type of scenario could be future work for this framework.

Additional work could be knowing a role signature for personalisation of interfaces to provide responsive and personal tools that could assist in supporting collaboration. This could be achieved as the tools and widgets available for different users can allow a user to demonstrate different behaviours - as shown in the Espresso-bility case study. This could be achieved by using a role signature so a tool could present functionality that aligns with the strongest behaviours. In conjunction with a social-tie a tool could also respond differently under different collaborative contexts as the strength of behaviours change based on context. For instance, a user may have a social-tie with a particular user when using a certain tool, however a user may also have a social-tie with a group the user belongs to. With this scenario a tool could present certain functionality for the individual user within the group, yet when using the tool at a group level it presents different functionality.

In addition to personalisation of interfaces, personalisation of resources for users could be provided based on the strength of certain characteristic behaviours within a role signature. For instance allocation of media or other users that share similar or complementing role signatures for collaborative tasks. As the framework determines role signatures as a composition of characteristic behaviour strengths, the type of resources allocated would be for behaviour based tasks, such as a type of media that leads to behaviour or media that needs action against it to move it forward in a collaborative work flow.

Given role signatures determined by the framework could be used for resource allocation, they could also be used for marketing or promotional purposes. A business, for instance, could target users with particular characteristic behaviours which are strong. For instance, using the *publisher* characteristic behaviour presented in this research, a company could provide information to them to disseminate through

their social groups for promotion knowing they have a strong behaviour that results in publication. The business would be using the role signature for targeted marketing or promotion.

Although a business could use role signatures for marketing, from the perspective of an [OCE](#) developer the use of a framework that can determine role signatures for users could provide valuable information for future design and development. Information that could cause focus of development on areas where weak characteristic behaviours are present amongst users. It could also be used to understand attrition rates or help identify reasons why users change behaviour over time within an environment. Regardless of how the information attained from user, the use of role signature gives the developer the ability to make informed decisions about an environment which could lead to better decisions when approaching design and development of features such as tools or interfaces.

As discussed the application and future work for the framework is extensive, and the need for this is growing as [OCE](#) continue to develop and be used for collaboration. The framework proposed in this study lay the foundations for a real world model of the social group role life cycle present in our daily ‘offline’ lives. With continual work the divide between the offline world and the online collaboration world will dwindle. This will allow us to collaborate online naturally within dynamic and flexible collaborative environments.

Chapter 8

Conclusion

The aim of this research was to provide a construct in the form of a research framework for automating the social group role model life cycle for the creation of social group role signatures within OCE. To achieve this aim, a constructive research methodology was used. This methodology resulted in a thorough literature review, from which a research framework construct was proposed and an implementation of the framework was presented and applied to three OCE case studies - two uniquely developed for this research - to demonstrate the framework's real world application. The output of the framework was then compared to participant use to demonstrate the frameworks capability to determine role signatures for collaboration within dynamic and flexible OCE. A discussion of the frameworks ability was then presented and compared against the research objectives. Also, a comparative discussion of the frameworks dynamic capability towards existing methods for role identification and formation was presented. Finally, a general discussion about the framework and future work was presented.

The use of this methodology has led to a viable construct to determine social group role signatures using the life cycle of social group role theory in an asynchronous way. This contribution is believed to be the first of its kind, as such it has laid foundational work that could open opportunities to form new OCE that are able to react to its user base. For instance, an OCE could identify and understand the role signatures present or formed over time without the need of pre-thought from the designer. Such an environment could provide resource allocation or interface personalisation based on a social group role signature.

The construct has also provided various new methods and approaches to user modelling for [OCE](#). The need for a predefined, structured or rigid approach to user modelling in a collaborative context is not always required. Using social group role theory in an asynchronous way has demonstrated that one can simulate real world concepts or social theories to observe, understand and predict social group role signatures or collaborative behaviours in a users social network online. With this asynchronous approach the possibilities of future work to understand, predict and create role signatures or to assist in successful collaboration online is extensive.

Reference List

- [1] Anne Croker, J Higgs, and F Trede. What Do We Mean by 'Collaboration' and When is a 'Team' Not a 'Team'? - A Qualitative Unbundling of Terms and Meanings. *Qualitative Research Journal*, 9(1):28–42, 2009.
- [2] Australian Bureau of Statistics. Household Use of Information Technology, Australia, 2012-2013. Technical report, 2011.
- [3] Facebook. Facebook Quarter Financial Report 2013. Technical report, Facebook, 2013.
- [4] L Dabbish, C Stuart, J Tsay, and J Herbsleb. Social coding in GitHub: transparency and collaboration in an open software repository. In *Proceedings of the ACM 2012 Conference On Computer Supported Cooperative Work.*, pages 1277–1286. ACM, 2012.
- [5] PM Leonardi. Social media, knowledge sharing, and innovation: Toward a theory of communication visibility. *Information Systems Research*, 25(4): 796–816, 2014.
- [6] E Gleave and HT Welser. A conceptual and operational definition of 'social role' in online community. In *42nd Hawaii International Conference on System Sciences*, pages 1–11. IEEE, 2009.
- [7] HT Welser, D Cosley, and G Kossinets. Finding social roles in Wikipedia. In *Proceedings of the 2011 iConference*, pages 122–129. ACM, 2011.
- [8] Richard L Moreland and John M Levine. Socialization in small groups: Temporal changes in individual-group relations. *Advances in Experimental Social Psychology*, 15:137–192, 1982.
- [9] BJ Biddle. Recent development in role theory. *Annual review of sociology*, 12(1986):67–92, 1986.

- [10] Rhys Tague, Anthony J Maeder, Corneel Vandelanotte, Gregory S Kolt, Cristina M Caperchione, Richard R Rosenkranz, Trevor N Savage, and Anetta Hinchliffe. Assessing User Engagement in a Health Promotion Website Using Social Networking. In *Global Telehealth*, pages 84–92. IOS Press, 2014.
- [11] Rhys Tague, Anthony Maeder, and Jim Basilakis. Adaptive Web Framework for Online Collaborative Environments. *Australian Journal of Intelligent Information Processing Systems*, 13(4):11–15, 2014.
- [12] Rhys Tague, Anthony Maeder, and Jim Basilakis. Classifying collaborative behavior in the form of behavioral stereotypes in collaborative mobile applications. In Aly A. Farag, editor, *Multimedia Technology IV*, chapter 165, pages 165–169. CRC Press, 2015.
- [13] GD Crnkovic. Constructive research and info-computational knowledge generation. In *Model-Based Reasoning in Science and Technology*, pages 359–380. Springer Berlin Heidelberg, 2010.
- [14] E Kasanen, K Lukka, and A Siitonen. The constructive approach in management accounting research. *Journal of management accounting research*, 5:243, 1993.
- [15] B Gray. *Collaborating: Finding common ground for multiparty problems*. Josse-Bass, 1989.
- [16] Ann Marie Thomson. *Collaboration: Meaning and measurement*. PhD thesis, Indiana University, 2002.
- [17] a. M. Thomson, J. L. Perry, and T. K. Miller. Conceptualizing and Measuring Collaboration. *Journal of Public Administration Research and Theory*, 19(1):23–56, nov 2009.
- [18] S Puntambekar. Analyzing collaborative interactions: divergence, shared understanding and construction of knowledge. *Computers & Education*, 47(3):332–351, 2006.
- [19] DL Schwartz. The productive agency that drives collaborative learning. *Collaborative learning: Cognitive and computational approaches*, pages 192–218, 1999.

- [20] J Roschelle and SD Teasley. The construction of shared knowledge in collaborative problem solving. *Computer supported collaborative learning*, pages 69–97, 1995.
- [21] C Collazos and L Guerrero. Introducing shared-knowledge Awareness. In *IASTED International Conference: Information and Knowledge Sharing*, pages 13–18, 2002.
- [22] Hiroaki Ogata, Kenji Matsuura, and Yoneo Yano. Knowledge Awareness: Bridging between Shared Knowledge and Collaboration in Sharlok-1996. In *Proceedings of Educational Telecommunications*, pages 232–237, 1996.
- [23] P Barthelmess and KM Anderson. A view of software development environments based on activity theory. *Computer Supported Cooperative Work (CSCW)*, 11(2):13–37, 2002.
- [24] V Kaptelinin. Computer-mediated activity: Functional organs in social and developmental contexts. *Activity theory and human-computer interaction*, pages 45–68, 1996.
- [25] a. Hemetsberger and C. Reinhardt. Collective Development in Open-Source Communities: An Activity Theoretical Perspective on Successful Online Collaboration. *Organization Studies*, 30(9):987–1008, 2009.
- [26] Sheizaf Rafaeli and Yaron Ariel. Online Motivational Factors: Incentives for Participation and Contribution in Wikipedia. In *Psychological Aspects of Cyberspace*, pages 243–267. Cambridge University Press, New York, NY, 2008.
- [27] J Bosman. After 244 years, Encyclopaedia Britannica stops the presses. *The New York Times*, page 13, 2012.
- [28] D.R Forsyth. *Group Dynamics*. Cengage Learning, 2009.
- [29] Miller Mcpherson, Lynn Smith-lovin, and James M Cook. Birds of a Feather: Homophily in Social Networks. *Annual review of sociology*, 27:415–444, 2001.
- [30] J.S Coleman. *Foundations of Social Theory*. Harvard university press, 1994.
- [31] Anita Williams Woolley, Christopher F Chabris, Alex Pentland, Nada Hashmi, and Thomas W Malone. Evidence for a collective intelligence factor

- in the performance of human groups. *Science (New York, N.Y.)*, 330(6004): 686–8, oct 2010.
- [32] Alan G Ingham, G Levinger, J. Graves, and V. Peckham. The Ringelmann and Effect : Group Studies of Group Size. *Journal of Experimental Social Psychology*, 10(4):371–384, 1974.
- [33] Gergely Palla, Albert-László Barabási, and Tamás Vicsek. Quantifying social group evolution. *Nature*, 446(7136):664–7, apr 2007.
- [34] Tadeusz Szuba. A formal definition of the phenomenon of collective intelligence and its IQ measure. *Future Generation Computer Systems*, 17: 489–500, 2001.
- [35] Nigel R Franks. Army Ants: A Collective Intelligence. *American Scientist*, 77(2):138–145, 1989.
- [36] Gerald Kane. It’s a network, not an encyclopedia: A social network perspective on Wikipedia collaboration. In *Academy of Management Proceedings*, pages 1–6. Academy of Management, 2009.
- [37] K. Crowston and J. Howison. The social structure of free and open source software development. *First Monday*, 10(2):1–27, 2005.
- [38] R.M Belbin. *Team roles a work*. Routledge, 2nd edition, 2010.
- [39] B. J. Biddle. *Role theory: Expectations, identities, and behaviors*. American Press, 1979.
- [40] S. Influence. Social norms, conformity, and compliance. In *The Handbook of Social Psychology: 2-Volume Set*, page 151. 1998.
- [41] Andrea Bonaccorsi and Cristina Rossi. Why Open Source software can succeed. *Research Policy*, 32(7):1243–1258, jul 2003.
- [42] T OReilly. What is Web 2.0: Design patterns and business models for the next generation of software. *Communications & strategies*, 65(17):1–48, 2007.
- [43] J Brown, AJ Broderick, and Nick Lee. Word of mouth communication within online communities: Conceptualizing the online social network. *Journal of interactive marketing*, 21(3):2–20, 2007.

