

# The Presentation of Self on a Decentralised Web

Amy Guy  
<https://rhiaro.co.uk/>



Doctor of Philosophy  
Centre for Intelligent Systems and their Applications  
School of Informatics  
University of Edinburgh  
2017

## Abstract

Self presentation is evolving; with digital technologies, with the Web and personal publishing, and then with mainstream adoption of online social media. Where are we going next? One possibility is towards a world where we log and own vast amounts of data about ourselves. We choose to share - or not - the data as part of our identity, and in interactions with others; it contributes to our day-to-day personhood or sense of self. I imagine a world where the individual is empowered by their digital traces (not imprisoned), but this is a complex world.

This thesis examines the many factors at play when we present ourselves through Web technologies. I optimistically look to a future where control over our digital identities are not in the hands of centralised actors, but our own, and both survey and contribute to the ongoing technical work which strives to make this a reality. Decentralisation changes things in unexpected ways. In the context of the bigger picture of our online selves, building on what we already know about self-presentation from decades of Social Science research, I examine what might change as we move towards decentralisation; how people could be affected, and what the possibilities are for a positive change. Finally I explore one possible way of self-presentation on a decentralised social Web through lightweight controls which allow an audience to set their expectations in order for the subject to meet them appropriately.

I seek to acknowledge the multifaceted, complicated, messy, socially-shaped nature of the self in a way that makes sense to software developers. Technology may always fall short when dealing with humanness, but the framework outlined in this thesis can provide a foundation for more easily considering all of the factors surrounding individual self-presentation in order to build future systems which empower participants.

## Lay summary

Many people express themselves online through social media, blogs, personal websites, and the like. Using these technologies affects our day-to-day lives, and sense of self. These technologies also change and develop in response to how people use them. Many of the tools we use come with constraints, and people often find ways to work around these constraints to suit their needs.

This thesis explores the different ways in which people express their identities using contemporary Web technologies. We conduct several studies, and show that there are many interdependent factors at play when it comes to online self-presentation, and that it is rare that all of these are considered when studying or designing social systems. We present a conceptual framework which will enable cohesive further research in this area, as well as guidance for future system designs.

In the second part, we discuss how these technologies are changing. We make contributions to an emerging alternative means of engaging with social media and similar technologies, and examine the implications of these new technologies on self-presentation.

## Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.



Amy Guy



## Acknowledgements

The process of assembling this thesis has been both social and decentralised, as well as distributed. I owe much to many, and can hardly begin to express my gratitude.

### *To PhD support groups..*

Jane and Felicity, cakes and unicorns. Paolo, Sergio and Michael, and all who passed through IF2.35. KitB and Tigo, following me around Edinburgh. TomSka, Bown; London friends and YouTubers who inspire and distract. Liz Powers; a year of sunshine and hammocktime and dramatic readings of my abstracts.

### *Invaluable collaborations..*

The BBC's Linked Data Platform team, for that 'real world' experience (and falafel). The SOCIAM research group, who took me in. Max van Kleek and Dave Murray-Rust, the fastest coauthors in the West. The IndieWebCamp community, for challenging me, teaching me, and most importantly... making me just ship it already.

The W3C Social Web Working Group; started out as procrastination, and the rest was history. The Decentralized Information Group at MIT, with whom I spent a year. Timbl is as nice as everyone says he is. Ilaria Liccardi's mentorship and unstoppable attitude is unbeatable.

Working for the W3C; a distant pipe-dream, suddenly reality. Thanks Sandro Hawke for having faith in me and being on the same wavelength when things are exasperating.

Lunatic scholarship; Sarven Capadisli takes you from stylesheet switching to revolutionising academic publishing in under a minute. #LinkedResearch, distractions of epic proportions; life will never be the same again. ☺

*Wandering enablers...*

I passed through 17 countries during the 4 years of my PhD. Credit to CISA's generous student travel budget, the Semantic Web conference circuit, and W3C meetings. And those who gave me a place to crash in their homes or hotel rooms across the US and Europe:

Ann Bassetti, what a role model. Chris Webber; Tantek Celik; Aaron Parecki (whos IndieWebCat Dora will always have a better website than me); Harry Halpin; Henry Story; and the roaming elf Pavlik.

In the last stretch of writing up, I escaped to the Southern Hemisphere:

I slept in my little brother Dave's closet for a month in Tokyo. Ann, my Malaysian Mum, took care of me for three months in Penang. I met Ninni at Bali airport, and we knocked out our theses together in Canggu cafes.

Tammy and Jon, the Wholey Wonders, who provided a continuous stream of vegan food and a wifi connection and let me hole up in their tiny apartment when my thesis was two days overdue. Things were getting desperate. I don't know what I would have done without you.

*For humouring me...*

Ewan Klein, I couldn't have imagined a better supervisor. Remember that one time I had a crisis and instead of sitting me in your office and giving me deadlines you took me to Peter's Yard and bought me coffee and pecan pie and we chatted about non-PhD things? Probably not. But I'll never forget. You gave me space when I needed it, and instruction when I needed it, and distraction when I needed it. It worked perfectly.

And my parents of course, who never once gave me the impression I was heading in the wrong direction.

# Glossary

**backstage**

Goffman's terminology for the role one can assume when one is no longer performing for an audience.

**decentralised**

A system in which multiple authorities control different components and no single authority is fully trusted by all others.

**federation**

The joining together of software instances such that activities on one are seen on another (usually by means of a protocol).

**frontstage**

Goffman's terminology for a persona, which is performed for an audience.

**interoperable**

The quality of being able to exchange data or trigger processes without any prior arrangement.

**monoculture**

A piece of software which can only interoperate with other instances of itself.

**online presence**

Traces of a person or persona which can be found around the Web, perhaps in the form of profiles.

**persona**

A role that one assumes or displays in public.

**profile**

A digital representation of a person or persona, made up of a subset of their attributes, activities, interactions, and generated data.

**protocol**

A set of possible communication actions between computer systems.

**self-presentation**

The act of performing a persona.

**silo**

A system which stores and/or generates data, but does not let any in or out.

**social system**

Web-based networked publics which offer individuals consistent and reusable access to an account which they can customise and use to interact in some form with others in the system.

**standards**

Technical specifications which formally define and describe software systems.

# Table of Contents

## 1 Introduction

- 1.1 Background and context
- 1.2 Research questions and contributions
  - 1.2.1 Conceptual framework
  - 1.2.2 Changing dynamics
  - 1.2.3 Impact on practice
- 1.3 Structure
- 1.4 Methods

## 2 The Presentation of Self Online

- 2.1 Introduction
  - 2.1.1 My perspective on this review
- 2.2 Performing the self
  - 2.2.1 The self in context
  - 2.2.2 The project of the self
  - 2.2.3 Extending the self
- 2.3 Offline to online... and back again
  - 2.3.1 Authenticity and integrity
  - 2.3.2 Dishonesty and deception
- 2.4 Networked publics
  - 2.4.1 Profiles and embodiment
  - 2.4.2 Social Network Sites
  - 2.4.3 Blogging and personal homepages
  - 2.4.4 Imagined audience
  - 2.4.5 Context collapse
- 2.5 Everybody knows I'm a dog
- 2.6 The ghost in the (social) machine
- 2.7 Conclusions

## 3 Social Web in the Wild

- 3.1 Introduction
  - 3.1.1 Overview of studies
- 3.2 What is a profile?
  - 3.2.1 Introduction
  - 3.2.2 Context and research questions
  - 3.2.3 Study Design
    - 3.2.3.1 Method
    - 3.2.3.2 Subjects
  - 3.2.4 Results
    - 3.2.4.1 What does a profile contain?
    - 3.2.4.2 How are profiles updated?
    - 3.2.4.3 How are people notified when a profile is updated?
    - 3.2.4.4 How is access to a profile controlled?
    - 3.2.4.5 How are profiles within a system connected together?
    - 3.2.4.6 How can profiles be exported from or imported into a system?
    - 3.2.4.7 What constraints are placed on a profile?
    - 3.2.4.8 How do profiles fit in with a systems apparent data model?
    - 3.2.4.9 What is the profile for?
    - 3.2.4.10 Who is the profile for?
  - 3.2.5 Features
  - 3.2.6 Discussion
- 3.3 Constructing online identity
  - 3.3.1 Introduction
  - 3.3.2 Context and research questions
  - 3.3.3 Study design

- 3.3.3.1 Method
- 3.3.3.2 Subjects
- 3.3.4 Limitations
- 3.3.5 Results
  - 3.3.5.1 Profiles and personas
  - 3.3.5.2 Connections
  - 3.3.5.3 Summary
- 3.3.6 Framework for understanding identities of social machine participants
  - 3.3.6.1 Applying the framework
- 3.3.7 Conclusion
- 3.4 The many dimensions of lying online
  - 3.4.1 Introduction
  - 3.4.2 Context and research questions
  - 3.4.3 Study Design
    - 3.4.3.1 Method
    - 3.4.3.2 Participant recruitment
    - 3.4.3.3 Limitations
  - 3.4.4 Results
    - 3.4.4.1 Self-reported frequency of deception/lying
    - 3.4.4.2 Reasons for Deception
    - 3.4.4.3 Pseudonyms
    - 3.4.4.4 Personas
  - 3.4.5 Discussion
  - 3.4.6 Conclusion
- 3.5 Computationally-mediated pro-social deception
  - 3.5.1 Introduction
  - 3.5.2 Context and research questions
  - 3.5.3 Study Design
    - 3.5.3.1 Method
    - 3.5.3.2 Participants
    - 3.5.3.3 Designing the vignettes
  - 3.5.4 Results
    - 3.5.4.1 Participants
    - 3.5.4.2 Case study: Privacy and people (P8)
    - 3.5.4.3 Case study: Honesty and self-image (P9)
    - 3.5.4.4 Case study: Privacy and technology (P13)
    - 3.5.4.5 Effort and Complexity
    - 3.5.4.6 Availability of Other Channels, Strategies
    - 3.5.4.7 Privacy and Control
    - 3.5.4.8 Authenticity and Crafting Personas
    - 3.5.4.9 Polite Social Signalling, Kindness, and Empowerment
    - 3.5.4.10 Access control and imagined audience
    - 3.5.4.11 Ethics and morality
  - 3.5.5 Discussion
    - 3.5.5.1 Morality of Deception
    - 3.5.5.2 Promoting Social Honesty
    - 3.5.5.3 Memory, safety, and plausible deniability
  - 3.5.6 Conclusion
- 3.6 Social Media Makers
  - 3.6.1 Introduction
  - 3.6.2 Context and research questions
  - 3.6.3 Study Design
    - 3.6.3.1 Participant Recruitment
    - 3.6.3.2 Method
    - 3.6.3.3 Limitations
    - 3.6.3.4 Analysis
  - 3.6.4 Results
    - 3.6.4.1 The network

- 3.6.4.2 (De)compartmentalisation and audience
- 3.6.4.3 Self-expression
- 3.6.4.4 Empowerment
- 3.6.4.5 Longevity
- 3.6.4.6 Ephemerality
- 3.6.4.7 Consent
- 3.6.4.8 Abuse and surveillance
- 3.6.4.9 Inspiration and triggers
- 3.6.5 Discussion
- 3.6.6 Conclusion
- 3.7 The 5 Cs of Digital Personhood
  - 3.7.1 Context
  - 3.7.2 Control
  - 3.7.3 Customisability
  - 3.7.4 Connectivity
  - 3.7.5 Cascade
  - 3.7.6 Conclusions and reflections

#### **4 Decentralising the Social Web (and other stories)**

- 4.1 Introduction
  - 4.1.1 Scope
  - 4.1.2 Decentralisation
  - 4.1.3 Social systems
  - 4.1.4 Selection method
- 4.2 When is a person not a person?
- 4.3 Standards and monocultures
- 4.4 Social (and) personal datastores
- 4.5 From status updates to microblogging, and beyond
- 4.6 If the app fits
- 4.7 The Social Web Working Group
  - 4.7.1 Implementations
- 4.8 Discussion

#### **5 Standards for the Social Web**

- 5.1 Introduction
- 5.2 Standards and self-presentation
  - 5.2.1 Standardisation as Context
  - 5.2.2 The standardisation process
  - 5.2.3 The Social Web Working Group charter
  - 5.2.4 Working Group participants and audience
  - 5.2.5 API Requirements
    - 5.2.5.1 Process
    - 5.2.5.2 Results
  - 5.2.6 Competing specifications
  - 5.2.7 Social Web Protocols
    - 5.2.7.1 Overview
    - 5.2.7.2 Reading
    - 5.2.7.3 Publishing
    - 5.2.7.4 Subscribing
    - 5.2.7.5 Delivering
  - 5.2.8 Protocols for Customisability and Connectivity
  - 5.2.9 Remaining problems
- 5.3 Personal data and self-presentation
  - 5.3.1 Introduction
  - 5.3.2 Background
    - 5.3.2.1 Personal Web Observatories
    - 5.3.2.2 Self-tracking and Quantified Self
    - 5.3.2.3 Social Publishing
  - 5.3.3 Building a Personal Web Observatory
    - 5.3.3.1 Architecture

- 5.3.3.2 Data
  - 5.3.4 Using a Personal Web Observatory
    - 5.3.4.1 Recording data
    - 5.3.4.2 Practical uses for small data
    - 5.3.4.3 If my website is down, I do not exist
  - 5.3.5 Discussion
    - 5.3.5.1 Personal and social impact
    - 5.3.5.2 Technical considerations
    - 5.3.5.3 Data context and integrity
    - 5.3.5.4 Limitations
  - 5.3.6 Personal datastores for Control, Customisability and Context
  - 5.3.7 Next steps
  - 5.4 Audience and self-presentation
    - 5.4.1 Introduction
    - 5.4.2 System design
      - 5.4.2.1 Vocabulary and data integration
      - 5.4.2.2 User interface
      - 5.4.2.3 Limitations
    - 5.4.3 Discussion
      - 5.4.3.1 Connectivity and cascade
      - 5.4.3.2 Control and customisability
      - 5.4.3.3 Context
  - 5.5 Conclusions
- 6 Conclusions**
- 6.1 Summary
  - 6.2 Review of research questions
    - 6.2.1 Methodological Recommendations
  - 6.3 Things to come
    - 6.3.1 Decentralisation considered harmful
    - 6.3.2 Future work

## Table of Tables

- 1 Overview of studies
- 2 Profiles study subjects: social systems analysed to generate the taxonomy
- 3 Profiles study features
- 4 Profile questions and features relation
- 5 Profiles study results: features of systems
- 6 Profiles study results: features of systems
- 7 Content creator subjects: Distribution of profiles for content creators in the study.
- 8 Summary of framework
- 9 Applying the framework: Examples of social machines along each dimension.
- 10 Focus questions: List of questions corresponding to subset of survey discussed in this analysis.
- 11 Q5 Tags: List of themes and categories resulting from analysis of Questions 5a and 5b.
- 12 Document Metadata
- 13 The top eight Social Web Working Group user stories
- 14 The API requirements and which specs they are met by
- 15 How each spec relates to the others
- 16 Post types
- 17 Document Metadata

## Table of Figures

- 1 Control, Customisability, Connectivity, Context, Cascade; and their relations to each other.
- 2 Interconnected social and technical systems necessary for publishing a vlog on YouTube.
- 3 Dane transparently links to 3 of his channels, two for alternative content types and one for a character he created. 2. Fred is a character played by Lucas, but the links on Fred's channel treat Lucas as a different person. 3. Andrew Lemming lists Khyan as "Uncle", although Khyan is the creator of the Andrew Lemming character.
- 4 Lucas Cruickshank was an early YouTube success story through his persona Fred Figglehorn. Here, different types of connections between various online accounts belonging to both Lucas and Fred are illustrated.
- 5 Self-reported use of social media, from 1=Never to 5=Often times a day. Medians: Facebook = 3, Twitter = 5, YouTube = 5, Tumblr = 3, Instagram = 2, Vine = 1
- 6 Responses to Q4b and Q4c on Likert scales b) How often do you tell lies or untruths on social media? (1=Never to 5=Often) c) How often do you think your friends lie on social media compared to you? (1=Vastly less to 5=Vastly more)
- 7 Tags and counts for responses to Question 4: Have you ever told untruths on social media, given fictitious info, omitted or distorted the truth online?
- 8 Tags and counts (N=286) for responses to Question Do you use pseudonyms on any social media platforms? Responses which were in the negative or unclear have been removed.
- 9 Tags and counts (N=267) for responses to Question 5b: Have you created any fictional personas (e.g. characters, alter-egos) to use on social media?. Responses which were in the negative or unclear have been removed.
- 10 Social Steganography: Diagram illustrating a social steganography tool for microblogging/SNS sites that hides 'real messages behind other, plausible status messages but allows certain people to recover the true meaning.
- 11 lieCal: Fictional interface for a tool which automatically generates excuses on behalf of the user, optionally including friends in the deception and strengthening alibis by posting on social media.
- 12 lieCation: Fictional tool to create a narrative of going somewhere (on holiday) or attending an event, along with images and social media posts to be sent out at preset times to corroborate the story.
- 13 lieMapper: Fictional tool for predicting the flow of information (e.g. a lie) across a person's social network starting from a single friend.
- 14 lieMoves: A fictional smartphone service for letting people obfuscate their location using various strategies, including blurring, substitution, past-replay and impersonation. (Based on Figure [liecloud-structures]e,f.)
- 15 Examples of websites by participants who used their own space to express themselves in ways they are unable to achieve with centralised services.
- 16 A view on how framework terms relate to one another, hierarchically.
- 17 How the Social Web Protocols specifications connect different (high level) parties together (all possibilities, click to show different configurations).
- 18 Notifying a hub of new content with LDN, using an AS2 Announce in the notification body. The object is the topic URL and the target is the hub itself. The hub can use this information to fetch new content for subsequent delivery to subscribers.
- 19 The high level architecture for my Personal Web Observatory. Parts in black are the core, and parts in grey are external or third-party services which interact or interoperate.
- 20 A screenshot of my homepage at the time of writing. The purple background shows I was at home. Note: I do not profess to be a designer.

# Chapter 1

## Introduction

## 1. Background and context

A goal of this thesis is to bridge established social science research about self-presentation and impression management, with a practical, developer-centric perspective on building Social Web applications. I seek to understand how and why contemporary social networking systems both help and hinder individuals' expressive needs, and what people do to work around technical and social constraints they encounter during day to day use. What we learn from this, we can use to design and optimise future social systems for individual and collective empowerment.

The society from which I write is at present enthralled with online social media. It is rare to encounter someone who does not have a profile on one or more of the major Social Network Sites, most of which are household names. These systems have shifted from the realm of techie early adopters to indispensable tools of daily life at a rapid pace, and continue to evolve. Yet when I speak to people about their social media use, I inevitably encounter grumbles or complaints. Some people are frustrated by unintuitive user interfaces; others feel trapped or pressured into using particular systems because everyone else seems to be there, or it's become the only way to get anything done. Yet others are hooked, distressed to find themselves whiling away hours by watching other peoples' lives go by, but unable to tear themselves away. Others are driven away from interactions they want or need by harassment and abuse.

I am far from a techno-dystopianist. I believe strongly in the Web as a force for good, as a means to communicate ideas and share experiences across the world. I've used the Web for almost my entire life; it has been an outlet for creativity, a forum to learn about myself through the experiences of others, a means to maintain relationships at a distance, a provider of remote serendipity and opportunities, and an invaluable asset to lean on when travelling the world with some dependence on the kindness of strangers.

I worry about the digital shadow of myself which corporations and governments have access to. I am sure it is thorough and accurate, and that they could use it for all sorts of mischief. I worry about being manipulated without realising, about being tracked, about being backed into a corner with nowhere to hide. I worry more about the countless people this is happening to who do not have my considerable privilege which stems from my country of birth, my stable

upbringing, my education, and the colour of my skin. The society from which I write is also undergoing some political upheaval. Ordinary people seem to be turning on each other, blaming people who are different from them for their problems; there is division and anger in the air. Or maybe it just appears this way through the particular lens of the Web through which I am witnessing current events unfold. Maybe it is somewhere in between. Maybe we have always been at war with Eastasia. It's not as bad as some parts of the world, nor many historical events, but it's not good. And social media, ever present, is playing a role.

Beneath the gauzy promises of democratized access to sociality, meaning, fame, and reputation, the business practices of the social platforms in and through which we self-present draw us all into privatized corporate strategies of social sorting, identity management, and control [147].

This thesis is not about politics, but it is about people. It is also about society and community, and how we interact with others near and far through the possibilities availed to us by the Web, and how we come to know and show ourselves in the process. For this reason I first look to the social sciences, and Erving Goffman gives me a place to start, with *presentation of self*. To catch these ideas up to the present day, I lean heavily on the work of danah boyd, whose contributions are foundational to subsequent work on online, as opposed to face-to-face, interactions. Seeing the insights to be found through ethnography - through talking to people and observing them - I proceed to learn a great deal from my various study participants.

I am a Web developer by trade, and so I inevitably fall back to trying to design and build software to help with social problems, and I am often surrounded by others with this outlook. Now, I can do so whilst mindful of the continuously turning wheels and shifting sands of society, aware on some level of every unique individual who might pass through a system. That is not to say I think it's possible to accommodate everyone, but I think with enough energy and consideration it is possible to build systems which do not do as much harm as the ones we have today.

There are many ways to approach improving society through the Web, and the one I have chosen is based on the idea of putting people's digital representations into their own hands. There are many ways of going about *this* as well, and now I take you from this high level painting of the state of the world to a suddenly specific and technical detail. My efforts are towards creating

standard protocols for Social Web activities, which allow other developers to build systems which can interact in an open and defined way. The Web Science Framework [25] expounds the cruciality of Web standards for the progression of the Web, as a process of social negotiation which yields tangible engineering outcomes. The Web itself has always been decentralised<sup>dns</sup>, and now I take inspiration from its inventor Tim Berners-Lee, who actively advocates for *re-decentralisation* of the social layer that has been built on top. Developers must build systems which respect the people who use them, because in decentralised systems people have the freedom to take their business elsewhere. I believe this changes the power dynamic in favour of the previously disempowered non-technical individual.

The process of creating Web standards, which I engaged with through the World Wide Web Consortium (W3C), involves nitty gritty technical work, understanding obscure specifications, practices, and web lore, and endless pedantic arguments. It is infuriating, but perhaps just as important to work through as it is to carry out the ethnographies I mentioned previously. We, along with other initiatives and open source projects, are inching towards progress. We are preparing the technical foundations on which social improvement can be built, when the time comes. Or at least, I hope so.

Throughout this work I advocate for particular visions of systems and technologies that I believe to be our best chance, however it is important to note that such systems are not an end in themselves. The social changes that technologies enable are what matters. Thus I want to emphasise that while Web technologies are constantly evolving, and subject to rapidly shifting political, economic and technological landscapes, that peoples' needs and desires, rights and responsibilities, are also reconfigured by changing technology. I hope to capture this feedback loop, and the enormity of an interconnected online and offline world, through this thesis.

Fortunately I am not alone in this. My work is situated in the field of Web Science, a multidisciplinary domain with a focus on the relationships between people and the technology that connects them. I lean somewhat on the narrative from the sub-topic of Social Machines, a concept which acknowledges the intrinsic co-dependence between humans and technology.

## 1.1. Why decentralisation?

I did not set out to explore the decentralisation of the Web specifically. When I started work on this thesis, I had a blog, and a healthy skepticism of social media, but no experience with federated or self-hosted social networks, Web standards, or decentralisation protocols. I stumbled across this world when I was searching for ways to empower content creators - people whose livelihoods were tied to (centralised) media platforms like YouTube and DeviantArt.

This thesis does not evaluate decentralised Web technologies and find them to be the most promising solution to the problems I described in this introduction. In fact throughout, I touch on new challenges that are a direct result of decentralising ordinarily centralised technologies or systems. In some cases I even go so far as to present possible solutions. In my concluding chapter, I summarise the most serious of these challenges. Rather, I decided that the decentralised Web vision of the future of online interactions was viable enough to be worthy of study; the timely formation of the W3C Social Web Working Group, and the specific focus of Tim Berners-Lee's MIT research group were contributing factors to this decision.

Accordingly, I do not mean for my focus on decentralised Web technologies to be read as a bias towards this as a solution. The bias which does exist, which of course influenced my choice of potential future to explore, is towards individuals and against for-profit companies; towards information access and transparency, and against manipulation and surveillance. Instead, the latter half of this work is a preemptive examination of a phenomenon which we may see come to pass in the near future, in relation to my core focus around presentation of self.

dns Don't mention DNS or certificate authorities.

## 2. Research questions and contributions

In this thesis, I ask questions about online self-presentation in the past, present, and future. I start by seeking to ease our understanding of what has gone before, or how we can make sense of historical research in this space. The landscape is changing under our feet, so I take a look at what is happening right now that will impact future research and evolving comprehension. Then I ask what we can take from this to prepare for and shape the road ahead.

### 2.1. Conceptual framework

#### **R1: How can we access the bigger picture when it comes to understanding the impact of networked publics on presentation of self?**

There is a multitude of work from the past two decades about self-presentation on the Web, from personal homepages and blogs to modern Social Network Sites. Ethnographic studies from social sciences and psychology investigate the impact of networked publics on people's everyday lives. Social media analysis studies from computer sciences look at network effects and find patterns in how people connect and the data they publish. Whilst there is usually some overlap, the former efforts tend to focus on people and the latter on technology. It would perhaps be unreasonable to expect anything else, for example, for social scientists to convey a profound understanding of the underlying systems their subjects engage with at every stage in their study. However, we are talking about socio-technical systems, and very complex ones at that. Networked publics cannot be properly understood without positioning an individual study against the contemporary background of what is occurring *both* socially and technically. Due to the sheer enormity of this task, much research fails to do this.

In order to make this easier, we need a device which allows scholars to access and organise concepts relating to broader socio-technical systems, so that they can situate their more in-depth niche or specialised analyses on particular topics. This can be achieved by means of a conceptual framework which captures a hierarchy of concepts that are applicable to understanding self-presentation in networked publics.

I begin by examining key concepts from relevant literature from both social sciences and computer sciences, and summarising a cross-section of findings from smaller scale studies about online social spaces in chapter 2. This sets the background for a series of studies of my own, which are described in chapter 3, and which are designed to tease out further diverse and novel considerations about social behaviour on the Web.

There are multiple interacting dimensions which ought to be considered when observing or designing online social systems, and until now it has been hard to find a coherent way of organising these. Chapter 3 proceeds to attempt to answer this research question with one possibility for a conceptual framework design, based on findings from existing literature and my own study results.

### **Contributions:**

- **C1a:** a novel conceptual framework - the 5Cs of Digital Personhood - that offers a consistent and comprehensive set of concepts and terminology for understanding the affordances and limitations of self-presentation in Web-based social systems from a user-centric perspective.
- **C1b:** an up to date survey of existing work, which relates studies of the Social Web back to pre-Web work on sociality;
- **C1c:** a survey of features offered by a specific set of contemporary technical systems when it comes to profile construction.

## 2.2. Changing dynamics

### **R2: How does self-presentation change depending on the power dynamics of the Social Web services they use?**

Next, I acknowledge the changing times of the Social Web. I believe (and hope) we are on the verge of an important transition from a world in which our personal data is stored and harnessed by powerful third-parties at great (but often unseen) cost to individuals, to the proliferation of technologies which enable people to be discerning about their choices of communication system. One route for this transition is by way of *decentralisation*, that is, by dispersing power from the few to the many. In terms of software for the Social Web, decentralisation entails making it possible for diverse systems to communicate without prior arrangement, to form spontaneous connections and pass data around seamlessly. Achieving this comes with both technical and social challenges, and to succeed would impact online social behaviour in ways we may not yet be prepared for.

The studies in chapter 3 contribute towards answering this question. On top of that, there is already much technical work towards decentralising the Social Web; in chapter 4 I describe a specific subset of this, and in chapter 5 I provide my own contributions to the field. Being directly involved in the technical work provides critical insight into the finer details of this question, from the 'other side' as it were, or the perspective of system creators rather than users who are at the heart of chapters 2 and 3.

### **Contributions:**

- **C2a:** a set of critical dimensions to consider when studying identity performance through creative media sharing;
- **C2b:** a description of people's habits and reactions to different kinds of deception on social media;
- **C2c:** an analysis of attitudes towards self-presentation by people who control the technology behind their social media presence;
- **C2d:** a critical look at the technical directions taken by the W3C Social Web Working group, and the social dynamics of group participants which underly them.

## 2.3. Impact on practice

### **R3: What can developers do to adapt to or accommodate self-presentation needs of individuals?**

Finally, I want to offer something of use to the designers and developers of future social systems. I seed some answers to this question in chapter 5 in an exploratory manner which is neither exhaustive nor conclusive. Nonetheless, my hope is that this prompts future work, theoretical and practical, around novel ways for the Social Web to empower its participants.

### **Contributions:**

- **C3a:** a primer for the technologies produced by the Social Web Working Group, and technical guidance on how to fit the different specifications together;
- **C3b:** a prototype implementation of a personal social datastore, and a report on the personal impact of long-term use;
- **C3c:** a speculative design for a novel system for indirect communication between a profile owner and their audience, used for learning and meeting audience expectations in the moment a profile is viewed.

### 3. Structure

This thesis is roughly in two parts, and each part comprises a background chapter and a chapter containing my novel contribution. Part one, chapters 2 (background) and 3 (new work), is about the past and present of online self-presentation. Part two, chapters 4 (background) and 5 (new work) are about the present and future of online self-presentation should the Social Web become decentralised. The first half has the feel of work from the social sciences domain, and the latter half is highly technical. These two parts in combination are necessary for providing both a novel and holistic take on online self-presentation.

We begin in chapter 2 with a literature survey of fundamental ideas from social sciences about the presentation of self, as well as more recent digital sociology work from social and computer sciences, psychology, and media theory, about how people use such systems to express themselves and connect with others. Grounded in this, we discuss the current state of the art of Social Networking Sites (SNSs), and how they meet (or fail to meet) peoples' social needs.

In chapter 3 I present five self-contained empirical studies which allow us to analyse different aspects of online identity behaviours 'in the wild', and further discuss the concept of an online profile as a tool for self expression. I use the results of these studies to ground the description and justification of the conceptual framework.

Chapter 4 takes a brief look at the history of implementation and standardisation efforts for decentralised online social interactions. I use this review as a lead-in for a deeper look at the socio-technical process of formal standards development, in the context of a W3C Working Group in which I participated, in chapter 5. I also provide a prototype implementation of the standards produced by the group, as well as a design for a novel interaction pattern that can be used alongside.

In the concluding chapter I draw together these findings, and suggest directions for future research.

## 4. Methods

The contributions of this thesis have been generated by a combination of surveying existing theoretical and practical work; conducting empirical studies; and practice through writing and implementing Web standards.

Empirical studies take the form of descriptive [249], whereby a detailed account of a particular situation is given; and ethnographic, whereby individual people are observed, surveyed, and interviewed [138]. These methods are a way of eliciting detailed insight into phenomena which create new ways of thinking about things or awareness of previously unknown possibilities, but do not necessarily provide a means to exhaustively map a problem space. For each individual study in chapter 3 I describe in more detail its particular method and limitations.

A conceptual framework is a useful cognitive tool, which "explains, either graphically or in narrative form, the main things to be studied - the key factors, concepts, or variables - and the presumed relationships among them." [201]. I developed mine through eliciting, aggregating and clustering many concepts from across my own study results and what I have learned from others.

Practice-led research is a way of immersing myself into the topic I'm studying, as well as directing outcomes and effecting change. Participation contributes a more complete understanding of an area than observation alone possibly can [238]. As such, I have been able to collaborate in designing standards, as well as report in detail on the process of standards-making. I have been able to build and use prototype systems, immersing myself in the perspectives of developer and user, and coming to a better understanding of the implications of both.

## Chapter 2

# The Presentation of Self Online

## 1. Introduction

Web users are engaging in computer-mediated self-expression in varying ways. The technology that enables this is developing fast, and how it does so is influenced by many more factors than just the needs of the people it touches. This chapter explores literature about the impact of computer-mediated self-expression on:

- people's everyday lives.
- individual self-expression and exploration.
- interactions and relations with others.
- interactions with and expectations of society and community.

We ground our discussion in established literature about non-digital self-expression and identity from the social sciences. This raises the key theme of individuals desiring control over how others see us, yet wanting to behave in a way that is authentic, or consistent with their internal identities. There is also emphasis on the collaborative and collective nature of identity formation; that is, our self-presentation fluctuates depending on the people we're with, the situation we're in, and norms of the society we're part of. The focus on face-to-face interactions and embodiment leads us to draw contrasts between online and offline experiences, and to look at the substitutes for the body in digital spaces.

The extent to which online and offline identities interact and overlap is hotly debated. Is creating an online identity a chance to reset, to reshape yourself as an ideal? Or are you simply using it to convey true information about what is happening in your daily offline life? Is it a shallow, picture of you, or a forum for deep self-exploration? How does the way one portrays oneself in digital spaces feedback to ones offline self-presentation? We explore these questions in section 3.

Section 4 examines social media and blog use, including how one's *imagined audience* affects self-presentation in public, and how *context collapse* might occur when the actual audience is different to expected. There are several examples of techniques for managing who sees which 'version' of oneself, and the types of 'versions' of self that are commonly seen to be constructed on social media, and with what degree of transparency they are linked together. Most of the longitudinal studies in this space are of teenagers and young people, who

have never known a world without social media, and who may incorporate it naturally and seamlessly into their daily practices, thus making it a core part of their identity during formative years. I draw a contrast between the relationship-driven architecture of contemporary social networking sites, and the more personal, customisable blogging platforms which preceded them. Studies of bloggers and blogging communities reveal some different priorities and habits than what is common practice today, and offer insight into how online self-presentation is evolving.

Throughout literature from both social and computer sciences, privacy is a common concern. In section 5 we look further at how tensions between users and the privacy settings of systems they use impact on personal information disclosure. Does self-censorship affect identity formation? How do people weigh up the risks and benefits of exposing themselves online? This is particularly pertinent for future systems development, as more and more people become aware of state surveillance, for-profit data collection, and their diminished rights over their personal data.

Finally we introduce the relatively new Web Science concept of Social Machines in section 6 in order to recapture the circular interdependencies between humans, technologies, and communities. We propose to build on current work of describing and classifying social machines to better account for the individual perspectives of participants.

Ultimately we posit that online is simultaneously a reflection, a distortion, an enhancement, and a diminishment of the offline world. They impact each other in complex ways, particularly with regards to self-presentation and identity formation. The various theories and studies described in this chapter form the basis for which we conduct the investigative and technical work in the remainder of this thesis.