- [44] CMK Cheung, PY Chiu, and MKO Lee. Online social networks: Why do students use facebook? *Computers in Human Behavior*, 27(4):1337–1343, 2011.
- [45] Damon Centola. The spread of behavior in an online social network experiment. *Science (New York, N.Y.)*, 329(5996):1194–7, sep 2010.
- [46] Saul Greenberg. Personalizable groupware: Accommodating individual roles and group differences. In *Proceedings of the second conference on European Conference on Computer-Supported Cooperative Work ECSCW'91*, pages 17–32. Springer Netherlands, 1991.
- [47] P Garcia and JM Balmaceda. Automatic detection of team roles in computer supported collaborative work. *IEEE Latin America Transactions*, 11(4): 1066–1074, 2013.
- [48] DD Curtis and MJ Lawson. Exploring collaborative online learning. *Journal of Asynchronous Learning Networks*, 5(1):21–34, 2001.
- [49] Moira Burke, Cameron Marlow, and Thomas Lento. Feed me: motivating newcomer contribution in social network sites. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 945–954. ACM, 2009.
- [50] Jamy Li and Mark Chignell. Birds of a feather: How personality influences blog writing and reading. *International Journal of Human-Computer Studies*, 68(9):589–602, sep 2010.
- [51] Y Zhao, G Wang, PS Yu, S Liu, and S Zhang. Inferring social roles and statuses in social networks. In *Proceedings of the 19th ACM SIGKDD international conference on knowledge discover and data mining*, pages 695–703. ACM, 2013.
- [52] Morgan Ames and M Naaman. Why we tag: motivations for annotation in mobile and online media. In *Proceedings of the SIGCHI conference on Human factors in computing*, pages 971–980, 2007.
- [53] Gerhard Fischer. Exploring , Understanding , and Designing Innovative Socio-Technical Environments for Fostering and Supporting Mass Collaboration. In *Mass Collaboration and Education*, pages 43–63. 2016.

- [54] Ricarose Roque, Natalie Rusk, and Mitchel Resnick. *Supporting Diverse and Creative Collaboration in the Scratch Online Community*. 2016.
- [55] AM Kaplan. If you love something, let it go mobile: Mobile marketing and mobile social media 4x4. *Business horizons*, 55(2):129–139, 2012.
- [56] KY Lin and HP Lu. Predicting mobile social network acceptance based on mobile value and social influence. *Internet Research*, 25(1):107–130, 2015.
- [57] GitHub. GitHub’s on your phone, 2013.
- [58] CH Ganoë and JP Somervell. Classroom BRIDGE: using collaborative public and desktop timelines to support activity awareness. In *Proceedings of the 16th annual ACM symposium on user interface software and technology*, pages 21–30. ACM, 2003.
- [59] M Heinrich, FJ Grüneberger, T Springer, and M Gaedke. Reusable awareness widgets for collaborative web applications—a non-invasive approach. In *International Conference on Web Engineering*, pages 1–15. Springer Berlin Heidelberg, 2012.
- [60] Dina Hussein, Ghada Alaa, and Ahmed Hamad. Towards Usage-Centered Design Patterns for Social Networking Systems. In *Future Information Technology*, pages 80–89. Springer Berlin Heidelberg, 2011.
- [61] D Boyd. Facebook’s Privacy Trainwreck: Exposure, Invasion, and Social Convergence. *International journal of Research into New Media Technologies*, 14(1):13–20, 2008.
- [62] Kees Van Der Sluijs and GJ Houben. Towards a generic user model component. In *Proceedings Workshop Personalization on the Semantic Web (PerSWeb’05), 10th International Conference on User Modeling*, pages 43–52, 2005.
- [63] Peter Brusilovsky. Adaptive hypermedia. *User modeling and user-adapted interaction*, 11:87–110, 2001.
- [64] Alfred Kobsa. Generic user modeling systems. *User modeling and user-adapted interaction*, 11(1-2):49–63, 2001.

- [65] Huizhi Liang, Yue Xu, Yuefeng Li, Richi Nayak, and Xiaohui Tao. Connecting users and items with weighted tags for personalized item recommendations. In *Processings of the 21st ACM conference on Hypertext and hypermedia*, pages 51–60, New York, New York, USA, 2010. ACM.
- [66] Heung-Nam Kim, Abdulmajeed Alkhaldi, Abdulmotaleb El Saddik, and Geun-Sik Jo. Collaborative user modeling with user-generated tags for social recommender systems. *Expert Systems with Applications*, 38(7):8488–8496, jul 2011.
- [67] JS Breese, D Heckerman, and C Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the Fourteenth cofnerence on Uncertainty in artificial intelligence*, pages 43–52. Morgan Kaufmann Publishers Inc, 1998.
- [68] J. Linden, G. and Smith, B. and York. Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1):76–80, 2003.
- [69] G. Adomavicius and a. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–749, jun 2005.
- [70] N Luz, R Anacleto, C Martins, and A Almeida. Lightweight User Modeling – A Case Study. In *Highlights in Practical Applications of Agents and Multiagent Systems*, pages 77–84. Springer Berlin Heidelberg, 2011.
- [71] B San Miguel, Jose M Del alamo, and Jaun C Yelmo. A Personal Data Framework for Distributed User Modelling in New Service Ecosystems. *IEEE Latin America Transactions*, 13(11):3604–3611, 2015.
- [72] CMK Cheung and MKO Lee. A theoretical model of intentional social action in online social networks. *Decision support systems*, 49(1):24–30, 2010.
- [73] RP Bagozzi and UM Dholakia. Intentional social action in virtual communities. *Journal of interactive marketing*, 16(2):2–21, 2002.
- [74] Christy M.K. Cheung and Matthew K.O. Lee. A theoretical model of intentional social action in online social networks. *Decision Support Systems*, 49(1):24–30, apr 2010.

- [75] AXL Shen, CMK Cheung, MKO Lee, and H Chen. How social influence affects we-intention to use instant messaging: The moderating effect of usage experience. *Information Systems Frontiers*, 13(2):157–169, 2011.
- [76] L Boratto and S Carta. Modeling the preferences of a group of users detected by clustering: A group recommendation case-study. In *Proceedings of the 4th International Conference on Web Intelligence, mining and Semantics (WIMS14)*, page 16. ACM, 2014.
- [77] J Masthoff. Group Recommender Systems: Aggregation, Satisfaction and Group Attributes. In *Recommender Systems Handbook*, pages 743–776. Springer US, 2015.
- [78] S. a. Golder. Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, apr 2006.
- [79] Elisabeth Joyce and Robert E Kraut. Predicting Continued Participation in Newsgroups Predicting Continued Participation in Newsgroups. *Journal of Computr-Mediated Communication*, 11(3):723–747, 2006.
- [80] Meeyoung Cha, Alan Mislove, Ben Adams, and KP Gummadi. Characterizing social cascades in flickr. In *Proceedings of the first workshop on Online social networks*, pages 13–18, New York, New York, USA, 2008. ACM.
- [81] Sarabjot Singh Anand, Patricia Kearney, and Mary Shapcott. Generating semantically enriched user profiles for Web personalization. *ACM Transactions on Internet Technology*, 7(4):22, oct 2007.
- [82] Sarah McBurney, Eliza Papadopoulou, Nick Taylor, and Howard Williams. Adapting Pervasive Environments through Machine Learning and Dynamic Personalization. *2008 IEEE International Symposium on Parallel and Distributed Processing with Applications*, pages 395–402, dec 2008.
- [83] J. R. Hauser, G. L. Urban, G. Liberali, and M. Braun. Website Morphing. *Marketing Science*, 28(2):202–223, mar 2009.
- [84] Heung-Nam Kim, Inay Ha, Kee-Sung Lee, Geun-Sik Jo, and Abdulmotaieb El-Saddik. Collaborative user modeling for enhanced content filtering in recommender systems. *Decision Support Systems*, 51(4):772–781, nov 2011.

- [85] E. Fias-Martinez, G. Magoulas, S. Chen, and R. Marcredie. Modeling human behaviour in user-adaptive systems: Recent advances using soft computing techniques. *Expert Systems with Applications*, 29(2):320–329, aug 2005.
- [86] B Hua, K.W. Wong, and C.C. Fung. Fuzzy logic based product filtering for web personalization in E-commerce. In *TENCON 2007*, pages 1–4, 2007.
- [87] BA Ojokoh and MO Omisore. A fuzzy logic based personalized recommender system. *International Journal of Computer Science, Information Technology, & Security*, 2(5):1008–1015, 2012.
- [88] PH Chou, PH Li, KK Chen, and MJ Wu. Integrating web mining and neural network for personalized e-commerce automatic service. *Expert Systems with Applications*, 37(4):2898–2910, 2010.
- [89] G Castellano and AM Fanelli. A neuro-fuzzy strategy for web personalization. In *Proceedings of the 23rd AAAI Conference on Artificial Intelligence*, pages 1784–1785, 2008.
- [90] G. Castellano, a.M. Fanelli, and M.a. Torsello. NEWER: A system for NEuro-fuzzy WEb Recommendation. *Applied Soft Computing*, 11(1):793–806, jan 2011.
- [91] E Frias-Martinez, SY Chen, and X Liu. Automatic cognitive style identification of digital library users for personalization. *Journal of the American Society for Information and Technology*, 58(2):237–251, 2007.
- [92] Fabrício Benevenuto, Tiago Rodrigues, Meeyoung Cha, and Virgilio Almeida. Characterizing User Behavior in Online Social Networks. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, pages 49–62. ACM, 2009.
- [93] Magdalini Eirinaki and Michalis Vazirgiannis. Web mining for web personalization. *ACM Transactions on Internet Technology*, 3(1):1–27, feb 2003.
- [94] L Nguyen. User model clustering. *Journal of Data Analysis and Information Processing*, 2(2):41, 2014.
- [95] G Klir and B Yuan. *Fuzzy sets and fuzzy logic Vol. 4*. Prentic hall, 1995.
- [96] Henry O Nyongesa, S. Maleki-Dizaji Tian Shicheng, Shr-Ting Huang, and J. Siddiqi. Adaptive web interface design using fuzzy logic. In *Proceedings*

- of the IEEE/ International Conference on Web Intelligence, pages 671–674. IEEE, 2003.
- [97] A Abraham. Neuro fuzzy systems: State-of-the-art modeling techniques. In *International Work-Conference on Artificial Neural Networks*, pages 269–276. Springer Berlin Heidelberg, 2001.
- [98] R Stathacopoulou, M Grigoriadou, M Samarakou, and D Mitropoulos. Monitoring students’ actions and using teachers’ expertise in implementing and evaluating the neural network-based fuzzy diagnostic model. *Expert Systems with Applications*, 32(4):955–975, 2007.
- [99] B Wu, Q Jin, X Zhou, and W Wang. Dynamically identifying roles in social media by mapping real world. In *Awareness Science and Technology and Uvi-Media computing (ICAST-UMEDIA)*, pages 573–579. IEEE, 2013.
- [100] D Yang, A Halfaker, R Kraut, and E Hovy. Who Did What: Editor Role Identification in Wikipedia. In *Tenth International AAAI Conference on Web and Social Media*, 2016.
- [101] B Gliwa, A Zygmunt, and J Koźlak. Analysis of roles and groups in blogosphere. In *Proceedings of the 8th International Conference on Computer Recognition Systems CORES 2013*, pages 299–308. Springer International, 2013.
- [102] A Zygmunt. Role identification of social networkers. In *Encyclopedia of Social Network Analysis and Mining*, pages 1598–1606. 2014.
- [103] O Arazy, F Ortega, O Nov, L Yeo, and A Balila. Functional roles and career paths in Wikipedia. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & SOcial Computing*, pages 1092–1105. ACM, 2015.
- [104] Ulrike Pfeil, Knut Svangstu, Chee Siang Ang, and Panayiotis Zaphiris. Social Roles in an Online Support Community for Older People. *International Journal of Human-Computer Interaction*, 27(4):323–347, feb 2011.
- [105] F Barcellini, F Detienne, and J Burkhardt. A situated approach of roles and participation in Open Source Software Communities. *Human-Computer Interaction*, 29(3):205–255, 2014.

- [106] V Junquero-Trabado and D Dominguez-Sal. Building a role search engine for social media. In *Proceedings of the 21st International Conference on World Wide Web*, pages 1051–1060. ACM, 2012.
- [107] HT Welser and E Gleave. Visualizing the signatures of social roles in online discussion groups. *Journal of social structure*, 8(2):1–32, 2007.
- [108] Barry Wellman. An electronic group is virtually a social network. *Culture of the Internet*, 4:179–205, 1997.
- [109] Charalampos Chelmiss and Viktor K. Prasanna. Social Networking Analysis: A State of the Art and the Effect of Semantics. *2011 IEEE Third Int’l Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third Int’l Conference on Social Computing*, pages 531–536, oct 2011.
- [110] E Bakshy, I Rosenn, C Marlow, and L Adamic. The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web*, pages 519–528. ACM, 2012.
- [111] C Buntain and J Golbeck. Identifying social roles in reddit using network structure. In *Proceedings of the 23rd International Conference on World Wide Web*, pages 615–620. ACM, 2014.
- [112] J Thom-Santelli, MJ Muller, and DR Millen. Social tagging roles: publishers, evangelists, leaders. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1041–1044. ACM, 2008.
- [113] M Kosinski. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15):5802–5805, 2013.
- [114] Y Bachrach, M Kosinski, and T Graepel. Personality and patterns of Facebook usage. In *Proceedings of the 4th Annual ACM Web Science Conference*, pages 24–32. ACM, 2012.
- [115] Blair Nonnecke, Jenny Preece, and R Voutour. Online lurkers tell why. In *Proceedings of the 10th Americas Conference on Information Systems*, number August, page 321, 2004.
- [116] D Ehls and C Herstatt. Open Source Participation Behavior - A Review and Introduction of a Participation Lifecycle Model. In *35th DRUID Celebration Conference*, 2013.

- [117] Sheizaf Rafaeli, Gilad Ravid, and Vladimir Soroka. De-lurking in virtual communities: a social communication network approach to measuring the effects of social and cultural capital. In *Proceedings of the 37th hawaii International Conference on System Sciences*, pages 1–10. IEEE, 2004.
- [118] Mario Bunge. A General Black Box Theory. *Philosophy of Science*, 30(4): 346–358, 1963.
- [119] Peter de B. Harrington. Sigmoid transfer functions in backpropagation neural networks. *Analytical Chemistry*, 65(15):2167–2168, aug 1993.
- [120] M Buscema. Back propagation neural networks. *Substance use & misuse*, 33(2):233–270, 1998.
- [121] Toby. Segaran. *Programming Collective Intelligence: Building Smart Web 2.0 Applications*. O’Reilly, 2007.
- [122] D H Anderson and L O Hall. MR. FIS: Mamdani rule style fuzzy inference system. In *IEEE International Conference on Systems, Man, and Cybernetics, (IEEE SMC’99)*, volume 5, pages 238–243, 1999.
- [123] Michael Negnevitsky. *Artificial Intelligence: A Guide to Intelligent Systems*. Pearson Education Limited, 2nd edition, 2005.
- [124] Gregory S Kolt, Richard R Rosenkranz, Trevor N Savage, Anthony J Maeder, Corneel Vandelanotte, Mitch J Duncan, Cristina M Caperchione, Rhys Tague, Cindy Hooker, and W Kerry Mummery. WALK 2.0 - using Web 2.0 applications to promote health-related physical activity: a randomised controlled trial protocol. *BMC public health*, 13:436, jan 2013.
- [125] Badrish Chandramouli, Justin J Levandoski, Ahmed Eldawy, and Mohamed F Mokbel. StreamRec: a real-time recommender system. In *Proc. ACM SIGMOD Int. Conf. on Management of Data*, pages 1243–1246, 2011.

Appendices

**A *Assessing User Engagement in a Health
Promotion Website Using Social
Networking***

Assessing user engagement in a health promotion website using social networking

Rhys TAGUE ^{a,1}, Anthony J MAEDER ^a, Corneel VANDELANOTTE ^b,
Gregory S KOLT ^c, Cristina M CAPERCHIONE ^d, Richard R ROSENKRANZ ^e,
Trevor N SAVAGE ^c, Anetta VAN ITALLIE ^b

^a School of Computing Engineering & Mathematics, Telehealth Research & Innovation Laboratory, University of Western Sydney, Sydney, Australia

^b Centre for Physical Activity Studies, School for Human Health and Social Science, Central Queensland University, Rockhampton, Australia

^c School of Science and Health, University of Western Sydney, Sydney, Australia

^d School of Health and Exercise Science, University of British Columbia, Canada

^e Department of Human Nutrition, Kansas State University, United States of America

Abstract. Remote provision of supportive mechanisms for preventive health is a fast-growing area in eHealth. Web-based interventions have been suggested as an effective way to increase adoption and maintenance of healthy lifestyle behaviours. This paper describes results obtained in the “Walk 2.0” trial to promote physical activity through a self-managed walking programme, using a social networking website that provided an online collaborative environment. Engagement of participants with the website was assessed by monitoring usage of the individual social networking functions (e.g. status post). The results demonstrate that users generally preferred contributing non-interactive public posts of information concerned with their individual physical activity levels, and more occasionally communicating privately to friends. Further analysis of topics within posts was done by classifying word usage frequencies. Results indicated that the dominant topics are well aligned with the social environment within which physical activity takes place. Topics centred around four main areas: description of the activity, timing of the activity, affective response to the activity, and context within which the activity occurs. These findings suggest that strong levels of user awareness and communication occur in the social networking setting, indicative of beneficial self-image and self-actualisation effects.

Keywords. Health promotion, online collaborative environment, physical activity, preventive health, social networking, user engagement, website, user profiling

¹ Corresponding Author: R. Tague, School of Computing, Engineering & Mathematics, University of Western Sydney, Locked Bag 1797, Penrith NSW 2751, Australia; Email: r.tague@uws.edu.au

Introduction

In the light of issues such as escalating costs, constrained resources, and the aspiration to improve longevity while maintaining the highest achievable quality of life, there is a global impetus for health system reforms. One area of emphasis in many reform agendas has been gaining increased value from health promotion and associated behaviour change supportive measures. For instance, greater use of preventive health interventions has been identified as one of the top national health priorities in Australia, as a means to reduce downstream utilisation impact on the health system [1]. Preventive health reforms can address issues arising at a population level, such as demographic changes and ageing profile, where there is much concern over increased prevalence of lifestyle related diseases, or infectious diseases and public health issues.

Recently, there has been much enthusiasm for remote provision of supportive strategies for preventive health and health promotion, making this a new and fast-growing area in eHealth and Telehealth. Web-based or online interventions have been suggested as an effective way to increase adoption and maintenance of healthy lifestyle behaviours, such as increased physical activity, improved nutrition, and reduction of alcohol and smoking habits [2, 3]. These interventions have the advantages of reaching a large user community, easily and at an affordable cost, while allowing individualised usage patterns to be developed to suit user preferences.

Online health applications are a fast growing area within the scope of Telehealth, both due to the remote nature of the computer-based source, and due to their connection with personal monitoring through self-reporting or logging data from wearable devices. The ease of access through conventional computers equipped with web browsers, or mobile devices such as computer tablets and smart phones, ensures that this area will continue to develop as consumer attractiveness and demand increases.

1. Social Networking Interventions

Within the web-based interventions domain, a particular area of focus has been the provision of social networking and social media interventions, to provide collaborative online environments in which users can interact and thereby mutually reinforce behaviour change activities. Reviews on the efficacy of these interventions [4] reveal that current evidence is not strong for the degree of behaviour change, nor can the influential factors affecting it be readily identified [5]. User engagement and motivation have been suggested as factors contributing to behaviour change [6, 7], despite a lack of clear and widely applicable definition of these aspects.

In social networking theory, the use of web-based online collaborative environments equates to the forming of social groups [8], which can provide certain influences for behaviour change [9]. These influences may be inferred from characteristic usage patterns observed within these environments [10] and the corresponding formation of social ties and norms [11]. The principles of collaboration as described theoretically in sociology [12] can be realised by establishing the existence of a number of collaborative usage patterns [13] in a certain social interactive setting, and classifying user habits accordingly.

An open issue is the choice of suitable models and metrics to describe these usage patterns, especially in a way that they can be causally linked with behaviour change. Simple measures such as degree of connectivity and volume of interaction traffic [14]

have been augmented by measuring communication, engagement and relevance [15]. The characterisation of engagement is of particular interest here, as the working hypothesis in this project has been that increased engagement through social network collaboration is the primary factor in achieving participant behaviour change. Indicators for engagement are typically related to stereotypical relationships developed between a user and the technology [16], and so the appraisal of engagement depends on the choice of the model used for describing such relationships.

2. “Walk 2.0” Project

The “Walk 2.0” study was established in 2010 with funding from the Australian National Health and Medical Research Council to develop and investigate the use and efficacy of Web 2.0 features in a physical activity promotion website to enhance self-managed programmes for daily walking and associated exercise. Using Internet or web-based physical activity interventions such as this, which incorporate innovative Web 2.0 features [17] including social networking support, have the potential to reach large groups of individuals and contribute to physical activity promotion.

The project aims to determine the effectiveness on participant attraction, engagement, retention and physical activity behaviour change in a 3-arm randomised controlled trial with sample size of more than 500 participants [18]. A further ecological trial component will explore the behaviour and experiences of users in the online aspects, based on open recruitment of a much larger number of users. It is expected that the findings from the two trials will enable generalisation of the functions provided in the web-based setting, and allow informed design, development and customisation of further online social collaborative environments for other health promotion purposes.

The study has involved the development and testing of a Web 2.0 based collaborative environment [19] (embodied in the Walk 2.0 website, www.walk.org.au) to investigate the effects of a “new generation” web-based application offering options such as blogs, posts and other social networking functions (see Figure 1). This intervention is being compared with an existing publically available Web 1.0 physical activity promotion website (the Australian 10,000 Steps website, www.10000steps.org.au) which provides a more conventional static environment, and also compared with manual logging of physical activity without computer assistance (via logbook).

By the end of the study, participants in the randomised controlled trial will have been monitored over 18 months using pedometers to log actual daily step counts, and presenting for detailed follow-up physical examinations and interviews. The information collected will enable the investigators to assess changes in levels of physical activity and other health indicators, as well as comparative impacts of utilization of the websites compared with each other and with the logbook, for user engagement and retention. It is hoped that understandings gained from analysing the characteristics of usage patterns and the influence of user interactions on the participants, will provide insights to inform the design of similar web-based interventions in the future.

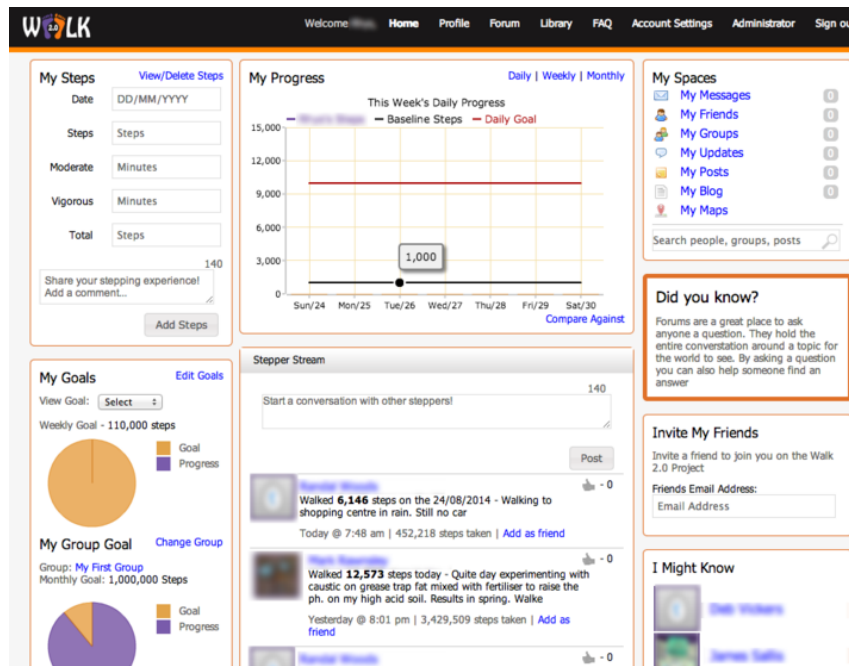


Figure 1. The “Walk 2.0” health promotion website, with social networking features.