### 1.1. My perspective on this review

I'd like to take a moment to note that whilst reading various studies about young peoples' reactions to and interactions with rapidly evolving digital technologies from the 2000s, it occurred to me that the subject of these studies is in fact my own age group. Some of the results are instinctively familiar to me; I was there, I experienced these things. Some are ridiculous. I don't know how my first-hand

experience of growing up with technology (I was born in the same year as the Web, and my parents were early adopters) affects my reading of these studies, or my ability to study others' use of technology, but it is something I ponder.

## 2. Performing the self

The obvious place to start when embarking on a discussion about self-presentation is Goffman [123]. In *The Presentation of Self in Everyday Life*, Goffman posits several, now well-established, theories using drama as a metaphor:

- Everyone is performing. The front-stage of our performance is what we create for others - the audience - to see, so that they may evaluate and interact appropriately with us.
- We also have a back-stage; how we act when there is no audience, or an audience of our *team*. Our team participate alongside and collude with us on the front-stage.
- Our performances have both conscious and unconscious aspects. That is, we consciously *give* information about ourselves to others in order to manage their impression of us, but we also unconsciously *give off* information that others may pick up on and take into account when deciding how to interact with us.
- Both actors and audiences are complicit in maintaining the cohesion of a situation. Performances break down if actors break character, deliberately or accidentally, or if there is a mismatch between parties' definition of the situation.

These theories emphasise the collaborative or social nature of self-presentation, and apply to face-to-face interaction.

Whilst Goffman's dramaturgy refers mostly to body language, a related theory is Brunswik's lens model [45, 120], part of which suggests that individuals infer things about others based on "generated artifacts", or things left behind. In [124] this model is used to study how personal spaces (offices and bedrooms) affect observers' assessments of the characteristics of the owner of the space. This study links individuals to their environments by:

- self-directed identity claims (eg. purposeful decorations like posters or use of colour);
- other-directed identity claims (eg. decorations which communicate shared values that others would recognise);
- interior behavioural residue (ie. "physical traces of activities conducted within an environment");

- exterior behavioural residue (ie. traces of activities conducted outside of the immediate environment which nonetheless provide some cues as to the personality of the environment occupant).

Self-presentation is largely unconscious in the physical realm and comes naturally to most people. People may also use in-crowd markers (like a shirt with a band logo on) consciously to send certain messages to people who will recognise them, whilst not drawing any attention from people who won't [40]. Later in this chapter I look at how our presence in digital spaces fail and succeed to take the place of the physical body when it comes to interactions and identity formation.

## 2.1. The self in context

By reflexively adjusting one's perception of self in reaction to society, people construct their individual identity. [40]

Development of personal identity is not only something that happens internally. We are strongly influenced by feedback (conscious and unconscious) from others around us, as well as the particular setting and culture in which we find ourselves. How we react to things outside of our control in part determines our identity construction, and some people adjust their behaviour in response to feedback more than others [253]. Thus identity is *socially* constructed, and often is dynamically adjusted according to context [40].

## 2.2. The project of the self

Giddens [118] looks at the relationship between macro and micro views of the world, acknowledging that broader effects of society impact individual behaviour, and vice versa, with neither one being the primary driving force. This suits well my ideas about online self-presentation, confirming the complex interplay between technological affordances, individual actions, and the place of both in a cultural and social context.

Giddens argues that self-identity is an aggregation of a person's experiences, an ongoing account, and a continuous integration of events. In contrast to Goffman's dramaturgy, Giddens downplays the role of an audience, and in contrast to Brunswik's lens theory, he downplays what we can learn from the traces someone leaves behind. Giddens argues that self-identity cannot be uncovered from a moment, but something which is ongoing, over time. Modern

society, according to Giddens, affords us more freedom to create our own narratives to determine our self-identity. In the past, rigid social expectations dictated our roles for us. However, increased choices about what to do with ourselves may also increase stress and prove problematic. Awareness of the body is central to awareness of the self, as the body is directly involved in moments we experience in daily life. As we are now explicitly constructing a narrative about our identities, rather than having one ascribed to us by society, the self is an ongoing project which takes work to maintain [119].

The focus on explicit actions and decision making about self-presentation is pertinent when it comes to digital representations of identity.

### 2.3. Extending the self

Early to mid 20th century philosophers and social scientists complicate notions of the 'self' by combining it and extending it with our physical surroundings, and this view emerged long before the Web. Heidegger expresses technology as coming into being through use by a human; when tools are used the tool and its user do not exist as independent entities, but as the experience of the task at hand (using the example of a carpenter hammering, unaware of himself or his hammer) [148]. McLuhan discusses media, literate and electronic, from the printing press and electric light to radio, TV and telephone, and its impact on how we communicate. He places communication technologies as simultaneously extensions of and amputations of our bodies and senses, which continuously and fundamentally re-shape the way we (humans) see and place ourselves the world [195]. More recently, Clark's Extended Mind Theory uses the example of a notebook as a means of externally processing information that would otherwise be carried out by the brain, drawing the external world in as party to our cognitive processes [66].

The next logical step is to consider how the modern digital technologies of Web and social networking can also be considered extensions of the self, and this is addressed in part by Luppincini's notion of Technoself [182]. Technoself incorporates (amongst other things) extension of the self through physical technology embedded in the body (cyborgology); in our changing understanding of what it is to *be*, as life is extended and augmented through advancing healthcare; but also in our relationships with our virtual selves. This is not a topic into which I will dive deeply from a philosophical standpoint, but the idea

of the Web and online social networks as extensions to the self rather than as separate entities or concepts is worth bearing in mind as this thesis proceeds to explore the complexities of intertwined digital and offline identities.

### 3. Offline to online... and back again

When people use digital technologies to communicate, they are passing a version of themselves through the filter of the platform they use. In this section I discuss the relationships between online and offline selves.

One might assume that the disembodied nature of the interactions lets people have more control over how they are "seen" by others; Turkle claims that it is "easy" to create and tweak a perfect self online [277]. In reality this control is affected by a great number of factors. Turkle writes that computer-mediated communication is predictable even in the ways it is unpredictable, and that people seek out this predictability in preference to face-to-face interactions, and in preference to facing emotional vulnerability. Turkle's argument assumes that we have full understanding and control of the digital systems we use and the audiences we are reaching through them. It presumes we know exactly how and where and when the data we input will be output in the short and long term, and how others will interpret it.

Turkle's argument neglects that at every crossroad in these 'predictable' systems are stationed unpredictable humans, perhaps with conflicting interests and motivations, from the conception of a social system, to its realisation and use. Turkle has been studying for decades how people explore, experiment and find themselves through technology, and her overriding narrative is of a desire to express an idealised version of the self; one that is not subject to any interpretation other than what the expresser desires. However, as discussed further in the next section numerous studies of social media users find a variety of other types of motivation for participation.

Keen [161] on the other hand emphasises the risks that individuals become trapped by technology of which they have neither understanding nor control. Being swept up in cultural technology trends, social media users may unwittingly become "prisoners" of a carefully curated digital "hyperreality", the importance of which supplants their offline lives. Similarly, [245] suggests that "fantasy gets in the way of real progress" when it comes to self improvement, but studies only examples of "catfishing", where individuals create exaggerated online profiles in order to deliberately mislead potential romantic interests. The idea that the online self *replaces* rather than supplements the offline self also

misses the nuances of how and why people use social media in the first place. Nonetheless, I agree that online worlds are certainly not just a mirror of the offline.

On the contrary, not only are online identities some form of reflection of offline identities, but the inverse can be true as well. In ethnographic studies of teenagers' use of mobile apps for socialising, it is reported that crafted online personas both reflect and feed back into teenagers' true sense of self [113]. Numerous studies from the field of psychology, reported in [29], demonstrate various ways in which playing a role online or in a virtual world reflects back and directly changes people's offline behaviour. Avatars can be used to build confidence and reveal suppressed personality traits. Relatedly, study participants who were asked to interact from behind avatars which conveyed different appearances of age, race, and body type expressed affinity with these previously unfamiliar experiences, and responded differently in personality evaluations before and after. A longitudinal study of teenage girls roleplaying online, a process through which they developed their identity through narrative, revealed positive impacts on their confidence, through new friendships [269].

There is a long history of assessments of online interactions which proclaim that the self-centered nature of social media makes people narcissistic, that competition for reputation isolates us [161], and that the construction of an ideal self or facade is damaging, especially to young people. However, [113] argues that narcissism is not created by the Web, but enabled by it, as an existing need for validation is more readily satisfied.

Several studies [2, 236, 290] find that people's self confidence or body image drops after viewing the online profiles of people who appear to be more attractive or more successful than them. [292] explains that passive use of social media is what appears to have a negative impact on people, whereas active use has a positive effect on well-being. This is shown in several studies, including [273], which explores the beneficial effects of browsing one's own Facebook profile; [210] which discusses how selfies can empower marginalised communities; and [214] which finds a positive impact on self-esteem of teenaged girls who engage in "auto-photography".

It is also worth bearing in mind that experience in the virtual world can cause physical reactions - laughter, tears - and the virtual and the physical blur together in the subject stream of experiences, adding to an identity which is made of virtual and physical events blended together [269].

### 3.1. Authenticity and integrity

In an interview, Facebook founder Zuckerberg said that "having two identities for yourself is an example of lack of integrity" [164]. This received public backlash at the time, and on several more occasions as Facebook and other social networking sites imposed real name policies, sometimes linked to an official ID [38]. This removes a level of identity control which many people take for granted. A particularly clear example can be seen in the reaction of a community of drag queens who were used to being able to interact online using chosen names. Real name requirements removes a vital aspect of fluidity from their interactions, glossing over the seams which provide poignant performance material and a route towards a group identity [179, 18]. Additionally the requirement for real names and the ability to report 'fake names' became a mechanism whereby a marginalised group could be harassed and silenced [197]. In general, an attitude that people should be happy to connect all of their identities together under a single legal name is an expression of social privilege: a result of having no features or proclivities which are socially censored.

Sandberg, whilst COO of Facebook in 2012, commented that profiles as detailed self portraits is a "shift towards authenticity" [161], but her organisation's notion of 'integrity' as a single complete version of oneself that is the same no matter to whom one is presenting is somewhat at odds with authenticity. Neither are people "intraviduals", caught between competing identities as claimed by Conley in 2009 [67], but expressing aspects of themselves appropriately and according to context. This is behaviour which we have already established via Goffman as ordinary offline, and so shouldn't be considered unusual online. One's 'authentic' self-presentation may be partial or moderated and no less genuine for that. Indeed, some individuals find they are more able to express their authentic selves online than they are offline due to oppression or social expectations which are disjoint from their core values.

However, in highly commercialised or competitive online environments, 'authenticity' is a quality to strive for, to maintain an audience. A Web search for 'authenticity on social media' will reveal a plethora of guides on how to craft an 'authentic' persona, how to maintain personal-but-not-too-personal ties with one's audience so that they see that you're just like them. For individuals who set out to explore and express their identity online, this can be a tricky world to navigate [91]. True authenticity in online communities is seen as disjoint with self-promotion and celebrity; popularity implies a reduction in authenticity,

perhaps linked to 'selling out' or 'pandering' to an audience [89, 100]. Whereas authenticity is often seen from an outside perspective as always something manufactured, an idealized reality [192].

The idea that online spaces are under control of their owners suggests others may be suspicious of their authenticity [300] but reinforcement of social identity from others can counteract this. Warranting theory describes how information that appears to be outside of the subject's control—for example, a message posted publicly by a friend on someone's profile—can reinforce the trustworthiness of the other profile data to an outside observer [301, 297, 299].

### 3.2. Dishonesty and deception

"Some argue that distinct contexts are unnecessary and only encourage people to be deceptive. This is the crux of the belief that only those with something to hide need privacy." - [41]

Most people like to consider themselves to be quite honest in their communications with friends, family and acquaintances. However, even honest people routinely modulate what they share, omitting and sometimes falsifying information in order to reduce social friction, avoid confrontation, defuse awkward situations, or to save face [51, 55]. Hancock et. al. introduced the term *butler lies* to refer to a common use of simple lies to manage communications, such as smoothly exiting from an unwanted conversation [144]. Online, the notion of who our 'friends' are has become increasingly blurred and difficult to define. In such settings, people commonly navigate different social spaces, projecting and varying self-presentation according to the ways they want to be perceived by each [192].

Whilst part of tailoring one's presentation to an audience is the ability to carry out some level of *deception*, with personal communications, there is an implicit expectation of authenticity [8]. However, online, the need to navigate multiple and uncertain audiences means that we may constantly vary our self-presentation. Authenticity becomes a social construct derived from the social context and how we wish to be perceived by a given audience [40]. We may be deceiving, at least to some extent, nearly constantly without even being conscious of it.

Deception has long been studied, both within and outwith the HCI community. Traditionally, deception has been cast in a negative light [30], to be used only if no other option is available. In the 1980s, however, communications researchers

began to investigate the positive aspects of lying, in particular *white lies* - socially acceptable lies which cause little or no harm to the recipient [58].

In 1992, McCornack cast deception as an understandable response to complexity: "[r]esearchers studying deception recently have begun to argue that deceptiveness is a message property that reflects a kind of functional adaptation to the demands of complex communication situations" [194]. People then manipulate the information which they share as a necessary part of participation in society. This has led to recent work on the positive aspects of deception in human computer interaction, in particular how butler lies are used to ease social situations [144], and how systems can deceive their users for beneficial reasons [1].

Several different taxonomies of lying and deceptive behaviours have been proposed [58, 80, 178]; Anolli et al. examined a family of deceptive miscommunications, including self-deception and white lies [7]. They look at *omission* of relevant information, *concealment* using diversionary information, *falsification* and *masking* with alternative, false information. Of particular interest is their claim that "a deceptive miscommunication theory should be included in a general framework capable of explaining the default communication", that is that deception should not be seen as a psychologically different activity than 'normal' communication. This tallies with the earlier approach of McCornack [194] who situates deceptive messages within the spectrum of *information manipulation*. This, combined with the lens of Gricean conversational maxims, allows for an explanation of deceptions where some of the truth is told, but information which the speaker knows is relevant to the listener is omitted or obscured [125].

Motivations for lying have also been extensively studied in social psychology. Turner et al.'s taxonomy included *saving face*; guiding social interaction; avoiding tension or conflict; affecting interpersonal relationships; and achieving interpersonal power [278]. Camden et. al. [58] develop a detailed categorisation of lies to do with basic needs, managing affiliation with others, self-esteem and miscellaneous practices such as humour and exaggeration.

Many malicious or undesirable behaviours are facilitated by the ability to create and alter identities. *Astroturfing* [64] has become common online [314], with corporations and governments employing sophisticated identity management software to carry out large scale operations. Possibly the most famous of this is the "50 Cent Party", hired by the government of the People's Republic of China

to post favourable comments towards party policy [310]. On a smaller scale, *sock-puppets* — multiple accounts controlled by a single person — are used to skew ideas of consensus and distort discussion in online societies, leading to attempts to automatically identify such accounts [48, 255]. Personas can be constructed for the purpose of *trolling*, whether it is overtly offensive in order to cause outrage or more subtle manipulation to trick people into wasting effort or taking caricatured positions, and correlations have been shown between enjoyment of trolling and everyday *sadism* [49].

Many of these activities are a form of *obfuscation*, in some way hiding the truth, polluting the data pool and diminishing trust. The ethical issues here are complex and contextual, with the viewpoints of different actors having considerable divergence [46].

Another strand of research borrows from information warfare, to look at the possibilities for disinformation. Disinformation tactics are most useful when a channel of information cannot be completely closed, but can be rendered useless by being filled with incorrect, but plausible, assertions in order to lower its overall signal-to-noise ratio [281]. The intended target of the lie may not be the official recipient of the message: lies can be directed at those who are eavesdropping on the communications channel or surveilling the participants [6]. Techniques used include *redaction* to remove parts of the message, *airbrushing* to blur parts of the message and *blending* to make the message similar to other plausible messages, as well as other forms of *information distortion* [6].

In chapter 3 I carry out two studies which aim to bring together these general theories of deceptive behaviour with a closer look at how and why people might engage in them online.

## 4. Networked publics

Social media technologies blur the boundaries between private and public, and this affects identity performance. Ten years ago, boyd hoped that educators and technologists would succeed in easing the cultural transition for young people into the networked era [34]. She describes social networking sites as a type of "networked publics", technologically-mediated spaces where people can virtually go to interact with their friends, and where they may be subject to observation or interjection by passers-by. Differently to offline public spaces, online publics may be persistent, scalable, searchable, replicable, and/or have invisible audiences. These features of networked publics affect how people express themselves and interact, however they do not directly dictate participants' behaviour [41]. Networked publics are not only spaces, but collections of people or "imagined community"; different publics can serve different purposes, but can also intersect with each other [39].

Over subsequent years, boyd and many others proceed to explore the effects of these differences on those who engage with online social media to different degrees. In this section I recount some of these studies and findings.

A benefit of participation in networked publics is that a wider variety of communities are accessible than offline. Niche identities don't have to be set aside to fit in [113]. Online interactions are "not simply a dialogue between two interlocutors, but a performance of social connection before a broader audience" [39]. boyd looks specifically at teenagers in networked publics, who she says have sought online spaces in recent years as they are not allowed to 'hang out' any more in physical spaces like malls [39, 190].

First I reflect on the digital substitutions for the physical body in online social interactions. Then, in comparing and contrasting 'old school' style blogging with contemporary (circa 2013-2017) social networking sites I look deeper into how differences in technological affordances impact peoples' interactions and self-expression.

Audiences for identity performance as well as the context in which the performance takes place are critical, but online both of these may be unknown or dynamic, or both. I'll introduce work around imagined audience and context collapse, both of which pioneer our understanding of identity behaviours in networked publics. When audiences and contexts are known, we can examine

how people connect with others and form communities; in the final section I look at trust, social reinforcement of identity, and studies of what social media participants choose to disclose or conceal.

#### 4.1. Profiles and embodiment

In Faceted Id/entity [40], boyd highlights several differences between self-representation offline compared to online. Embodiment is a key factor in self-presentation and she claims that there is considerable difference between performing one's identity through appearance, eg. fashion and body language, when walking into a room, compared to explicitly describing oneself by entering attributes and other personal information into an online form to create a profile. The disembodied nature of online interactions means that people must find new ways to express themselves, and manage the impressions other people have of them, or "a new type of body" [269]. A lack of control over one's online self presentation is compounded by the inability to visualise - or perhaps even be aware of - the data that is collected by the systems we use. Online activities are logged over time to an extent that most individuals are not aware of; these activities, an individual's *expressions given off* (Goffman), are used, largely unknowingly, for the commercial benefits of third parties; this constitutes a kind of implicit or unconscious profile. boyd suggests that visualising all of one's personal information that is available online, as well as visualising one's 'audience' or social network connections, would provide an individual with better awareness of, and so better control over, their online image. In boyd's prototype interfaces, users are explicitly asked for personal data in order to build a profile of themselves, and boyd does point out the problematic nature of this, compared with the unconscious or implicit identity performance one conducts in offline social settings.

Counts in [70] explores the impact of profile attribute selection on self presentation, and finds that upon completing the values for 10 attributes, participants converge on their "ideal" representation of themselves. This study also finds that free-form attributes are better than ones with preset choices for participants' satisfaction with how they have portrayed themselves. This study does not take into account that most online profiles are created in a particular context, with a particular purpose in mind. Asking participants to express their ideal self-presentation 'in general' vastly oversimplifies reality. Participants are

not told who the consumers of the profile they are creating are expected to be, or how it is to be presented; nor are participants given an opportunity to indicate who their expected audience is or what they think the profile is for.

Since boyd's prototypes were designed, social media gained widespread popularity. Most, if not all, mainstream systems request input of explicit personal data to build an initial profile, despite the discord of this activity compared with offline identity expression. However, unlike in boyd's prototypes, it doesn't stop there. Such systems encourage ongoing engagement through adding and messaging contacts with various degrees of publicness, creating status updates to broadcast a current situation, production of creative media content, and feedback on content and updates created by others. As we will see in chapter 3, contemporary social media builds one's profile from various combinations of these online activities, and typically use far more than the explicit data entered by the user to generate a representation of a person. This increases the likelihood that individuals may not have an accurate impression of what this representation looks like to others.

Recent studies confirm that visuals are a key part of expressing identity online. Many focus on selfies as a modern substitute for the body [285, 173, 246, 110], but [270] examines self-presentation through other kinds of photos. Examples include humorous images from popular culture or photographs of other things with an overlayed caption, coupled with a tag (eg. #currentstate) that indicates the poster relates to this concept; as well as photos of items that people carry with them day to day. In [220], self-expression is performed through use of Twitter hashtags, and [172] suggests that food photography is a means of self-presentation.

Pointing at something and saying that one has chosen it as self-representative makes the assemblage of tags, text, and image a culturally intelligible self-representation [270].

A lack of embodiment can also have a distinct advantage. In [257] several studies of people with disabilities who use online social systems are reviewed, and reveal findings about increased control over disclosure of disability (which may not be possible offline) and reduced isolation when people are able to interact online.

Next we look more closely at the behaviour of users of Social Network Sites, of which "profiles" are a key feature [33].

## 4.2. Social Network Sites

In 2007 boyd and Ellison defined Social Network Sites (SNS) to be Web-based, bounded, public or semi-public, and afford creating and viewing connections with or between other users [33]. They note that users of these systems tend to connect with others with whom they already have a 'real life' social relationship, and present a fairly thorough history of SNS from 1997 onwards, which I won't recount here. This definition is pertinent to this thesis due to its emphasis on profiles, implying self-presentation, as a core feature of SNS. In 2013 they updated their definition to incorporate different types of content and data into profiles; to de-emphasise the traversal of connections (as this became more important to machines than humans); and to emphasise participants' interaction with streams of user-generated content [98]. I will proceed to focus on case studies and experiments which were carried out since these definitions, and due to the rapid pace of change in this area, prioritise those from the last four years.

Many studies of identity formation on social media focus on young people and teenagers. One reason is because this is a crucial point in life for understanding oneself and asserting a personal identity. Other studies approach teenagers as somewhat alien "digital natives", born into a world of social technology which is expected to fundamentally change how they interact with the world compared with older generations, who don't or can't distinguish between online and offline [17]. I will relate the results of these studies, but note that I disagree with the notion of a "digital native" because being born in a particular year or even raised around modern technology does not automatically give one a natural instinct for identity expression in digital spaces, and not even necessarily more opportunity to experiment and reflect than older SNS users [39].

Digital communication technologies can help or hinder identity formation. *The App Generation* [113] provides a balanced argument between the pros and cons of teenagers socialising through mobile applications. They find that some applications provide a "prepackaged identity" for users to adopt rather than encouraging experimentation. The affordances of applications shape the forms of expression that are available, and so identity formation is in a way controlled by the application designer. *Born Digital* [17] suggests that teenagers experiment with identity online, but aren't fully aware, or don't care about, the traces that are left behind when they do so.

In some cases, for example fan communities, self-presentation shifts between a more playful fictional identity performance, and an identity which is closer to 'real life' [14]. In others, such as professional self-presentation, individuals lean on automatically generated metrics by the system they use to convey a positive image, with gamification or commodification of the self becoming commonplace [136]. When SNS provide a platform for professionalising passions such as content creation, [92] notes that participants may be even more vulnerable to the consequences of performing and maintaining one's self-presentation in an exposed online space, as well as the "labour of visibility" that goes into it.

Most people occupy multiple roles offline, find ways to establish and maintain boundaries between them, and continue to do so to different degrees when taking representations of these roles to online spaces. SNS increase the permeability of boundaries, but users employ various tactics to manage their identity when a one-identity-per-person model is imposed on them [230].

[250] describes how Twitter users subvert features of the system to express themselves in new ways, as well as reflecting on how changes to the functionality of the system affect how people use it. This supports [220] which, through content analysis of trending hashtags, also describes how people work around technical constraints of Twitter to meet their self-presentation needs.

Even in the early days of SNS (specifically Facebook) beginning to rise to popularity when use was overflowing into the workplace, the access control settings offered by Facebook were considered too complicated to enable most people to realistically manage connections with both professional and personal contacts from a single profile, despite the potential advantages of connecting with colleagues through the platform [85]. More recently, we see that SNS users manage tensions between their multiple roles and the affordances of systems by segregating their audience across using multiple platforms. The interview study in [315] found that sharing decisions across multiple sites are made primarily based on the known audiences of the different sites, and the content being shared. This study also recounts previous work on motivations for using different SNS, including to connect with old friends, and share pictures, which feed into decisions taking regarding content sharing. A similar study found that family was a crucial audience to whom more private sharing was desired [104], and findings in [293] indicate that Facebook users desire to re-asssert their offline boundaries when online, and concurs that managing this through the

tools that Facebook provides is cost-intensive. Facebook itself compounds this issue by using identity information as a "social lubricant" which encourages people to make new connections [99].

### 4.3. Blogging and personal homepages

Personal homepages and blogs have been around for considerably longer than SNS, though remain a comparatively specialist practice. It is widely accepted that blog or website owners have more control over their online space than do users of SNS, [237, 189], including freedom to innovate with the site's appearance and thus explore more individualistic aspects of the online self [274]. Relatedly, communities of bloggers are not owned or controlled by a single entity [79].

Through observations of over 200 blogs within a particular community and semi-structured interviews with 40 bloggers, [79] identifies five aspects which affect how bloggers build their identity: name and blog title; descriptive attributes; post content; voice; affiliations; and visual design. All of these are subject to change over time, and sometimes major offline transitions can cause a shift to a new pseudonym or blog altogether; often the audience is invited along however. Blogs are often designed to reinforce community norms, to enhance a sense of belonging; as a result, the community develops and evolves its own identity, which in turn influences how newcomers choose to present themselves. Bloggers' contributions are fragmented across different domains, and where their writing style and topics constitute a part of their identity, so it fragments their identity.

Blogging communities are traditionally more accepting of pseudonyms but [79] notes that distrust is not of other community members, but rather of personally-known community outsiders who may accidentally stumble across blog entries.

Earlier in this chapter I mentioned Brunswik's lens model which describes a way in which identity can be constituted through physical traces left behind. This model has also been used to understand how observers make personality judgments about people based on the traces left in their *digital space*, ie. personal homepages [103, 189, 291, 219].

The importance of themes and designs of blogs and homepages is emphasised by [79], who mentions that whilst some blog consumers use a feed reader to receive new content from the blogs they are interested in, they often click through to the original post to view it in the context of the author's own space.

On the other hand, [244] takes a snapshot of a random sample of blogs in 2003 and maps the state of the blogosphere through analysing visual elements in depth and tracking commonalities. The conclusion is that significant customisation of blog templates was in fact relatively rare, with most people only slightly tweaking colours or adding custom images. A likely explanation for this is that bloggers lacked the technical expertise to do so.

Studies of blogging communities outside of the US demonstrate that blogging is not a uniform practice that can be understood as a whole [237]. Certain communities (in this case, Muslim ones) which are seen by outsiders as homogeneous use blogging to highlight their uniqueness and individuality. Others (for example in China) emphasise their ethnicity and culture as a key part of their identity. Blogs from the Paris Banlieues in fact had a direct impact on how the mainstream media portrayed their plight; an example of how personal identity expression in networked publics was able to affect a broader social understanding of that identity. The overriding message from these studies is cultural taboos and offline societal context affect narrative about identity, and this is reflected online.

#### 4.4. Imagined audience

The audience to whom one performs is critical in forming the context in which one is performing [41]. On SNS, people are often expressing themselves to multiple audiences simultaneously. When people are aware of this, they take different strategies when it comes to navigating what they share; individuals with many followers on Twitter practiced self-censorship (only posting things they are happy for the worst-case audience to read) and practicing coded communication (strategically targeting some posts at some audience members, and others at others, to maintain overall interest) [192].

However, given the many possible ways in which Tweets can be discovered and consumed ([192] questioned people who post publicly) it is virtually impossible to determine the actual audience for one's content. Thus, people imagine who their audience is likely to be, and express themselves accordingly. Obviously these imaginings, which may stem from understanding of the affordances of a particular platform, or a particular community or topic of discussion, impact how people express themselves online.

[180] theorises about how the imagined audience is synthesised, and draws in Giddens' structuration framework, noting a combination of macro- (social roles, technical affordances) and micro-level (individual motivations, technical skills) factors. [180] ultimately concludes that asking people about their imagined audience is prone to errors or misinterpretations, as imagined audience is a concept which is both difficult to measure and difficult to express.

Relatedly, as people perform in networked publics, they must contend with a "networked audience," who are not connected only with the performer, but also with each other [192].

#### 4.5. Context collapse

I have so far discussed how people attempt to map boundaries from their offline lives into their online interaction spaces, and the notion of imagined audience. Context collapse occurs when boundaries come down and personas intended for different audiences are merged [192]. The consequences of this may range from slight social awkwardness, to direct breaches of privacy and potential danger, and have been examined in a variety of different circumstances, such as [74, 303, 95, 94].

Thanks to the properties of networked publics such as searchability and persistence, contexts may also collapse when information is consumed later, or through a different systems, whereby it may be interpreted differently by the consumer than how it was originally intended [41].

As we look forward to how SNS and online self-presentation in general will evolve, we must consider how the lines people have drawn around their contexts are tethered to particular (versions of) systems. What happens when these systems change, merge, or disappear? As designers of new systems, we must be cognisant of the role technical affordances play in creating, enabling, and destroying social boundaries.

## 5. Everybody knows I'm a dog

"While once viewed as a set of technologies built in resistance to the ugliness of the dot-com era, social media is now intertwined with neoliberal capitalism and data surveillance" - [36]

We are rapidly moving into a world where information about nearly every aspect of our lives is becoming sensed, recorded, captured and made available in digital form. Data is captured and shared voluntarily, as tools invite ever more intimate participatory surveillance [5]. While the abundance of information traces has unlocked a wide range of new kinds of applications (eg. [4, 69]), the creation and potential for disclosure poses new threats to individual privacy and autonomy. The overall lack of transparency by manufacturers regarding how they are capturing and handling personal information has created a heightened sense of unease among many, in addition to the potential threats dealing with their unintentional disclosure or misuse [106, 198, 105].

Various data and surveillance scandals involving private companies and governments [132, 209] that gained media attention mean that awareness of surveillance and personal data collection is growing amongst the general public. There are many studies examining peoples' awareness of and attitudes towards privacy and surveillance on SNS, but I will not detail them here. During studies in the 1990s, Westin defined three categories to describe how ordinary people feel about privacy: "pragmatists", "fundamentalists" and "unconcerned." Privacy pragmatists accept that there may be tradeoffs between benefits to information sharing and the intrusiveness of requests for information. Fundamentalists distrust organisations which request personal data. The unconcerned are comfortable with sharing personal data with organisations in exchange for services [170]. Westin found, prior to the Web becoming mainstream, that approximately half of the general public are pragmatists; just over half of the remainder are fundamentalists, and a minority are unconcerned. Suffice it to say that people *do* care about privacy, and are just finding new ways to manage it [41], contrary to what certain tech executives might claim [158, 102]. My main concern in terms of this thesis is how privacy infringement might impact presentation of self. We have already seen that online identity performance may alter a general understanding of oneself, which reflects in the offline world. So I must ask: when people self-censor online due to privacy concerns, how does this stifle self-expression, and in turn impact internal identity construction?

Furthermore, implications of our online sharing decisions affect more than just ourselves; "interpreted selves" are created by recognising patterns across millions of people [42].

There is often an asymmetry about the collection and use of data. To take a relatively prosaic example, Facebook introduced 'read receipts' on messaging, which indicate when a user has seen a message. This feature has been shown to cause anxiety when present in email systems, as users seek to maintain their *responsiveness image*, the impression which they project to others about how they respond to input and partition their attention [279]. Once ambiguity about attention has been removed, a whole class of white lies - 'The internet was bad, I couldn't check my messages' - are no longer possible, and people develop alternative strategies, such as not opening messages until they feel prepared to respond. The key difference in the context of SNS is that the user does not have the same degree of control over the channel - email receipts can be switched off, but SNS offer different levels of control.

The social aspects of privacy relate to what DeCew terms *expressive* privacy - a freedom from peer pressure and an ability to express one's own identity [78]. Nissenbaum's contextual integrity [212, 213] seeks to understand "appropriate sharing", looking at the ways in which flows of information are governed by norms, which may be easily violated as technological systems repurpose and share data.

## 6. The commoditised self

Social systems which involve content creation (like YouTube) or knowledge generation (like Wikipedia) are commonly seen as cooperative communities, whose participants generate value both for each other and also for the organisation behind the system. In [83], van Dijck et. al. contest several uncritical manifestos for the business and communal interests of revolutionary Web 2.0 peer-production. They point out that seemingly open co-creation platforms are still profit-driven commercial entities. These entities do not provide tools out of benevolence, but in order to harvest metadata about their users, which they can process and resell. The balance of power between individuals and corporations is not swinging back towards the individual, as proponents of user-generated content sites claim, but the illusion is created that it is.

Users of systems often have little understanding of how their activities are being exploited - or nudged [84]. Even as users are empowered by technology to create media, products, or services they desire, [84] calls into question their *agency* when participants are being used and manipulated by commercial entities under the guise of community formation or participatory culture. [21] describes SNS profiles as "commodities, both produced and consumed." In more recent years, awareness of this fact has spread. A popular refrain from advocates of less commercial alternatives is that "if you're not paying for it, you're the product."<sup>1, 2, 3</sup>

In a similar vein to the previous section, we must also wonder about the impact of external commercial and economic forces which shape the tools and systems people are using to express themselves online. As a contrast, in the second half of this thesis, I focus on decentralised systems, which are potentially much less likely to exploit user metadata for profit.

## 7. The ghost in the (social) machine

Social Machines are systems for which the human and computational aspects are equally critical. In most cases, humans do the creative work whilst machines do the administrative tasks [22]. Up to now, systems have not been designed *to be* Social Machines; rather, the concept and definition of a Social Machine is derived from observations of existing (usually Web-based) systems. Many are products of the contemporary social web, on many different scales and in many different domains, often evolving, responding to technological and social developments, and interacting with each other [149, 56, 77, 261]. Social Machines can be identified within and across social media networks, within and across online communities, and within and across technological spaces.

Social Machines are pertinent to our work here as they provide a lens through which we can examine sociotechnical phenomena which emphasises the interdependence between humans and technology. In studies of social networks discussed previously in this chapter, humans are considered as *users* of systems, and discussions focus around how people react to technology, how people behave in the context of particular technical or social constraints, or how people's lives are changed in response to their interactions with and through digital environments. It is important to also reflect upon the ways in which technology evolves or is reconceptualised as a result of passive (mis)use and active (mis)appropriation by humans. We similarly must recognise technical systems in the wider context of society, and include in our dialogue the developers who design and build technical systems, the organisations and legal entities which finance and drive them, and the cultural and economic climate in which they are situated.