3. User Engagement

The project methodology was not designed to specify a particular model for user engagement with the website, so a simple set of web usage measures were adopted, loosely based on the work of Burke et al. [20]. Posting was chosen as the surrogate for engagement on the basis that it offered users options for communicating with each other comparable to traditional offline conversational communication. The raw data for assessing user engagement was derived from a total of 5,481 user posts generated by 254 active and unique users on the Walk 2.0 website over a 3-month period determined by a user’s registration date, observed via transaction counts for all sessions logged in the application database. The transaction counts indicated an average of 21.6 posts per user (with standard deviation of 123.3), based on a total of 132,185 words posted by all the users. The average number of words per post was 24.1, equating to an average of 520.4 words per user overall.

Five separate posting functions were provided for users in the application: Progress, Status, Private, Blog, and Forum. Progress is a public posting of an automated fixed single line message of user step count combined with open text situational information (comment) provided by the user, which is broadcasted to the user base upon posting (users keep track of daily number of steps taken using a pedometer provided by the research team). Status is an open-text user-entered short public posting of current user situational information. Private is an open-text user-entered private posting of a message sent to another user profile. Blog is an extended open-text user-entered long public posting to an individual user-designated posting area through the user profile. Forum is an open-text user-entered public posting in one of six public topic categories, to allow discussion in a series of interleaved messages from multiple users responding to each others’ comments. Only Forum posts can be publicly replied to by all users.

Table 1 shows the distribution of posting function usage across the 5,481 user sessions. The automated Progress posts made up by far the greatest volume of session traffic on the website at 79.4%; it is surmised that this is because of the ease of generating the posting automatically. The next highest source of traffic was Status posts at 13.3%, a form of public post with wide visibility without initiating further direct interactive public communication. Private posts were at 4.0% and Blog posts at 1.8%; this indicates that users were disinclined to post where exposure was limited and/or less social in nature. Use of Forum posts was negligible at 0.5%. These results reveal that users generally preferred contributing through broadcasting of semi-automated open text information about their individual physical activity, with some lesser amount of private communication with friends and reflective blogging, than seeking a highly collaborative mode of communication.

Table 1. Overall posting function usage across all users.

Posting Function	Posting Count	Posting Percentage
Progress	4350	79.4
Status	727	13.3
Private	220	4.0
Blog	96	1.8
Forum	26	0.5

Next, we considered the habits of users who were most active in the online collaborative environment to establish whether they showed similar usage patterns. It was established that 50% of the posts (2,781) were made by 21 users out of 254 (8% of all users), and 25% of the posts (1,416) were made by 8 users (3% of all users). The highest number of posts made by the top user was 259, almost 5% of the total number of posts and equating to a rate of approximately 3 posts per day. The average number of posts was equalled or exceeded by 65 users (25% of all users). Fewer than 5 postings were made by 126 users (50% of users), and only one posting was made by 77 users (30% of all users). This highly skewed distribution is typical of social networking sites [20]. Figure 2 shows the posting rate, ordered (left to right) from most to least active users.

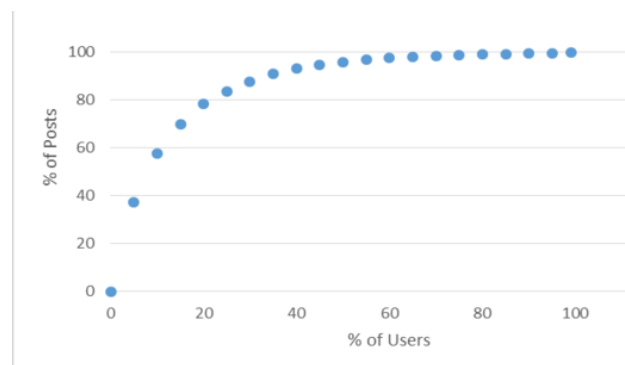


Figure 2. Cumulative posting rate for all users.

Table 2 shows the distribution of the number of postings of all types made in the top 50% (25%) of postings. The patterns of usage are comparable with those in the whole population of users shown in Table 1, with some discernible increase in Progress postings and decrease in Status, Private and Blog postings, which is even more marked in the top 25%. It appears that more online-active users wish to spend less effort communicating in open text and instead concentrate on the easy automated broadcast mechanism offered by Progress postings.

Table 2. Posting function usage by most active users making 50% (25%) of all posts.

Posting Function	Posting Count 50% (25%)	Posting Percentage 50% (25%)
Progress	2251 (1261)	80.9 (89.1)
Status	408 (116)	14.7 (8.2)
Private	78 (22)	2.8 (1.6)
Blog	37 (14)	1.3 (1.0)
Forum	3 (2)	0.1 (0.1)

4. Topic Analysis

Further indication of engagement and the emergence of collaborative social norms between users can be gauged by analysis of task-related topics covered during postings. Following a simplified probabilistic topic modelling approach [21], a frequency count was made of all unique words occurring within the 132,185 words contributed by user open text. A total of 240 words were found to occur 50 or more times, with the most frequent words occurring almost 1,000 times (day = 974; steps = 885). Trivial words such as prepositions, which did not convey meaningful concepts, were excluded from the analysis.

A random sample of approximately 10% of the text in the 5,481 postings was read and four distinctive topics were identified, associated with the typical user behaviour and general social environment within which the main task (i.e. physical activity) takes place. These topics were:

- Description of the activity (e.g. intensity, timing, variety)
- Timing of the activity (e.g. time of day, duration, repetition)
- Affective response to the activity (e.g. emotions, reflections, attitudes)
- Context in which the activity occurs (e.g. other forms of physical activity, companionship, lifestyle, location).

Words associated with each of these topics were grouped according to similarity of meaning or intent. The similarity criteria were developed from the postings read, by establishing links between frequently occurring words. Table 3 provides details for some of the more frequent words in these groupings. There are relatively few groupings in each topic, indicating that the posts are generally narrowly focussed around topical themes and users are not directing their comments to matters outside the online social collaborative environment.

Table 3. Selection of frequent words for the four topics.

Topic	Words
Description of the activity	step/steps/walk/walked/walking (2,245); pedometer (198); goal (121); extra (97); exercise (84); count (75);
Timing of the activity	time/hour/day/week/month (1,224); night/morning/early/late (359); today/tomorrow/yesterday/daily (279);
Affective response to the activity	few/some/more/short/long/many/much/again (1,023); good/better/hard (418) ; active/busy (234); hope/happy (115);
Context in which the activity occurs	house/home (286); weather/rain (174); gym (114); group/family/couple (203); dog/dogs (162); shopping (98); bike (71); beach (70); down/up/around/back (696);

Considering the size of word counts for major groups of concepts associated with the undertaking of the desired physical activity, it can be inferred that users were posting comments that were highly relevant to their activities in meeting their goals. Very few words indicating distraction from purpose were detected. These findings suggest that strong levels of user awareness and communication of intent occur in the collaborative social media setting, which are indicative of beneficial self-image and self-actualisation effects. This claim is supported by the presence of a number of strong affective responses displaying positive attitudes and messages of mutual encouragement.

5. Conclusion

This paper has presented some initial findings from a project which aimed to explore the effect of a Web 2.0 based social collaborative user environment, specifically for influencing user engagement for increasing and maintaining their level of physical activity. Without a specific model for the emergence of social relationships which such an environment could support, a basic analysis was undertaken using simple models of social collaboration behaviour.

The patterns of use for website functions that were provided for posting information revealed that broadcast, non-interactive communications (in the form of comments) were preferred to interactive or private discussions. This may be a consequence of the focussed nature of the task, which is a highly individual and personal challenge endeavour, or the approach to recruitment being a randomised controlled trial where participants were socially unrelated. The topic analysis showed that users' attention is very well aligned with the task, as most high frequency words are directly related to the undertaking of the physical activity. Further insights could be gained by conducting interviews with users to establish reasons for their preferences in use of the functions, and whether they were subject to peer pressure when using those functions which encouraged them to adopt a different priority than they naturally would.

It is acknowledged in the literature [22] that web based interventions are complex to evaluate, and that their effectiveness is subtle to measure. It would be beneficial to develop and apply specific models to analyse such situations, based on the conjunction

of available functions for user communication and interaction, and the user goals which exist in the underlying intervention. Use of such models could provide more consistent and comparable quantitative results concerning user engagement and reinforcement.

Acknowledgement

This research was undertaken with partial support of the Walk 2.0 project funded by the National Health and Medical Research Council (Project Grant GNT0589903). The contributions of all members of the Walk 2.0 project are noted here with gratitude.

References

- [1] Australian National Preventive Health Agency (ANPHA). (2013). *State of Preventive Health 2013*. Report to the Australian Government Minister for Health. Canberra; ANPHA.
- [2] Webb, T., Joseph, J., Yardley, L., & Michie, S. (2010). Using the internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *Journal of Medical Internet Research* 12(1):e4.
- [3] Korda, H., Itani, Z. (2013). Harnessing social media for health promotion and behavior change. *Health Promotion Practice* 14(1):15-23.
- [4] Nakhasi, A., Shen, A. X., Passarella, R. J., Appel, L. J. & Anderson, C.A.M. (2014). Online social networks that connect users to physical activity partners: A review and descriptive analysis, *Journal of Medical Internet Research* 16(6):e153.
- [5] Kohl, L. F. M., Crutzen, R., & de Vries, K. (2013). Online prevention aimed at lifestyle behaviors: A systematic review of reviews. *Journal of Medical Internet Research* 15(7):e146.
- [6] Maher, C. A., Lewis, L. K., Ferrar, K., Marshall, S., De Bourdeaudhuij, S., & Vandelandotte, C. (2014). Are Health Behavior Change Interventions That Use Online Social Networks Effective? A Systematic Review. *Journal of Medical Internet Research* 16(2): e40.
- [7] Laranjo, L., Arguel, A., Neves, A. L., Gallagher, A. M., Kaplan, R., Mortimer, N., Mendes, G. A., & Lau, A. Y. S. (2014). The influence of social networking sites on health behavior change: a systematic review and meta-analysis. *Journal of the American Medical Informatics Association* amiajnl-2014-002841.
- [8] Backstrom, L., Huttenlocher, D., Kleinberg, J., & Lan, X.. (2006). Group formation in large social networks: membership, growth, and evolution. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '06)*. ACM, New York, 44-54.
- [9] Cao, L. (2010). In-depth behavior understanding and use: The behavior informatics approach. *Information Sciences* 180(17):3067–3085.
- [10] Benevenuto, F., Rodrigues, T., Cha, M., & Almeida, V. (2012). Characterizing user navigation and interactions in online social networks. *Information Sciences*, 195:1–24.
- [11] Cheung, C. M. K., & Lee, M. K. O. (2010). A theoretical model of intentional social action in online social networks. *Decision Support Systems* 49(1):24–30.
- [12] Wood, D. J. (1991). Toward a comprehensive theory of collaboration. *The Journal of Applied Behavioral Science* 27(2):139–162.
- [13] Thomson, A.M., Perry, J.L., & Miller, T.K. (2007). Conceptualizing and measuring collaboration. *Journal of Public Administration Research and Theory* 19(1):23–56.
- [14] Mislove, A., & Marcon, M. (2007). Measurement and analysis of online social networks. *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement* 29–42.
- [15] Drula, G. (2012). Social and online media research – data, metrics and methods. *Review of Applied Socio-Economic Research* 3(1):77-86.
- [16] O'Brien, H., & Toms, E. G. (2008). What is user engagement ? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information and Technology* 59: 938–955.
- [17] O'Reilly, T. (2007). What is Web 2.0: Design patterns and business models for the next generation of software. *Communications & Strategies* 1:17-37.
- [18] Kolt, G. S., Rosenkranz, R. R., Savage, T. N., Maeder, A. J., Vandelandotte, C., Duncan, M. J., Caperchione, C. M., Tague, R., Hooker, C., & Mummery, W. K. (2013). WALK 2.0-Using Web 2.0

- applications to promote health-related physical activity: A randomised controlled trial protocol. *BMC Public Health* 13:436.
- [19] Tague, R., Maeder, A., & Basilakis, J. (2014). Adaptive web framework for online collaborative environments. *Australian Journal of Intelligent Information Processing Systems* 13(4).
 - [20] Burke, M., Kraut, R., & Marlow, C. (2011). Social capital on Facebook: Differentiating uses and users. [In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems](#), ACM, New York 571-580.
 - [21] Blei, D.M. (2012). Introduction to probabilistic topic models. *Communications of the ACM* 55(4):77–84.
 - [22] Murray, E. (2012). Web-based interventions for behaviour change and self-management: potential, pitfalls, and progress. *Medicine 2.0* 1(2):e3.