Social Machines which have been studied so far have been described and categorised in terms of purpose [77], motivations and incentives, technology used, goals and processes, quality assessment of outputs, and user participation and interaction [252]. The emphasis in this work is on discussing Social Machines in collective terms; that is, 'a' Social Machine - wherever its boundaries happen to have been drawn (so far these boundaries are typically drawn around the edge of a "service" [252]) - is considered as a coherent whole. The circumstances of the individual human participants not been given extensive consideration. For any given Social Machine, individual participants are diverse and participate in different ways, with varying goals, motivations

and outcomes. As we have seen through studies of social media mentioned previously, they manipulate their online presence(s) so that they may behave in different ways according to different contexts, or may work together to construct a single image controlled by multiple people [71]. Such behaviour impacts our understanding of roles, autonomy and awareness, incentives and attribution, and accountability and trustworthiness of participants. Overlooking unique individual perspectives when observing a Social Machine as a whole can cause incorrect assumptions, for example: believing that participants who lie about who they are have negative intentions in a Social Machine whose overall "purpose" is to strengthen social ties. We must also bear this in mind when designing systems, so that a system may grow in response to unexpected actions of participants rather than hampering their explorations.

I argue that due to the complex nature of online identity, understanding nuanced individual behaviours of participants in a more granular way is crucial for Social Machine observation. I advance this argument in the next chapter through an empirical study of a Social Machine centered around creative media production.

## 8. Conclusions

Present day social media has dramatically increased participation in publishing and sharing online content. Easy-to-use services lower the barrier to entry for connecting with and pushing thoughts out to an audience. Identities expressed through social media are inherently collaborative; every interaction is pushed to a network, and part of a dynamic cycle of consumption and creation feedback. Commenting on someone else's post automatically links the post to your own profile, and often it appears there as well, accessible from two different contexts. Yet SNS permit little customisation, providing preset options for content or reaction templates, and consistent inflexible designs for profile pages. This tips the balance away from the individual aspect of identity construction performed by bloggers in the earlier years of the Web. Yet blogs and personal homepages are left wanting for a dynamically constructed and low barrier to entry network, which impedes the collaborative aspects of identity construction.

People manage shortcomings with the affordances of both blogging platforms and SNS in different ways—contending with invisible audiences and collapsed contexts, as well as reduced expectations of privacy—through carefully crafted personas, strategically omitting or amending the information they post online, or simply using different platforms for different purposes. We see that there are a multitude of factors which affect people's presentation of self online, which vary according to broader cultural or technological contexts, as well as personal motivations and abilities.

Along with their updated definition of SNS mentioned earlier, [98] calls for social media scholars who are studying individuals or communities online to systematically describe the technology in which their participants are situated, and the practices of the users. Technologies are changing rapidly still, so studies which are a snapshot in time can be linked to a broader discourse or overview, and remain relevant as time progresses, if they situate themselves appropriately. One way of accessing this bigger picture is through an organising framework that helps to record the background state of the environment being studied, and so surfaces connections between work that is otherwise perhaps not directly comparable [84]. In chapter 3 I use several studies of my own on diverse identity behaviours across various social network sites in order to propose such a framework.

<sup>1</sup> Facebook is “deliberately killing privacy”, says Schneier (Information Age)  
<<http://www.information-age.com/facebook-is-deliberately-killing-privacy-says-schneier-1290603/>>

<sup>2</sup> If You're Not Paying for It; You're the Product (LifeHacker)  
<<https://lifehacker.com/5697167/if-youre-not-paying-for-it-youre-the-product>>

<sup>3</sup> If You're Not Paying For It, You Become The Product (Forbes)  
<<https://www.forbes.com/sites/marketshare/2012/03/05/if-youre-not-paying-for-it-you-be-come-the-product/#1a01a99a5d6e>>

# Chapter 3

# Social Web in the Wild

## 1. Introduction

The previous chapter discusses an individual's relation with their online representation; how users understand profiles; how the affordances of a profile impact the culture of an online community, including how users interact and relate to each other, and how users understand themselves as part of the community. This chapter contains original studies; two which examine online profiles from an outside perspective, by looking at what systems offer and how individuals appear to be making use of this. Three of the studies go behind the scenes to actually ask profile owners about their participation in the social Web ecosystem.

Each study resulted in a small taxonomy useful for categorising the participants' experiences in each particular scenario. A core contribution of this thesis is to coalesce the results of these new studies, along with knowledge from existing literature, into an overall framework consisting of five concepts. This framework - the 5 Cs of Digital Personhood - constitutes the key components for describing online self-expression experiences. The framework is summarised here for reference, and I discuss its derivation in more detail in the conclusion of this chapter.

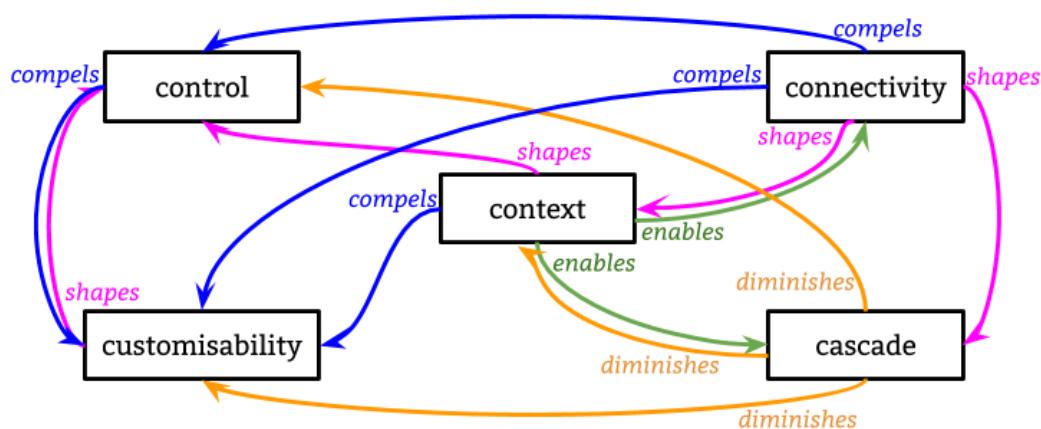


Figure 1. Control, Customisability, Connectivity, Context, Cascade; and their relations to each other.

Each component encapsulates a variety of different parts or aspects which are revealed through the studies in this chapter, as well as prior research:

- **Control:** over persistence or ephemerality of identities, attachment or not to real names, traceability between different identities (eg. Can I delete my profile?).
- **Customisability:** of the data that is included in an online representation, the extent to which this is available to others, and how it is presented (eg. Can I change the name that appears on my profile?).
- **Connectivity:** to others and an audience, known or imagined, and how impressions by this audience can be managed (eg. Do I know how this profile appears to my mother?).
- **Context:** the social/cultural expectations of a platform or community; personal motivations and use cases; technical constraints of systems; offline cultural norms or biases which affect or constrain online behaviours (eg. Are the people who control this platform obliged to adhere to the same laws as I am?).
- **Cascade:** of personal information throughout a network, perhaps unknown; 'profiles' generated by algorithms, data passed around by third parties or collected through surveillance; expression 'given off' over which individuals have little knowledge or control (eg. Is my data being used to recommend products to me?).

Whilst all five components influence each other in complex and shifting ways, I illustrate key relations with the following terms:

- **compels:** the existence of aspects of one necessitates the involvement of aspects of the other.
- **diminishes:** aspects of one reduce the effect of aspects of the other.
- **enables:** aspects of one increase the effect of aspects of the other.
- **shapes:** aspects of one feed into aspects of the other; the latter is formed according to or depending on variations in the former.

## 1.1. Overview of studies

Table 1 summarises the methods, inputs and outputs of the five studies in this chapter.

The previous chapter established that there are various different (potentially overlapping) perspectives that need to be taken into account when discussing online self-presentation:

- **Active users** of a system, who maintain a profile.
- **Passive users** of a system, who may not have a profile of their own.
- **System designers** and developers, who must model and display data about their users.
- **Third-party developers** who build additional services using data from another system.
- **Outside bodies** which seek to influence or direct how systems are used for legal, ethical or economic reasons.

The five empirical studies in this chapter touch on each of these perspectives to some degree.

The first study sets a baseline for describing and categorising online profiles by asking the question "what is a profile?" and takes an objective look at 18 online systems which employ user profiles in a social capacity to classify their features. Subsequent studies focus on the people behind the profiles, or behind the systems themselves.

As hinted at in the previous chapter, individuals are rapidly and often intuitively developing coping mechanisms and practices to improve their handling of online self-presentation and impression management despite the constraints of the tools they use. The studies build on this background, first by observing system users from the outside (in the case of creative content producers on YouTube), and then by asking them questions and exploring their feelings and experiences with online profiles, with regards to: deception and lying on social media; imagining social systems as tools for mediating reality; and designing and building one's own customised social systems.

Table 1. Overview of studies

| Study  | Type                       | Participants | Publication | Perspectives | Resulting terminology/themes  |
|--|----------------------------|--------------|-------------|--------------|---|
| What is a profile?                                     | Descriptive, observational | 18           | n/a         | A S T        | flexibility, access control, prominence, representation, portability  |
| Constructing online identity                           | Empirical, observational   | 10           | WWW14       | A            | roles, attribution, accountability, traceability  |
| The many dimensions of lying online                    | Survey                     | 500          | WebSci15    | A P S        | system, authenticity, audience, safety, play, convenience   |
| Computationally mediated pro-social deception fictions | Interviews                 | 15           | CHI16       | A P O        | effort & complexity, strategies/channels, privacy & control, authenticity & personas, access & audience, social signalling & empowerment, ethics & morality |
| #ownYourData   | Interviews                 | 15           | n/a         | A S T O      | self-expression, persistence/ephemerality, networks & audience, authority, consent  |

**Perspectives:** A — Active users; P — Passive users; S — System developers; T — Third party developers; O — Outside bodies

- Publications:**
- WWW14: Guy A. & Klein E. (2014) *Constructed Identity and Social Machines: A Case Study in Creative Media Production*. Proceedings of the 23rd International Conference Companion on World Wide Web - WWW'14 Companion.
  - WebSci15: Van Kleek, M., Murray-Rust D., Guy A., Smith D., O'Hara K., & Shadbolt N. (2015). *Self Curation, Social Partitioning, Escaping from Prejudice and Harassment: The Many Dimensions of Lying Online*. Proceedings of the ACM Web Science Conference. 10:1-10:9.
  - CHI16: Van Kleek, M., Murray-Rust D., Guy A., O'Hara K., & Shadbolt N. (2016). *Computationally Mediated Pro-Social Deception*. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 552-563.

## 2. What is a profile?

This is a descriptive study of 18 systems which employ profiles in a social capacity. This study results in five features and each system is scored according to the degree each feature is present. Using these features we can cluster similar systems together, or differentiate them, and future studies can use these features to create baseline descriptions or characterisations of systems for comparison. The features are: *flexibility, access control, prominence, portability, representation*.

### 2.1. Introduction

In order to build on our understanding of the role an online profile plays in self-presentation, identity and interaction we need a more nuanced understanding of what a ‘profile’ *is* in a general sense. What is the *meaning* of *profile*? I carried out an empirical analysis of digital representations of users of 18 different online systems. From this analysis I derive a set of constructs to capture features of profiles in online systems. I propose this for assessing the benefits and drawbacks of how profiles are implemented in existing systems in such a way that takes into account the scenarios in which they are used, as well as groundwork for deriving requirements for profiles when designing new systems which need digital representations of their users. Once we have a characterisation of a particular *type* of profile a system enables, we can use these as control features when comparing systems side by side. Interesting future study would be to determine how the features of a profile influence actions of users or community formation, and vice versa.

For the purposes of this thesis, I define social systems to be Web-based networked publics which offer individuals consistent and reusable access to an account which they can personalise and use to interact in some form with others in the system.

### 2.2. Context and research questions

Profile generation is an explicit act of writing oneself into being in a digital environment (boyd, 2006 [35])

boyd's definition of profile generation above is based on teenagers' use of Friendster and MySpace in 2006. Today, online social systems use profiles in a variety of different ways, and present them in a variety of configurations. Profile generation is not only explicit, but can occur implicitly, without necessarily even the consent or awareness of the profile subject. As discussed in the previous chapter, studies of online profiles tend to focus on oversimplifications or very specific (unrealistic) use cases, which do not take into account the broader system in which the profile exists. This approach often reduces an individual's representation in the system to a single document or webpage, and neglects the rich array of interactions and activities in which they engage in order to create a presence for themselves. In reality, profiles vary in how they are constructed and the roles they play.

This study serves to introduce a formal classification of profile features, and asks the following questions:

- What are common features of the ways users are represented in online social systems?
- How do these features vary between systems?

Kaplan & Haenlein (2010) categorise social systems into six groups [160]:

- *Blogs* are "special types of websites that usually display date-stamped entries in reverse chronological order"
- *Social Networking Sites* are "applications that enable users to connect by creating personal information profiles, inviting friends and colleagues to have access to those profiles, and sending ... messages between each other"
- *Collaborative Projects* "enable the joint and simultaneous creation of content by many end-users"
- *Content Communities* are for "the sharing of media content between users"
- *Virtual Gaming Worlds* are "platforms that replicate a three-dimensional environment in which users can appear in the form of personalized avatars and interact with each other ... according to strict rules in the context of a massively multiplayer online role-playing game (MMORPG)"
- *Virtual Social Worlds* "allows inhabitants to choose their behavior more freely ... there are no rules restricting the range of possible interactions"

The subjects of this study (see Table 2) are a cross section of these, but there are also some which do not fit into this framework. Since Kaplan & Haenlein's categorisation, (at least) two new types of system have emerged:

- *Quantified Self*: life-logging or self-tracking; automated or manual recording of minutiae of daily life;
- *Transactional*: networks that exist for exchange of goods or services.

## 2.3. Study Design

This is a descriptive study [127], which aims to gather and present information about the current state of social systems with regard to how their users are represented. I do not try to determine causal effects between features of systems, nor do I hypothesise about how these features impact users. Rather, I provide a characterisation of a set of systems as a foundation for future exploratory research.

### 2.3.1. Method

I started with the following areas to investigate:

- Data contained within a profile.
- How profile data may be accessed by others (within and outside of the originating system).
- How profile data may be distributed or pushed to others (within and outside of the originating system).
- The role of profile data within the broader system.

The starting point for a 'profile' was typically a unique identifier for an entity (which could be an individual or group) such as a URL or username. After initial explorations of the profiles in a few systems, these areas were refined into specific questions:

1. What does a profile contain?
2. How are profiles within a system connected together?
3. How are profiles updated?
4. How are people notified when a profile is updated?
5. How is access to a profile controlled?
6. How can profiles be exported from or imported into a system?
7. What constraints are placed on a profile?
8. How do profiles fit in with a systems apparent data model?
9. *What* is the profile for?
10. *Who* is the profile for?

I took one system at a time, and answered all of the questions by logging in (where applicable) to my own account and observing the behaviours of the system in response to interactions with my own and other users' profiles (where necessary), and took screenshots. I also read systems' terms of service, "About" pages, introductory descriptions or statements of purpose, and leaned on my own background knowledge of how the systems are used by myself and others.

Having answered all of the questions about each system, I passed through each one again to confirm, and add more detail if necessary, and I noted similarities and differences between systems. From the results, I derived a set of potential features for profiles, and ranked each system according to the presence of features. This allowed some clustering of similar systems into a general categorisation framework.

### *2.3.2. Subjects*

18 social systems were selected for the initial analysis phase. Most are ordinary websites which one uses by registering, then logging in and out. Some include or require self-hosted software.

Popular systems which I have personal experience were chosen, in order to take advantage of latent background knowledge when navigating the systems.

The information in Table 2 serves to give a feel for the diversity of the social systems being studied.

Table 2. Profiles study subjects: social systems analysed to generate the taxonomy

| System        | URL               | Type               | Specialisation | Overview   | Category <sup>k</sup> |
|---------------|-------------------|--------------------|----------------|--|-----------------------|
| AirBnb        | airbnb.com        | website            | travelers      | Accommodation renting                              | T                     |
| CouchSurfing  | couchsurfing.com  | website            | travelers      | Accommodation, cultural exchange, new connections  | T                     |
| Facebook      | facebook.com      | website            | general        | New and existing connections                       | SNS                   |
| Friendica     | friendica.ca      | website / software | general        | New and existing connections                       | SNS                   |
| Github        | github.com        | website            | developers     | Collaborate on software                            | CP                    |
| Indieweb wiki | indieweb.org      | website            | developers     | Collaborate on ways to develop social web presence | B, CP                 |
| LinkedIn      | linkedin.com      | website            | professional   | New and existing connections                       | SNS                   |
| OkCupid       | okcupid.com       | website            | relationships  | New connections                                    | SNS                   |
| PeoplePerHour | peopleperhour.com | website            | professional   | Hiring freelancers                                 | T                     |
| Pump.io       | pump.io           | website / software | general        | New and existing connections                       | SNS                   |
| Quora         | quora.com         | website            | general        | Q&A (any topic)                                    | CC                    |
| ResearchGate  | researchgate.net  | website            | academic       | Advertise/find research publications               | CC                    |
| RunKeeper     | runkeeper.com     | website            | sports         | Track sporting activities                          | QS, CC                |
| StackOverflow | stackoverflow.com | website            | developers     | Q&A (tech)   | CC                    |
| Tumblr        | tumblr.com        | website            | general        | New and existing connections                       | CC, SNS, B            |
| Twitter       | twitter.com       | website            | general        | New and existing connections                       | B, SNS, CC            |
| YouTube       | youtube.com       | website            | general        | Consume/create media                               | CC                    |
| Zooniverse    | zooniverse.org    | website            | science        | Citizen science                                    | CP                    |

**Categories from Kaplan & Haenlein:** B — Blog (including Microblog); SNS — Social Networking Site; CP — Collaborative Project; CC — Content Communities

**Additional:** QS — Quantified Self; T — Transactional

### 2.3.3. Limitations

As with everything in this thesis, this study is limited by a Western, English-speaking perspective on the systems in question. The observations were conducted from an IP address in either the UK or the US, and I did not attempt to find out how each system differs based on the language preferences or geographical location of users.

Significantly these systems change over time, often rapidly, in response to changing markets, legislation, and available technologies. Most of the data was collected and screenshots captured in the summer of 2015. Some data points were verified to be largely in line with the original findings, but not deeply verified, during writeup in spring 2017. It is important to note that the results are a dated snapshot which cannot be assumed to hold true indefinitely.

I will emphasise again that the nature of a descriptive study does not give any indication of cause-effect relationships between any of the results. Similarly, I can only describe systems as they appear, and not speculate as to *why* they appear such.

## 2.4. Results

Here I summarise the findings of the study.

The most distinct of the systems is the Indieweb wiki, which largely functions as an ordinary wiki except that one identifies oneself with a domain name (logging in with the IndieAuth authentication protocol <<https://indieweb.org/indieauth>>) and thus the 'profile' is tied to one's personal blog, website, or homepage. As a result, profiles are highly custom and diverse; even though they are not hosted centrally by the wiki software they are the main source of identification between users of the wiki, so they are considered here in the same way as the profiles in other systems. In order to study them without visiting the domains of every single user, I also make use of the contents of the wiki itself, which is focused around documenting and recommending best practices for creating a social Web presence; that is, I assume that practices relevant to profile creation described in the wiki are adopted by a majority of users.

### 2.4.1. *What does a profile contain?*

Profiles contain some combination of: **attributes** (key-value pairs of data); **content** (text or media) created by the profile owner; a list of **activities** or interactions the profile owner has carried out in the system; **links to profiles**

with which they are connected; **links to content** the profile owner has interacted with (e.g. 'likes'); links to **collections** of content curated by the profile owner; **statistics** about the profile (e.g. 'member since'); automatically generated **rankings** or ratings of the profile owner; reviews, messages or content **left by other** members of the network.

All of the 18 systems use attributes in the profile, and none use *only* attributes. Attributes may be generic (such as *name*, *bio*, *location*), as well as tailored to the specific system (*countries I've visited* on CouchSurfing; *knows about* on Quora; *looking for* on OkCupid). Some attribute values are offered as a fixed set to choose from, and others permit free-text input. Some systems may require a minimum input of certain attributes, and some leave everything entirely optional.

Facebook has the broadest array of possible attributes, including the possibility to create your own keys, and use ones that others have created. CouchSurfing and OkCupid make extensive use of free text input, prompting users to write short essay-style answers to certain questions. Most systems encourage an avatar or display picture, and several also permit uploading a prominent header image (also known as 'banner' or 'cover photo'). The Indieweb community bases attribute-style profile content around the microformats h-card <<http://microformats.org/wiki/h-card>> specifications, which provides a fixed set, all optional.

Indieweb profiles tend to be the homepages of blogs (although they may be a more static 'about' page) and are heavy on the content and activities aspects. SNS like Twitter, Facebook, Tumblr, Pump.io and Friendica, also lend prominence to content (typically text-based status updates; often photos) and a feed of activities on the site. YouTube incorporates videos created by the profile owner, and how these are organised is highly customisable. For users who have not uploaded video, YouTube profiles contain mostly attributes and activities, and elevate interactions with other content on the site, such as commenting on videos, adding to playlists, and subscribing to channels.

Activity feeds in general vary in their level of detail. Quora displays if someone edited a question or answer. Pump.io distinguishes between 'major' and 'minor' activities, displaying them in separate feeds. Mixed in with posts by the profile owner, Twitter includes a heavily algorithmically curated subset of activities, such as recent follows or likes. Most sites do not include a complete log of all of the possible interactions however. For example, CouchSurfing enables a rich

array of activities, from offering to host a guest, to posting in group forums and arranging events; but none of these are displayed on a user's profile. Similarly, most systems do not display a feed of changes to attributes of the profile, which could also be considered activities.

On the other hand, when users interact with content on a system, for example by liking or favouriting it or adding it to a collection (a playlist on YouTube), reblogging it on Tumblr, voting on it on Quora or StackOverflow; this content becomes part of the profile.

StackOverflow, GitHub, PeoplePerHour, ResearchGate, Quora and RunKeeper are very statistics-oriented. RunKeeper focusses on a feed of offline activities, calculating for example how many calories you lost this week from logged exercise, or how far you ran. GitHub visualises code commits and 'contributions' (helpful interactions with projects) in a coloured grid. ResearchGate and Quora display statistics about how much others have interacted with the profile owner's content. OkCupid also generates statistics based on answers to short, multiple-choice personal questions, and these statistics are dependent on who is viewing the profile, e.g. percentage romantic match, and things like '30% more social'.

Sites which make heavy use of content left by others on a profile are CouchSurfing, AirBnB, and PeoplePerHour. Each of these display reviews of the profile owner by other users, typically in a way that cannot be amended or removed. Facebook allows one to 'write on the wall' of another profile, but users can disable this. However, comments and likes by other users commonly show up alongside activities or created content on a profile as well. LinkedIn prompts users to 'endorse' one another for particular skills, and these endorsements are prominent on profiles. StackOverflow and Quora aggregate ratings left by others on content into overall numbers or rankings to display on profiles.

Many systems give prominence to the connections with other users in the system; LinkedIn displays neither likes nor status updates on the profile, but emphasises contacts and the network around them; Twitter displays followers and following; YouTube, ResearchGate, Pump.io, Friendica, and Quora display subscriptions and subscribers.

#### *2.4.2. How are profiles within a system connected together?*

Connections between profiles may be uni- or bi-directional; some systems permit both. **Bi-directional** connections need to be mutual; triggered by one user and confirmed by the second. **Uni-directional** connections may or may not need approval from the second user, depending on either the system as a whole or individual user preferences. Some systems contain more than one kind of unidirectional connection, which may be named or displayed differently, and carry different connotations. Systems vary in whether or not they notify other users (than the ones involved in the connection) about new connections.

Systems with uni-directional connections are Twitter, Tumblr, Pump.io, Facebook, Quora, LinkedIn, ResearchGate, Friendca and GitHub ('follow'), YouTube ('subscribe'), OkCupid, PeoplePerHour ('like/bookmark/favourite'). Systems with bi-directional connections are CouchSurfing, Facbook, Friendica, and RunKeeper ('friends'), LinkedIn ('connect'). The intersection of these (systems with both) is Facebook, LinkedIn, and Friendica.

Some Indieweb profiles include a list of others the profile owner follows using XFN markup [62], but this is not necessarily widespread. StackOverflow, Zooniverse and AirBnb do not have a means of creating persistent connections between profiles, besides leaving references in the case of AirBnb.

Systems which permit more specific information or **categorisation of connections** are CouchSurfing (specify 'hosted', 'surfed', 'traveled with' or 'never met' as well as the closeness of the relationship), and Facebook (can specify specific relationships, e.g. 'brother'). When a follow request is sent on Friendica, the recipient can accept it as uni-directional (the follower is labelled a 'fan/admirer') or bi-directional, so the recipient also sees the follower's updates. Bi-directional connections on LinkedIn require a reason or more information as 'proof' of a mutual connection, before the request is even sent.

YouTube connects profiles together through subscriptions to channels, however it also explicitly provides input for profile owners to link to other profiles without creating a subscriber relationship. This lets content creators list, for example, other users they admire, or the people they collaborate with. Many YouTubers use this feature to link to other profiles they have on the site. The system gives users free text fields to name this list, as well as each individual link in the list. This particular phenomenon is examined in more detail in the next study, Constructing Online Identity.

OkCupid connections are uni-directional, and only revealed to the recipient if and when a mutual action is made. On Twitter, following another user sometimes (not consistently) appears as an activity in your timeline; notifications are also sometimes sent to your followers to advertise the new connection.

#### 2.4.3. How are profiles updated?

Profiles may be updated by profile owners via a system's user interface, programmatically through an API (Application Programming Interface; the means through which data can be read or written by third-party software). The latter is relevant because programmatic access suggests that third-party applications (outside of direct control the system itself) can also influence a profile owner's view on the possibilities of the profile.

Most systems provide a **Web form** to add or update *attributes*, or a similar UI in a native mobile application. The editing interface and the profile display may be tightly coupled (Twitter, Quora, LinkedIn, ResearchGate) completely divorced, or a combination (Facebook, OkCupid). Indieweb profiles are updated with custom editing interfaces, or simply by editing static HTML; there are currently no specific recommendations for protocols or UIs to edit profile attributes.

For the non-attribute data which makes up a profile, separate, often **specialised interfaces** for both Web and mobile exist, e.g. for posting status updates or media content. For data like statistics and activities, this content is generated by **algorithms or sensors**, with no explicit input from the profile owner. In a few cases it may be hidden by the profile owner, but rarely changed. An exception is RunKeeper, where one can edit an automatically generated GPS trace after the fact, which can correct distance and speed records. On CouchSurfing, AirBnB and PeoplePerHour, one may respond to a review left by someone else, but not remove it.

Only Pump.io, RunKeeper and GitHub provide **APIs** to update all *attributes* of a profile. Facebook and Zooniverse provide limited update access to certain attributes. Most systems provide write APIs to create, follow and like (or equivalent) non-attribute content.

#### 2.4.4. How are people notified when a profile is updated?

The attention a system draws to profile updates could affect how people engage with their own profiles. When profile attributes are updated by the profile owner, most systems **do not notify** other users of the system at all.

Facebook however pushes updates to friends' timelines along with status updates and content interactions, though the extent to which it does this for each friend depends on their **arbitrary** content distribution algorithm, and from a user perspective is hard to predict. The most reliably seen attribute updates are changes to profile pictures, cover photos, and relationship status. Whenever the profile owner updates an attribute on Facebook, they are asked to make it a 'story', which sustains a reference to the fact the attribute changed. Friendica notifies about changes to profile pictures only.

OkCupid and LinkedIn provide the **option to enable sharing of changes** to profile attributes. In the case of LinkedIn, updates are pushed to contacts' feeds, but may also be displayed to non-immediate contacts in the network as a form of promoting connections. OkCupid may display updates to other users in their activity feeds according to whether the system thinks these people might be interested in your profile. How either of these are decided is opaque to the user.

#### *2.4.5. How is access to a profile controlled?*

Systems may provide all-or-nothing access to profiles, make everything public but all optional, provide access control on the basis of groups or networks, or individual users, and provide granular access to individual aspects of profiles.

Systems which have limited or **no access control**, but make all or most data **optional** to enter include OkCupid, Quora, CouchSurfing, AirBnB, Friendica, Zooniverse, Pump.io and GitHub. OkCupid and CouchSurfing allow profile visibility to be restricted to other logged-in users. CouchSurfing permits users to hide their full name, and GitHub permits users to hide their email address.

Quora permits users to answer or ask questions as '**anonymous**' whilst logged into their account. These questions/answers do not show up on the user's profile. Otherwise, the only other control profile owners have is disabling their online presence. Friendica permits connections to be hidden, as well as certain aspects of content. On AirBnB, profile attributes are optional but *hosts* can automatically decline users who omit certain attributes.

Systems with more **granular concepts of audience** than public/private include Pump.io, LinkedIn, Facebook, Twitter, and ResearchGate. In Pump.io individual objects can be 'addressed' so that only particular groups (which can be created by the profile owner) or individuals can see them. LinkedIn permits visibility of some individual profile attributes to 'everyone', 'my network' and 'my connections'. The profile can be set to publicly visible, with certain attributes individually excluded. Connections can be private or public, and content and interactions can be designated different levels of visibility from entirely private to entirely public, with 'network' and 'connections' in between. ResearchGate enables hiding certain statistics, certain attributes, and certain content. Uploaded papers can be visible to 'everyone', 'mutual followers' or 'ResearchGate members'.

Twitter allows users to 'protect' their profiles, which means only those **requesting access** can see content and connections; however, all *attributes* are visible to anyone regardless. Profile owners can block other users, preventing them from seeing everything but their name, display picture and profile banner.

Systems with granular access control across several different aspects of the profile include YouTube, Facebook, RunKeeper and ResearchGate. YouTube provides granular access controls for various attributes, interactions, links to content, some statistics (like number of subscriptions) and content. RunKeeper attributes can be assigned levels of visibility individually ('everyone', 'friends', 'just me').

Facebook has complex granular access controls, including individual attributes, content, interactions, connections and links. Defaults can be set, as well as updated on a per-object basis at the time of posting/creating. Content can be restricted to include or exclude individuals, groups, particular networks. Read and write access controls are distinct; that is, one can create a post that is publicly readable, but comments on that post may be restricted or disabled completely.

Tumblr's use of 'primary' and 'secondary' blogs is interesting; where a blog constitutes a profile, users can essentially have as many profiles as they want attached to one login. Primary blogs (one per login) are always public, but secondary blogs (unlimited) can be **password protected**. There are no automatic **links between** a user's primary blog and secondary ones, including through the API. There is also no way to tell if a particular profile is primary or secondary, or the account to which a secondary blog is attached. Secondary blog

owners may also grant *write* access to other system users, enabling multi-user profiles. Blocking users prevents the blocked user from interacting with or seeing content.

#### *2.4.6. How can profiles be exported from or imported into a system?*

In the Indieweb model of profile ownership, all data is assumed to be on a server controlled, or at least trusted, by the profile owner. As such, they can move it however they please. Similarly, Pump.io and Friendica are open source software platforms which allow people to either opt to use an instance on a server they trust, or install their own instance for complete control. They both use the standard ActivityStreams 1.0 data model [12] (Friendica has extensions); while Friendica provides **import/export** functionality in the UI, Pump.io doesn't, however the database or JSON feed is compatible across instances.

Twitter, Facebook, YouTube, LinkedIn, ResearchGate, and RunKeeper provide a **download link** for an archive of content. In most cases these are a snapshot of current profile attributes, without a history of changes, except for Facebook, which provides a comprehensive activity log. All exports are proprietary schema in JSON, HTML or CSV.

StackOverflow profiles are **reusable** across different StackExchange sites; there is no export, however there are public dumps of Q&A data. GitHub data is available through an **API**.

Tumblr, CouchSurfing, Quora, OkCupid, PeoplePerHour, AirBnB and Zooniverse provide neither an export nor an API to access all profile data.

#### *2.4.7. What constraints are placed on a profile?*

In this section I examine the terms of service of systems to determine how users are *expected* to engage. In some cases these are enforced by technical constraints.

Twitter, CouchSurfing, Facebook, OkCupid, LinkedIn, PeoplePerHour, AirBnB and GitHub state that a user **may not have multiple** accounts. Twitter qualifies this with "overlapping use cases".

Tumblr users cannot create two primary blogs with the **same email address**, and can create 10 secondary blogs per day on the same login with no overall limit. Secondary blogs are somewhat constrained in their functionality compared to primary blogs.

Couchsurfing, Facebook, Quora, StackOverflow, LinkedIn, PeoplePerHour, AirBnB, GitHub and RunKeeper explicitly **disallow 'fake'** profiles; the profile owner must be a single 'real' person, and **not be impersonating** someone else.

#### *2.4.8. What is the data model of a profile?*

To answer this question, I have examined wording in systems' documentation around profiles, in user interfaces as well as APIs. Where possible, I have also looked at internal data models of the software.

Accounts and people are roughly equivalent for Twitter, Indieweb, Pump.io, LinkedIn, Facebook, Quora, PeoplePerHour, ResearchGate, OkCupid, AirBnB, Zooniverse, RunKeeper, and GitHub profiles. That is, a profile sufficiently **identifies a person**; for example the "name" attribute of a profile is the name of the profile owner (rather than the name of the profile). Activities associated with these profiles (e.g. "distance ran" or "commit made") are assumed to have been carried out by the profile owner.

Tumblr and YouTube equate an account - or username/password combination - with a person, but each account may be attached to multiple profiles: secondary blogs in the case of Tumblr, channels in the case of YouTube. Profile owners can carry out interactions from behind one of these profiles at a time.

Friendica permits a user of one account to create multiple profiles with different attributes, and set up access control so that certain people see a particular profile. Different profiles are different '**views' on one person**'. Profile owners can also assign a 'type' to their profile which automatically sets some defaults for privacy and access control settings.

#### *2.4.9. What is the profile for?*

This question looks at the purpose of the profile within the system, rather than any purpose of the system itself, though the two may be similar.

In Twitter, Tumblr, YouTube, Quora, StackOverflow, Indieweb, ResearchGate, Zooniverse, RunKeeper and Github, profiles serve as a central hub for **aggregation of content** by the profile owner. In the cases of Twitter, Tumblr, Pump.io, CouchSurfing, Facebook, LinkedIn and Friendica, a profile serves as an **endpoint for connections** and relationships within networks where connections are important.

In systems with high levels of interaction and often some concern about trust or reputation, profiles provide a **face behind content** so that statements may be evaluated against the backdrop of 'who said it' (e.g. Twitter, Tumblr, Pump.io, YouTube, CouchSurfing, Facebook, Quora, StackOverflow, ResearchGate, Friendica, Github). Systems which are particularly geared towards building trust or **reputation** as a foundation for future relationships and interactions within the system are CouchSurfing, Quora, AirBnB, OkCupid, StackOverflow, LinkedIn, PeoplePerHour, ResearchGate and Zooniverse.