B *Adaptive Web Framework for Online Collaborative Environments*

Adaptive Web Framework for Online Collaborative Environments

Rhys Tague¹, Anthony Maeder¹, Jim Basilakis¹

¹School of Computing, Engineering and Mathematics

University of Western Sydney Australia

Email: r.tague@uws.edu.au, a.maeder@uws.edu.au, j.basilakis@uws.edu.au

Abstract. Online social networks are online collaborative environments where users can interact with other users to create, annotate, communicate, modify and delete media together. However, online social networks do not provide adaptive interfaces, by which one user might be presented with tools which are relative to their collaborative behaviour with another. This work presents behavioural stereotypes derived from the literature to help understand and predict user behaviour in an online social network underpinned by an adaptive interface framework. A user's behaviour stereotype will be determined by a proposed classification framework which processes usage data in real time.

Keywords: semantic web, collaborative work, adaptive systems, user interfaces, human computer interaction

1 Introduction

With the continual rise of online social networks, which are essentially software driven user environments which are “personal” in nature, users naturally desire more personal interaction experiences. We do see some personalisation in content [1], and it is the predominant adaptive medium today, but tool or interface adaption is limited.

With the use of interface adaption in an online collaborative environment, such as a online social network, a user may be more efficient and engaged in interaction with another user compared to a static interface. This paper presents five behaviour stereotypes which are derived from the literature to be used in a web personalisation framework to provide adaptive interfaces in a online social networking environment. The framework will adapt an interface based on a user's behaviour over time when interacting with another user. By providing interface adaption based on user to user links in combination with behaviour, a user can be represented by a behavioural stereotype to increase collaboration.

2 Background

2.1 Social Groups and Collaboration

When one looks at online social networks as an environment for collaboration one also needs to look at the dynamics of groups and the different components which make them successful. Online social networks tend to be made up of social groups as social networks fulfill some need or user interest [2]. The formation is not designated but continually grows and evolves over time as more people become interested or believe the group fulfils a need. A result of this is social groups tend to be fluid and dynamic and they lack explicit structure. Instead, their structure is derived by the interactions the members carry out with each other over time. Moreland's Model of Group Socialisation [3] and Biddle's characteristic behaviours [4] show us ways we can represent how the structure of such social groups form. Figure 1 shows Moreland's model in combination with Biddle's model. From this figure the circular and evolving nature of the group's structure throughout its life is apparent. Each member of a social group applies this model to assign roles to other members of the group. By so doing this the member is capable of understanding another members role and their own characteristic behaviours towards the other member. This dynamic role assigning of social groups is how online social networks can create environments which are made up of millions of users and still provide successful collaboration. Wikis are an example of this. Bryant *et al* [5] give an example of how a Wiki's environment based on the different behaviour of users can lead to successful collaboration. Without the different user-adopted behaviours users take on the collaboration would be unsuccessful.

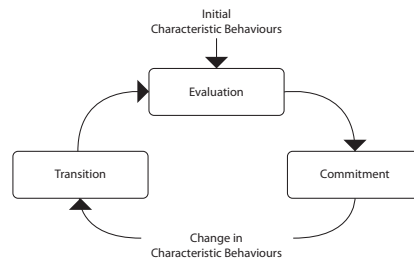


Fig. 1. How social group roles are implicitly created. Moreland and Biddle.

2.2 Adaptive Interfaces

Adaptive methods for web interfaces have been around since the early 1990s starting out as adaptive hypermedia [6]. Adaptive hypermedia was the method of adapting a user's navigation links based on their navigational path. Over the years it has moved into content [1] and web user interfaces [7]. For instance Linden *et al* [1] shows how product recommendations are made based on past media consumption, and the effects of presenting such recommendations. Schmidt *et al* [7] presents a method of personalising a user interface in a rich internet application which the user has been interacting with and the consequence of that personalisation.

There are two general components of adaptive environments. These components work together to achieve an approximation of what a user is doing or will do in relation to user interaction. The first component is a *user model*. A user model is a storage of user characteristics which are collected over the life of the user profile. The characteristics stored are used to build knowledge about the user, therefore it is a fundamental component to any successful adaptive interface. There are many methods for developing user models from researchers intuition to formal structure based methods such as Pohls [8]. However, the end user model and its design is determined by its purpose. The more data is collected in user models over time the more accurate in predicting and understanding different characteristics of users. The second component of adaptive environments is *soft computing* techniques which are used to data mine or to find patterns of behaviour in the user model. There are many ways to carry out data mining and Frias *et al* [9] presents a detailed review of the different techniques one can use. For the framework, neural networks and fuzzy logic are used for the soft computing component.

Typically user models are used to represent users in singular interaction patterns, meaning the focus of the user model is on the tools which the user has used or the content which they have consumed. There has not been much work on collaborative user behaviour modelling, such as taking into account the user profile link and behaviour. However, there is still a lot of literature on modelling users in online social networks [10, 11]. Schifanella *et al* [10] work looks at the characteristics of user profiles in online social networks to determine how a user's actions are similar to the users close to their user profile. However, this work doesn't look at the collaborative behaviour between the two user profiles, but instead focuses on the user's interests and folksonomies. Benevenuto *et al* [11] also look at user characteristics, but their work focuses more on click streams and various similar variables such as login rates which they mine to try and represent user workloads when using online social networks.

What is to be adapted tends to be part of the design process of the adaptive interface framework. Each tool or interface component has to be designed so that it can change over its life or provide another type of action, for instance [7] uses Asynchronous Javascript and XML to dynamically build new components of the interface during run time, where [12] get the user to fill out questions to determine their cognitive style and then *morph* the web-page by replacing some elements with others according to the user's cognitive style. Both methods are vastly different however they still provide an adaptive interface based on the user's characteristics.

2.3 User Behaviour In Online Social Networks

In user modeling terms, behaviour can be represented as interests, tasks, goals, beliefs and/or interactions. For this work we will be focusing on the behaviour of interaction where two or more profiles are involved in the interaction. Currently a majority of the literature does not focus on collaborative behaviour user modelling, however there are some typical behaviours, or usage patterns, which are described in the literature based on observations of online social networks [13, 14]. These behaviours and usage patterns are not true representations of a user's behaviour. For instance Hussein *et al* [14] presents various usage patterns which are seen in online social networks, however they are traditionally used more in software design and development. For this reason the general purpose of these usage

patterns need to be examined to determine the underlying behaviour, as some of these usage patterns when refined encompass many tools which fall under a general behaviour.

There is a lot of literature in regards to psychological behaviour which is determined by behavioural tests or Natural Language Processing, such as a *The Big Five* behaviours and shyness [15]. They discover these behaviours through questionnaires and tests the user completes. The results are more subjective compared to quantitative methods such as using click stream data and other means.

Overall literature on collaborative behaviour in online social networks is very limited. Through the analysis of usage patterns presented by [14] one can generalise behaviour by looking at the purpose of the tools available in the online social network.

3 Behaviour Stereotypes in Online Social Networks

As there is very limited collaborative behaviour literature in online social networks which does not just focus on psychological aspects, we have to look at the underlying roles or behaviours users carry out when using tools in online social networks. The work by Hussein *et al* [14] can be used as a general starting point, but we have to look at the tools of these usage patterns and their purpose in a collaborative context to determine a general behaviour. The general behaviour for this work will be called a behavioural stereotype, as the behaviour is not a true representation of the user's behaviour but a generalisation of interaction the user has performed.

If we look at tools such as microblogging/status updates, file uploads, or blogs we can state that the user is publishing content when using these types of tools. For instance microblogging allows users to publish small messages with a time stamp to be shared in another tool for user's who are lurking [16]. The underlying behaviour stereotype here is publishing so the user is a *publisher* when using these tools.

Tools which allow users to annotate content by providing tagging, comments, and ratings give more meaning to the original media and makes it easier to retrieve [17]. The behaviour stereotype chosen for these types of tools is *annotator* - a user who uses annotating tools to add more meaning to other user's content.

Very recently tools are appearing whose functions are dedicated to sharing content within online social networks from other user profiles. We have seen similar action in microblogging where people repost the message on their own profile instead of producing an exact copy [18]. However the new tools are more dedicated more towards creating exact copies of the content, which then allows the user to annotate the original. This allows content which has been already published to be passed on internally through an online social network. The general behaviour chosen here is sharing, where the user shares content which they have not created. The behaviour stereotype chosen for these distinctive tools is *sharer*.

Majority of tools which exist on social networks involves asynchronous communication, however there are some tools which allow communication in a synchronous manner, such as chat or video conferencing. Even though there are not many tools which adopt this type of behaviour they are still a vital component of any online social network. The behaviour stereotype for these types of tools is *communicator*.

Our final behaviour stereotype model chosen is *lurker* [13]. The lurker behaviour stereotype represents any user who consumes media but doesn't contribute or engage in the media. Lurkers are the most common type of behaviour seen in online social networks today [13].

These behaviour stereotypes will be used to add semantic meaning to the interactions user's carry out when collaborating with other users. They will be used by the framework to represent state and function of the interfaces presented in a online social network.

4 Framework

4.1 Framework Design

The adaptive interface framework is made up of various components which exist on the client side and server side of the online social network. Figure 2 gives a representation of the adaptive interface framework. Showing the various components: Fuzzy Logic System, Neural Network System, User Model Storage, Experience Storage.

The Fuzzy Logic system and the Neural Network System are the components of the framework that provide the approximation and determination of the behaviour stereotype. The Fuzzy logic determines the stereotype during an interaction session by using Fuzzy rules to determine fuzziness in regards to the user being engaged or if the tool is the dominate tool during the session. The Fuzzy Logic system is programmed in Javascript, so it produces a fuzzy set which will be passed back to the Neural Network via AJAX to determine the behavioural stereotype.

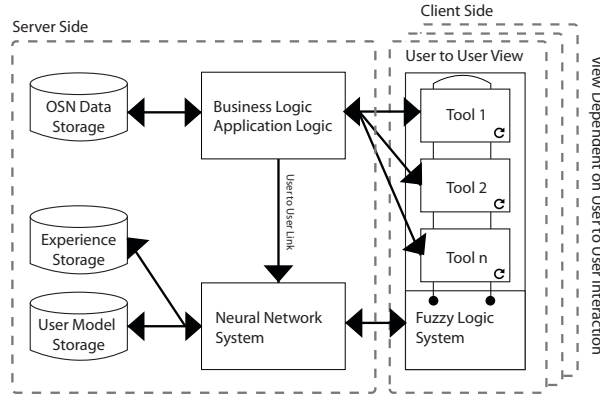


Fig. 2. A representation of the adaptive interface framework.

The type of Neural Network which will be used for this framework is a dynamic node allocation network. This type of Neural Network was chosen as it can dynamically update the nodes of the network when the user starts to interact with another user in their social network. With this the neural network is capable of providing a separate approximation for all the users the user interacts with. It is built out of the Python programming language with a noSQL database to store the experience information which the neural network outputs when it is put to training.