Profiles which are geared particularly towards **self-expression**, or establishing a presence, are Indieweb, Twitter, Tumblr, Facebook, Pump.io, YouTube and Friendica.

#### *2.4.10. Who is the profile for?*

Often who a profile is intended for is related to the profile's purpose within the system. In some cases, the **audience is known** (e.g. you know who follows you on Twitter; Tumblr, Pump.io, YouTube, Facebook, Quora, LinkedIn, Friendica, RunKeeper, Github) and in other cases **imagined** (you have an idea of who OkCupid might be promoting your profile too, but no sure evidence; the same for CouchSurfing, StackOverflow, Indieweb, PeoplePerHour, ResearchGate, Zooniverse) and in some cases both (your Twitter profile is public, so people who aren't your followers will see it; also similar for Tumblr, Pump.io, Facebook, Quora, LinkedIn, Friendica, RunKeeper, Github).

In cases where a profile is constituted of an aggregation of personal data, content, and online interactions, the **profile owner** is a member of the audience, as they can use it for self-reflection or self-expression (Twitter, Tumblr, Pump.io, YouTube, Facebook, Quora, OkCupid, Indieweb, StackOverflow, LinkedIn, ResearchGate, Friendica, RunKeeper, Github).

Systems like Quora, CouchSurfing, OkCupid, StackOverflow, LinkedIn, PeoplePerHour, Twitter, Facebook, YouTube and AirBnb use data from user profiles as input to core **algorithms** which enable the system to function,

providing a service to profile owners.

Similarly, systems such as Twitter, Facebook, Tumblr, YouTube, CouchSurfing, LinkedIn and RunKeeper use profiles as input to algorithms which sustain the companies behind the systems, for example through selling data to **third-parties** like advertisers.

## 2.5. Features

From this analysis, five features of profiles were derived and are described below, and summarised in Table 3.

Table 3. Profiles study features

| Feature        | Strongly applies (1)   | Does not apply (0)  |
|----------------|--|---|
| Flexibility    | Profile owners have choice about the kinds of content associated with their profile and how it is presented. | Profiles are generated as a side effect of owner's activities or automatically (e.g. from sensor data) and owners cannot amend. |
| Access control | Profile owners have control over which parts of the profile others see.                                      | Profile owners have no control over what others see.  |
| Prominence     | Profiles are integral to functioning of the system as a whole.   | Profiles are a side-effect of some other function of the system, and/or not necessary to use the system.                        |
| Portability    | Profile owners can move their data in or out of a system.  | Profile data cannot be imported or exported.  |
| Representation | The profile <i>is</i> a person, as far as the system is concerned.   | The profile is a document describing some aspect of a person(a).  |

**Flexibility** is a function of the different types of content/data which make up a profile, and the relationship the profile owner has with those who see or use their profile. As some times of content are under more control of the profile owner than others, we consider the proportion to which they make up the profile, and weighting given to each. Flexibility also considers the systems technical or policy constraints around profile contents.

**Access control** involves the granularity of the controls, the extent to which profile owners can opt into or out of publishing certain aspects, and the awareness of the owner of who their audience is.

**Prominence** takes into account the extent to which a system would function were users' data (of the various kinds) not aggregated into profiles. Prominence of profiles may depend on the role a user is playing in the system, so the

potential varying roles are also taken into account. Systems with a high emphasis on connecting people feature profiles prominently, whilst systems with lots of interactions but little need for reputation do not necessarily require consistent profiles to be useful.

**Portability** considers how easy it is to get profile data out of a system, as well as how reusable that data is in other systems. This includes whether data is exported into a known standard data model, and standard file format, and the extent of additional processing that may be required to port it elsewhere.

**Representation** connects the systems' model of users with its purpose. Systems with the possibility or expectation of personas or partial representations of individuals are not considered representative, whilst systems with emphasis on 'real people' and one-to-one mappings between profiles and profile owners have high representation. Systems in which the real-life human is required for legal or transactional purposes (e.g. to make a payment or provide a service) make a distinction between the profile and the person, and this lowers representation.

An overview of the questions which contributed to the derivation of each feature is in this table and the rankings of each system against each feature are in the following table.

Table 4. Profile questions and features relation

| Feature        | Questions |   |   |   |   |   |   |   |   |    |
|----------------|-----------|---|---|---|---|---|---|---|---|----|
|                | 1         | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Flexibility    | X         | X | X |   | X |   | X |   | X |    |
| Access Control | X         |   | X | X |   |   |   |   | X |    |
| Prominence     |           |   | X |   | X |   |   | X | X | X  |
| Portability    |           |   |   |   |   | X |   | X |   |    |
| Representation |           |   |   |   |   |   | X | X | X | X  |

Table 5. Profiles study results: features of systems

| System        | Flexibility | Access Control | Prominence | Portability | Representation |
|---------------|-------------|----------------|------------|-------------|----------------|
| AirBnB        | 0.2         | 0.2            | 1.0        | 0.0         | 0.9            |
| CouchSurfing  | 0.5         | 0.1            | 1.0        | 0.0         | 0.9            |
| Facebook      | 0.3         | 0.9            | 0.7        | 0.5         | 0.8            |
| Friendica     | 0.8         | 0.8            | 0.7        | 1.0         | 0.2            |
| Github        | 0.2         | 0.1            | 0.3        | 0.4         | 0.8            |
| Indieweb wiki | 1.0         | 0.0            | 0.5        | 1.0         | 1.0            |
| LinkedIn      | 0.3         | 0.8            | 0.8        | 0.5         | 0.9            |
| OkCupid       | 0.5         | 0.1            | 1.0        | 0.0         | 0.7            |
| PeoplePerHour | 0.5         | 0.1            | 1.0        | 0.0         | 0.9            |
| Pump.io       | 0.8         | 0.2            | 0.7        | 1.0         | 0.5            |
| Quora         | 0.5         | 0.2            | 0.3        | 0.0         | 0.9            |
| ResearchGate  | 0.6         | 0.4            | 0.3        | 0.8         | 0.9            |
| RunKeeper     | 0.7         | 0.5            | 0.1        | 0.5         | 0.4            |
| StackOverflow | 0.5         | 0.1            | 0.3        | 0.3         | 0.9            |
| Tumblr        | 0.8         | 0.7            | 0.2        | 0.0         | 0.1            |
| Twitter       | 0.6         | 0.5            | 0.2        | 0.4         | 0.5            |
| YouTube       | 0.7         | 0.7            | 0.2        | 0.7         | 0.5            |
| Zooniverse    | 0.5         | 0.1            | 0.1        | 0.0         | 0.8            |

## 2.6. Discussion

Five features of online profiles were derived from observations of the functionality and uses of a set of existing social systems. We can use these features to cluster similar systems and give us a better understanding of online profiles in the social web ecosystem today. In this section I discuss some noticeable clusters. When I use 'highly' in reference to a score, I mean the score was greater than 0.5.

Though much of the literature around studying user profiles only acknowledges *attributes* [132, 70] we can see that profiles are constituted of much more than just descriptive attributes about an individual. Content that makes up a person's profile may be input directly by the profile owner, generated or inferred from their online or offline activities, combined with content of others in the system and/or generated directly by other users of the system. Different systems emphasise different aspects of a person's online presence and allow users to adjust this to varying degrees.

The features which enable greatest control over self-representation for users are flexibility, portability and access control. Flexibility means that users have freedom to choose which information and contents make up their profile; portability means that they can move this data around or repurpose it easily; and access control means that the profile owner can choose who sees what. These things in combination are particularly empowering. Thus, the systems which give users the greatest control are Friendica and YouTube, which score highly for all three, and Tumblr, which scores highly for flexibility and access control. To a lesser degree, Pump.io, Indieweb and ResearchGate score highly for flexibility and portability, but with limited access control. This means that profile owners must employ strategies of omission or self-censorship to effectively manage what their audience sees. Facebook and LinkedIn on the other hand score very highly for access control, but lower for flexibility and portability; that is, you don't have much control over how your profile is constructed, but at least you can control who sees the information.

Systems with high prominence scores tend to also have high representation scores. However Friendica has a very high score for prominence, as profiles are crucial in a network where making connections is the end goal, but it has a low score for representation, as the expectation is that profile owners present *personas*, and may have more than one for different aspects of themselves. The high-prominence and high-representation systems (CouchSurfing, PeoplePerHour, AirBnB, OkCupid, Facebook, LinkedIn) have strong ties to 'real life', for example in-person meetings, employment, or service exchange.

Low prominence systems are geared towards an end purpose that is not oriented around user profiles, such as content creation, collaborative projects or information aggregation (Zooniverse, YouTube, Twitter, Quora, StackOverflow, ResearchGate, Github, Tumblr, RunKeeper). Profiles are useful, but not an end in themselves. On top of being low prominence, Tumblr and RunKeeper are not very representative; Tumblr permits multiple profiles and the community generally expects anonymity or pseudonymity; RunKeeper contains a very small subset of information about a person. Zooniverse, StackOverflow, Quora and Github nonetheless score relatively highly for representation, since unique profiles for individuals is necessary for establishing reputation or standing, a key element in these communities.

To be able to classify systems according to these features it is necessary to consider multiple perspectives: those of the profile owner, others who will see the profile, and the organisation which runs the system itself. As such, the

classification process gives a holistic view of a system, but only at a surface level. It misses out on the finer details of how the system is situated in the context of a society, how profile owners use one system alongside others, and the multiple possible uses of a system by different people, or different roles people may play. Nonetheless this provides a baseline idea of how people *could* use a system, in order to carry out more detailed studies about how individuals actually *do* use a system.

In particular, in future studies of users of a particular system, researchers can refer back to the features of the system (perhaps scoring systems which have not been covered here, or updating scores for ones which have changed) in order to put the users' actions in the bigger picture.

Throughout the remaining studies in this chapter, where specific systems are highlighted, I refer back to these features.

### *2.6.1. Contributions to the 5Cs*

Different systems require different levels of engagement with one's own profile. The prominence of a profile within a system, as well as how representative a profile is (or should be according to system rules) of its owner indicate that individuals may have different levels of **control** over their self-presentation. Relatedly, if one can take all of one's data out of a system and even move it elsewhere (portability), this may influence decisions about persisting or maintaining profiles.

Systems may be flexible about what data appears in a profile, how that data is presented, and how it is accessed by other users. I consider both of these features to contribute towards the **customisability** of self-presentation.

Access control and flexibility both indicate an awareness of the profile owner's audience. These, along with the prominence of a profile within a system, indicate that we must pay attention to the links between participants within a system, or the **connectivity**.

Users of systems are affected by both technical and policy constraints in terms of flexibility and portability of their profiles. The purpose of the system itself also influences the prominence and representation of profiles. These outside constraints and goals constitute the **context** formed by a system, as well as being influenced by the overall context in which a system exists (eg. legal frameworks, business interests).

Representation and access control together can drive or inhibit linkability between profiles in different systems, and offline identities. The spread and aggregation of information about an individual, possibly without their knowledge or consent, is part of the **cascade** of information beyond where it originated.

### 3. Constructing online identity

In the previous study we took a high level look at 18 social systems; in this study, we zoom in on one of them — YouTube. According to the previous study, YouTube channels are relatively flexible, access controlled, and portable, but not very representative, and even less prominent. Users participate in different roles on YouTube, from passive, possibly anonymous consumption, to engaged consumption with comments, interactions and curating playlists, to active content creation. The latter group also vary the level to which they participate; some users spontaneously or casually post videos for a small localised audience; some engage across multiple channels, manage branding, collaborate, nurture a fanbase, and create videos on a professional level.

The high flexibility and low prominence of YouTube profiles gives users a chance to be creative when expressing their identities. The following study empirically examines some different ways identities are expressed through YouTube channels, including a closer look at the affordances of the system and how individuals work within and outside of these.

Whilst YouTube is at the core of the online presences of the subjects of this study, their activities span a variety of other systems, not wanting to fall into the trap of imagining a system exists in isolation, I discuss these as well.

I identify four concepts that are useful for understanding individuals in a system with flexible self-presentation opportunities: *roles*, *attribution*, *accountability*, *traceability*.

This section has been adapted from work published as *Constructed Identity and Social Machines: A Case Study in Creative Media Production* (2014, Proceedings of WWW, Seoul).

#### 3.1. Introduction

In chapter 2 I described existing work in understanding socio-technical systems as *social machines*. Due to the complex nature of online identity, understanding nuanced identity behaviours of social machine participants in a more granular way is crucial. First I will briefly describe *creative media production social machines*, then present the results of a study of profiles portrayed by

participants in one of these. The contribution is a set of dimensions along which a social machine can be classified in order to better understand human participants as individuals, as opposed to participants in aggregate.

Amongst the plethora of user-generated content on the web are a huge number of works of creative media, and behind these are independent content creators pushing their work to a global audience and actively seeking to further their reach. Within this ecosystem we can see creative media production social machines on a variety of different scales. The definition of creative media production social machines encompasses a class of systems where:

- humans may use a purely digital, or combination of digital and analogue methods, and a degree of creative effort, to produce media content;
- the content is published to be publicly accessible on the web;
- a global audience may consume, curate and comment on this content in technologically-mediated environments.

These social machines exist both within and across content host platforms (e.g. YouTube) and within and across online communities and social networks. Many, if not all, media types and genres are represented among the media artefacts that emerge from these systems, and the content and the reception it receives can have a sometimes profound effect on media and culture in the offline world.

Figure 2 shows the interconnected social and technical systems engaged when a simple vlog (video blog) is uploaded to YouTube. These processes would be further expanded if the creator was to branch out and produce different types of content, collaborate with another creator, cross-publicise, share audiences or even co-own a YouTube channel or other website profile.

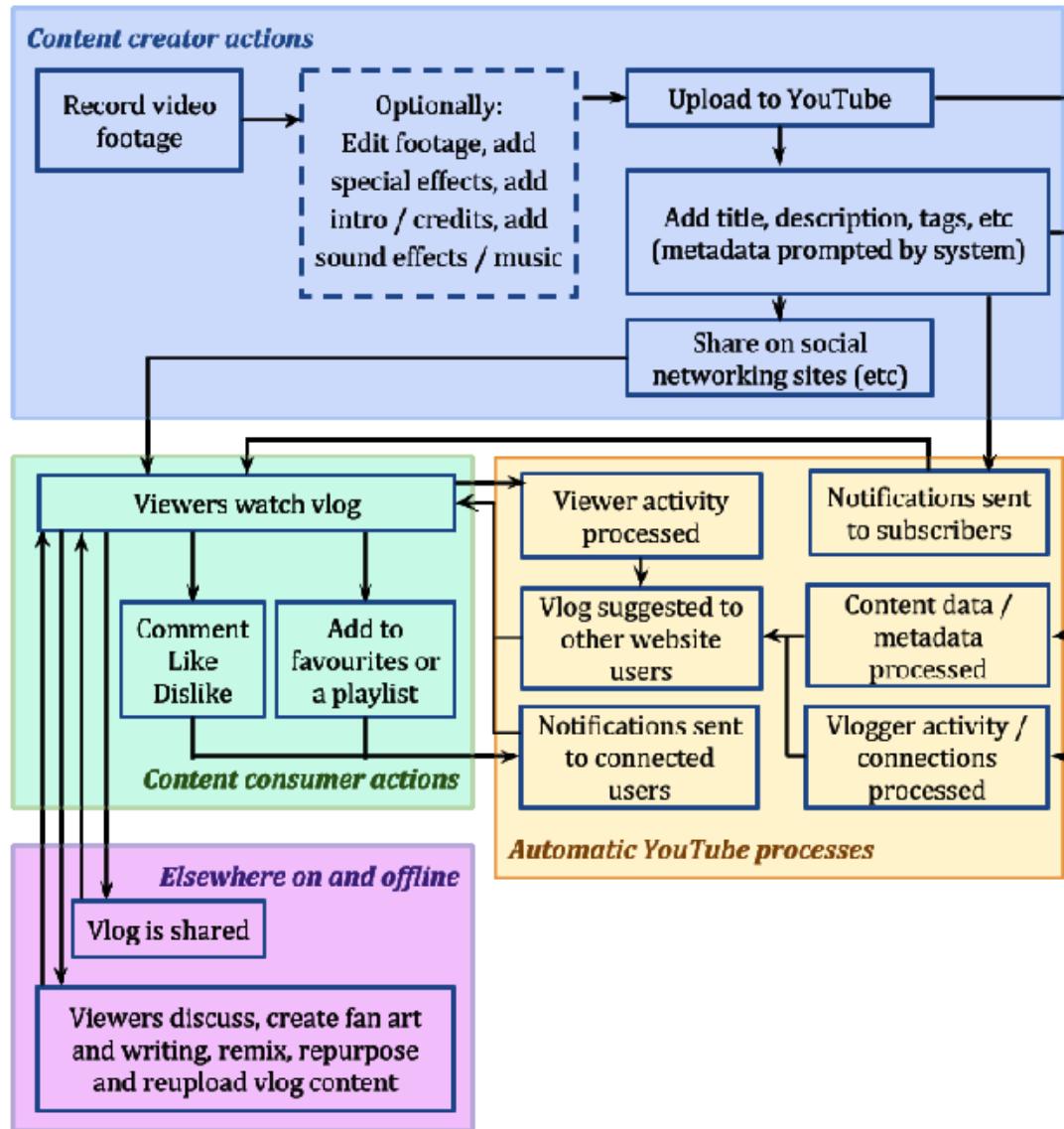


Figure 2. Interconnected social and technical systems necessary for publishing a vlog on YouTube.

Creative media production social machines create an environment in which content creators of all backgrounds and abilities are able to publish outside the constraints of traditional media channels. These creators are actively vying for attention from massive audiences; competing for views, likes and shares on a global scale. How they present themselves to their audience can be critical to their success, but also a ground for playful experimentation.

### Motivations for participation

It is worth noting that there are a variety of motivations or incentives for content creators to participate in creative media production social machines.

Some content host sites provide direct financial incentive for popularity (e.g. YouTube's Google Adsense). Others facilitate a commission based model, where creators show off their work and take paid requests for custom pieces from the community (e.g. DeviantArt). For content creators who publish primarily on such systems, their activity on other systems is usually tied to driving traffic back to the content which makes them money, or entertaining the fanbase from whom they thrive (e.g. a creator who publishes sketch comedy on YouTube might use their Twitter account to tell original jokes to maintain interest between video releases).

But for many content creators, the financial rewards from their chosen content host sites might be a convenient side-effect of doing something that they love. Reputation as a creator of high quality content, as a talented artist or as a particularly funny comedian might be their primary driver. There are also social cues in many communities that affect content creator behaviour. Sometimes creators don't want to be accused of 'pandering' to their audience or losing their artistic integrity, and regulate their behaviour accordingly.

The visibility of quantitative data collected by a content host site – such as how many views a piece of content has, how often a participant is referred to as a co-creator, or how often a participant responds to viewer comments – may also impact behaviour. Technical factors are often highly conflated with the social norms in a community.

Thus, the core reasons for creating content can affect both the content created and how creators present themselves to their audience in the process.

### 3.2. Context and research questions

To recap some background from chapter 2, the nature of identity and anonymity in online spaces is well discussed [90, 133, 231, 234]. Humans naturally adjust the way they present themselves according to the context, and different online spaces may afford different levels of flexibility in doing this. Systems which don't require any kind of registration to post content, allow people to adopt and discard personas as needed, and to create social cues to identify each other that are not designed as part of the system [26]. Entirely different behaviour occurs in systems that strongly encourage or even try to enforce usage of real names.

Often it is trivial for people to create multiple accounts under different pseudonyms anyway, but there may be an increased expectation of honesty from other users of the system, which itself affects the culture of communities within.

In many cases the fact that people present themselves differently in different contexts is unconscious; a side effect of their participation in a particular system according to the social norms or even technical affordances (e.g. their desired username may be unavailable resulting in the forging of new branding around an alternative). In other cases, the creation of alternative personas is engineered and deliberate, either from the outset or as something that has evolved over time. Multiple individuals may also participate in the portrayal of a *single* persona [71] and one individual may present versions of themselves through *multiple* personas.

The public profiles of content creators were examined with the following questions in mind:

1. How do content creators present themselves within and across communities?
2. To what extent are content creators' online presences consistent across platforms, and how is their content distributed across different online presences?
3. How, and to what extent, do content creators present connections between their own online presences?

To add depth, I also take note of their audience, the type of content they create, and the capabilities of the platforms on which they publish their content.

### 3.3. Study design

This is an in-depth empirical study in which publicly visible data about individual social media users are analysed. The data includes content created by the subjects, attributes from their profiles, and links between profiles. We use only human-led, in-browser exploration of the profiles, and employ no scripts or API access to gather data.

#### 3.3.1. Method

I first familiarised myself with the different ways of updating and modifying the data that appears on a YouTube profile (also known as a channel), so I could understand the actions that profile owners had to undertake to build their

presence on YouTube.

The starting point for data collection was a particular YouTube channel per subject. The different types of profile information that were present were noted. Links from the profile content were gathered, and ones which were determined to connect to other profiles, within and outside of YouTube, were followed. The information on these profiles was similarly logged. I collected:

- The types of profile data visible.
- The number of inbound and outbound connections to other profiles.
- What kinds of other profiles belonging to the channel owner were linked to from a YouTube channel.
- How these links were labelled or described.
- How the data on these additional profiles differed from or overlapped with each other.

### *3.3.2. Subjects*

Ten content creators were selected from a subset of creators with whose content I have a passing familiarity through encountering it online over prior months to years. This resulted in a broad spectrum of content types (video, animation, music, art, written word) genres (comedy, game commentaries, educational, political), popularity, well-knownness and activity levels. I deliberately examined content creator profiles from the perspective of a content consumer, or casual audience member. Thus, for the purposes of this study, we do not have access to deeper insight about the personas beyond what is accessible publicly through the web. To identify each subject for the remainder of this study I use short non-anonymised nicknames.

### *3.3.3. Limitations*

The results are based upon a very small (albeit diverse) sample, and cannot be considered representative of content creators in general. I seek to describe a subset of behaviours within content creation social machines, but do not claim to be exhaustive.

I have no doubt that content creators have more online profiles which are *not* linked from their YouTube channels, however I was obviously not able to discover and study these.

### 3.4. Results

#### 3.4.1. Profiles and personas

For ten content creators, 93 profiles were discovered. Of these, 23 were YouTube channels, 16 Twitter profiles, 13 Facebook, 9 Vimeo, 7 Tumblr, 6 personal websites, 5 Instagram and 4 Vine, 3 Google Plus, 2 Bandcamp and 2 DeviantArt and 1 each of Patreon, FormSpring, BlipTV, and Newgrounds. Table [6] shows how the profiles are distributed. As we can see, in the domain of creative content production identities are not site- or community-specific. Creators spread their activities across a number of networks in order to shape a more complete identity.

Table 6. Content creator subjects: Distribution of profiles for content creators in the study.

| <b>Creator</b> | <b># profiles</b> | <b>Mean profiles per site</b> |
|----------------|-------------------|-------------------------------|
| Dane           | 18                | 2.3                           |
| Khyan          | 13                | 1.9                           |
| Bing           | 13                | 1.3                           |
| Lucas          | 11                | 1.4                           |
| Bown           | 9                 | 1.5                           |
| Todd           | 7                 | 1.2                           |
| Arin           | 7                 | 1.0                           |
| Suzy           | 6                 | 1.2                           |
| Ciaran         | 5                 | 1.3                           |
| Chloe          | 4                 | 1.0                           |

'Second channels' are common on YouTube. Creators who focus on one type of content (e.g. sketch comedy) publish this on their main channel as well as using their main channel identity for interactions on the site. On their second channel they publish content that they may consider to be of interest to only a part of their main audience, such as vlogs about their lives, out-takes from main channel content, or experimental pieces. Most content creators with second channels post explicit links to them on their main channel, and often publicise them within content metadata or as part of the content directly. In some cases, including those where the connection between two channels is explicit and obvious, the creators behave differently towards their audience through second channel content. This varies greatly depending on the type of content produced. In some cases, second channels may be perceived to be more reflective of the creator's 'true' personality, if they project themselves as more serious or honest, and publish more personal content like vlogs or behind-the-scenes footage.

Whether or not this is accurate is impossible to know without intimate knowledge of the creators' offline life. The significance is that persona variations exist, and creators do not necessarily hide these alternative presentations of themselves from their audience.

Additionally, there are profiles which are *not* directly linked from the (self-identified) 'main' profile, or the links are treated as though the profile belongs to a different person. Figure 3 shows three screenshots of different YouTube channels showing different ways creators link out to other versions of themselves.

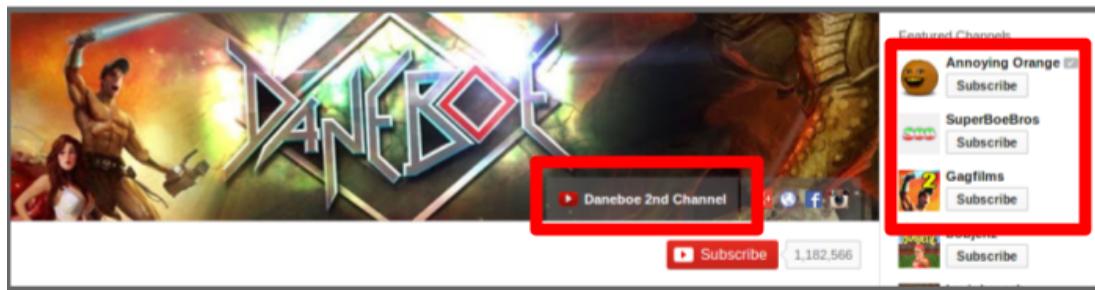


Figure 3. 1. Dane transparently links to 3 of his channels, two for alternative content types and one for a character he created.

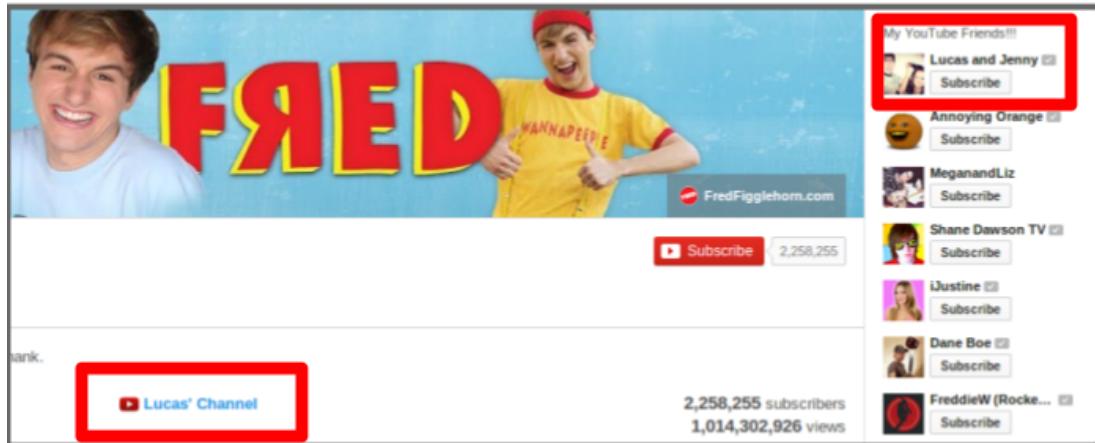


Figure 3. 2. Fred is a character played by Lucas, but the links on Fred's channel treat Lucas as a different person.

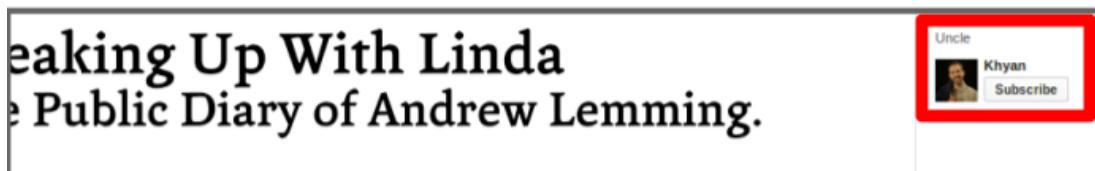


Figure 3. 3. Andrew Lemming lists Khyan as "Uncle", although Khyan is the creator of the Andrew Lemming character.

Creators also used their profiles to link to shared channels (where either multiple creators post content independently of each other, or creators collaborate to produce joint content, or a both), and channels of others with whom they regularly work.

Most of the platforms discovered which host profiles for the subjects of this study offer limited options for customisation, and the use of consistent branding between different systems was intermittent. This mostly took the form of

identical or similarly styled display pictures, similarly phrased introductory paragraphs, and similarly styled content.

Some creators have profile sets across different platforms which are distinctly grouped into alternative personas. This was evident from the branding, content and connections between them.

### 3.4.2. Connections

How connections to other people were represented varied depending on the technical system. We can differentiate between mutual relationships between accounts (e.g. 'friend') and one-directional relationships (e.g. 'follower', 'subscriber'). Some systems offer both types of relationship, some one or the other. For YouTube channels, popularity ranged from over 3.5 million subscribers for Dane's character channel *realannoyingorange* to 118 for Bown's secondary *bowntalks* channel.

The importance of these connections varies depending on the system as well as on the attitude of the system user. Mutual connections may initially be presumed to indicate a closer relationship, but this is not always the case. Some systems allow users to accept all friend requests en masse, which they may do to please fans, resulting in a lot of essentially meaningless mutual connections. Instead, outbound one-directional connections come in far smaller numbers, and indicate the content creator is particularly interested in the outputs of the other creators they choose to follow. It appears normal for content creators to follow other creators with whom they have collaborated.

Although their use is to some degree shaped by community norms, such connections are strongly influenced by the architecture of the particular website. However, most of the websites examined allow enough control over the textual content of a profile that profile owners can manually create links to other documents on the web, potentially circumventing the site's built-in connection mechanisms. Creators may also be able to adapt the content publishing interfaces to add additional connections (e.g. adding links to Twitter and Facebook accounts in the description of a YouTube video), and often do. These connections necessitate extra effort on the part of the content creator, and tell us more about their relationships with other online accounts. Figure 4 shows different types of connections between profiles and personas for one content creator.

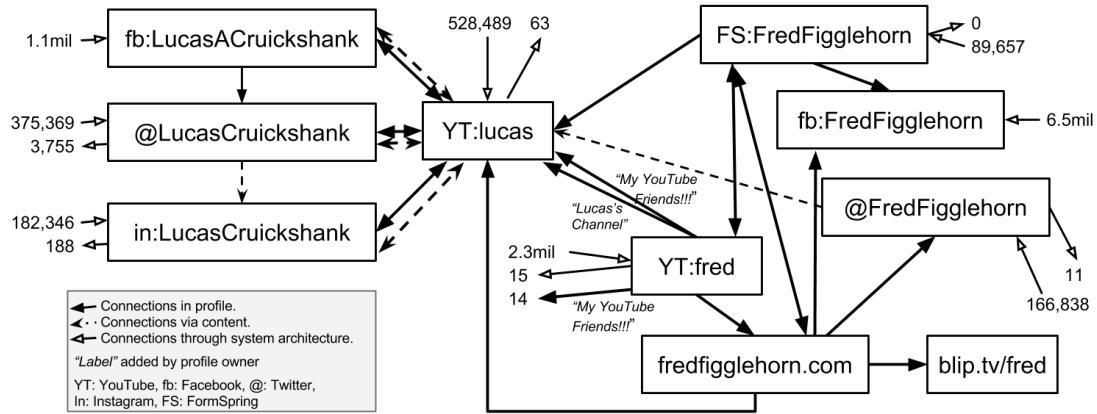


Figure 4. Lucas Cruickshank was an early YouTube success story through his persona Fred Figglehorn. Here, different types of connections between various online accounts belonging to both Lucas and Fred are illustrated.

### 3.4.3. Summary

Content creators at all levels of activity do not have straightforward relationships with the systems they use for publishing and publicising their content. Through manually examining profiles, it is possible to identify personas, and connections between creators, and learn about the likely explanations behind them. Currently there is no way to formalise these deductions, so in the next section I propose a small taxonomy for describing the experiences of individual participants in social machines.

## 3.5. Taxonomy

Based on the findings previously described, I propose four closely linked but distinct concepts that are useful in a granular discussion of identities of social machine participants: roles, attribution, accountability and traceability. I will explain each in the context of creative media production social machines, and show how they can be used as dimensions to assess the nature of individual identity in a social machine.

Table 7. Summary of taxonomy

| <b>Dimension</b> | <b>Description</b>  | <b>Degree</b>                    |  |   |
|------------------|---|----------------------------------|--|---|
|                  |   | <b>0</b>                         | <b>0.5</b>                                     | <b>1</b>  |
| Roles            | the ease with which participants can change the role they play in a system  | one role,<br>everyone<br>equal   | multiple<br>roles,<br>participants<br>play one | multiple<br>roles,<br>participants<br>move<br>between<br>them |
| Attribution      | whether or not crediting participant contributions is important   | unimportant                      | sometimes<br>important                         | very<br>important   |
| Accountability   | whether the provenance of the inputs make a difference. In a Social Machine where this is critical, regulating identities to ensure trustworthy data would make sense | unimportant                      | sometimes<br>important                         | very<br>important   |
| Traceability     | the transparency or discoverability of connections between different profiles and personas  | required, or<br>mostly<br>useful | optional,<br>may be<br>useful or<br>harmful    | not<br>required, or<br>likely<br>harmful                      |

## Roles

A creative media production social machine contains at least consumers, commentators, curators, and creators [183]. These roles are interchangeable, and content creators may wish to adopt different personas according to the role they are playing. Plus, content creators are often multi-talented and they may wish to put on a different face according to the different types of content they publish. How easily this is accomplished - according to the social expectations and technical affordances of systems that are part of a social machine - can impact the behaviour of participants.

## Attribution

In content creation communities, contributions to media output are directly connected to building reputation, so content creators generally desire to have their name attached to work they produce. If the publication system does not allow this directly, as is often the case for sites that host collaborative works (a video published on one YouTube channel may contain contributions from several creators, each with their own channels but formally linked with only the uploader), then creators adapt the system as best they can, eg. the uploader may list links to the channels of all contributors in the video description [184]. Even when a content host site provides automatic linking to other user profiles -

common in remixing communities – this isn't necessarily enough. [202] finds that human-given credit means more, and so free-text fields for content metadata are often used anyway.

## Accountability

In many of the commonly-discussed social machines, like Wikipedia, Galaxy Zoo, Ushahidi, and the theoretical crime data social machine in [56], accurate data is critical to the usefulness of the output of the system(s). Thus, accountability through identity is important. It is reasonable then to want to regulate participants somehow. But this is not universally applicable.

The production of creative content is a domain that exemplifies the need for taking a more flexible approach to identity understanding and management. On the one hand, creators wish to be accurately credited for their work and plagiarism may even result in a financial or reputational loss. On the other hand, creators may appear under multiple guises, engage in diverse behaviours and make contradictory statements about their participation in a creative work, all in the name of entertainment. Creators may also engage in some activities under an alternative identity in order to avoid any effect on the reputation of their main persona. These are valid uses of the anonymity provided by online spaces – a core feature of the World Wide Web. These activities won't necessarily even result in diminished trust. A content consumer may fully enjoy a series of vlogs, unaware that the vlogger is a character and the life events portrayed are entirely fictional, and be none the worse off for it.

## Traceability

We consider traceability in terms of the settings in which an individual might interact with others. A person participating in a creative media production social machine may exist behind a different persona when participating in a scientific discovery social machine, and yet another in a health and well-being social machine. The discovery that other participants in the health and well-being social machine are aware of their alternate persona in the creative media production social machine may cause them to amend one or both of their personas. If the risk of their multiple identities being 'discovered' is high they may adjust their behaviour accordingly, whether this is ceasing all attempts at 'deception', or taking steps to decrease the overlap of the communities of which they are a part.

Well known content creators often appear at offline events to meet their fans. Those who star in popular live-action video content are recognised in the street. They are interviewed by journalists and contracted to produce viral adverts by marketing companies. Only with careful control of their online persona can they successfully engage in offline interactions like this. A content creator who believably portrays an undesirable character across multiple platforms online may not be considered a candidate for a job in broadcast media thanks to the blurred lines between reality and fiction, online and offline.

In 2017, video game commentator Felix Kjellberg (PewDiePie) lost a lucrative contract with Disney and Google for using racist language in his voiceovers<sup>p</sup>. In 2014 vlogger Alex Day was widely renounced by his *online* community (as well as his record label) because of *offline* allegations of sexual assault and abusive relationships with fans<sup>a</sup>. Different worlds interact; contexts collapse, and the repercussions are felt through them all.

An example in which the traceability of personas was crucial is the DARPA Network Challenge [266], for which participants needed to provide their 'real life' identities to win the cash prizes. Even if they had operated under pseudonyms during the competition, in order to validate their claims they needed to make known these personas and consolidate them with an identity that would allow them to receive the prize money.

Since a YouTube profile is not assumed to be *representative* of a single complete individual, profile owners must find other ways to establish and moderate the relationships between their profiles. How they do this will depend on the roles they play, and their motivations in taking part. Knowledge of others present - audience and colleagues - in the online and offline spaces in which someone spends time may influence how they establish their personas in these spaces. An evolution of these spaces or a change in the individual's circumstances over time may cause them to revise their personas.

<sup>p</sup>PewDiePie: YouTube megastar's N-word outburst sparks developer backlash (The Guardian)

<<https://www.theguardian.com/technology/2017/sep/11/pewdiepie-youtube-racist-developer-campo-santo-backlash-felix-kjellberg>>

<sup>a</sup>Vlogger admits 'manipulative relationships with women' (BBC)

<<http://www.bbc.co.uk/newsbeat/article/26664725/vlogger-admits-manipulative-relationships-with-women>>

, YouTuber 'sorry' for 'manipulative' relationships (BBC)

<<http://www.bbc.co.uk/newsbeat/article/29506320/youtuber-sorry-for-manipulative-relationships>>

### 3.5.1. Applying the taxonomy

We can apply these concepts to social machines in order to understand the significance of individuals' identities within them. We use some well known social machines as examples for each dimension, in Table 8.

Table 8. Applying the framework: Examples of social machines along each dimension.

| Dimension      | Examples                | 0 | 0.5                       | 1                                |
|----------------|-------------------------|---|---------------------------|----------------------------------|
| Roles          | ReCAPTCHA               |   | The Obama campaign        | Creative media production        |
| Attribution    | ReCAPTCHA               |   | Wikipedia                 | Creative media production        |
| Accountability | GalaxyZoo               |   | Creative media production | A crime reporting social machine |
| Traceability   | DARPA network challenge |   | Creative media production | Mental health support forum      |

Refer to Table 7 for descriptions of each dimension, and what the numbers mean for each dimension.

## 3.6. Discussion

I have demonstrated through an empirical study that participants in social machines often have complex relationships with their own self-representation, and with their connections to others in a system. Individuals may have one-to-many or many-to-one relationships with online personas, for a number of different reasons, and with different levels of transparency. This section includes a taxonomy of four dimensions: roles, attribution, accountability and traceability. We can use these to better understand individuals in a social machine in relation to the whole, despite this complexity.

### 3.6.1. Contributions to the 5Cs

The role(s) taken on by an individual are affected by the extent to which an one is able to create and discard identities. Whether participants can be attributed or held accountable for their contributions, and the extent to which one identity can be traced to another, are affected by whether identities are persistent, and whether anonymous contributions are accepted. These are all aspects of the **control** someone has over their online self-presentation.

Roles arise through, and may be enforced by, either the technical affordances of a system, or the social expectations of a community (or both). The role(s) an individual chooses to take on may also be affected by their personal motivations, desires or needs. Thus understanding roles requires us to account for the **context** in which a system is being used.

Through Attribution and traceability we discover the **connectivity** of a system. Participants may see each others' contributions, and may build reputation accordingly and present a particular impression to their audience. This reputation and impression can translate to other technically disconnected systems if identities are transparently linked.

The degree to which connections between identities are traceable affects the spread of information about an individual. Intended or unwitting links between personas contribute towards an automatically generated or inferred aggregate profile. This spread may feed into unknown systems on and offline, and have unforeseen consequences. I label this the **cascade**.

Deliberate traceability may be created between profiles on different systems through consistent visual branding, as well as actual hyperlinks placed in profiles and annotated. This is only possible to the extent that systems permit participants to **customise** their profiles.

## 4. The many dimensions of lying online

In the previous study we see some of the creative ways in which individuals work around constraints of even flexible profiles in order to meet their expressive needs. We learned that misrepresenting one's real-life identity is not necessarily in conflict with the functioning of the system, and may even be a culturally important aspect of participation.

I expand on these observations with a survey of social media users who reveal the ways in which they bend the truth in their online profiles, and why, and how they feel about others doing so. Portraying matters as other than they truly are is an important part of everyday human communication. The survey enquires into ways in which people fabricate, omit or alter the truth online. Many reasons are found, including creative expression, hiding sensitive information, role-playing, and avoiding harassment or discrimination. The results may suggest lying is often used for benign purposes, and conclude that indeed its use may be essential to maintaining a humane online society.

The results are a set of categories which characterise the spectrum of lying and deception practices routinely used online: *system, authenticity, audience, safety, play, convenience*.

This section was adapted from work originally published as *Self Curation, Social Partitioning, Escaping from Prejudice and Harassment: the Many Dimensions of Lying Online* at ACM WebSci 2015 with Max van Kleek, Dave Murray Rust, Daniel Smith and Nigel Shadbolt. I participated equally in the design of the survey, participant recruitment, and coding and analysis of the results.

### 4.1. Introduction

People avoid telling the “full, open, and honest truth” in many situations, whether it involves simply the omission or falsification of information, to more substantial forms of deception and lying. Such behaviours have been shown to amount to, by some accounts, nearly a third of offline interpersonal communications [80,31].

This study is aimed at prolific internet users, who spend a substantial part of their daily lives in social encounters online, therefore likely to engage in the widest variety of such behaviours. We are particularly interested in how such

practices arise or are used differently across contexts, situations, and spaces. We are interested in the intent behind the deception, but we do not examine the moral or ethical dimensions of such practices, as these can be highly subjective and grounded in particular personal philosophies.

As described in the following sections, our analysis found that while there are a wide range of reasons people use deception online, few reasons for doing so are self-described as malicious (or comprised of “dark lies”); in fact, a majority of the reasons involve impression management, conflict avoidance, and in order to fit in to groups.

## 4.2. Context and research questions

As individuals increasingly manage multiple social contexts of growing complexity in their daily lives, techniques are required for navigating the interlocking and often antagonistic demands placed on them. Examination of deceptive practices has shown that they often serve as coping strategies for managing and mitigating these complex social situations. Examples of such reasons include protecting one or another's reputation or identity, to preserve particular relationships or ties, avoiding confrontation, showing solidarity with another, and covering up accidental transgressions, among others [140, 54]. Various background concepts relevant to interpersonal deceptive practices are discussed in chapter 2. Of particular interest are butler lies, to ease social interactions, and subconscious adjustments to self-presentation to remain authentic in context.

As the prevalence of computationally mediated socialisation increases, so does the need to understand the role and use of lying and deception in online interaction, and to uncover the kinds of social tensions and attendant complexities that arise from the new social affordances that the Web provides [143]. People now conduct their interactions and curate their identities across a large number of online spaces whilst attempting to balance their privacy, reputation and roles throughout. Deception is a tool used to cope with this complexity, and a lens through which their difficulties and needs can be observed [140, 54]. We attempt to characterise peoples' online behaviour through the simple question: Why and how do people lie on social media?

## 4.3. Study Design

### 4.3.1. Method

We took a several-step approach to designing the survey questions. First, we looked for precedent in previous surveys (e.g., the *Questionnaire on Academic Excuses* for student lying behaviour [235], elicitation method for daily lying studies [80]).

Second, we iterated on the phrasing of the questions by consulting other colleagues as experts in the process to shape the specific foci and wordings. We initially considered several methods besides survey, including semi-structured interviews, and artefact examinations, but fell back to a web-based survey to be the most appropriate for getting a wide sample from our target population. In order to characterise the broad class of behaviours we wanted to examine, we first showed a list of candidate terms including terms such as *deceptions*, *lies*, *falsifications*, *omissions* and *untruths* to several experts, alongside a small but diverse list of example behaviours we wished to seek. Our colleagues, comprising two Web Science doctoral students and three postdoctoral researchers, gave us feedback about which term(s) they considered most appropriate, and then discussed the range of behaviours we were seeking to elicit. The outcome of this process was to break out three distinct questions: one pertaining to the use of untruths, one pertaining to the use of pseudonyms, and finally to the use of fictional personas, which are identities for characters that were entirely fabricated.

The survey was delivered via the web, and comprised 12 sets of questions including one set of demographic questions, and 8 open-answer free responses. In this analysis, we focus on the subset of the questionnaire delineated in table 9.

Analysis of free-response questions was done using a *grounded theory* [260] approach; themes were identified across responses through a process starting with open coding process by each of three researchers separately, followed by a discussion process where themes were refined and combined. Multiple themes were permitted per entry. Once consensus was achieved on themes, all responses for a given question were re-coded against the final set.

Table 9. Focus questions: List of questions corresponding to subset of survey discussed in this analysis.

| No. Question   | Answer type    |
|--|----------------|
| Q4a Have you ever told lies/"untruths" online? Why?                          | free text      |
| Q4b How often do you tell lies/"untruths" on social media?                   | 5-level Likert |
| Q4c How often do you think your friends lie on social media compared to you? | 5-level Likert |
| Q5a Do you use any pseudonyms online? Why?                                   | free text      |
| Q5b Have you created any fictional personas? Why?                            | free text      |

### 4.3.2. Participant recruitment

The survey was published online, with no restrictions on participation. General recruitment was carried out by handing out flyers with the URL, and the researchers' social media presences (primarily Facebook and Twitter). This was augmented by enlisting two people with popular twitter accounts (@TheTomSka, 191k followers, and @DameWendyDBE, 4k followers) to promote the survey.

In order to ensure a good selection of passionate internet users — people who live a lot of their life on the Web, and care about their online presence — additional recruitment was carried out in person at two events in London: *ComicCon* and the *WebWeWant Festival* during summer 2015.

### 4.3.3. Limitations

Among the limitations of the study, the self-report of lying behaviours may be different from actual practices for several reasons; retrospective bias effects may cause consistent under-reporting (e.g. "I think I am a mostly honest person, therefore I really must not lie that much"). A second reason that self-report is challenging here is that, due to the degree to which lying practices may be ingrained, there may be classes of behaviours that people may not consider, realise or think of as lying or deception at all. Indeed, a major class of *butler lies* were not even perceived as lies by participants of a prior study [144]. In order to mitigate this effect, we iterated on the wording of the survey questions to try to elicit as wide a variety of relevant behaviours as possible, as described in the method. Second, as with all surveys, selection-bias effects may have affected the results; in particular, those that volunteered (or, indeed, took any notice to begin with), were perhaps more likely than not to have a pre-existing interest in these topics.

Another limitation of this study is that it is reflective of only one specific demographic that we targeted; young, Western, social media enthusiasts comprising YouTubers and other 'web nerds', as these individuals have been shown to have complex, entangled online social lives [39, 181, 175, 176, 186]. As such, the kinds of concerns and experiences people reported may not be representative of other Web demographics; for example, some demographics may be less likely to maintain separate fictional personae online, or have any need to keep separate their social media fanbases. However, studies of specific

online groups, such as gamers on MMORPGs [311] have demonstrated that demographics were considerably more diverse than previously suggested, particularly in specialised online communities [63].

Finally, this study is an exploration of the kinds of untruthful practices carried out rather than an attempt to rigorously determine how often they are used. As such, we have not leaned heavily on any quantitative analysis — frequency counts have been used as an organising principle rather than a means of comparison or a fundamental part of our claims.

#### 4.4. Results

Out of the 500 survey responses, 39% (198) provided a gender; 50.2% responded female, 49.8% male, and 1% transgender. With respect to age, 59% responded, 91% were between 18 and 25, 7% 26-35, and 2% 36+. The age distribution skew was reflective of, and likely due primarily to, the predominantly young audiences at the two festivals.

Nearly all respondents were very active social media users, although use of particular platforms varied significantly. Figure 5 shows the self-reported Likert scores per platform for six social media platforms. The popularity of YouTube and Twitter for respondents was likely influenced by the method of recruitment (via Twitter), and the fact that one of the popular Twitter users who disseminated news of the survey is a popular YouTuber. The other platforms, meanwhile, were more divided, with Tumblr being the most divided between highly active (125, 27%) and those that never used it (144, 32%). Vine was used the least overall with (422, 91%) reporting having either never or rarely used it.

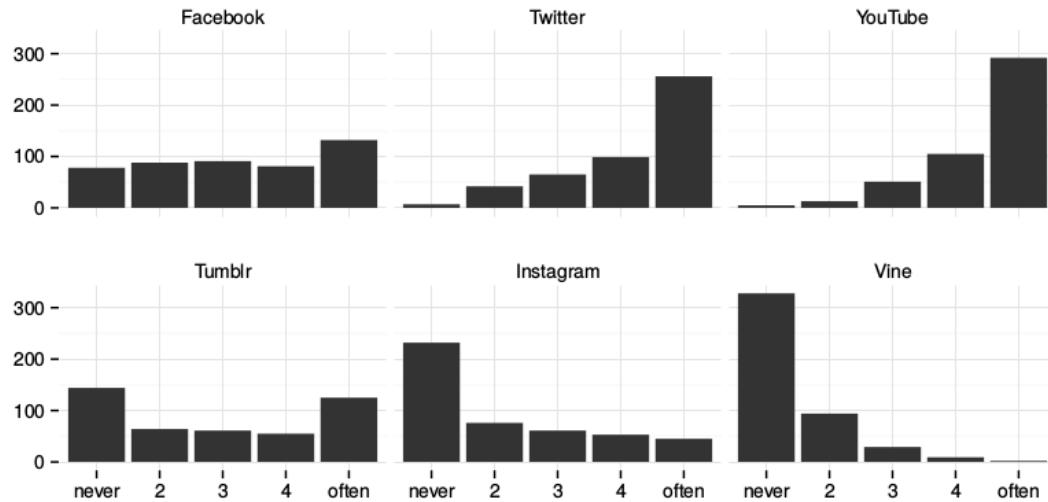


Figure 5. Self-reported use of social media, from 1=Never to 5=Often times a day.  
Medians: Facebook = 3, Twitter = 5, YouTube = 5, Tumblr = 3, Instagram = 2,  
Vine = 1

#### 4.4.1. Self-reported frequency of deception/lying

In terms of frequency of lying, 77% of participants (N=387) responded to Q4b, *How often do you lie on social media?* the distribution of answers is displayed in figure 6-a]. The median response was 2, with a majority (N=330, 85%) of responses answering either a 1 or 2.

Question 4c asked *How often do you think your friends lie on social media compared to you?*, and 77% (N=386) again responded overall (figure 6-b]). The median value was 3, with (N=87, 22%) responding with a value that their friends lie less than they do (e.g. 1 or 2), while (N=119, 30%) responded that their friends lie more (e.g. 4 or 5).

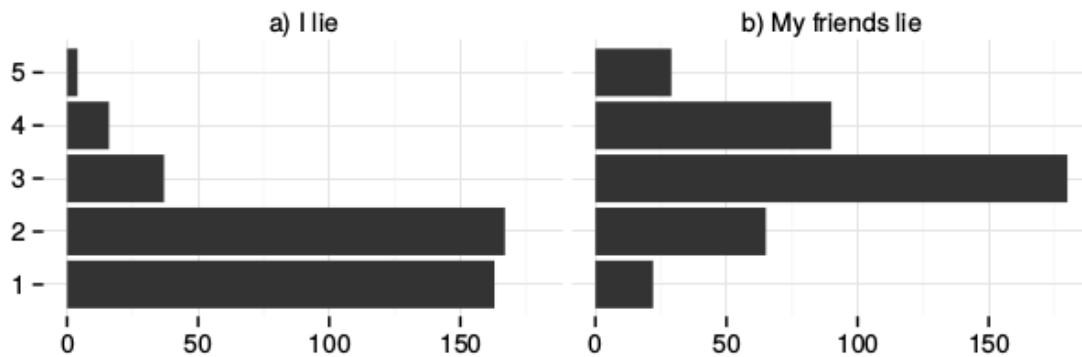


Figure 6.

Responses to Q4b and Q4c on Likert scales

b) *How often do you tell lies or untruths on social media? (1=Never to 5=Often)*c) *How often do you think your friends lie on social media compared to you? (1=Vastly less to 5=Vastly more)* This figure is missing?

#### 4.4.2. Reasons for Deception

A total of N=134 responses were received for Q4, which asked people to explain whether they remembered telling lies (or “untruths”) online and to explain the circumstances. Out of the total respondents a quarter (N=34, 25%) answered that they had or did not lie or use any form of deception online. The rest of the respondents admitted to performing some form of deception regularly.

Thematic coding of the remainder of the responses revealed 12 themes, listed in Figure 7, including an extra for *yes*, a category standing for responses admitting participating in deception with no explanation, and *no* for responses that denied using deception on social media.

The most prominent theme was **playup** (N=35), which corresponded to the rationale of wanting to be more appealing, interesting or attractive to others. There were several subtypes of this activity, starting with simply falsifying personal attributes (height, weight, age) towards what they perceived would make them more attractive, to *exaggerating* details of stories, to making things “seem relatable”. Four respondents mentioned aspects relating to making one's self seem popular or important by filling their social calendar to appear busy, while two discussed fabricating stories, such as of having met celebrities. Contexts ranged from online dating to social interaction with strangers.

Less common, although present were responses about fabricating or creating fictional events and situations (N=3), while two respondents described appropriating other people's content, including "funny tweets" and status posts, as if they had been their own.

Far less common (N=9, 7%) was the opposite reason, in which participants reported deliberately distorting or omitting information in order to not *attract* attention or in many cases to prevent disclosure of illness or situation to protect their reputation. This theme, coded as **playdown**, included the following responses:

| Lied about my mental health countless times, denied depression and suicidal thoughts. (354)

| I very selectively curate my online personae, particularly on Facebook, where I am careful to hide my mental illness, my frustrations, and my negative emotions. (461)

| I tend to lie about how sick I am so people don't worry/employers don't get anxious. (49)

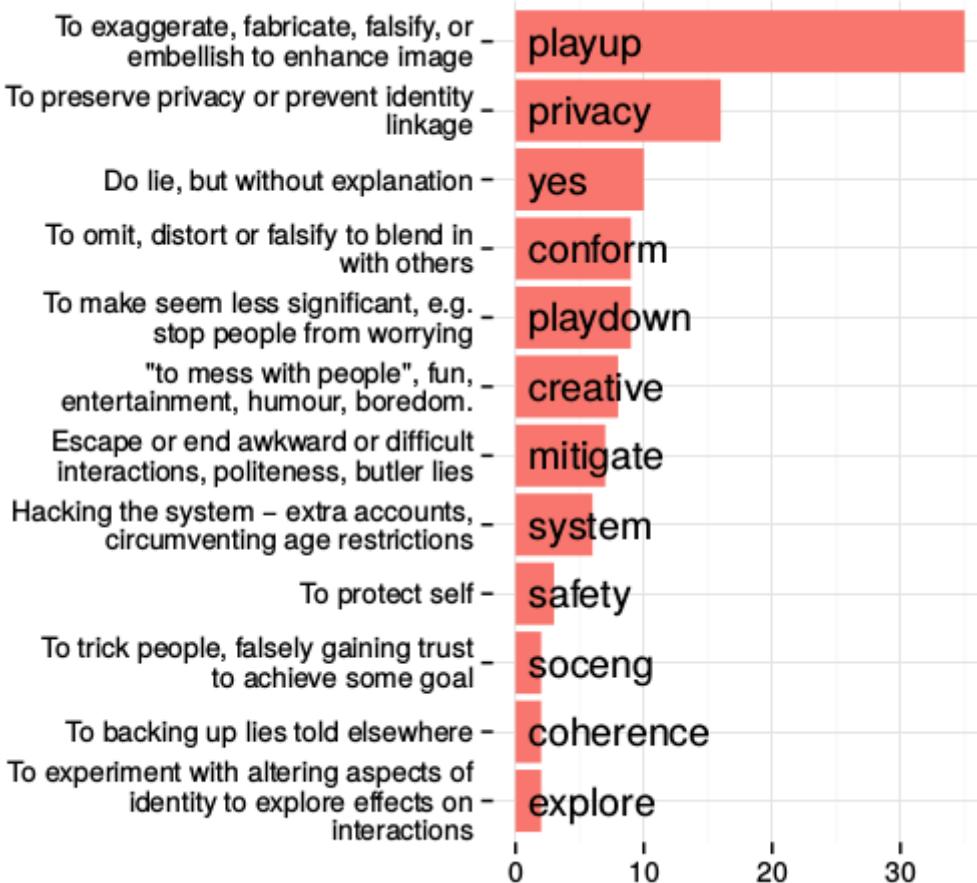


Figure 7. Tags and counts for responses to Question 4: *Have you ever told "untruths" on social media, given fictitious info, omitted or distorted the truth online?*

The second most prominent theme after **playup** was **privacy**, a theme used to encompass a variety of privacy related concerns. Respondents reported explicitly withholding information often, and, where information was required, providing false values about themselves. The attributes most often mentioned were age (N=17), real name (N=13), physical location (N=6), gender (N=3) and birth date (N=2) to web sites that they did not trust. Four mentioned that this was in order to prevent identity linkage to their real-world identities, e.g.:

On fetish sites, I will lie about my birthday (displacing my age by a few months to a year in the process) and my hometown, making my identity there harder to connect to my real identity. (461)

Others said that they adopted the strategy of falsifying attributes when social networks asked for information that they felt to be unnecessary, for example:

Whenever a social media asks me to provide personal details which are not directly necessary for them to deliver the service (e.g. Facebook asking for my workplace), I constantly feed them wrong information. First and foremost to stop them asking me for such information while at the same time keeping my personal data private. (500)

A different reason given for falsified attributes was coded as **conform**, when falsification was done in order to fit in, in particular to avoid harassment and discrimination. Such behaviour including avoiding potential confrontation surrounding personal beliefs (pertaining to religion or politics), or to personal attributes including gender, age, race or sexual orientation. One participant described her choice of declaring herself as male improved her position in debates online which were often predicated with *ad feminam* attacks on her gender:

if I pretend to be a man my sayings won't be regarded through the bias of my gender, while if I say opinions (completely disconnected from gender issues) as a woman, it will probably be the 1st thing my opponents will use in a debate. (301)

A smaller category (N=6) involve tricking the system in some way (**system**), predominantly falsification of age in order to circumvent controls on age-restricted websites.

Another set of responses (N=6) corresponded to deception or lies told for fun, humour, or “just messing about”. The tag **creative** was used for this group, which included examples such as pretending to have a twin, pretending to have met someone famous, or permuting another person's words.

Lies used to diffuse, or bring an end to, unwanted social situations we called **mitigate**. This class (N=7) was a superset of butler lies; while butler lies serve primarily to terminate and divert unwanted social interactions, the lies in this category included those which were told to be polite, such as agreeing with a person to avoid an argument. Meanwhile, **safety** (N=3) corresponded to the responses describing omission or falsification to avoid compromising one's physical safety, or from potential litigation for potentially illegal activities.

Some users described the use of deception in order to deceive, trick or manipulate situations to the individual's advantage; such reasons were coded **soceng** (N=2) because it reflected the common notion of “social engineering”.

These responses described falsification of academic credentials for jobs and posing as another person online and attempting to attract her partner's attentions as this fake identity in order to test her partner's loyalty.

Finally, **explore**, and **coherence** each had two responses. The first, **explore**, pertained to responses that discussed experimenting with aspects of their identity, in particular to "test the reactions of others". Meanwhile, **coherence** was the act of lying in order to maintain consistency with other lies told elsewhere to prevent lies from being discovered.

#### *4.4.3. Pseudonyms*

Table 10. Q5 Tags: List of themes and categories resulting from analysis of Questions 5a and 5b.

| <b>Code</b>     | <b>Description</b>   |
|-----------------|--|
| bespoke         | Several online identities kept separate.   |
| character       | Role-playing an obviously fictional character.   |
| conform         | Conform to community norms, fit in with others.  |
| creative        | For entertainment or creative purposes.  |
| discoverability | Use of a pseudonym to connect identities or be discoverable.   |
| discrimination  | Avoid being judged unfairly.   |
| disnomia        | Dislike real name.   |
| experiment      | Role-playing different real-world identities to experience the way they are treated, and/or trying to get someone else's viewpoints.                       |
| expression      | Saying things without fear or repercussions.   |
| habit           | Force of habit.  |
| hide            | Hiding activities from everyone.   |
| identity        | Online identity more closely matching true self.   |
| intimate        | Posting intimate thoughts and feelings.  |
| no              | "No", with no reason given.  |
| nothide         | Use of a pseudonym, but not trying to hide one's identity.   |
| plus            | Mentioned the Google+ "real names" policy.   |
| privacy         | General feeling of not wanting to reveal stuff.  |
| reuse           | Used a nickname or variation of offline names.   |
| safety          | Protection from other people.  |
| separation      | Separate concerns (professional, family, between friends).   |
| sex             | Anything about sex.  |
| soceng          | Tricking people or gaming the system e.g. falsely gaining trust, fake qualifications, circumventing age restrictions, using sockpuppets, and spam control. |
| spy             | The system is spying on me, merging my accounts, and sharing data.   |
| yes             | "Yes", with no reason given.   |
| -               | Theme unclear from answer.   |

Question 5a is *Do you use pseudonyms on any social media platforms? Why do you do this? Do you try to hide your real name/identity?* and 5b is *Have you created any fictional personas (e.g. characters, alter-egos) to use on social media?*

A total of N=286 responses were received for Q5a, which asked for information about whether participants had used pseudonyms, and why. A group (N=82, 27%) claimed not to use pseudonyms online, and a further group (N=5, 2%) gave answers which were unclear. This left 70% of respondents claiming to have used an online pseudonym.

The most common reason for pseudonym use was tagged as **separation** (N=63, 22%). This covers several different lines of division. The three most prevalent reasons were i) separating online and offline lives; ii) separating personal and professional identities; and iii) maintaining distinction between groups of friends or family:

... It was mainly done to slightly separate my identity from reality and the internet. (266)

... I also do not want future employers and such to be able to find all of my social media straight away and making judgements based on it. (79)

... I used to have a nerdy YouTube channel which I did not want my peers finding out about, so almost all of my online activity connected to that was under a different (screen) name. (150)

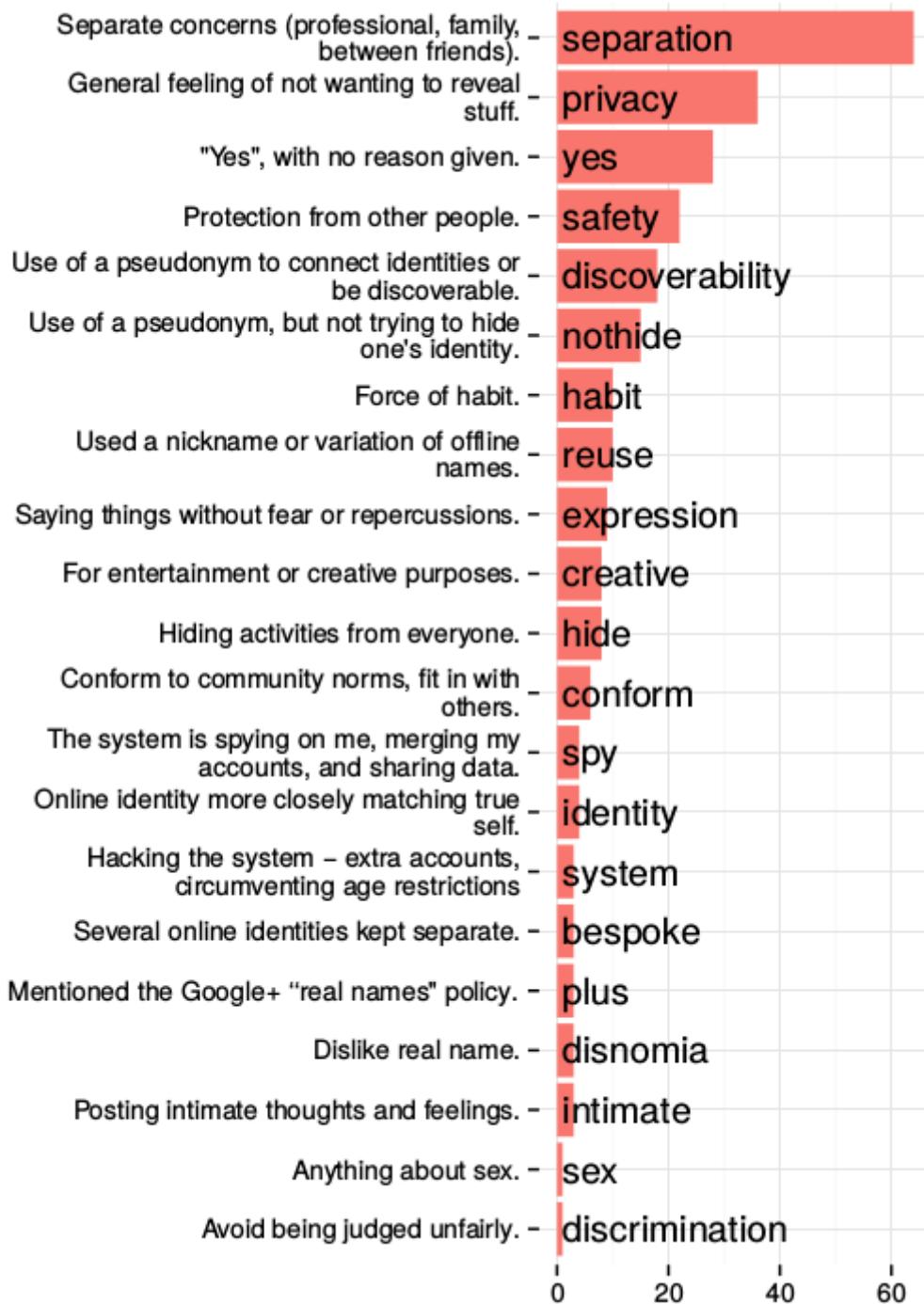


Figure 8. Tags and counts (N=286) for responses to Question *Do you use pseudonyms on any social media platforms?* Responses which were in the negative or unclear have been removed.

Related to **separation**, several people used pseudonyms to **hide** (N=8) their activities online. This is distinct as it covers activities that they would like *no-one* to know about, rather than seeking to separate different identities. Most commonly, this had to do with pornography:

| Yes, especially when using pornographic sites such as Chaturbate. (183)

However, there were also examples of more general hiding:

| I have to do things that people don't need to know about but I don't hide my real persona. (398)

The next most common reasons were **privacy** and **safety**; while these codes are related, there are some distinctions in the meanings we found. **safety** (N=22) related to a fear of repercussions spilling out of that particular online world. Some of these threats were specific ideas of violence:

| As a person on the internet with (rather unpopular) opinions I find myself constantly subjected to pretty severe harassment such as very graphic rape and death threats, so I feel it would be safer to reveal little to no identifying information on certain platforms. (168)

Many people were concerned about the idea of being stalked, of what might happen if people could find them 'in real life', while others had a general sense that one should be safe or careful online:

| Yes, I do, because I am concerned that people might stalk me if they know my real name. (184)

| ... tends to involve a lot of total strangers, so I feel I need to be more careful. (169)

This is distinct from the responses concerned with a more general notion of **privacy** (N=36). This code was used for responses which simply mentioned privacy, or a desire for one's data not to be shared. This ranged from a passive sense of not wanting to share more than necessary to an active, explicit desire to maintain privacy:

| ... I just don't feel the need to have that info on there at the moment... (151)

Some users were also change names in order to reduce the ability of systems to **spy** on them, or share their data unnecessarily (N=4).

Not all uses of pseudonyms related to hiding or privacy. A significant number of people (**nothide**, N=15) explicitly stated that they were not using a pseudonym in an attempt to hide, while several carried on using pseudonyms out of **habit** (N=10).

I use pseudonyms because they're fun, I don't use them to hide my identity, I'm not batman. (383)

A slightly larger number (N=18) used the pseudonyms to aid in their **discoverability**, by having a common name across several platforms, or to **conform** to the norms of the community (N=6).

Similarly, several people (N=10) **reused** real-world identities, often in order to allow people they know offline to find them. There is often an exclusive component to these responses, that only the desired set of people will be able to find them:

Normally just a username which is based on my real name because if you know me then you will know it is me otherwise you would not (223)

People also used pseudonyms to support **creative** activities, or simply for amusement (N=8). They also allowed the **expression** of parts of their personality without fear of repercussions (N=9), sharing of **intimate** content (N=3), and a presentation closer to their internal **identity**:

I really identify as a guy, so I go by a male name. Nobody IRL knows about that though. I do this cause I just want to be... Who I really am inside? Cheesy, but true. (44)

Some people (N=3) had a dislike of their civil name, and simply wanted a change, or had a desire to create **bespoke** identities for certain activities (N=3).

Finally, a few people used pseudonyms in order to have multiple accounts to manipulate the sociotechnological **system** (N=3) — avoiding copyright issues, or tracking who sends spam mails.

Three people explicitly mentioned Google+'s insistence on real names or merging accounts, one person created a pseudonym to escape **discrimination**, and one in the pursuit of **sex**, and .