The user model storage of the system is a storage of all the behaviour stereotype fuzzy sets and defuzzified results with timestamps. This user model is used to train the neural network on intervals and the results of the training stored in the Experience storage. The user model can be represented visually and can be used to adjust the results of the adaption. Figure 3 gives the visual representation that the user can use to increase or decrease a behaviour stereotype for a particular user they interact with.

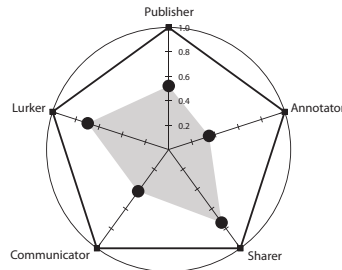


Fig. 3. A visual representation of the behaviour stereotypes one user may have with another user.

The framework understands which behaviour stereotypes the user is undertaking as the behaviour stereotype is applied to the interface tools as semantic data. This way the system will be able to record clicks and interactions from the user and provide meaning to each event which occurs. These type of semantic events are similar to the work by Schmidt *et al* [?].

4.2 Framework Testing

The framework is still being implemented into a social network to determine if the adaptive interface components in the online social network do indeed improve the user's engagement in a collaborative environment. There will be two trial arms which will be run to compare against, one being a control group where the user's behaviour stereotypes are collected but not acted on, and the second the experimental, where the behaviour stereotypes of the users are collected and the interface is adapted based on the users being interacted with.

5 Discussion

There is a great need for understanding collaborative behaviours in online social networks and modelling those behaviours so environments are not generic across every user profile. With an adaptive interface framework like the one presented one can provide an interface moderated not merely on context but on actual collaborative behaviour, even though it is only a stereotype. The usage patterns presented by Hussein *et al* [14] do provide insight into users' use of online social networks, but as mentioned they are more for software design applications than user modelling applications.

The number of behaviour stereotypes may grow in size over time as new tools are designed and implemented. This is normal for social networks which are dynamic in their nature as the user's characteristic behaviours determine the networks structure. As a result, any framework which applies any type of adaptive technology requires the flexibility handle this change. As shown this framework takes into consideration the dynamic behaviours displayed in social networks as well as account for the growing membership of interactive users.

6 Conclusion

An examination of existing social groups show how a social group's structure is implicitly defined by its members' characteristic behaviours. In addition, it has been discussed how online social networks are online collaborative environments, and yet the interfaces are generic for all users. To overcome this limitation this paper presents an adaptive interface framework that can be employed in an online social network which adapts a user's interface based on their behavioural stereotype and the user they are interacting with. We have also used semantics on interface tools to determine a user's behaviour stereotype to help understand the type of behaviours a user is undertaking with another.

References

1. Linden, G, Smith, B, Your, J., Amazon.com recommendatoin: Item-to-item collaborative filtering, *Internet Computing*, IEEE, pg. 80–89 (2011)
2. Forsyth, D., *Group Dynamics* Wadsworth Pub Co., (2009)
3. Moreland, R., Levine, J., Socialization in small groups: Temporal changes in individual-group relations, *Advances in experimental social psychology*, 15, pg. 123–192 (1982)
4. Biddle, B., Recent development in role theory, *Annual review of sociology*, pg. 67–92. (1986)
5. Bryant, S., Forte, A., Bruckman, A., Becoming wikipedia: transformation of participation in a collaboriatve online encyclopedia, *Proceedings of the 2005 International ACM SIGGROUP conference on upporting group work*, ACM, pg. 1–10. (2005)
6. Brusilovsky, P., Adaptive hypermedia, *User modelling and user-adapted interaction*, 11:1, pg. 87–110, (2001)
7. Schmidt, K., Dörflinger, J., Rahmani, T., Sahbi, M., Stojanovic, L., Thomas, S., An user interface adaption architecture for rich internet applications., *The Semantic Web: research and nApplications*, pg. 736–750, (2008)
8. Pohl, W., Logic-based representation and reasoning for user modeling shell systems, *User Modeling and User-Adapted Interaction*, 9:3, pg. 217–282, (1999)
9. Frias-Martinez, E., Magoulas, G., Chen, S. Marcredie, R., Modeling human behavior in user-adaptive systems: recen tadvances using soft computingn techniques, *Expert Systems with Applications*, 29:2, pg. 320–329, (2005)
10. Schifanella, R., Barrat, A., Cattuto, C., Markines, B., Menczer, F. Folks in folksonomies: social link prediction from shared metadata, *Proceedings of the third ACM international conference on web search and data mining*, ACM, pg. 271–280, (2010)
11. Benevenuto, F., Rodrigues, T., Cha, M., Almeida, v., Characterizing user behavior in online soical networks, *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference*, ACM, pg. 49–62, (2009)
12. Hauser, J., Urban, G., Liberali, G., Braun, M., Website morphing, *Marketing Science*, 28:2, pg. 202–223, (2009)
13. Nonnecke, B., Preece, J., Shedding light on lurkers in online communities, *Ethnographic Studies in Real and Virtual Environments: Inhabited Information Spaces and connected Communities*, Edinburgh, pg. 123–128, (1999)
14. Hussein, D., Alla, G., Hamad, A. Towards usage-centered design patterns for social netowrking systems, *Future Information Technology*, pg. 80–89, (2011)
15. Ryan, T., Xenos, S., Who uses facebook? an investigation into the relationship between the big five, shyness, narcissism, loneliness, and facebook usage, *Computers in Human Behavior*, 27:5, pg. 2658–1664, (2011)
16. Zhao, D., Rosson, M., How and why people twitter: the role that micro-blogging plays in informal communication at work, *Proceedings of the ACM 2009 international conference on supporting group work*, ACM, pg. 243–252, (2007)
17. Ames, M., Naaman, M., Why we tag: Motivations for annotation in mobile and online media, *in Proceedings of the SIGHI conference on Human factors in computer systems*, ACM, pg. 971–980, (2007)
18. Starbird, K., Palen, L., Pass it on?: Retweeting in mass emergency, *International Community on Information Systems for Crisis Response and Management*, pg. 1-10, (2010)

**C *Classifying Collaborative Behavior in the
Form of Behavioral Stereotypes in
Collaborative Mobile Applications***

Classifying collaborative behavior in the form of behavioral stereotypes in collaborative mobile applications

Rhys Tague, Anthony Maeder & Jim Basilakis

School of Computing, Engineering and Mathematics, University of Western Sydney, Australia

ABSTRACT: Online Social networks empower users to collaborate with other users through complex interfaces. They enable users to take on various behaviors to achieve an objective or goal together. However, with the rise of smart devices and their small view ports these interfaces have been restricted. This results in the user having to wait until they have access to a desktop version before they can interact with these complex interfaces again. This paper presents a framework for classifying collaborative behavior in the form of Behavioral Stereotypes. In addition it presents initial results of a implementation of the framework in a collaborative mobile application to demonstrate its ability to help understand user behavior and how it changes from social ties users establish.

Keywords: classifying behaviour; collaborative environment; user modelling; mobile computing

1 INTRODUCTION

Actions speak louder than words. This old proverb can be mapped to a user's interactions with an interface. When a user clicks, slides, scrolls, or manipulates an interface their actions not only result in system feedback, but also behaviors and intentions to achieve an overall goal or desire.

Today, with Online Social Networks (OSNs) making up 66% of online usage within Australia alone [1], users are able to create new behaviors and have new intentions with other users through interfaces. This has derived from OSNs enabling users to connect and collaborate together through interfaces to carry out collaborative tasks together. Users are encouraged to create social ties with other users or nodes, such as media. For example, GitHub, an online programming code sharing and publishing environment, enables users to achieve code production and maintenance through the use of various interfaces [2]. In this environment users are able to take on varying stereotypical behaviors you would normally associate within a software development company. For instance, a user is able discuss future development releases, identify bugs, provide patches and also create documentation. In addition, they are able to do this for multiple projects with different subsets of users.

Although these environments can provide such interfaces, their ability to provide the same variance through smart devices is limited. The viewport smart devices possess can be a fraction of the size of its desktop counterpart. For this reason, interfaces are more focused and sub-level interactions are found deeper within the interface or don't exist at all. It is up to

the designer of the interface to choose the most crucial or important features for their environment [3]. Therefore, the capability of the user to carry out stereotypical behavior, like that found in complex interfaces, diminishes. Instead, they are left only with simplistic behaviors that may or may not coincide with their normal behavior and current needs.

By understanding a user's stereotypical behavior within a smart device environment, users could be provided with personalized and frictionless interfaces to overcome this barrier found within smart device environments.

This paper presents initial results for classifying stereotypical collaborative user behavior within a smart device environment. It demonstrates this through the use of a collaborative mobile application for the capture of user interactions with other users or social ties. These interactions were classified through the use of a framework implemented within the mobile application environment.

2 BACKGROUND

Today with OSNs the type of information that can be collected is endless. This information can consist of user preferences or their interactions within an environment. This information though can be difficult to understand without a formal approach. One approach is user modeling [4]. It provides methods and theories on how to understand user behavior and capture data in a meaningful way to create a model that represents a user. Although effective, a user model's characteristics are based on the purpose of the user model's existence,

which leaves it open for interpretation. For this reason, in recent times there have been many approaches to user modeling, and more importantly, classifying user behavior, or roles, within OSNs [5][6]. These approaches focus on singular users and their overall behavior within an OSN as an entity, or their position within the user population, or a sub-set of users. The behaviors, or roles, used for these studies are general role or *status* based, which is dependent on their content types and the number of items published. There are other approaches for classifying which exists using online social network analysis and is more aligned with this work [7]. Instead of how much content a user publishes, it focuses on the connections made and actions taken on those to find patterns of behavior and usage. The approach is after the interaction has taken place, and therefore, can only be used as future steering instead of assisting the user at the point of interaction.

Many user-adaptive techniques have been used to prevent this lag time seen with these approaches to OSNs and other collaborative environments [8][9]. They achieve this through machine learning techniques and other logic based systems. However, their focus is either on a individual user or the whole population, such as in collaborative filtering [9].

3 METHOD

3.1 Environment

This study involved the development of an online collaborative environment focusing on tools and actions towards collaboration. The resulting environment was a mobile application built using the PhoneGap Software Developer Kit. The mobile application was provided in the Google Play Store as a free download with public access. Upon installation a basic registration form with personal details, such as email address, was presented to all participants before they could be part of the environment. The only criteria for users to enter the environment were that they were 18 years of age or older.

Within the environment the tools available promoted social interaction and collaboration for the purpose of content creation and discovery. The theme for content was around establishments that served coffee or food. The tools provided in this collaborative environment were simple traditional tools found in other similar environments such as Yelp [10]. Every tool within the environment had three major objectives, one, to promote collaboration between users for the purpose of content creation, two, awareness of other users and locations and their interactions, and finally, social tie creation. This involved users being able to communicate with each other and also be notified when a user carried out an action towards a social tie that is relative to another. Users were also able to favorite other users and locations to “keep tabs” on their interactions.

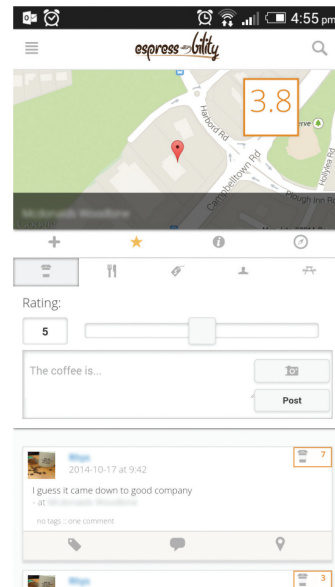


Figure 1. Screenshot of content publication tools.

The sets of tools available to users were: *content publication*, *content reply*, *content annotation*, *expanding views*, *location creation*.