#### 4.4.4. Personas

A total of N=267 responses were received for Q5b, in which participants were asked if and why they had created any fictional personas for use on social media. 65% reported that they do not or never have; 5% responded in an unclear manner or described pseudonyms (just changing their name) rather than personas. Of the remaining third, the most common reason was for **creative** purposes (N=21), including to entertain themselves or others. Related to this are those who explicitly state they're role-playing a fictional **character** (N=11) and those creating **bespoke** identities (N=1).

I just role-play characters I like to escape from my everyday hell hole. (44)

I use another persona to have fun telling fictional stories. (256)

I have a blog that I update in the voice of a character but that's for my own personal use as it's helping me to write a book (443)

I have and I did it because I created a fictional character and I wanted to give the illusion that the character was real (449)

I did so to make fun of some naive friends on a facebook group. (482)

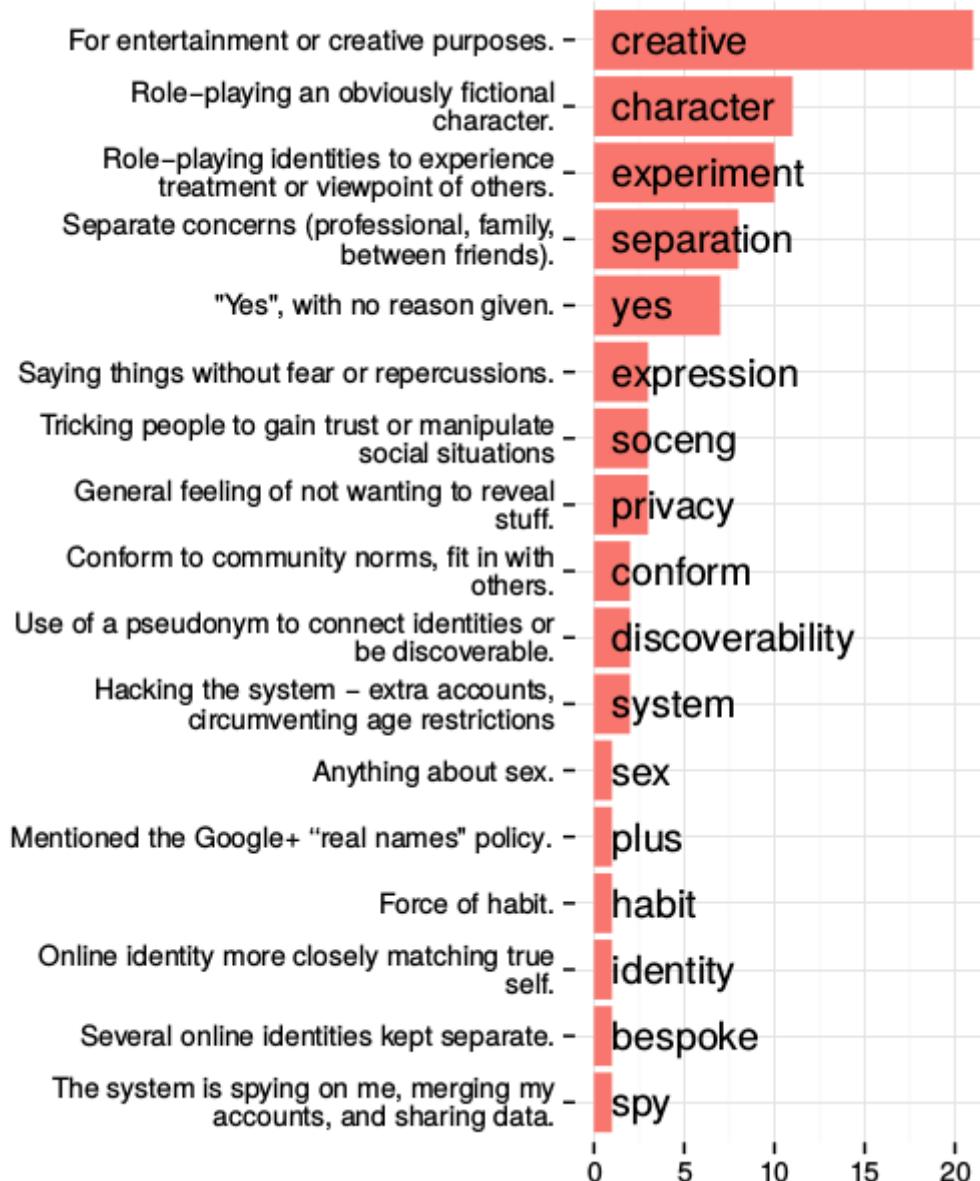


Figure 9. Tags and counts ( $N=267$ ) for responses to Question 5b: *Have you created any fictional personas (e.g. characters, alter-egos) to use on social media?*. Responses which were in the negative or unclear have been removed.

The next most common response ( $N=10$ ) was to **experiment**, including testing the reactions of others to different ages, genders or political views, or for self-exploration.

I use to when I was younger on tchat to see How people talk to different kind of people (male, female, younger, older etc...) (112)

yes. many... i do this to role play different personalities online and sometimes learn more about my actual persona by doing so. i like the act. (303)

...I have created two alter-egos. One was a short-lived novelty account that posted in the voice of a fictional character, while the other is a member of a hate group whom I used as a kind of psychological experiment in empathy — by performing as a member of that group, I came to a fuller understanding of what compels their bigotry. (461)

N=8 responses were tagged with **separation**, where respondents created personas to separate work and social lives or posting of different content types.

Yes, to comment on Youtube, because I don't want Google+ on my regular upload account. (381)

i've got accounts to post on when i feel annoyed so that friends/family dont see and it doesnt affect their impression of me (444)

Yes, I have 2 different twitter accounts that I use, one for general Fan base use which I am an overactive mad sloth and one which is for school people to think is my only one (492)

Some users took on pseudonyms for **privacy** (N=3) or to aid their self **expression** (N=3) finding it gave them the power to give voice to parts of their personality:

Yes, it helps me be more confident and say things to people that I would otherwise be unable to say. (371)

Social engineering was also a motivation (**soceng**, N=3), typically pretending to be someone new to gain trust or find out people's private opinions:

I once created a fake persona to ingratiate myself with an online community and see what they were saying about me in private. (473)

Finally there was one respondent with each of the following motivations: resistance to the system **spying** on them, or explicitly fighting the Google+ real names policy (**plus**); force of **habit**; presenting an **identity** closer to their 'true self'; and for **sex**.

#### 4.4.5. Themes

Examining themes common to all of the questions we analysed, we consolidated them into five groups, which are discussed here.

#### 4.4.5.1. Audience

Several of the themes can be considered reflections of offline social practices. Impression management behaviours such as *playup*, *conform* and *mitigate* commonly occur in day-to-day life. The online performances aimed at impressing friends and attempting to diffuse awkward social encounters seemed largely analogous with their face-to-face equivalents.

Similarly, a number of participants attempted to *playdown* or not disclose problems they were having — they described their motivations as not wanting to worry others, or not wanting employers to find out. These participants are essentially using lying to manage how others perceive them, effectively giving them more control over their illnesses, rather than being forced to disclose them, and having to deal with potential consequences of that disclosure. This particular use goes beyond the butler lies phenomenon discussed previously, and instead enables control of psychological projection and public perception of self online.

Pseudonyms and personas, meanwhile, were commonly used as mechanisms for preventing *context collapse* [151, 40, 192], maintaining a *separation* of concerns between different facets of respondent's lives. Identity was partitioned based on both the content posted and the intended audience. This included having separate Twitter accounts for personal vs. professional posts; 'secret' accounts used to interact with fandom communities away from the judgemental eyes of peers; and pseudonymous Tumblrs which allow the solicitation of advice from strangers regarding their non-parent-friendly *intimate* secrets.

While it is apparent that many of the deceptions discussed are neither new nor malicious, and complement or mirror pre-Web forms of social mediation, some were self-reported to be less innocent. Responses in the *soceng* category included creating fake accounts to stalk an ex-partner or to test the faithfulness of a current one; gaining trust to see what people were saying about them behind their backs; and manipulating social situations for personal advantage.

Another reason to construct a persona or mislead others about certain aspects of themselves was to *conform* or fit into a particular community. Online communities quickly develop cultural norms and expectations and participants tend to engage most successfully if they follow these.

#### 4.4.5.2. *Authenticity*

Some respondents reported being able to project their true selves online in a way that they cannot elsewhere. This ranged from simply using a name they felt more comfortable with, to being able to disclose attributes, ask for intimate advice, or engage in activities that they did not feel they are able to do in their offline lives. This is contrary to what systems such as Facebook claim: that 'authentic' users tell the whole and only truth about themselves [158]. This shortsightedness misses that some individuals are forced to play roles in their 'real lives' to meet others' or societal expectations, but which are discordant from how they really wish to live.

#### 4.4.5.3. *Convenience*

Consistent pseudonyms were reported as useful for allowing others to track individuals across platforms (*discoverability, coherence*), or link certain aspects of their persona together whilst excluding others, without requiring the sharing of any personal details. This would perhaps not be required if disparate systems cooperated with one another to permit some kind of data sharing or account linking. This connects to the concept of traceability from the earlier Constructing Online Identity study.

Others reported they maintained pseudonyms or personas out of *habit*, something which they presumably would not continue do if, despite what certain systems want users to do, behaving closer to their 'true' identity dramatically improved their interactions with the system.

#### 4.4.5.4. *Play*

Lies in the form of impersonations, parodies, role-playing, or storytelling were used creatively to entertain others and alleviate boredom — just as joking around in person would do. The behaviours reported are extensions of ways in which people construct the multiple facets of their identity offline. This is consistent with findings reported by boyd following ten years of ethnographic studies of social media use by teenagers [39], that the primary attraction of social media to young people is the ability to claim a social space of their own, in which they can 'hang out' when restricted from being physically co-located with their peers. boyd argues that privacy norms have not changed as technology executives like Eric Schmidt and Mark Zuckerberg would have us believe, but rather that young people are continuously evolving new ways to maintain much-desired control over social situations [39, 158].

Some of the reported behaviours serve to highlight differences between online and offline practices. While role-playing is used in the real world in order to help people work through difficult or novel situations, the malleability of identity on social networks enables participants a greater control over how they present. This allowed several people to put themselves in the shoes of others, to experience the treatment given to women, or the feeling of being part of a hate group.

#### 4.4.5.5. Safety

Most people are told from a young age not to talk to strangers in the street, the even more uncertain nature of the audience of online interactions seem to make many of our respondents innately wary. Altering or omitting personal details was considered 'the done thing' by many, who either feared for their physical safety or just wanted to avoid nasty comments. Some had a sense that they would be stalked by strangers if they revealed their location, regardless of whether or not they considered their online activities provocative.

A small number of respondents said they could alter their identity to avoid discrimination, allowing an ease of engagement which was otherwise not available. This illustrates empowering potential of the Web, where the ability to control information about oneself can be a positive force for good.

#### 4.4.5.6. System

There is distinction between respondents concerned about maintaining their privacy from other *people*, and those concerned about privacy from the *platforms* they use. From those who felt that systems simply did not need to know all their details, or were suspicious of advertising tactics, to those who were specifically concerned about the context collapse that might result from social networks which merge or cross-post to each other (e.g. Google+ and YouTube).

Another observation from our study relates to how platform restrictions become barriers to the kinds of activities we described. Platforms can limit control over identity accidentally or deliberately, through policy or technically. In particular, it is clear that several of the deception strategies described were deployed in order to preserve safety, privacy, or separation of identities in the face of platforms that were designed to thwart such separation and/or anonymous use. Common examples include providing false attributes to platforms that required personal info "it had no business asking for" and creating separate identities

where platforms provided no means of opting out of advertising or tracking. Perhaps the most irksome to the participants of our study was the consolidation of YouTube and Google+ identity namespaces with the introduction of policies requiring the use of real names. Opposition to this policy gathered over 240,000 signatures in a petition in 2013 when the change was made [88], indicating the widespread desire to maintain separate, controllable identities. Examples of careful and deliberate control over public profile information on YouTube are documented earlier in this chapter [129], showing that strategies for persona management continue despite attempts by Google to reduce the fluidity of identities of their users.

## 4.5. Discussion

In summary, this study found that people self-reported many routine kinds of lying, deception and omission strategies, reflecting a variety of needs and coping strategies for sustaining healthy, safe, and fun social interactions online. Only a small proportion of responses found deliberate attempts to socially manipulate others, while the vast majority corresponded to instances of trying to make oneself look good, maintaining separation among one's personal, professional and other social roles, fit in with others, avoid harassment, avoid causing others' worry, and to protect themselves from potentially harmful violations of privacy.

Despite not asking about platforms in Q4 or Q5, many participants mentioned adopting behaviours for specific platforms, for example, to separate their 'intimate' content on Tumblr, or to mitigate potential privacy concerns with trolls on Reddit or YouTube.

The fact that users must take active steps to circumvent the default behaviour of systems to maintain their online presence(s) suggests that current social media platforms have some way to go to provide a service that sufficiently affords the complex self-representation needs of users. The variety of benign and positive reasons users had for creating untruths indicates that these representations should be supported in order to maintain vibrant online spaces.

Developers of emerging systems can consider how they expect their users to engage, and then reflect on the types of reasons individuals fabricate or modify their personal information online: for *playful* reasons, for their own *safety*, for *convenience* relating to how they currently use or have used other systems in

the past, to be *authentic* to their true selves, and to mitigate against intrusive *systems* — and then decide which of these they want to facilitate, rather than work against, to provide a better experience for their users.

In the next study, we design some systems explicitly for enhancing 'deceptive' possibilities in online interactions and explore how people respond to these.

#### 4.5.1. Contribution to the 5Cs

In considerations of *audience*, we see peoples' offline social interactions reflect into their online spaces; we witness behaviours like diffusing awkward situations and managing context collapse. On the one hand, people claim and explore malleable fluid identities thanks to the opportunities provided by disembodyment (per *play*). On the other, they seek to prevent consequences of online interactions overflowing dangerously into the offline world (per *safety*). These scenarios each imply **control** over their online presence(s), in different ways and for different reasons.

Reflecting on one's *audience* also highlights the **connectivity** of networked publics. Those who knowingly manipulate others online, or unconsciously engage in impression management, consider to some extent by whom they are seen.

Constructing one's online presence according to community norms and expectations, as we see in *audience*, also suggests the **customisability** of online profiles. So too do the actions categorised for this study as *play* and *authenticity*.

Managing context collapse or dealing with the seams between different facets of life also contributes to our understanding of the **cascade**. People are not infrequently thrown off by data filtering through and across systems in unexpected ways. This is particularly evident from the mitigating behaviours described in *system*.

Finally, we have several contributions to our concept of **context** from this study. Fitting in with a community or particular audience forms part of the digital context in which one interacts. External, cultural or societal context are reflected in peoples' concerns about *safety* and *authenticity*. Personal, immediate practical context comes through *convenience*. So too does the technical context of a system itself and its constraints; this is also seen with peoples' anti-*system* defense mechanisms.

## 5. Computationally-mediated pro-social deception

Building on the previous survey, which broadly classified people's motivations for engaging in deceptive behaviour in their social media profiles, we conducted an in-depth interview study to better understand individuals' thoughts about deceptive behaviour in digital social systems. We particularly focus on the themes *safety*, *system*, and *audience* as we draw out how deception facilitates social behaviours in networked publics.

This time rather than asking about participants' current habits with real systems, we designed vignettes of five fictional but feasible systems which deliberately exaggerate deceptive functionalities, and participants reflected on and reacted to these in semi-structured interviews. The following themes resulted: *effort & complexity*, *strategies/channels*, *privacy & control*, *authenticity & personas*, *social signalling & empowerment*, *access control & audience*, *ethics & morality*.

This section has been adapted from work published as *Computationally-mediated pro-social deception* at CHI 2016, with Max van Kleek, Dave Murray-Rust, Keiron O'Hara and Nigel Shadbolt. Beyond early brainstorming I did not contribute directly towards the designs of the vignettes themselves, but participated equally in the questionnaire design, carrying out interviews, and coding and analysing the responses.

### 5.1. Introduction

Relevant background about deception in the context of online self-presentation can be found in chapter 2. The use of deception as a technique for system designers has been discussed previously within the HCI community. For example, manipulation of users' mental models of systems in ways that benefit both systems' designers and end-users were documented by Adar et al. [1]. Ambiguity, often promoted through deception, gives people space for flexible interpretation [116], and to tell stories they need to in order to preserve face and reputation [8, 28]. However, the complexity of modern social software dictates that a growing cast of actors be considered, both human and computational, as targets, confederates, dupes and adversaries for any action.

We base our use of the term 'deception' on McCornack's information manipulation theory [194], which encompasses both falsification and selective disclosure, such as for the purpose of creating ambiguity, or identity management.

## 5.2. Context and research questions

Here, we are interested in exploring the complex contexts in which deception might take place, to consider not just cases where the system lies to a user [1] or computer mediated communication where one user lies to others, but situations where systems lie to each other about users; where a user needs to lie to one audience but not another; where tools or systems might protect a person from disclosure to other systems or tools. As Nissenbaum puts it:

Those who imagined online actions to be shrouded in secrecy have been disabused of that notion.

As difficult as it has been to circumscribe a right to privacy in general, it is even more complex online because of shifting recipients, types of information, and constraints under which information flows.

We have come to understand that even when we interact with known, familiar parties, third parties may be lurking on the sidelines, engaged in business partnerships with our known parties. [212]

The actors involved now include not just the people who are being immediately addressed, but others who are peripheral or incidental to the interaction as it occurs. Many systems include silent 'lurkers', who observe without speaking. Others will discover and read conversations later, outside the contexts of their production. Beneath the visible surface of the communications tools people use, a growing series of actors mine the interaction data which occur on their platforms, and still others use the results of this mining. Many of these actors are computational systems of increasing power, sifting, sorting, re-purposing and inferring from the full spectrum of communicative data.

How might sophisticated privacy tools in the future facilitate greater end-user control of personal information through obfuscation and deception? What might be the personal, moral, and ethical implications of the use of such tools online? In this section, we explore these questions, and provide the following contributions:

- An expansion upon previous models of *computer-mediated social deception* with new configurations, in which tools conduct or facilitate deception towards other people/systems/tools;
- A description of a speculative design experiment in which reflections on fictional tools for social deception were elicited;
- A characterisation of the practical and social perspectives on the use of such tools, along with design guidelines for future tools employing deception in social contexts.

## 5.3. Study Design

### 5.3.1. Method

We sought to elicit diverse perspectives and experiences from people from a variety of backgrounds, around various deception configurations. Drawing inspiration from critical design [16], we adopted a speculative design method in which we first generated a series of speculative design proposals [96] consisting of realistic depictions of imagined, 'near future' privacy tools. These fictional privacy tools, with accompanying descriptions, which will henceforth be referred to as *vignettes*, were then showed to participants in semi-structured interview settings.

Interviews were conducted in person and via video chat. At the start of interviews, participants were asked an opening question, "How do you feel about your privacy online?" which was used to gauge general attitudes and sensitivity towards privacy online. Then, two framing questions were asked during the interview for each vignette; the first was whether the individual would consider using a tool like the one described (and why/why not), and second, whether the ways they perceived others and information they saw online would change if they found out their friends were using a tool like the one described. Finally, participants were encouraged to share thoughts or personal experiences that they were reminded of by the vignette.

Audio from sessions was recorded, transcribed and anonymised for identifiers of people, places and entities. Inductive thematic analysis was carried out on the transcripts by analysing and coding them for themes, by three researchers independently. Themes were then compiled, combined into a single pool, and

discussed to derive a final coherent set of themes. Related themes were then clustered into groups. We organise our discussion of results according to these clusters.

### 5.3.2. Participants

We recruited participants via Twitter, open Facebook groups, and word-of-mouth through personal connections. Those interested first answered demographic questions covering age, gender, employment status, frequency of use of social media, and self-perceptions of honesty. Fifteen participants (aged 18+) were selected in a way that maximised diversity over the attributes collected.

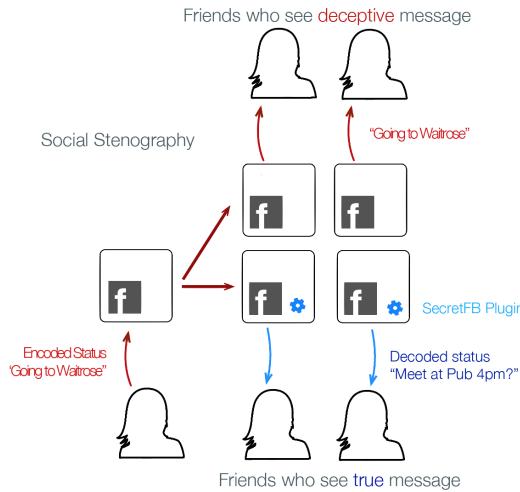
### 5.3.3. Designing the vignettes

The vignettes were selected from an initial pool of sketches according to:

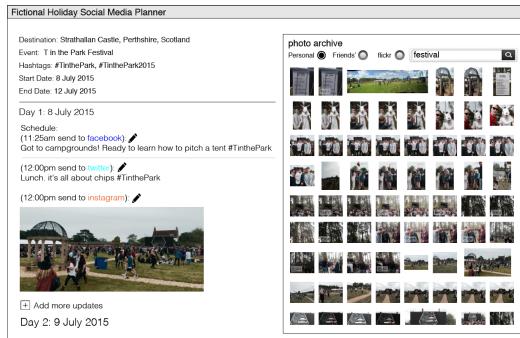
- the degree to which machines mediated the deception;
- the "balance between concreteness and openness" per Gaver's *Conceptual Design Proposals* [114].

We wanted to aim for tools that would be realisable in the near future, inspired by Auger's *speculative designs*: "speculative designs exist as projections of the lineage, developed using techniques that focus on contemporary public understanding and desires, extrapolated through imagined developments of an emerging technology" [13].

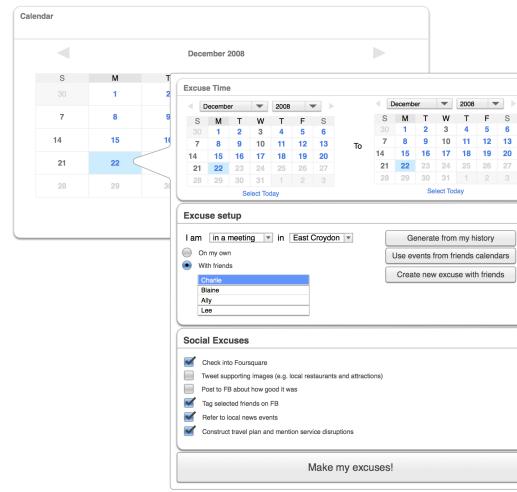
We preferred simpler, plausible vignettes to encourage participants to focus on implications rather than the tools themselves. See figures 10-14 for the vignettes used.



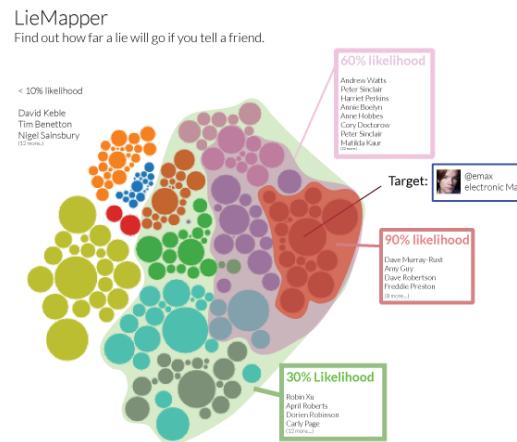
**Figure 10. Social Steganography:** A tool for microblogging/SNS sites that hides real messages behind other, plausible status messages but allows certain people to recover the true meaning.



**Figure 12. lieCation:** Create a narrative of going somewhere (on holiday) or attending an event, along with images and social media posts to be sent out at preset times to corroborate the story.



**Figure 11. lieCal:** A tool which automatically generates excuses on behalf of the user, optionally including friends in the deception and strengthening alibis by posting on social media.



**Figure 13. lieMapper:** Predict the flow of information (e.g. a lie) across a person's social network starting from a single friend.

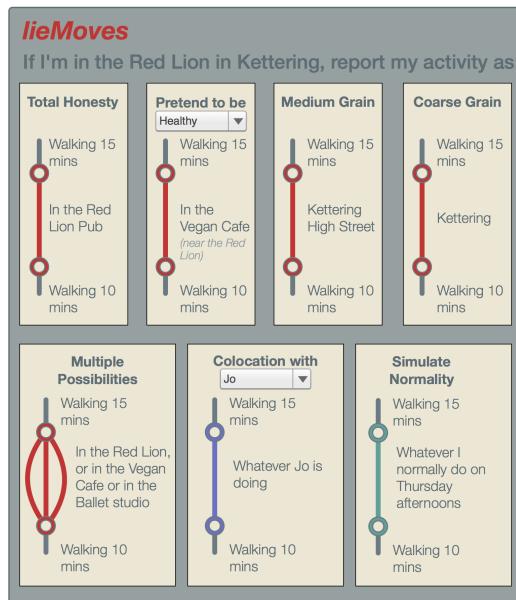


Figure 14. *lieMoves*: A smartphone service for letting people obfuscate their location using various strategies, including blurring, substitution, past-replay and impersonation (inspired by the real Moves

<<https://itunes.apple.com/us/app/moves/id509204969?mt=8>>

app).

- **Social Steganography**, inspired by danah boyd's studies of networked teens [39] who used in-group codes to discuss activities so that they were inscrutable to their parents. Here, the steganography is performed automatically: a trusted set of people see the 'real' message, while everyone else sees an 'innocent', socially plausible message.
- **lieCal** can automatically or semi-automatically fill one's shared calendar with fictitious appointments based on past (and typical) daily schedules, to create ample opportunities for butler lies. Friends can be enlisted to give support to the lie, and additional corroborating evidence can be posted on social networks.
- **lieTinerary** draws on Merel Brugman's *Same Same But Different*, enabling the pre-curation of a fictitious trip or fictional event attendance through pre-scheduled, coordinated posts across multiple social media platforms.

- **lieMapper** shows the interconnectedness of communication channels. Extending Facebook's 'this post will go to X people' functionality, it works across multiple networks to visualise all those within one's friend networks likely to hear about a particular piece of information.
- **lieMoves** is a fictional service for mobile phones that replaces the user's actual location with data from user-selectable and customisable deception strategies: blurred (low-grain), superposition of locations, past replay, or 'typical herd-behaviour or individual simulation.'

## 5.4. Results

In the following sections, we first present detailed case studies of three participants (P8, P9 and P13) to illustrate how individuals' attitudes towards privacy influenced their answers to some of the vignettes. We follow these descriptions with a presentation of themes derived from all participants.

### 5.4.1. Participants

Assuming they reported truthfully, the 15 participants we selected covered most of the major attributes in our demographic categories. One notable exception is that all participants identified as either male or female, and almost half of the participants were males aged 22-30. We did not collect information on race, sexuality or any other attributes which might be used to identify marginalised groups.

11 participants reported that they used social networks several times a day, and all but one believed that half or less of their real world activity was represented on social media. 11 agreed or strongly agreed that they saw themselves as honest, but only seven agreed or strongly agreed to seeing themselves as honest online. Nearly half agreed that they thought their friends were honest.

As additional background, we wanted to gather the 'paranoia' levels of our participants and used Westin's categories (see chapter 2) to understand the responses. 13 reported being at least *somewhat* concerned about their privacy online. According to responses to the opening question, slightly over half fell into the Westin category of *privacy pragmatists*, while two fell into the category of *privacy fundamentalists*, and the remaining four were *unconcerned* about privacy. (High inter-rater agreement was achieved for this category; Fleiss's  $k=0.624$  for 3 raters and  $N=15$  participants). These results show that in comparison to Westin's large survey of the American public [170], which had a

respective breakdown of 55%-25%-20%, we had relatively few privacy fundamentalists among our participants, and slightly more of those in the unconcerned category. However, a meta-survey of privacy indices show that our proportion is comparable to more recent results [171]. In other words, we have a sample fairly reflective of the general US populace.

#### 5.4.2. Case study: Privacy and people (P8)

P8 is a former gradeschool teacher who has returned to university to get her Ph.D. She started using social media ten years ago when she was still working at the school, and her role as a teacher strongly shaped how she managed her exposure online. Specifically, her role led to caution in disclosing too much personally identifying information, but she acknowledged that disclosure itself was important for fostering relationships and participation online.

When I was a teacher, I was very careful about what I said about teaching in school because at that point I'm not just 'me', personally; I'm also 'me' as a teacher, representing that school I was working at. Since I've stopped being a teacher, I unlocked my Twitter feed, but still try not to post too much personal stuff online. But really, if you don't share some personal information then you miss out on so much interaction stuff, so it's a real balancing act.

She kept her Twitter feed primarily for her professional colleagues, and her Facebook contacts for her offline personal friends. She believed that, as a result, most of her interactions were with honest people, and tried to be as honest in her interactions online as possible, just as in real life.

When discussing *lieTinerary*, she described discovering that her ex-partner was fabricating extravagant holidays after their breakup in order to make her jealous.

[H]e wants me to think, 'Oh, I should have stuck with him — he's having a really good life!'. So there were pictures he was putting up [on Twitter] which were supposedly where he was on holiday, but of course once you know how to scrape people's Twitter data, you could see all of his posts were made in the UK. And at that point it became really obvious that that's what he was doing, so that made me smile. But that's the only case [...] he's doing it purely because he likes to think I'm reading them, and of course I have dipped in and have had a look and had a bit of a laugh about it.

She described wanting greater controls to be able to block said partner from getting around creating new profiles to look at her information:

I do know that, if he really wanted to he could easily set up another account. So in the end, although he's blocked [on Twitter] I don't assume he can't see what I'm saying; I assume that he can, and that's another reason that I'm a bit careful with what I say. So I wish it was easier, to stop people from being able to see what you're doing — how that would happen I don't know — but that would be really helpful.

#### 5.4.3. Case study: Honesty and self-image (P9)

P9, a 22-year-old recent graduate, confessed he was very concerned about the availability of the data he gave out online due to a mistrust of companies. Valuing honesty, however, he said he would feel guilty using tools that would cause other individuals to be deceived, especially if those tools left digital interaction traces that could serve as later reminders of such acts:

I feel like I'm told that I have a certain level of privacy, I don't quite ... know enough about comp sci or technology to properly have faith in that. Like Facebook, Microsoft ... all tell me I'm safe online, and I might understand a bit of what they mean, but there isn't a great deal of explanation and I still think there are people out there who can get access to this stuff if they really wanted to.

With respect to how he manages his personal information, he prefers to be honest and transparent when the asking party is a person, even strangers online, but adopts a strategy of omission or falsification when the asking party is a company.

I'm quite an honest person, [...] like if I was on a forum and I was talking to someone I'd tell the truth. But if a company were to ask me for my number or my name — I won't bother.

I imagine [lieCal] would be useful because it would give me an excuse if I wanted to do something, but I would probably feel worse ... because it would serve as a reminder that I lied

However, he was confident there were others online that consider using tools like *lieTinerary* to promote themselves and make themselves appear popular or cool, such as by pretending to go to exclusive events:

Well they might use [lieTinerary] to come across as fashionable or trendy — they might put up a post like 'oh yeah I'm at London Fashion Week' when they're not really [...] I could say I'm at Glastonbury for the weekend, and immediately my cool points would go up.

P9 believed that such fabrication was widespread already even without such tools, alongside acts of playing one's self up:

I know people who have paid for likes and followers and stuff and they hashtag everything to death because they're so desperate for attention [...] there are lots of people nowadays who just want quick success and they'll take all of these cheap, cheating routes.

#### *5.4.4. Case study: Privacy and technology (P13)*

P13 is a postgraduate student in his mid-twenties; technologically savvy and uses social networking sites every day. He is acutely aware of the volumes of data being collected through his web use, but finds himself weighing up the practicalities of taking steps to preserve his privacy with his immediate communication needs, often concluding that "life's too short" to act on his discomfort around third-party software.

I say what I'm doing on my Facebook because otherwise no-one will ever talk to me [...] I try and use small bits of privacy enhancing stuff, to whatever extent they actually work [...] So in the past I've had Facebooks where they're not tied to my... my lying even extended to that and all the information on them was fake. Nowadays I tend not to do that because the net effect of that is no-one talks to you.

He takes steps to manage who sees his data on social media, by segregating his audience by platform, choosing who to share which aspects of his self with, and using privacy settings built into the platforms themselves. Sometimes this leads him to obtain information by proxy:

I don't connect to my mum's stuff and I don't want to connect to her stuff [...] but I wanted to find something out and so I remember asking my sister to look it up for me.

He is also resigned to data leakage, and being surveilled, by both the government and advertisers.

if it's online it's public to a certain degree right, it's... you can try and use all these controls to a certain extent, but they don't... there's always a way around things. It's like, when you're a kid and you're trying to listen to the radio and you're trying to store it cos you want to listen to whatever the thing is again. You know, theoretically they've got these measures that say oh no you can't copy this, but you stick a mic out into a cassette recorder and you've got the Hitchhikers Guide, and hey.

I don't think I'm under any illusions about web stuff. If it's out there, it's out there. If someone wants to find it and knows the information or ways to get the information then they can get it. It's annoying, but it's a fact of life.

This does not stop P13 from providing false information to services whenever he has the opportunity, under the impression that the data many services ask for is superfluous. He speculated that tools could be useful to generate more believable false data on his behalf.

So for instance airport wifi. I spend a large amount of my time in airports. So I think I'm listed as John Smith ... in Edinburgh airport, different email address, different contact information, and yeah, so we start to lie about [...] So mostly it's whenever these anonymous websites want some personal information that they don't tend to have, then I tend to lie [...] But I always sort of wonder, should I be able to generate this?

In general, he was concerned about the social risks of using tools to aid online deception, "especially when you can do this social ways, just going, oh I forgot to use the Google calendar again" but was also skeptical about how much he could trust the tools themselves.

If [social steganography] was something that I could run on my computer and I'd have it disconnected from the network then maybe.

Despite his concerns, P13 expected that he would follow the status-quo if many people began using these tools, and expressly supported other people's right to use them, reasoning that the more people did so, the more effective they would become. However, he also anticipated that the output of the tools may be prone to detection and thus rendered useless.

You could imagine someone attacking these kinds of things and trying to start to write distinguishers for when is this posted by a human or is this posted by a social media bot.