Content publication empowers users to create new and original content for a geographical location. In addition, they were also able to create new geographical locations for other users to interact with. Fig. 1 presents a screenshot of the tools available for this category. They allowed a user to segment their publication into one of five service areas an establishment possesses: Coffee, Food, Price, Customer Service, and Location. They were also able to compound the publication with images and a review of 140 characters or less. Upon completing the content publication the content created is accessible globally within the environment for others to discover. This is shown in Fig. 2. A global stream of all user content is presented for user's to discover and be aware of other user behavior and interactions.

Content annotation is the annotation in the form of content replies to and tagging content published. Annotation allowed users to establish a dialogue around content a user may have published. Users were also able to annotate locations in the form of small brief summary statements so others users could discover when finding a new location. Fig. 3 provides a screenshot of tagging user content. The tags result in hyperlinks allowing users to “follow breadcrumbs” to discover new content. This category of tools also allowed user's to add comments to existing content in the form of additional 140 character replies.

Expanding views enabled users to explore additional content for user publication or geographical location. Expanding views are a common element used in smart device applications to cater for the small view port they possess [3]. Fig. 4 gives an example of an expanded view for the user to investigate additional content about a location.

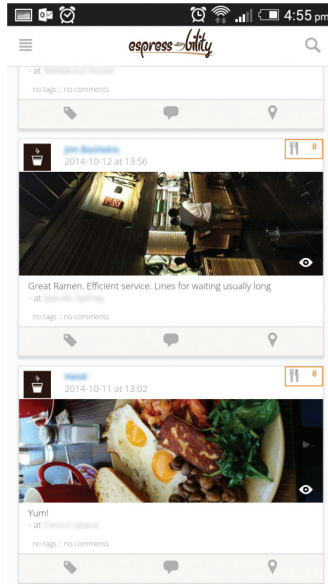


Figure 2. Screenshot of global status stream.

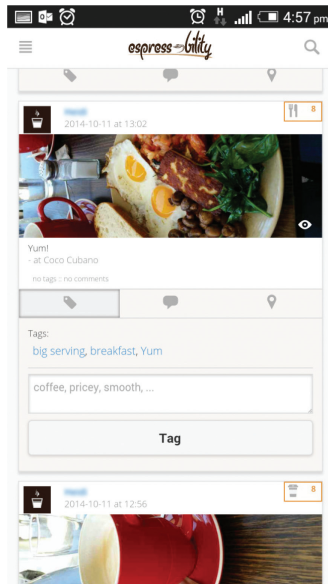


Figure 3. Screenshot of available annotation tools.

Location creation was the final tool set available. It allowed users to use their geographical location to generate new locations within the environment for other users to annotate and publish new content to.

3.2 Framework

Fig. 5 presents the model of the framework and its components implemented within the application. The framework extends between client and server side of the application environment to achieve classification. Having the framework on the client enables real time logging of user interaction that may or may not result in content publication. For instance, a user carries out large amounts of content scrolling that doesn't result in content creation or manipulation.

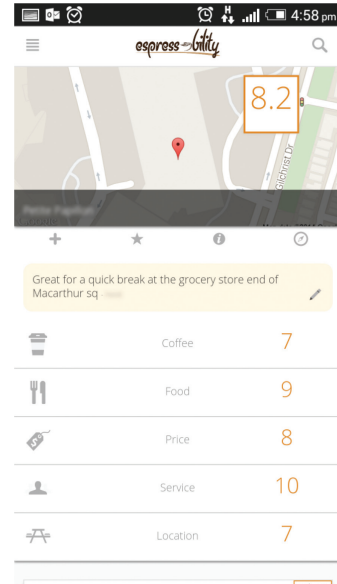


Figure 4. Screenshot of an expanding view.

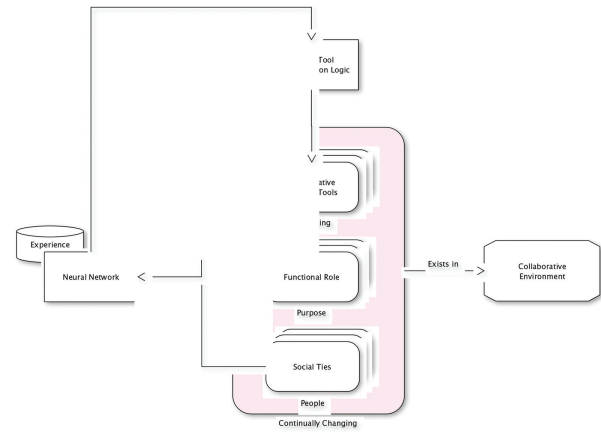


Figure 5. A simple representation of the framework components.

Although the framework extends to the client, or smart device, it is limited to meta-data and basic interaction identification data. The meta-data consists of the tool's identification, and the stereotypical behaviors the tool are categorized in. The behaviors used are derived from behavioral stereotypes a representation of common behavior when interacting with social ties [11]. Only three behavioral stereotypes were chosen for this environment: *publisher*, *annotator*, and *lurker*. These three behavior stereotypes align with the tool sets used through out the environment. A *Publisher* stereotype results when a user predominantly creates content with or for a common social tie. For instance, when they continually visit a particular coffee location or user profile and publish new content for that particular user or location. An *Annotator* stereotype results when a user predominantly annotates a common social tie's interactions or content. The final stereotype, *Lurker*, is a result of a user consuming or discovering content that a social tie creates or annotates, yet does not interact with.

The server side of the framework composes of a dynamic node neural network for the purpose of generalizing behavior towards social ties and user collaboration. This type of artificial neural network was chosen because of its ability to be trained over a period of time, and also the ability to not know and adapt to the environment's social network that will exist in the future.

To maintain performance, the training of the neural network was carried out for every user session that was 24 hours old or more. The training consisted of back propagating the cached interaction packages sent from the client. Below is an example of such an interaction package as a JSON string:

```
{'inputs': [1, 'review-creator', 'u3'],
'outputs': ['publisher', 'annotator', 'lurker'],
'train-key': 'publisher'}
```

From this example the package is made up of three objects: Inputs, outputs and a training key. These objects are used for the neural network during back propagation (training). Inputs consist of a user id, interface tool id, and social tie id. The nature of the dynamic node neural network enables the outputs to be any number of behaviors. This flexibility enables new tools that may present new behaviors to be integrated in the environment without having to alter the environment or framework in the future. Finally the last segment of the package is the training key. This is the behavior stereotype shown by the user during the interaction with the tool, and it has to be one of the elements within the output object.

The final component of the framework is the use of the classification returned by the framework. This is in the form of a behavior weighting. Each behavior that a user can undertake with a social tie is given a weighting from -1 to $+1$. This weighting can be used to personalize the interface based on the social tie and current interface present.

4 RESULTS

The recruitment of participants involved the posting of brochures and dissemination of material for the purpose of downloading the application from the Google Play Store. Registration was open to anyone who was able to download and install the application on a compatible device. Upon registration users were asked for their email address, and basic details. These details were used to provide daily emails to remind them to use the app.

The initial recruitment phase resulted in 22 participants downloading and registering for the application. Of these 22 participants, 16 have carried out initial interaction in the environment. This initial interaction ranges from browsing to content creation. They have also created 54 geographical locations for others to review and interact with. This has lead to 185 content producing interactions. For instance, location reviews and annotation of those reviews.

Packages Per User

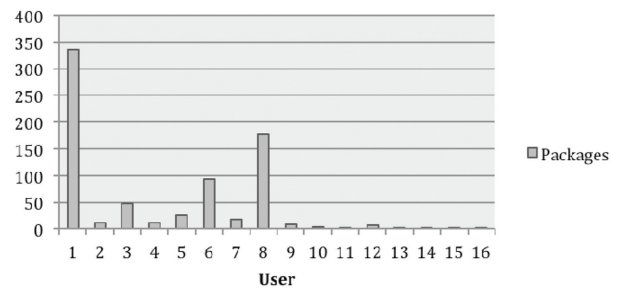


Figure 6. Distribution of behaviour packages created by the framework for each user.

Table 1. Results created by the framework for behavioural stereotype.

User	Publisher	Lurker	Annotator
1	0.2695	0.7133	0.0608
2	0.286	0.6489	0.0335
3	0.239	0.6281	0.0271
4	0.3796	0.6761	0.0659
5	0.1664	0.7048	-0.0158
6	0.3821	0.5998	0.0386
7	0.054	0.7625	-0.0327
8	0.1134	0.7154	0.1057
10	0.2508	0.6109	0.0569
11	0.0771	0.5476	-0.0045
12	0.3017	0.5127	-0.0073

From these standard interactions the framework has created 745 interaction packages with an average of 46.56 packages per user. Although the average number of packages per user is high the standard deviation is 87.12. This is a result of dominant users within the environment producing the majority of the interactions.

Figure 6 shows the distribution of these packages. The users are ordered in the date of the registration. The first user, being the user who has been in the system the longest has the largest number of packages. This is a result of continual use over the period of time. The most recent users, 13, 14, 15, and 16 only have one behavior package consisting of lurking. For this reason they were not passed through the neural network for classifying.

Table 1 presents the results produced by the framework of the dominant behavioral stereotype each user presented when interacting in the environment with no social ties. These results represent the weighting of behavior between -1 and $+1$. The closer the value is towards $+1$ the more dominant the behavior. After analyzing these results the users have shown a dominant behavioral stereotype of *Lurker* consistently throughout the environment. This coincides with many other studies when users enter a new environment where they *lurk* and explore before they participate [12].

Although all users present *Lurker* globally, each user's behavioral stereotype changes when the framework uses a user's social tie to determine their behavioral stereotype. For brevity, analysis of only a few examples is given.

User 1, User 6 and User 8 have the most behavior packages. These users have established social ties amongst each other. When analyzing User 1 to User 6, User 1's behavioral stereotype remains as a *Lurker*, however their *Publisher* value increases considerably; *Publisher*: 0.5431, *Lurker*: 0.7268, *Annotator*: 0.0400. The inverse to this relationship though, User 6 to User 1, results in User 6's dominate behavioral stereotype changing to *Publisher*; *Publisher*: 0.7007, *Lurker*: 0.6613, *Annotator*: -0.092. User 1 to User 8 results in a similar to that of User 1 to User 6: *Publisher*: 0.5014, *Lurker*: 0.6759, *Annotator*: 0.0419.

5 CONCLUSION

This paper has presented a framework for classifying user behavior in a collaborative mobile application in the form of behavioral stereotypes. In addition, initial results have been presented showing early stages of generalized stereotypical behavior and more personalized interactions with social ties changing such behavior. These results will help establish direction for future work in using the framework to provide personalized collaborative environments within smart devices.

REFERENCES

- [1] "8146.0 Household Use of Information Technology, Australia, 2012–2013." Australian Bureau of Statistics, 2014.

- [2] L. Dabbish, C. Stuart, J. Tsay, and J. Herbsleb, "Social coding in GitHub: transparency and collaboration in an open software repository," *Proc. ACM 2012 ...*, 2012.
- [3] L. Findlater and J. McGrenere, "Impact of screen size on performance, awareness, and user satisfaction with adaptive graphical user interfaces," *Proc. SIGCHI Conf. ...*, 2008.
- [4] P. Brusilovsky, A. Kobsa, and W. Nejdl, *The adaptive web: methods and strategies of web personalization*. 2007.
- [5] S. Angeletou, M. Rowe, and H. Alani, "Modelling and analysis of user behaviour in online communities," *Semant. Web-ISWC 2011*, 2011.
- [6] M. Forestier, J. Velcin, and D. Zighed, "Analyzing Social Roles using Enriched Social Network on On-Line Sub-Communities.," *ICDS 2012, Sixth Int. ...*, 2012.
- [7] U. Pfeil, K. Svangstu, C. S. Ang, and P. Zaphiris, "Social Roles in an Online Support Community for Older People," *Int. J. Hum. Comput. Interact.*, vol. 27, no. 4, pp. 323–347, Feb. 2011.
- [8] E. Friasmartinez, G. Magoulas, S. Chen, and R. Macredie, "Modeling human behavior in user-adaptive systems: Recent advances using soft computing techniques," *Expert Syst. Appl.*, vol. 29, no. 2, pp. 320–329, Aug. 2005.
- [9] J. Linden, G. and Smith, B. and York, "Amazon.com recommendations: Item-to-item collaborative filtering," *IEEE Internet Comput.*, vol. 7, no. 1, pp. 76–80, 2003.
- [10] M. Luca, "Reviews, reputation, and revenue: The case of Yelp. com," 2011.
- [11] R. Tague, A. Maeder, and J. Basilakis, "Adaptive Web Framework for Online Collaborative Environments," *Aust. J. Intell. Inf. Process. Syst.*, vol. 13, no. 4, 2014.
- [12] B. Nonnecke and J. Preece, "Online lurkers tell why," in *Proceedings of the 10th Americas Conference on Information Systems*, 2004, no. August, pp. 1–7.