#### 5.4.5. Effort and Complexity

A common reason why participants wouldn't use these tools related to the amount of effort required to use them. P8 observed that the effort-of-use barrier is a challenge even for tools already available today, and postulated that platforms were exploiting the lack of adoption of these tools to their advantage:

The thing I've noticed is that people will always do the easiest [thing]. That's why nobody encrypts. I don't. You know, for all my concerns about privacy, I don't encrypt anything, [...] very few people take the extra security steps they can because it's convoluted. And the minute you ask people to do that, they'll just take the easiest route. Providers like Facebook and Twitter and all the apps out there know that, and that's why it's so easy for them to collect data — they know people will just take the easiest route. — P8

However, for some vignettes the extra effort was seen to pay off as an opportunity. For instance, in response to *Social Steganography*, P6 contemplated that by broadcasting different status updates to distinct subsets of his friends on Facebook, he could control multiple identities simultaneously:

I think essentially at this point you are projecting two identities simultaneously and you really would want to manage both. [...] it almost becomes twice the task. But the really interesting thing would be if different groups all had different keys — so you'd send a single status but they'd all see different ones. That would be sort of be neat, [to be] projecting multiple identities at once, because you can't really do that offline. Finally, technology would give us a chance to BETTER control our identities! — P6

A second aspect that was mentioned was not the direct effort of use, but the effort that would be indirectly required to stay on top of the wake of deception left by using such tools. In some settings, participants noted specific compensatory measures that would be required to prevent being found out, and noted the complexity and effort of these measures.

If I used a tool like this and said I had been in meetings but then actually NOT logged the hours against the project, what the meeting was about or anything like that, it would make my accounting for my own time very hard. — P7

#### 5.4.6. Availability of Other Channels, Strategies

The most common reason given for not needing to use a tool was the availability of alternative approaches to achieving the objectives for which the fictional tools were imagined to be most useful. A common such strategy was for individuals to simply *omit* or *suppress* information they did not wish to share; this strategy was used for a variety of privacy-related concerns as an alternative to use of the tools depicted in the *Social Steganography* and *lieMoves* vignettes. A second common strategy was the use of *other channels* and *access control features*. For instance, P13 discussed the use of encryption to both help control scope of a message and for unwanted leakage by platforms. Several mentioned Facebook and Google+'s built-in access control features for limiting the scope of a particular message as an alternative to using a steganography approach.

In some cases, participants identified that alternate strategies were imperfect, and sometimes the fictional tool offered a better solution. For example, the alternate strategy of suppressing location leakage by turning location tracking off, was perceived as worse than *lieMoves* by both P6 and P9, because doing so would cause apps that needed the user's location to simply refuse to work.

There were fewer alternative strategies given for the other vignettes; "simply being honest", and in particular "blocking off time" was given as a common strategy for situations where *lieCal* would be useful (P4, P8, P9).

#### 5.4.7. Privacy and Control

Several participants cited potential benefits to privacy control and management. The leaking of location information was a concern; six participants reported keeping location services on their smartphones turned off by default for reasons such as to prevent apps from sending their location to third-parties without their consent.

[lieMoves] would mostly catch out apps that were taking my location without even asking, because if I want to tell the truth when I think it matters, I can still do that, but those that are just spying on me gets crap! And that appeals, because they shouldn't be able to collect in the first place! — P6

P8 asked whether *lieMoves* was available for use, because she wanted it immediately to keep Google from tracking her.

I want to install it immediately and keep using it for the rest of my life! I wouldn't have any ethical worries about it because I wouldn't be lying to anyone, I would be lying to Google, and that's exactly what I want to do!

Because they shouldn't have this information in the first place, so giving them wrong information is perfect. As I said, can I have this today, please? — P8

Others pointed out that a remaining impediment to adoption of such tools is still a remaining lack of awareness of how services operated and used people's information.

People can't make value judgements about the systems they interact with because they don't understand them well enough yet, especially what's going on behind the scenes. They don't actually feel the need to deceive system and platforms because they don't even know they're being spied upon. — P6

#### *5.4.8. Authenticity and Crafting Personas*

Participants reflected on how the data they shared affected other people's perceptions of them, as well as their perceptions of others on social media. P11 (in agreement with P1, P2, P3, P6, P7, P8, P9, P12 and P15) assumed that her friends engaged in "image-shaping" by "being quite selective or trying to present a particular kind of persona", and described an occasion when a contact's online presentation was at odds with what she knew to be happening offline.

People will always seem like they're having a really good time and post about how great everything is but then you talk to them and things aren't actually quite how they're made to be portrayed on social media. [...] So like one of my friends, her sister was just posting about her one year anniversary of getting married, and how brilliant it was, and they were both posting about the presents they got for each other. Within a month they were separated [...] I know more about that from talking to my friend personally, but in terms of what's presented online to a different audience, to a much wider audience, that was not what was going on. — P11

P12 described a friend who, unable to withhold information or resist questions from an inquisitive audience, made up stories about her life to satisfy them, thus creating a persona.

'Cos of the following that some fanfiction gets, she gets asked a lot of personal questions and she doesn't want to feel rude so she just lies, so she answers these very personal questions so she feels connected to her audience but she deliberately lies 'cos she finds it sometimes a bit invasive. — P12

P8 and P15 similarly mentioned deception used to protect privacy without alienating people. In contrast, others saw total openness in their sharing as important for presenting their "authentic" selves on social media, and thought less of those who they perceived to be engaged in deliberate image-shaping.

I wouldn't be friends with people who would be lying all the time or who make up stuff just for attention. [...] if I found out that there was someone I was interested in doing this the faith I put in them or the fact that I was being very genuine would take a hit. — P9

#### 5.4.9. Polite Social Signalling, Kindness, and Empowerment

Though sometimes in conflict with attempts at authenticity, a number of respondents echoed the sentiment that degrees of deception are crucial for maintaining a well-functioning society.

I think that not telling people — everyone, everything — is a central aspect of being kind in the world. — P15

It's about empowerment — little lies, like "I'm just too tired and you're quite a taxing person" could be the truth but that's a bit mean, and you didn't want to say that! versus "oh no sorry I have plans with my boyfriend" which might be a lie, but it's nice. — P6

Often you lie to save people's feelings or — to stop someone finding out about a surprise party. Like there are really nice reasons to lie, and if you could help people make nice lies safer, that would be awesome! — P14

P6 commented that this could be a subtle method of signalling violations of personal privacy online:

The idea of being able to put massively sarcastic calendar appointments just so that, when someone looks at my calendar to see what I'm doing, they know I don't want them to know, and they should just stop asking. — P6

Such methods were also viewed as a form of social empowerment; a way of giving people freedom to block off time (*lieCalendar*) or send a message (*Social Steganography*) in situations where the honest approach would be awkward due to shyness, introversion, or differences in social positions, e.g. having to contradict a superior or respected senior.

Somebody younger, less experienced, less confident might find that this is a nice, straightforward way of blocking time out for themselves and feeling good or comfortable about it. Because it can be quite difficult saying “no, I’m not free” to someone senior. — P8

#### 5.4.10. Access control and imagined audience

Many participants discussed their expectations of who could access their social data and messages. P15 and P11 mentioned assuming private Facebook messages could be read only by the recipient; all but four participants segregated their friends using platform privacy settings.

If I wanted only certain people to know something I’d just send them a private message rather than put it as my status anyway. — P11

Very occasionally I will post things only on Facebook and not Twitter because then only my friends can see it. — P6

Why would I need such a tool when I can define on Facebook, for every single message exactly who sees it? — P7

Contrarily, several participants also rationalised that they must assume that anything they post online could be made completely public at any moment, and the safest technique is not to share at all.

If there’s something you don’t want people to know, then you just don’t tell. — P7

P12, meanwhile, said *lieMapper* would potentially improve her ability to maintain separation among her separate personas online by showing her when information from multiple identities linked up.

This one is more just a way for you to control your privacy [...] Cos I don’t actually have anything linked directly [between identities], but probably they have certain things links other ways, so that would be quite interesting to see. — P12

P2 and P14 were among those who considered themselves not interesting enough for anyone to want to invade their privacy, and P4 even found it “felt quite good” when he found his private facebook profile had been accessed by someone outside of their network because he “was of interest to someone”.

#### 5.4.11. Ethics and morality

Finally, many of the participants volunteered their views on ethical or moral reasons of why they would or would not use these tools in specific ways. Perspectives varied in general and according to the vignette presented.

The technology vignettes could be seen as ethically neutral, with the ethics coming from the manner of their use:

If your intention is to use these tools to harm someone, then that's the individual's own decision to make and you can decide for yourself whether that's morally right or wrong. But simply using the tools themselves doesn't imply you're going to do something that is harmful or morally wrong. — P5

However, in some cases, there was such a strong correlation between the design of the tool and the kinds of lies which it facilitates that the morality of the tool became the morality of the action:

Well as someone who's considered murdering people before, this is exactly how I would do it. I would create a fake social media presence so I could go off and do something illegal or even ... I could commit adultery, I really can't see much of a practical application for ethically good things... — P14, discussing lieTinerary

To P6, whether deception was moral was contextually dependant on whether the recipient had a legitimate need for the truth and why.

If someone has a right to know something for some reason [...] then lying to them there is more problematic than if they didn't have a right to ask you, or to be looking for that information. [...] that's their own fault; they should have known they shouldn't have looked. — P6

Some participants suggested that they would need a really good reason to use deception tools. P14 felt that a better alternative to having to lie was to get out of situations in which one felt the need to lie.

And if you're in a situation where you have to lie to people about where you are, then that's a situation you need to get out of cos that's a creepy situation [...] The only time I can see this being good is like if you're in an abusive marriage and you're going to a divorce lawyer in secret. — P14

There was often a moral distinction made between friends and platforms as the targets of deception. A majority (11) took issue with deliberately deceiving friends and there was also widespread consensus on wanting not to deceive a general audience on social media. By contrast, there was a feeling that lying to platforms is not dishonest.

well if I'm talking to my friend I always tell the truth; I'm quite an honest person ... but I don't think lying to Facebook is unethical [...], because it's not affecting any of your friends or anyone on your list, so it has no effect — so you're not really lying to anyone? [...] I don't trust these companies enough, to be honest, with the information I supply them. — P9

P6 took the position that lying to platforms should be the moral choice, even part of one's civic duty.

I think lying to Facebook is to be encouraged! [platforms] spend so much effort in deceiving users into thinking they're doing one thing when they're doing another, that giving users some control seems fine. Its sort of like the debate whether minorities can be racist against white people — like, whether the power imbalance seems to negate any meaningful argument, certainly when it comes to lying to services. — P6

## 5.5. Discussion

Deception is a long-established strategy for informational self-determination, and it is not a surprise to see the practice in online behaviour. The study reported here is a necessary preface to the deep study of deception, and establishes interesting lines of enquiry which mark out a descriptive vocabulary and a potential design space. Nissenbaum outlined the importance of contextual integrity for online design, the idea that individuals bring a set of expectations and meanings to their online interactions that are often derived from offline analogues, appropriately or otherwise [213]. A designed interaction that leaves no space for someone to present themselves creatively for non-malevolent purposes fails to preserve contextual integrity, and would consequently produce an asymmetry of understanding between user and system of which the user may be unaware.

Deception is often an expensive strategy, involving some creativity, the avoidance of passivity and the maintenance of consistency in an alternative narrative. In all but its simplest forms, it is not something that most people do lightly. Particular strategies and opportunities for deception were common to

many of our subjects, who were often concerned with the balance between the moral injunction against lying, and their own interests. Mitigating factors were sought: for example, if the counterpart in the interaction is non-human (a platform, for instance), or if the interaction provided an opportunity for malign activities (e.g. could be used by a stalker), or if the counterpart did not have a good reason for requesting the data, then these were seen as justifications for using deception for protection.

### *5.5.1. Morality of Deception*

Our participants, like the majority of people, like to think of themselves as being generally honest, but this has a nuanced relationship with their reported behaviour. There was a common feeling that deceiving platforms and corporations was acceptable, or even a moral imperative. Nomenclature was significant: casting activities as 'lying' provoked responses which paid more attention to the ramifications of being found out, and a greater sense of ethical violation. However, 'hiding information' was generally seen as acceptable, as was partitioning information for different audiences, especially in the context of avoiding unwanted attention. Politeness was often cited as a valid reason for performing white lies, a variety of kindness.

Akerlof and Schiller's account of phishing [3] focuses on deception from the point of view of corporations, and therefore helps explain the existence of situations in which our participants were motivated to deceive. In the information economy, data subjects are beguiled, misled or strongarmed into giving away more data than is required for the service they wish to access. However, perhaps because their focus is wider than the information economy, Akerlof and Schiller fail to consider the possibility of the individual creating counter-asymmetries by manipulating the data they provide to corporations. Their recommended counter-measures are all intended to support truthfulness — standards-setting, reputation, regulation. Yet these all require concerted action, while deception is a strategy open to the individual.

### *5.5.2. Promoting Social Honesty*

One viewpoint is that mendacious impulses are indicative of a problematic situation: that fixing the socio-technical context would remove the need to deceive, and the community could become more socially honest. Systems requesting excessive information frequently provoked anger, and a feeling that feeding back fictitious information was justified. One lens for designers to

engage with this issue is Grice's conversational maxims [125]. Typically, these are used to define one side of a social contract: the quantity, quality, relation and manner of information production.

A complementary view applies to requests for information. This accounts for many of the indignant responses we received — systems were asking for *too much* information, or *irrelevant* information. Providing clarity here, relating information demands to the current context, limiting information to that which is necessary can guide designers towards upholding the platform's end of the social contract. Our *lieMapper* vignette asked how far through our social networks personal information was likely to diffuse, alerting the user to social information violations; similarly, when designers illuminate the hidden pathways which our data takes — or doesn't — it provides a grounding on which trust can be built.

Legal identities, and the problems which they cause, highlight the multifaceted aspects of life, whether online or off. The general trend is towards a collapse of context, the joining of identities across sites and networks, but the attitude that people should be happy to connect all of their identities together in this way is an expression of social privilege. Tools exist to aid the management of multiple personas, typically used by astroturfing organisations [169, 121]. As a provocation, what would design for multifaceted life look like? Are there ways to support participants in plural presentation, helping them to understand and maintain their context bounds, rather than attempting to force a homogenisation. How can we support radical self-expression and support marginalised groups? What about systems which acknowledge that there are parts of users' lives which they don't want to share publicly, but which they still need to express in order to connect with similar people? Designing for contextual authenticity rather than imposing singular identity pushes back against marginalisation.

### 5.5.3. *Memory, safety, and plausible deniability*

It was clear from responses that being reminded of one's lies can be upsetting, especially for people who consider themselves honest. On one hand, this suggests that systems might automatically remove, or reduce the visibility of, digital traces that could serve as reminders of one's past deceit. The recent growth in messaging apps that automatically delete messages after a single viewing [93] might, in fact, be related to this perceived design need. On the

other hand, visibility of such actions can lead people towards greater honesty — knowing how often one was deceptive could clearly be a powerful push towards veracity.

A second major theme addressed the effort, both of using the tool, and dealing with its potential consequences. It was clear that any tool that required more time and effort than customary was perceived as too burdensome. There was also the consideration of the side-effects caused by such tools, and the degree of effort required to ensure such repercussions would not cause deceptions to be discovered. But having to explicitly act at all was also viewed negatively; that is, having to engage with a tool in order to carry out a deception, such as with *lieCal*, was viewed less favourably than something that could do it automatically, such as *lieMoves*.

An additional problem with requiring users to carry out an explicit action is that doing so often leaves little space for plausible deniability: it becomes often difficult to maintain that such an action was taken accidentally or unintentionally (assuming the individual is of sound mind). If we instead imagine tools that *deceive by default*, the possibility that a deception was simply a side effect of being busy or forgetting to make the system tell the truth would remain. For example, a deceive-by-default variation of *lieCal* might automatically fill the person's calendar with false but plausible appointments, allowing its user to quickly identify and replace them with real ones as needed. Such designs would additionally support many of the goals of *privacy-by-design* [243].

Another significant barrier to the use of such tools is related to safety and discovery. The first: ensuring that deceptive actions would not have unintended consequences, while the second pertains to the effort and actions necessary to ensure deceptions would not be discovered. Such concerns suggest that there is a potential space for future tools that are able to support *safe deception*, both in terms of highlighting potential hazards, and towards mitigating the burden of covering up active lies and their effects. Tools such as *lieMapper* that are able to provide situational awareness about social information flow could help individuals tell certain lies, especially *nice* ones (as described by P8), with less risk of exposure.

#### 5.5.4. Design implications

Despite the preliminary nature of this study, the results suggest many questions for consideration by system designers. Those providing services for data need to identify, respect and avoid the factors which lead users to deception. The act of deception creates a situation in which data minimisation is in the interests of the platform — the less that it asks for, the more likely it is to be trusted, and the less likely the deception strategy is to be invoked. In particular, contextual integrity is preserved if users are able to represent themselves differently in different contexts, and it is clear to them that the more data that is demanded, the easier it is to resolve these personas. Similarly, there is a set of deceptions, such as butler lies, which are adapted to specific communication situations, and facilitating these will also help transfer and preserve expectations in the digital context.

Systems which facilitate deception will have both positive and negative potential. Most obviously, their wide uptake would reduce trust in data generally. On the other hand, it is clear from our study that for most people, deception is a last resort, that is, the majority self-image is one of general honesty so that deception would demand ad hoc justification. A rather more calculated invocation of a deception system might, if such attitudes were widespread, be a step too far. Framing the objective of the system will be key — for example, classifying such systems as privacy-enhancing, rather than deceiving, might increase their acceptance. However, software that maintains a consistent, false record of events might remove the burden of understanding for users that their behaviour is deceptive, thus making it easier to deceive. Such divergent potential outcomes require investigation.

#### 5.5.5. Contributions to the 5Cs

The themes that emerged from these interviews mostly serve to expand our understanding of the **contexts** in which people interact online. The ethics or general acceptability of deception varies depending on the moral standpoint of the respondent; which is likely developed by their immediate and cultural environment. Technical contexts were raised regarding the ability of tools to retrain traces of lies as well as the fact people often don't understand how the tools they use actually work. Tradeoffs between being fully honest online and just not participating at all are also a function of the social environment (ie. it may be acceptable neither to avoid online interactions, nor to be fully authentic, leading to some form of deception as a necessity). The alternate *strategies* discussed reminds us that people use many tools and systems in conjunction, and these uses influence each other. This is part of a personal context.

Wishing to have plausible deniability, as covered by *effort & complexity*, as well as the strategy to *omit* information or use privacy controls, all feed into our notion of **control**.

Engaging in image-shaping and other *social signalling* may require some level of **customisation**.

Participants demonstrated their awareness of audience - **connectivity** - in responses about access control and the morality of deception depending on who is being deceived.

Concern about information flow around a system was particularly highlighted by the lieMapper vignette. The idea that everything one puts online might become public, concern about being tracked by third-parties, and traces of deception being persisted by systems, all feed into the **cascade** aspect.

## 6. Social Media Makers

The previous three studies demonstrate that considerable effort is made by users of mainstream social media to circumvent constraints of the systems they use in order to better engage with other users, or to protect themselves from perceived top-down threats. But what of avoiding mainstream social media altogether? One of the systems examined in the first study is the Indieweb wiki. This is a particularly flexible, portable and representative way of managing profiles which does *not* rely on a centralised service or authority. The Indieweb community are amongst a growing number of Web users who are replacing or supplementing mainstream social media use with DIY personal social platforms. I call this kind of Webizen 'social media makers', and in order to compare this approach with our findings from centralised social media users I take a closer look at their activities and motivations in the following in-depth interview study. The results are that they opt-in to highly flexible and portable profiles despite technical costs, influenced by the following factors: *self-expression, persistence/ephemerality, networks & audience, authority and consent*.

### 6.1. Introduction

The Web today is a very different place than the one imagined by its creator, Tim Berners-Lee [24]. Instead of a vast network of individuals running their own web servers to host homepages or share information, most people simply navigate to Twitter to tweet, log in to Facebook to post a status update, use Wordpress.com to write up their thoughts. With a daily active user population of over 1.13 billion<sup>1</sup>, Facebook alone constitutes a full quarter of all Web traffic<sup>2</sup>.

However, there are individuals who in certain ways reject the massive social platforms that have swallowed the Web. Instead, they embrace 'home-grown' approaches to building their own web presences, much like the 'old days' of the Web. But unlike the old days, when such a presence might have comprised a homepage, 'DIY Web' hackers now piece together their own bespoke social and data management platforms, akin to the kinds of services offered by social platforms, for managing their interactions and identities online.

<sup>1</sup> Facebook reports second quarter 2016 results, investor.fb.com, 2016  
<<https://investor.fb.com/investor-news/press-release-details/2016/Facebook-Reports-Second-Quarter-2016-Results/default.aspx>>

<sup>2</sup> Facebook is eating the internet, The Atlantic, 2015  
<<http://www.theatlantic.com/technology/archive/2015/04/facebook-is-eating-the-internet/391766/>>

Often, such capabilities are realised by using an ensemble of open source tools and standards supported by developers with like interests. However, the various motivations for these individuals, and the perspectives they have gained from doing so have not, thus far, been studied extensively.

## 6.2. Context and research questions

In the previous studies we examined how individuals who participate in massive, centralised online communities present themselves and manage their interactions with an audience. In this section, we present a study which seeks to understand how individuals who *avoid* mainstream social platforms find other ways to present themselves online, as well as a descriptive characterisation of their digital spaces. We targeted a broad class of self-described *digital makers*: those who identify with taking a hands-on 'DIY' approach to meet their own immediate online social interaction and self-presentation needs. We conducted semi-structured interviews supported by live demonstrations of participants' own systems and their social media profiles, to address the following questions:

1. What are the main motivations of digital makers in replacing or supplementing mainstream social media profiles with their own personal systems?
2. How do their uses of (if any) and feelings about mainstream social media compare and contrast with their own personal sites?
3. What challenges do digital makers face regarding competing discourses from different social contexts, and how do they address these?

## 6.3. Study Design

With the wide availability of different social platforms, people often tend to use one or more to manage their online social activities. Different platforms are tailored to various types of preferences, philosophies and purposes, and target different communities and individuals' needs (social, professional, leisure). In this research we are interested in identifying and investigating individuals who desire the same 'type' of interactions that come through using mainstream social networking sites, but maintain their own platform (e.g. blog or website) as their primary online profile.

We designed interview questions to encourage participants to reflect on their activities, rather than just recount them. The intention was not to compare their experiences with particular systems, but rather elicit their motivations and habits, and ideas and feelings about the ways in which they interact online.

### *6.3.1. Participant Recruitment*

From the existing literature (see chapter 2) we can see similarities and differences in habits and motivations of bloggers and social media users. What we haven't heard about are social media 'makers', who occupy some space in between.

Such individuals must be technically competent or willing and able to learn. They share the DIY attitude with maker communities who engage in physical/hardware hacking but in the purely digital realm. In the same way that hardware hackers seek to understand and control their physical day-to-day environment, digital makers who see social media as a core part of their everyday lives are engaging in similar practices online.

Social media makers are different too from the open source software developers who work on decentralised social platforms like GNU Social, Friendica, pump.io or Diaspora (these platforms are discussed further in chapter 4). That is not to say they are mutually exclusive and indeed many participate in several projects which are relevant to their interests in this space; but makers focus primarily on building systems which affect their own lives, and only secondarily address use by others.

I recruited suitable participants through distributing an online signup form in IRC channels and online forums known to be frequented by individuals engaged in building personal websites and social media systems, and asking those who responded to refer others they know. The signup form included a brief description of suitable participants so that people were able to self-select for the study.

The signup form asked for demographic information (age, gender, occupation, ethnicity, country of residence) as well as a list of personal websites and social media sites they use on a regular basis.

### *6.3.2. Method*

Participants were asked seven opening or closing questions, and a set of five questions about each of their personal sites and each of their social media profiles, so the total number of questions depended on the number of personal sites and social platforms they used. I enquired about their motivations for building their own platforms, the particular functions they use them for, and

about their audience. I asked similarly about how they use social networking sites, their audiences there, and how the functionality and audience overlap or differ between their personal sites and different social networking sites.

I conducted semi-structured interviews in order to gain a first-hand account of participants' experiences with building and using their personal social media systems. I used the list of sites and systems gathered during recruitment as a starting point and encouraged participants to click around their websites and profiles during the interview, and we recorded a screencast of the process. This served as a prompt for both the interviewer and participant which allowed me to tailor the conversation around the participant's particular experiences. It also aided participants in accurately recalling the systems they use, as well as backing up their anecdotes with specific examples from their personal sites or social streams. I also allowed participants to show and discuss sites they had not initially reported if they wanted to do so.

The interviews took place across a variety of different locations convenient to the individual participants. All but one participant used their own laptops when viewing their websites and social media profiles, so things were set up in the way that they were used to day-to-day. Participants were permitted to pause the screen recording and/or audio at any time during the interview if it made them more comfortable.

I used open ended questions as a guideline, but allowed participants to deviate freely to other topics if prompted by one of the questions, or something on their screen.

Participants were rewarded with a 15 USD gift voucher for their time.

### 6.3.3. *Limitations*

While a qualitative semi-structured interview is the appropriate method to gather people's various technology usages, motivating factors and associated examples, it also presents several challenges. Qualitative data gathering may suffer from a lack of detailed or accurate recollection of events; participants might report their perception of general trends instead of specific descriptions of their activities, and may be subject to unconscious influences or motivations. Participants may also deliberately withhold or distort information. Using this method some information can be misinterpreted or overlooked.

I mitigate against these issues as follows:

- Detailed recollection of use: I asked participants to visit their sites and profiles and answer questions based on what was on screen. This was done in order to ground their reasons and preferences in concrete examples, to be able to interpret their responses in context, and explore further issues on the basis of what was visible if necessary.
- Perceptions vs. actual actions: In order to gather users' activities accurately I asked them to provide specific examples to support their responses, which they were able to show us on their screen.
- Withholding or distorting information: Participants were informed that they could decline to answer any question, or could answer questions without being recorded, in order to mitigate against their feeling obliged to provide any response to questions they may uncomfortable with. Participants were able to answer vaguely if they preferred, rather than giving granular detail that may be inaccurate.
- Detailed information required vs. study length: In order to avoid a fatigue effect I asked participants to start with their personal sites, and then prioritised the social media platforms they felt they used the most. The study session was scheduled to run for 60 minutes, with an additional 30 minutes buffer for participants who wished to talk for longer.

Since this target community is niche and at an early stage of development, and since my recruiting options were correspondingly limited, it is inevitable that the conclusions I draw from these results cannot be generalised to a broader population.

#### 6.3.4. Analysis

I take a grounded theory approach to analysing the data gathered [260].

Immediately following each interview, I recorded pertinent words or phrases and notable highlights from the discussion, as well as a general impression. These notes were compiled into a preliminary set of codes in order to begin the process of identifying potential themes. I carried out inductive thematic analysis on the interview responses in several stages:

1. Listening to the interview recordings and transcribing to re-familiarise with the responses. Taking snapshots from the screen recording to include in the transcript when something on screen was explicitly referred to.

2. Coding the transcripts, beginning with the list of phrases from the initial interview notes, but adding to this list throughout. Each transcript was passed through at least twice and re-coded to account for new codes which emerged later during the process.
3. Noting relationships between the codes and how they co-occur, and categorising the codes to identify broader themes.
4. Reviewing themes to understand how they overlap or relate to one another, and refining them to make sure themes are distinct. Identifying related topics which are missing from the data.

I organise the results section according to the derived themes.

## 6.4. Results

13 interviews were conducted in person, and 2 over video chat. 10 of these took place during or after one of three technical events, conveniently over the same week in nearby cities, which were of interest to and therefore well attended by our target participants during June 2016.

All participants work or study in the technology industry, which is typical of the “social media makers” target, though not all work in web development. 11 participants identified as male, 3 as female and one declined to answer. A majority of participants are white; two listed their ethnicity as Hispanic or Latino, one as Jewish and one as Asian, and two declined to respond. All are resident in North America or Europe. These biases are reflective of the technology industry and the routes through which we were able to recruit participants.

Interviews lasted between 45 and 90 minutes, and participants discussed between 1 and 4 personal websites, and between 1 and 9 social networking sites (mean 4.7). Some paused screen recording when typing in passwords or if they were interrupted by incoming chat messages when browsing social media. None paused the recording in order to show the interviewer something ‘off the record’.

Participants talked about a combination of personal social experiences both on and offline, including how their online activities affect or are affected by their every day lives; their feelings about others working or experimenting in the decentralised social web domain; things they have accomplished and things they

want or plan to accomplish in future; and technical details of systems they have built themselves. In many cases, participants naturally covered answers to the guideline questions without explicit prompting.

Across all of the participants there was diversity in both systems used and the main emphases of conversation, however there are many common threads around *control* and *audience*, which we discuss here.

"As long as I've known that it's possible to publish creative works I have ... once I realised that I can't trust others to have my own best interests in mind I started having my own websites." (L)

#### 6.4.1. The network

All participants maintained personal sites, and all used profiles on one or more major centralised social media services. All but one participant cross-posted content from their personal site to social media sites to some degree, referring to this process as "syndication", which may be manual or automatic. All participants said they need centralised services in order to reach their social network(s), as they did not expect their friends and family to go out of their way to read their personal website on a regular basis. Participant K had been involved in various open source social network projects, but missed his regular contacts: "I was originally in like the GNU Social and Diaspora... they didn't make a ton of sense because I didn't know anybody that was on them." Participant J agreed that he had looked at some open source projects, but didn't like their "wholesale rejection of silos" because his friends were still using centralised systems and he wants them to read the things he writes.

"Most of my friends are in silos, I want to be in those conversations and interact with them there, but I also want the control of both the data and the presentation... It's about interacting with people in silos while not being locked up in silos" (M)

However, most participants did not simply copy all content to all networks indiscriminately but employed a variety of policies when deciding which posts to syndicate where.

**Contents:** is the media or data or length of text contained in this post consistent with the type of content generally posted to this site? This could be a cultural constraint, or a technical restriction. Participant M says content must "fit in" so any inconsistencies from what's expected don't distract from the

content itself. Participant E makes collages of photos to post because “this is generally what people do on Pinterest” and F points out the “unspoken rules about what goes on there and how people interact”.

Does the content cover a subject appropriate to discuss on the third-party platform? Participant N wouldn't write academic things on Facebook for fear of boring people.

**Frequency of posts:** Almost 75% of participants were worried about creating undue “noise”. For example, only carefully selected photos are generally posted to Instagram, but a whole, uncurated set is uploaded to Flickr.

On twitter nobody cares if you're too noisy, but on facebook they really do, they start complaining. (B)

I don't want to post too much and like dominate someone's feed... I wanted to post 20 photos today and I definitely feel a pressure not to do that because it's [being syndicated to] Instagram (I)

**Who will see it:** is the content of the post appropriate for the connections they know they have on the third-party platform? Will the content be distributed publicly or privately on this platform? Is the anticipated level and type of engagement with the content by others desirable?

Participant B uses Medium to reach to people beyond his normal circles of “affluent white men who work in tech”. Participant F withholds certain content from Facebook to avoid social-media-novice family members making unrelated personal comments which all of her other contacts can see. Participant I keeps his social media profiles fully private, even though they mirror the public content posted on his personal site, considering the interactions on Twitter “too abusive and spammy” for it to be worthwhile letting strangers comment on his posts.

**What will it look like:** 70% think about how their content would be rendered on other networks when deciding whether to syndicate there; for example, since short text-only notes look unappealing on Facebook, but images and link previews are presented well, they only syndicate content to Facebook when it includes the latter.

Even when posting primarily on their own website, the importance of the network means that participants are still bound to some extent by the norms and expectations of the communities and platforms used by their friends and

contacts. Most participants found more freedom in posting to their own sites with regards to types of posts, content and posting frequency. However those who were more committed to cross-posting, or did not have tooling available to allow them to be selective about cross-posting, were strongly influenced by the other destinations for their content when making posting decisions, in some cases self-censoring their content or amending how it is presented.

#### 6.4.2. (De)compartimentalisation and audience

One third of participants said they do not think about their audience, but seven described how they are very aware of who might be reading their content, including that they revise content until they feel it is appropriate for multiple audiences they imagine might see it. Five people say they primarily post for themselves, and four said that whilst they selectively cross-post subsets of content based on the norms and audience of third-party platforms, they actively want to collapse these contexts on their own site.

I'm ready to collapse everything... I want to be as complete as possible... It'd be fun if I had a blog that's part let's say ... Swift coding, and also part cute beefcake pictures. (J)

Some would consider filtering based on who was looking, if it was technically straightforward. However most don't care at all about audiences from different aspects of their lives coming across their posts, or imagine the audience for their site is so small that context collapse is unlikely. Some people even saw this as positive, describing it as "healthy" (M, I) or "more human" (I). Four acknowledged their privilege as non-vulnerable members of society which allowed them to feel this way.

Despite this, seven participants said they post pseudonymously or anonymously in other systems which are completely disconnected from their primary online identity. This is usually about sensitive topics that they are not willing to share more broadly, and don't trust their own technical expertise to build sufficient access control into their own systems.

#### 6.4.3. Self-expression

Despite the uses for centralised systems in terms of network reach and audience management, one aspect of *control* desired by participants is how they are able to express themselves.

**Self-expression through visuals:** Participants cared a lot about what their websites look like. Three quarters said that having creative control over the appearance of their content was one of their main motivations for publishing on their own site. For some, this stemmed from wanting 'clean and simple' visuals as opposed to the 'noisy' interfaces of Facebook and Twitter. For others, it was important that they had freedom to experiment.

I like playing with the form. I think the thing about cookie cutter sites is that they are one size fits all. And I think form in some ways dictates content and so if you're publishing on a network things you are writing are led by what's already on the network. If you've got your own site it's like starting form scratch you can just dictate exactly what's there, you can choose your own identity, you can have a huge say over what you're publishing. (B)

Participant L showed other peoples' sites he was inspired by visually, including one which uses a unique design for each individual blog post. Participant N publishes his art on his site — alongside his technical and academic essays — which involves executing code, so he is unable to use existing platforms.

Participant I has archives dating back over a decade, and for many years experimented with a different design every month. These designs are frozen in time, so clicking back through these archives reveals radical changes in visuals which capture moods, ideas and experiments from the time the posts were written:

I could easily try out a new style knowing that I wasn't committing all my archives or the future to this style. (I)

He was convinced these temporal visuals were "worth preserving" by his own memories of blog posts he'd read looking a certain way, and the associations or even nostalgia that come with that when articles are re-discovered. The same participant periodically makes time-limited or post-specific updates to his visual style, and compares this with the way people change their profile pictures on social media to support a particular cause.



Figure 15. Examples of websites by participants who used their own space to express themselves in ways they are unable to achieve with mainstream services.

When interacting with others in their community, participants may post a message on their own site which triggers a notification to the addressee, who will often display the reply. Despite their strong feelings over how their posts are presented on their own site or on social media, no participants whose replies are displayed on another *homegrown* site (as opposed to a silo) were concerned with how their message is presented there, so long as the content itself is unchanged. The domain owner is within their rights to display incoming content in their own space however they see fit.