D *Espressobility Recruitment Brochure*



"A new dimension to...

...let's grab a coffee"

Calling all Cafeinates!



see the vibe.



enjoy the hit.



discover.



who will reign supreme.

espressobility.com.au

espresso  *Gility*



Scan to go to the app

E *Squashies Recruitment Brochure*



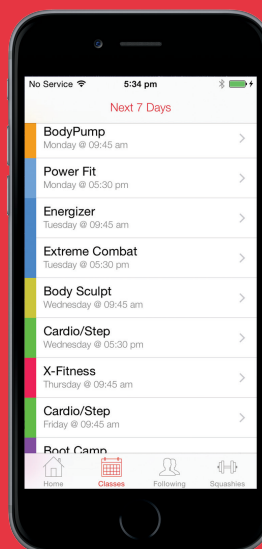
Get the free Squashies App!

You can:

- See upcoming classes
- Book your class spot
- Socialise with other members
- Follow other members
- Keep up to date with gym news
- See opening times on holidays
- ... with lots more to come!



Don't worry, Android is coming soon



squashies.com.au

**F *Ethics Consent with Terms and Conditions
of Participant Use***

(This information will be represented in electronic form presented on either a web page or a mobile window/dialog/box. It will be formatted according to the device upon access)

Participant Information

Human Research Ethics Committee
Office of Research Services, University of Western Sydney

Hi!

You're invited to participate in a study conducted by Rhys Tague, PhD Student of the School of Computing, Engineering, Mathematics at the University of Western Sydney.

Project Title:

A framework for understanding and predicting the take up and use of social networking tools in an online collaborative environment.

Who is carrying out the study?

The principal researcher of this project is Rhys Tague. He is a current PhD student at the University of Western Sydney and supervised by Professor Anthony Maeder, Doctor Jim Basilakis. If you have any questions about this project please use the contact form or send mail to one of the following contacts:

Rhys Tague
Bld. 21.G.37
Campbelltown Campus
Locked Bag 1797
Penrith South DC
NSW 1797 Australia

Principle Research Supervisor
Professor Anthony Maeder
Bld, 26.1.24
Campbelltown Campus
Locked Bag 1797
Penrith South DC
NSW 1797 Australia

Co-Supervisor
Doctor Jim Basilakis
Bld. 26.1.21
Campbelltown Campus
Locked Bag 1797
Penrith Bag South DC

NSW 1797 Australia

What is this study about?

The project involves an online social network where users can collaborate and produce content for reference, sharing, socialising, and general interaction. The online social network will have a framework applied to test if we can predict and understand how you use the network. If we are successful we will achieve a more engaging and entertaining environment that allows for easier collaboration with others.

What does this study involve?

It involves users who are in an online social network who collaborate together. This means be yourself. Explore. Discover new things. It is like any other mobile app or web site you visit.

How much time will the study take?

It's up to you. You can use the site/mobile apps as much as you like or you can use them once and never use them again.

Will the study benefit me?

The study involves entertainment and knowledge creation, so it might! You may use it and find it does something that you like. It's really up to you.

Will the study involve any discomfort for me?

No.

How is this study paid for?

This study is getting funds from the principle researchers involved. If there needs to be more funding we'll seek it from the School of Computing, Engineering and Mathematics at the University of Western Sydney

Will anyone else know the results? How will the results be disseminated?

Once we've finished will publish a generalised form of results. This will be in the form of a blog article or similar.

Can I withdraw from the study?

Yes. At any point where you don't want to be part of the project just stop using it. It's that simple. You have the power. The anonymous information we collect for research up until your last login will still be used for research.

Can I tell other people about the study?

Yes PLEASE! The more users we have the better, also it will make the environment just that little bit more fun. In fact, you can even share our apps and site on other social networks.

What is being collected?

Even though you signed up with some basic details such as your name and email address, we won't be using that in our research. We are collecting anonymous information on how you use the web site, mobile and tablet applications. What

we collect has nothing to do with who you are, so you can be sure your personal information is safe. The anonymous usage information we collect will be owned by the researchers involved, so if you pull out at some point we will still use the information collected up until that point.

What will be done with the collected information?

We'll be analysing the information for the purpose of research. We will use the results of the analysis in publications and also in Rhys' PhD Thesis. Of course, all this information will be anonymous.

How much does this thing cost?

Availability and access to this project is free. There are no costs associated with it.

What if I require further information?

If you would like to find out more about this project please get in contact with the principle research. The principle research is: Rhys Tague and you may use this contact form [click here](#) or use the mailing addresses above.

What if I have a complaint?

If you have any complaints or reservations about the ethical conduct of this research, you may contact the Ethics Committee through the Office of Research Services on Tel +61 2 4736 0229 Fax +61 2 4736 0013 or email humanethics@uws.edu.au

Any issues you raise will be treated in confidence and investigated fully, and you will be informed of the outcome.

To be registered and take part in this project to use the web-site/mobile applications you will give consent during registration.

Terms and Conditions:

To be part of this project, meaning have access and be allowed to use the online services, mobile and tablet applications, you must agree to and consent to the terms and conditions of this project.

The following terminology applies to these Terms and Conditions and any or all Agreements: "User", "You" and "Your" refers to you/I, the person accessing this website and accepting the researchers terms and conditions and giving consent to the researchers research for which you will be a participant. "The researcher", "Ourselves", "We" and "Us", "Our", refers to researchers or associated investigators, . "Party", "Parties", or "Us", refers to both the User and ourselves, or either the User or ourselves. Any use of the above terminology or other words in the singular, plural, capitalisation and/or he/she or they, are taken as interchangeable and therefore as referring to same.

Privacy Statement

We are committed to protecting your privacy. We constantly review our systems and data to ensure the best possible service to our users. We will not sell, share or rent your identifiable information to any third party or use identifiable information for unsolicited mail. Any correspondence by us will only be in connection to the agreed services we provide.

Your Content and Information:

You are the owner of all the content and information you share. The content you share is your responsibility. When you share content it will be publically available. If you do not want content or information to be publically accessible do not share.

We are not held responsible for the views, opinions, comments, statements, and beliefs you provide through sharing, publication, or in any form of media upload or download. If you defame, or cause loss to any party because of your actions with us you understand you will be held responsible for that loss in its entirety.

For content that is covered by intellectual property (IP content), such as videos, images, or audio you here by assign/give us a non-exclusive, transferable, sub-licensable, royalty-free, worldwide license to use any IP content that you share with us to fulfil requests within our web site, mobile and tablet applications. Deletion of this content when requested by you will result in the end of this license.

If content is copyrighted and not solely owned by you, you must seek permission from original copyright holder before use with us. Failure to do so will result in the offending content removed and deleted from our servers and repositories with out notice with the possibility of suspension or deletion of your account.

Data Use Policy:

The type of information we collect from you may be personally identifiable. This information will not be sold, distributed or rented. This information is collected for normal system operations to provide a service to you.

Examples of personally identifiable information may include, but not limited to, First Name, Last Name, Country, Age, Gender, Social networking identifiers.

We will collect information on how you use our web site, mobile and tablet applications. This information includes, but not limited to, the type of social networking/collaborating tools you may or may not use; how you use these tools (click data); the social ties (profiles) you interact with; your Internet browser type; your operating system type; your computational device type. This information will be stripped of personally identifiable information when used for the purpose of this project.

For successful operation of our web site, mobile and tablet applications we may or may not use cookies and/or sessions to store information. This information includes, but not limited to, date and time of access; social tie; profile id; the web address (URL); IP address; browser type; the device type used to interact;

geographical location (mobile device). The collection of this data will allow us to fulfil requests by and for you. This data is system operational data and may or may not be used in analysis of the research project. This information, if needed for project, will be stripped of identifiable information.

Online Behaviour:

The Internet should be a place where you are free to express yourself and respect others. You're free to express yourself, however trolling, offensive, discriminative, or vulgar behaviour or behaviour that disrespects other members of our service will result in possible suspension or deletion of your account without notice.

Other types of behaviour that will result in suspension or deletion of your account without notice are:

- Uploading content or behaviour that results in crashing or halting the service provided. E.g. Denial of service attack.
- Posting/distributing content that is threatening, pornographic, incites violence, or contains nudity.
- Promoting unlawful multi-level marketing, such as pyramid schemes for the purpose of recruiting or other reasons.
- Publishing advertising material considered spam.
- Intentional or non-intentional defaming of a location, person, or other.

Mobile and Tablet Devices:

You understand the usage of our service consumes resources, such as battery life, Internet download and upload. It is your responsibility for any fees you incur with your service provider (telecommunications provider) to access our service. Our service is free, yet your service provider may charge you for use of our service. We are not responsible for incurred usage fees or any fees associated with operation of your device.

Account Termination:

If you no longer want your account to be associated with us you must write an email to us to have your account deleted. When your account is deleted it will be like emptying your trashcan on your computer. The data may be deleted, yet it may still be recoverable for a period of time. Once your account is deleted it will not be recovered for any reason. All information will be lost and considered unrecoverable.

Changes to terms and conditions:

If there are any changes to these terms and conditions or your participation in the research you will be notified by a viable contact method, such as email. If the changes are unacceptable to you stop using the web site, mobile and tablet applications or terminate your account.

Agreement:

By registering or participating in any use of our web site, mobile and tablet applications you agree and give consent to the research project, your participation in the research, the use of anonymous information that may or may not originate from you, and the terms and conditions outlined above.

(This information will be represented in electronic form presented on either a web page or a mobile window/dialog/box. It will be formatted according to the device upon access)

Consent:

Project Title:

A framework for understanding and predicting the take up and use of social networking tools in an online collaborative environment.

The following terminology applies to this ensuing consent statement. “I”, “me”, “my”, refers to you, the person accessing this online content.

I consent to participate in the research project titled: A framework for understanding and predicting the take up and use of social networking tools in an online collaborative environment.

I acknowledge that:

I have read the participant information page and have been given the opportunity to make contact and discuss the project information and my involvement in the project with the researcher/s.

The procedures required for the project and the time involved have been outlined to me, and any questions I have about the project have been answered to my satisfaction.

I consent to the terms and conditions outlined of the services provided by the researchers.

I understand the anonymous information that is gained during my involvement in the study may be published but no information about me will be used in any publication or research output of the study that reveals my identity.

I understand I can stop using the services provided to by the researchers at any time, without affecting my relationship with the researcher/s now or in the future.

I understand upon registration of the services provided by the researchers for this project that I give consent on the date of registration to the participation information, terms and conditions, and this consent.

This study has been approved by the University of Western Sydney Human Research Ethics Committee.

The Approval Number is:

If you have any complaints or reservations about the ethical conduct for this research, you may contact the Ethics Committee through the Office of Research Services on Tel +61 2 4736 0229 Fax +61 2 47360013 or email humanethics@uws.edu.au. Any issues you raise will be treated in confidence and investigated fully, and you will be informed of the outcome.

Sample Registration form:

Each form will have the same fields, but rendered differently dependent on device for access

(All copy that is italic is a link)

All fields are required:

First Name:

Last Name:

Email:

Password:

Birth Year:

Country:

☐ I agree and give consent to the *participant consent*, *participant information* and *terms and conditions*.

Register