**Self-expression through voice:** Over half of participants were worried about their self expression being compromised through censorship or not being able to use their own voice. One participant, only recently returned to personal publishing, stopped posting on his own site and on social media for several years after a post about his gay relationship resulted in an unpleasant message from his father. As he became increasingly reclusive online, this started to reflect into his offline life as well, and he withdrew socially. Recently he realised this wasn't healthy and made a concerted effort to become more expressive, more confident in his identity and to be "forward" about his experiences as a queer software developer. Following this, he very deliberately intends not to conceal parts of himself going forward. He now finds "joy in intersections" of the different parts of his life and adds that on your own domain, "no-one can tell you that you can't call yourself whatever you want."

Participant L cares deeply about the visuals of his site, but ultimately sees them as something that will change: "content will always represent me more than any visual design will, because visual designs come and go." As a result, he does not display any content created by others: "I want every pixel on my site to be mine."

**Self-expression through functionality:** 60% of participants appreciate the ability to create types of content that no other social systems allow, and to mix and match the types of content they post all in one place, agreeing with the sentiment: "it's my site I can post whatever I want" (I). 11 participants described things they do with their own systems that they cannot get elsewhere, including adding licenses (L), editing or deleting posts (I, O), posting events or RSVPs (A, I), custom lists, logs or channels (A, E, G, K, L).

#### 6.4.4. Empowerment

Participants varied in how they felt their personal sites empowered them, beyond self-expression. For some it was about ownership. Participant J expressed concern that his generation rarely own things, from cars and houses, to music, and his personal site was a way of claiming something back: "my personal website at the center or origin of my ideas... it's like expanding my own real estate."

However participant B pointed out that "owning your own stuff is only useful if you can actually control it... if you're not a developer that's actually disempowering" and participant L describes personal data ownership as "a tool toward autonomy. Just owning a lot of stuff doesn't give you control or freedom or agency."

75% want a place under their control to be the *canonical* source of their content; the definitive location for their online persona. Participant I's personal site is entirely public, yet his social media profiles — to where all of his posts are syndicated — are locked down, "to encourage people, if they wanted to share something that I tweeted, to then share my original copy instead on my website." (You can't retweet private tweets, but the original link could be posted in a new tweet).

#### *6.4.5. Longevity*

12 participants expressed the importance of being able to archive their content and data. They wanted copies. If a centralised service disappears or bars their access — something which has happened at some point to every participant — most now have their content and often the context of conversations with others, on their own servers. Participant G “had no idea Geocities [a popular 1990s HTML hosting platform] would go away.”

For some, this is also about personal development, reflection, and spotting patterns in their behaviour over time. “I like being able to have a record of everything I've published over the last ten years and being able to come back to it and go oh this is what I thought about this back then” (B). Participant L thinks we should take a 2000 year view of our digital lives, and that we all have a right to store personal data “cradle-to-grave... that will live on untouched after our death” to benefit future societies.

#### *6.4.6. Ephemerality*

In contrast to longevity, seven participants create content with the expectation that it will disappear, and see value in being able to do this, and 4 participants explicitly consider social media to be a place for ephemeral content. “I feel that I'm much more responsible for what I post [on my own site]... whereas on Twitter you're part of the faceless hoard”, said F, explaining why she posts throwaway or snarky comments to Twitter without bothering to archive them in her own space. Many participants treat Facebook replies and likes the same way: “if it's a comment on someone else's post I'm assuming it's pretty much throwaway. I write assuming it could last forever, but I also write assuming that if it got deleted for any reason I wouldn't care... doublethink” (I). However several participants reported that they would archive certain types of content if they could, but at present the technical barrier is too high, and their priority for doing so is too low.

#### *6.4.7. Consent*

Another aspect of control is consent. Many members of the community use common tooling to fetch replies from social media to the syndicated copies of their posts. There are privacy safeguards in the tooling that prevent private or access controlled posts from being exposed publicly, but 4 participants either

didn't do this, despite a desire for archives, or expressed misgivings about the fact they were: "it feels a little bit weird to be pulling people's stuff without their consent" (B) because most contacts on social media were unaware of the possibility that their content would be copied to somewhere else on the Web; somewhere potentially more likely to be indexed by search engines.

#### *6.4.8. Abuse and surveillance*

Participant B once posted something which triggered a flood of reactions from an infamously abusive online community. His tooling automatically pulled these responses from social media through to his personal site. Rather exercising his ability to remove these posts, he "decided to keep it up and sort of showcase the idiots." Other participants consider themselves lucky not to have experienced this, and give little thought to how they would handle it.

Although 7 participants "don't write sensitive stuff on digital technology" (N) none expressed concern about centralised systems mining the data they syndicate, or the terms of service they are agreeing to in doing so.

#### *6.4.9. Inspiration and triggers*

Almost all participants said they were inspired by others in the community, and other personal sites they see on the Web. Some took specific ideas to do with visuals or the types of content they can post; others were just inspired by the movement towards data ownership in general. Some participants said they replicated features they like from centralised services.

Half of participants said they built new functionality into their systems when their current way of doing something became too painful or inconvenient. Others want to keep up with the trends in the community in general, so they implement new features in order to continue interoperating with others, or just to try things out. Many said they update their systems when they have enough free time to do so, and have long todo lists of things they want to achieve.

Most were triggered to update their bios on both their personal sites and across their social media profiles when something changed in their life. Except for participant O, who updated specifically when he realised he would need to give his URL(s) to someone and didn't want his information to be out of date.

### 6.5. Discussion

The “digital makers” we interviewed revealed their primary motivations in replacing or supplementing mainstream social media with their own personal systems are control of their online representation, and over the longevity (or not) of their content, and decompartmentalising or making a canonical source for all aspects of their online presence.

Across all of the participants there was diversity in both systems used and the main emphases of conversation, however as mentioned previously there were many common threads around *audience* and *control*.

Participants who feel over-constrained by the limitations of social media with regards to the kinds of content they can create and how it must be laid out or displayed have developed their own completely custom publishing environments in order to more freely express themselves. They demonstrate many creative ways of displaying types of content similar to that ordinarily found on social media, as well as innovating with new “post types” or ways of sharing information that they cannot do elsewhere at all. This freedom to experiment leads them to reflect on and perhaps better understand their identities. Further, participants feel empowered by the ability to archive their content for life, or hide or remove content from their own space as they like.

Participants still wanted to reach their **networks** on mainstream social platforms, and were influenced by the norms and expectations of these platforms when deciding what to post. A result of this is that the content they post is still influenced by the platforms they know it will end up being seen on. A technique to mitigate the risk of violating norms on other platforms or encountering technical barriers is to be selective when cross-posting. Though there is no hard and fast formula to follow, participants commonly consider audience, visuals, content types, posting frequency, and topics as part of a gut intuition when making these decisions. Thus personal social systems cannot be studied understood in isolation.

Individuals who are preoccupied with their own ability to control their profiles also think about how they interact with the content of others, discussing **consent** when it comes to re-displaying posts or data belonging to others.

Data ownership for these participants is helped by the use of a single personal platform, and as a result participants need to find new ways to manage audience and context collapse. Several participants actively desired context collapse in

their personal systems, to create a complete image of themselves, no matter who is viewing it, even though they segregate their audience across different mainstream platforms, and selectively *cross-post* accordingly.

Three quarters of participants want to control the **authoritative source** of their content. Three quarters said that having creative control over their **self-expression** through appearance of their content is one of the main motivations for publishing on their own site. Participants feel empowered by the ability to archive their content for life, or remove content from their own space as they like (**persistence** vs. **ephemerality**). If a third-party service disappears or bars their access most now have their content and often the context of conversations with others, on their own servers.

We can see that the *social media makers'* prioritisation of the more individualistic aspects of identity management (per blogging) combines with the goals of retaining audience, interaction and *network* to allow collaborative identity construction (like contemporary social media); perhaps this is the beginnings of a more complete presentation of self online.

#### 6.5.1. Contributions to the 5Cs

What it means to **control** one's online self-expression was emphasised by this study. A priority of these participants was having the ability to choose whether content is persistently archived or temporary, ephemeral. They also wished to choose where their data shows up (eg. through cross-posting), and expressed concern that others may not have that option (per *consent*). Another aspect of control is to be the canonical or authoritative source of one's online presence.

Many interviewees cited *self-expression* as a primary motivation. Their online spaces tended to be highly **customisable** as a result, in contrast with mainstream SNS.

The isolation of running a personal site was mitigated by hooking into the *network* of mainstream social media. This demonstrates a novel means of **connectivity** which shows both a hyper-awareness of *audience* as well as some degree of disregard for who reads their content from different contexts.

Critically, despite avoidance in principle of centralised social systems, such systems strongly impact the **context** in which our makers operate. Community norms *and* technical constraints of alternate platforms influence content and

presentational decisions. It is especially pertinent that these individuals may be considering cross-posting something from its source location in their own system to *multiple* third-party systems at once.

## 7. The 5 Cs of Digital Personhood

The studies in this chapter span a variety of different perspectives, technical systems, and use cases. Each bring to light certain considerations around online profiles or self presentation. Here, the results of each study are summarised and briefly discussed. I clustered the results of each study according to similar components (see Table 11). The clusters make up the overall framework of five concepts which can be used to organise ideas around digital self-presentation whilst keeping track of different perspectives and influences, and acknowledging the interconnectedness between them.

*What is a profile?* describes affordances of systems which integrate online profiles in a social capacity and raises five features of systems with regards to their representations of users: *flexibility, access control, prominence, portability, representation*. These features in different combinations may empower profile owners more or less, and they reflect on how much authority profile owners have over the data that makes up their profile, what it looks like, and who sees it. The perspective of the profile owner is considered alongside the system owners/designers/developers, and other developers or third parties who can access and potentially influence profile information.

In *Constructing online identity* the focus is on cross-system profiles within a creative media production social machine. Where participation centers on generating and interacting with content, and participants have both more ability and more desire to experiment and entertain through their online profiles, we identify *roles, attribution, accountability* and *traceability* as critical dimensions along which to discuss self-presentation in these spaces. These dimensions reflect on the links between profiles within and across systems, the creative ways in which profile owners can visualise their identities, and how these representations are reused, remixed, and propagated throughout online and offline systems.

Studies about deceptive practices in online social interactions, *The many dimensions of lying online* and *Computationally-mediated pro-social deception*, a survey and interview study respectively, yield two sets of related themes about common usages of mainstream social media. The former results in *system, authenticity, safety, audience* and *play* as reasons people limit or modify their online self-presentations when compared to their offline selves. When

individuals are asked to think more deeply about how they and others might mediate social interactions through technologies that help them to customise the 'truth', the latter uncovers concern about *effort & complexity*, *privacy & control*, *authenticity & personas*, *access & audience*, and *social signalling & empowerment*, when making decisions about when and why they would employ social deception techniques, as well as discussing alternative *strategies* for achieving the same ends without technological help, and the *ethics & morality* of doing so. These themes reflect on peoples' relationships with other users of systems as well as with the systems themselves — the extent to which people can oversee or manipulate how others view, access, and interpret their personal information; as well as the day-to-day social norms and expectations they are surrounded by before, during, and after their online engagements.

Finally, interviews with *Social media makers* delve into how individuals are maintaining online profiles without centralised services, motivated by *self-expression*, *persistence/ephemerality*, and *authority*, and constrained by the effects of their *networks* and audience *consent*. These themes reflect the importance of visual expressiveness in self-presentation and control over where data is stored and how it is maintained.

All of these studies demonstrate that online self-presentation is both constituted and affected by *who* sees a representation of an individual, and *what* it is they see, both of which are encompassed by the situation *whereby* it is seen (see fig 16). Next, I present the five terms which cover the possible lenses through which we must look at online profiles in order to understand them fully.

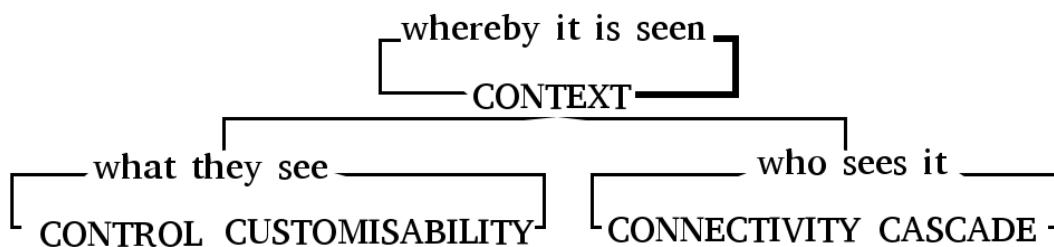


Figure 16. A view on how framework terms relate to one another, hierarchically.

Table 11. How results from each study make up aspects of the 5Cs.

| <b>Concept</b>         | <b>Aspect</b>                                | <b>Study result</b>  |
|------------------------|--|--|
| <b>Context</b>         | technical affordances                        | S1 flexibility, S1 portability, S2 roles, S2 attribution, S2 system, S5 networks             |
|                        | social expectations, including participation | S2 roles, S2 attribution, S3 audience, S3 authenticity, S4 authenticity, S5 networks         |
|                        | personal motivations                         | S2 roles, S2 attribution, S3 convenience   |
|                        | policy constraints                           | S1 flexibility, S1 portability   |
|                        | purpose of system                            | S1 prominence, S1 representation, S3 system  |
|                        | avoiding danger/discrimination               | S3 safety  |
|                        | ethics and morality                          | S4 ethics  |
|                        | being kind to others                         | S4 social signalling   |
|                        | using multiple tools together                | S4 strategies  |
|                        | create and discard identities                | S2 roles   |
| <b>Control</b>         | persistence vs emphemerality                 | S2 attribution, S2 accountability, S2 traceability, S4 effort, S5 persistence & ephemerality |
|                        | if a profile is required                     | S1 prominence  |
|                        | ease of moving data                          | S1 portability   |
|                        | how much of a person is a profile            | S1 representation  |
|                        | preventing context collapse                  | S3 audience  |
|                        | diffusing awkward encounters                 | S3 audience  |
|                        | malleable identities                         | S3 play  |
|                        | avoiding danger/discrimination               | S3 safety  |
|                        | omitting information                         | S4 strategy  |
|                        | understanding options                        | S4 privacy   |
|                        | being kind to others                         | S4 social signalling   |
|                        | consent (self and others)                    | S5 authority, S5 consent   |
|                        | authoritative source of personal info        | S5 authority   |
|                        | decompartmentalisation                       | S5 networks  |
| <b>Customisability</b> | visual branding                              | S2 traceability, S5 self-expression  |
|                        | links to other profiles                      | S2 traceability  |
|                        | how data is presented                        | S1 flexibility, S5 self-expression   |
|                        | which data is shown                          | S1 flexibility   |
|                        | to whom data is shown                        | S1 access control, S3 audience   |
|                        | being oneself                                | S3 authenticity  |
|                        | malleable identities / image shaping         | S3 play, S4 authenticity, S5 self-expression   |

being kind to others

S4 social signalling

| Concept             | Aspect                                    | Study result   |
|---------------------|---|--|
| <b>Connectivity</b> | audience known/unknown                    | S1 access control, S1 flexibility, S2 attribution,<br>S2 traceability, S3 audience, S3 authenticity,<br>S4 audience, S5 networks |
|                     | reputation                                | S2 attribution, S2 traceability  |
|                     | purpose of system                         | S1 prominence  |
|                     | claiming social space                     | S3 play  |
|                     | multiple/cross-network audiences          | S1 portability, S5 networks  |
|                     | aggregate profiles                        | s2 traceability  |
| <b>Cascade</b>      | spread of information                     | S2 traceability, S4 privacy  |
|                     | connections to other profiles / real life | S1 representation  |
|                     | access to data by others                  | S1 access, S4 ethics, S4 effort  |
|                     | context collapse                          | S3 audience, S3 safety, S3 system  |

Study numbering: (S1) What is a profile?; (S2) Constructing online identity; (S3) The many dimensions of lying online; (S4) Pro-social deception; (S5) Social media makers.

## 7.1. Context

Individuals are situated in societies according to geographical, cultural, and familial boundaries. These societies vary in size and have different political, legal, and economic factors as well as social norms and expectations. We all navigate an intersection of different societies daily, some people more than others. Our identities are strongly influenced by what our societies expect (and demand) of us, and how we react to these expectations. Expectations of different societies can conflict, for example, when a woman is raised in a conservative religious household but in a country with a liberal non-theist culture, she may need to navigate different identities pertaining to her home life and work life. How we engage with identity online is of course impacted. A lack of geographical boundaries and a blurring of political and legal jurisdiction can complicate how individuals want to and are permitted to present themselves.

Despite the lack of geographical boundaries, the Web is not equal in every country. Governments censor particular systems, filter content, and surveil populations; organisations must adhere to different types and degrees of regulation around for example privacy and data protection. This too makes up part of the broader context in which digital representations of people exist.

These factors also serve to influence how technical systems are developed. From company revenue streams to subconscious bias of engineers, every technical decision — every feature added or removed — is framed by societies. This in turn impacts system users, who may exist in an entirely different social setting. The technology fundamentally affects what profile owners can and can't do; behaviors they are coerced or driven in to.

Not least, the motivations and day-to-day needs of individual profile owners constitute personal contexts. A single system may be used in ten different ways for ten different reasons — or even in ten different ways by the same user — depending on personal circumstance; profile owners are never homogeneous, and their situations are always changing. Similarly, people use multiple systems in conjunction, and these uses influence each other.

## 7.2. Control

Once an individual has established an online profile, how much authority do they have over the information collected and presented there? Determining whether aspects of a profile are ephemeral or persistently and reliably archived are important aspects of control of that self-presentation. Data entrusted to a third part may be lost — or sold — and **cascade** out of reach. Throwaway remarks may be indexed by search engines or snapshotted by Web archival systems, increasing the likelihood that they are accessed devoid of the context of the system or conversation of which they are a part and making it much more difficult to let go of or conceal a particular representation of oneself.

Whether certain personas or aspects of oneself are traceable, either to an offline 'real world' identity or to other online representations, is a function of how much control profile owners have over the **connectivity** of their profiles. Control is not increased by anonymity nor by a blue tick<sup>t</sup> of authenticity; either of these (or anything in between) could indicate greater or lesser control depending on other things (like the **context**). The amount of control available affects the extent to which profile owners need to **customise** their self-presentation, with regards to its presentation and access.

<sup>t</sup> Twitter users who have been 'verified' as 'real' by the Twitter company/platform (usually by means of an application process and/or credit card transaction) are marked with a blue tick by their name.

## 7.3. Customisability

For any given digital representation, an individual may or may not be able to adjust the information contained within. Profiles are composed of a variety of different data, and profile owners potentially have limited awareness of the data collected and processed about them. Certain data may be editable or deletable, or perhaps so for a limited time. Furthermore, profile owners may or may not be able to customise *who sees which* elements of a particular profile, depending on the levels of **connectivity** of the system they are using.

Online profiles can act as a proxy for an individual's physical presence. The customisation of online profiles also describes what the contents of a profile look (or feel or read) like, and can determine the impression others have of the profile owner, as well as impact how the profile owner feels about themselves. Just as one adjusts one's dress, posture, or facial expressions in person according to the particular role one is playing at the time (behaving differently with your friends compared with your teacher, for example), online profiles can be customised (not necessarily accurately) according to both known and imagined audiences. Customisation of profiles is strongly tied into the technical constraints or affordances of a particular platform (part of the **context**); less ability to customise may reduce a profile owner's ability to express themselves effectively, which in turn limits their control, or perhaps understanding, of the impression(s) given to their audience(s). Customisations can help profile owners to express individuality, or to demonstrate that they are part of a particular community or in-group.

Customisability also gives individuals the power to explore and experiment with different identities, whether these are closer to or further from what they feel to be their 'true self'. This freedom is important and healthy for many populations, including vulnerable or oppressed people, and minorities seeking connection and support from a geographically dispersed community. On the other hand, the potential to imitate others or hide one's identity in order to behave in malicious ways is also available.

On the other hand, more ability to customise self-presentation results in more decisions that need to be made by an individual, which can result in cognitive overload and perhaps be disempowering after all [209].

In contrast to the physical world, different aspects of an individual can be represented simultaneously by multiple online profiles. Whilst most people would avoid being on a date and at work at the same time, a dating site profile and an employment site profile can co-exist, and even be opened side by side by

the same viewer. As one's self-presentation changes over time or with circumstance, snapshots of versions of oneself may be stored by search engines or Web archival systems, limiting **control** the profile owner has over their customisations.

Data from profiles may be displayed differently in different contexts, or in different systems, perhaps in ways which do not match what the profile owner intends or agreed to. Customisability may be lost through the **cascade**.

## 7.4. Connectivity

Systems have different affordances when it comes to connecting profiles together, both between different users and between multiple representations of the same individual. How connections are used and displayed within a system affects the imagined audience of a profile owner, and may impact how they need or want to **customise** their profile contents. Changes to profile information may propagate through a network to different degrees depending on the purpose of connections, and so connectivity is related to the **cascade**.

One's connections or social network can say a lot about a person, and so the connections themselves also constitute part of a profile. The prominence and semantics of these relationships varies between systems and communities, part of the **context** of profile use.

## 7.5. Cascade

The cascade corresponds to Goffman's expression 'given off' [123]; subconscious side effects of social interaction which nonetheless affect how one is seen or understood. In offline interactions, interpretation of expression 'given off' is typically restricted to the physically co-present (though they might pass their observations on later). With regards to online profiles, information about individual representations are propagated through and across systems: processed by algorithms, packaged, remixed, interpreted, correlated, aggregated, re-packed and oftentimes sold on, given away, leaked, or stolen. Many people know about the cascade to some degree, but most ordinary social media users are unaware of its extent. People have come to accept that Facebook sells their profile attributes to advertisers, but may not consider that Facebook can also make use of their mouse movements and clicks, visits to

other websites entirely, learn about their life through running analysis on the text of their status updates, and may be legally required to hand over the contents of their 'private' messages if asked to do so in court (see **context**).

The cascade is also a function of **connectivity**; as friends, fans and followers get hold of one's profile, they can potentially share or repurpose the information beyond the original owner's control or knowledge. Thus unknown effects of the cascade can cause a loss of both **customisability** and **control**.

## 7.6. Conclusions and reflections

Through the five terms described in this concluding section, we can come to a well-rounded understanding of factors influencing the presentation of the self online. It is easy to forget or ignore the multitude of angles from which individuals are impacted when navigating networked publics. I propose this framework as a guideline for future work in both studying and designing digital social spaces, and hope that this helps both in terms of avoiding over-generalisations of social media users as we study them academically, and in terms of taking into account to a fuller extent individuals' circumstances as we make technical decisions about systems. The nuances of each term in the framework are explored in more depth in each of the studies presented here, although these by no means cover every possible angle; other angles are covered by other studies, reviewed in the previous chapter.

Of particular interest when it comes to building new systems (and developing existing ones) are the ways people 'misuse' features, or use them in ways other than system developers intended. Individuals and entire communities can appropriate particular features of systems to meet their own unanticipated needs; similarly, techniques for circumventing technical or policy constraints, or just 'breaking the rules' are widespread. Developers can learn from these activities, particularly if they attempt to understand their users as individuals who exist *beyond* and *outside* of a single system, and beyond the digital as well. I only hope that a greater understanding of system users leads developers to strive to better meet their users' needs, rather than improve their models of oppression.

On that note, the remainder of this thesis looks to *decentralisation* as a means for empowering individuals, and approaches this from a technical perspective. That is, taking the power out of the hands of centralised entities like the companies behind contemporary mainstream social networks, and putting it

back into the hands of profile owners. Work in this chapter and the previous one illustrates to some extent how non-centralised, self-hosted, or individually-controlled personal systems (like blogs) increase the possibilities for individual aspects of identity construction, but potentially make it more difficult to integrate collaborative aspects, which are similarly critical for a complete digital *self*. Thus, we proceed to investigate the role of standard protocols for federating social interactions. Common protocols allow otherwise un-associated systems to work together without prior agreement, avoiding the *lock-in* that comes with the current centralised model. This brings, of course, a new set of challenges to address.

# Chapter 4

## Decentralising the Social Web (and other stories)

Given the large influx of new users into the fediverse, you may be unaware of the tradition that we all vehemently hate each other and refuse to work with each other because the other guy is a twat.

If you use Hubzilla, you hate Diaspora, Ostatus, and RedMatrix.

If you use RedMatrix, you hate Hubzilla.

If you use Diaspora, you hate Friendica, Hubzilla, and RedMatrix.

If you use Friendica, you hate Diaspora.

If you use ostatus, you hate everyone else who uses ostatus.

This is the law of the fediverse, please bear this in mind. When picking fights with random strangers, make sure it's always based on nothing other than their software choices.

This has been a public service announcement on behalf of the fediverse.  
Thank you, and good day.

- Thomas Willingham, decentralised Social Web developer. Posted on Friendica, seen through Mastodon  
<<https://soc.beardyunixer.com/display/0c5b5b901858fce51629a5e267815437>>

# 1. Introduction

The previous two chapters had social-science leanings, in both the literature surveyed and the studies carried out. This chapter marks the beginning of the more technology-focused half of the thesis. It is necessary to define jargon, discuss software specifications, and talk about the details of specific technologies used in particular systems. This is where I try to align my theoretical findings with outcomes that are practical and meaningful for software architects and developers.

First of all, this chapter presents a survey of decentralised social systems: systems which store data about a person, their attributes and activities, and encourage sharing of this data and interactions with others in the network. These systems are analysed along various axes to determine their intended function, fitness for purpose and how they are or were ultimately used and to what extent. The core focus is on treatment of individuals as users of or participants in a system, and I classify systems according to their approach to handling user identities and profile data. I derive modules of decentralised social systems and the contexts in which they are likely to work well and benefit users; I also attempt to identify gaps and common pitfalls in existing theoretical and practical work.

I finish by describing the most recent work (still ongoing at the time of writing) of the W3C Social Web Working Group, which is producing standards for decentralised social interactions on the Web.

## 1.1. Scope

I would love for this chapter to be a complete history of decentralised social networks, but that would take many years to research and write. There exist many partial histories, glimpses into the past through the lens of some masters thesis or an ancient blog post found through archive.org by someone who was There At The Time. These people are even around today, and could be interviewed, if one was so inclined. Efforts towards decentralising social networks have been going on for as long as there have been social networks, and making use of a wide variety of different technologies, such as SMTP, XMPP, peer-to-peer architectures and the Blockchain. Many efforts in academic, commercial or FOSS environments are driven by producing secure private

systems (ie. keeping messages transmitted over the network hidden from all but the sender and recipient), or by producing optimised, fast, highly scalable architectures.

Given the topic of this thesis, I constrain the contents of this chapter in two ways:

- to systems on the *Social Web*; that is, which operate over HTTP, and
- to reports which center the 'user' as a human being, rather than a node; that is I will not cover projects which focus on a purely network architecture perspective.

For broader or differently-scoped surveys of the decentralised social networking space, as well as the motivations for pursuing this work and advantages and drawbacks of various approaches, see [72, 132, 135, 9, 199, 224, 209, 65].

## 1.2. Decentralisation

Decentralisation is a fairly contentious term with different definitions or understandings depending on the background of the person talking about it. In this thesis, I use the definition from [135]:

A system in which multiple authorities control different components and no single authority is fully trusted by all others. Decentralized systems are a subset of distributed systems.

This differentiates decentralised systems from *distributed* systems, which are not necessarily decentralised, in that the latter "may be managed by a single root of trust or authority."

## 1.3. Social systems

I defined the type of social system I am interested in in Chapter 3, and repeat it here for convenience:

Social systems: Web-based networked publics which offer individuals consistent and reusable access to an account which they can customise and use to interact in some form with others in the system.

Other concepts which will be useful throughout this chapter are:

### **Protocol**

A set of possible communication actions between computer systems.

### **Implementation**

A piece of software, possibly designed according to a protocol.

## Instance

An implementation which has been installed and is running somewhere, probably a server. One implementation can have many instances.

## Federation

The joining together of software instances such that activities on one are seen on another (usually by means of a protocol).

### 1.4. Selection method

There are numerous wiki pages and blog posts which list decentralised social systems<sup>lists</sup>, and academic literature reviews, and project reports which reference related work. I followed links to all projects listed (performing additional Web searches for information if necessary) and performed an exhaustive search on the literature through reading reference lists until I was no longer seeing new citations on the relevant topics. I excluded projects which are out of scope; which are clearly abandoned (or have an unknown status) with no public information or documentation; which claim a desire to federate in the future but haven't actually achieved it; and projects which have pivoted in a different direction.

There are a great deal of abandoned projects in this space, spanning more than a decade. It is my hope that one day I or someone else manages to locate their founders and conduct a more thorough post-mortem.

I have, however, included projects which have wound down or are no longer actively maintained if they still provide documentation or blog posts about the work, and ideally running instances are still available, as these are still worthwhile to learn from. As a result, there may be some bias towards abandoned academic projects, since these are fairly well documented in peer-reviewed literature.

lists A Distributed Social Network: You're Doing It Wrong (benwerd.com)  
<<http://benwerd.com/2010/06/04/building-a-distributed-social-network-youre-doing-it-wrong/>>

, Comparison of microblogging services (wikipedia)  
<[https://en.wikipedia.org/wiki/Comparison\\_of\\_microblogging\\_services](https://en.wikipedia.org/wiki/Comparison_of_microblogging_services)>, Comparison of Internet forum software (wikipedia)  
<[https://en.wikipedia.org/wiki/Comparison\\_of\\_Internet\\_forum\\_software](https://en.wikipedia.org/wiki/Comparison_of_Internet_forum_software)>, Comparison of instant messaging clients (wikipedia)  
<[https://en.wikipedia.org/wiki/Comparison\\_of\\_instant\\_messaging\\_clients](https://en.wikipedia.org/wiki/Comparison_of_instant_messaging_clients)>, Comparison of software and protocols for distributed social networking (wikipedia)  
<[https://en.wikipedia.org/wiki/Comparison\\_of\\_software\\_and\\_protocols\\_for\\_distributed\\_social\\_networking](https://en.wikipedia.org/wiki/Comparison_of_software_and_protocols_for_distributed_social_networking)>

, Social networking service (wikipedia)  
<[https://en.wikipedia.org/wiki/Social\\_networking\\_service](https://en.wikipedia.org/wiki/Social_networking_service)>, Social software (wikipedia)  
<[https://en.wikipedia.org/wiki/Social\\_software](https://en.wikipedia.org/wiki/Social_software)>, List of social networking websites (wikipedia) <[https://en.wikipedia.org/wiki/List\\_of\\_social\\_networking\\_websites](https://en.wikipedia.org/wiki/List_of_social_networking_websites)>, List of defunct social networking websites (wikipedia)  
<[https://en.wikipedia.org/wiki/List\\_of\\_defunct\\_social\\_networking\\_websites](https://en.wikipedia.org/wiki/List_of_defunct_social_networking_websites)>, Comparison of social networking software (wikipedia)  
<[https://en.wikipedia.org/wiki/Comparison\\_of\\_social\\_networking\\_software](https://en.wikipedia.org/wiki/Comparison_of_social_networking_software)>, Federated Social Web Platforms (w3.org)  
<<https://www.w3.org/2005/Incubator/federatedsocialweb/wiki/Platforms>>, Projects (indieweb.org) <<http://indieweb.org/project>>

## 2. When is a person not a person?

Identity and reference are a critical part of broader Web architecture [25, 134], and in some systems built on top of the Web, this distinction matters. This topic was the subject of the infamous 'httpRange-14' issue debate, a lengthy discussion (over a decade of W3C mailing lists exist) and a supposed resolution about the *meaning* of an HTTP URI.

Q: When is a person not a person?

A: When they're a document.

HTTP URIs can be used to represent both Web-based documents (information resources, IRs) - which can be transported over HTTP - and everything else (non-information resources, NIRs, like people and places) [233]. The latter, needless to say, cannot be transported over HTTP (yet). In this case, some systems separate *people* from the *documents about* them by using fragment URIs (fragments aren't passed to servers, so the server can only return the resource represented by the URI up to and not including the fragment) or 303 redirects (the server says "sorry this URI identifies a NIR which I can't return, so I have to send you to an IR about it instead").

Not all systems make this distinction however. For some, conflating metadata about a webpage and metadata about a person is simply not a problem, or not one worth solving. However, when we are thinking about federated systems, architectural differences in data models like this may cause interoperability problems. The representation feature of profile-hosting systems which was elicited in the first study of the previous chapter demonstrates some variations in how people and profiles are modeled in various contemporary social systems.

Other ways systems vary their data models are by using different vocabularies and syntaxes. In many cases, open data models for representing people and their activities within social networks have been created over the years independently of software implementations that use them, and are sometimes picked up by separate projects. Common vocabularies, or mappings between them, significantly *improves* interoperability options. The following are well known, and may be referred to as in-use by other protocols or implementations throughout the remainder of this chapter:

- FOAF: An RDF vocabulary with various properties for representing attributes of people, as well as the links between them using the `knows` predicate. A cornerstone of any Semantic Web based social projects [44].
- Relationships: An extension to FOAF to provide more granular relationships than `knows` [73].
- SIOC: An RDF vocabulary for representing online discussions, such as blogs, forums and mailing lists. Complements FOAF [222].
- vCard: A standard for representing attributes of a person that might be found on a business card. A common export format from email clients and mobile address books, and the core of many 'contacts' type applications [226].
- Atom: An XML based standard for representing streams of content from blogs [215].
- ActivityStreams 1.0: An XML or JSON syntax for representing social interactions as activities, which consist of `actor`, `verb`, `object`, and `target`, as well as other more domain-specific properties; widely deployed, in particular as part of the OStatus stack. Predecessor to ActivityStreams 2.0 [12].
- microformats2: A set of classes which can be used with HTML elements to provide semantics about the value of the element. The classes cover types and properties for a core set of social objects (eg. people, organizations, events, locations, blog posts, products, reviews, resumes, recipes) as well as a specification for how to parse them. Built on vCard. [163].
- XHTML Friends Network (XFN): A set of values for the HTML4 `rel` attribute which indicate a relationship between the person who created the link and the person being linked to (eg. "met", "coworker", "parent") [62]. Works with microformats2. This one is not bothered by httpRange-14, ie. my homepage can be friends with your homepage.

### 3. Standards and monocultures

Certain more modular (ie. only do one thing) open standards are used by multiple projects. I describe them here for convenience (references are therefore not inline when they are mentioned again):

- Webfinger: is a discovery mechanism for user profiles using the .well-known URL pattern and `user@domain.tld` style identifiers [159].
- WebID+TLS: an authentication mechanism that uses FOAF files and browser certificates [145].
- OAuth 2: a commonly used authentication mechanisms which uses bearer tokens [146].
- JSON-LD: a JSON syntax for RDF [256].
- Semantic Pingback: a federation protocol which alerts a resource when it has been linked to on the Web [275].
- PubSubHubbub: a federation protocol for subscribing to publishers and delivering content to subscribers [107].
- Salmon: a federation protocol for passing responses to decentralised threaded conversations 'upstream' so the originator of the conversation is notified [217].

Some projects publish the specifications of their systems as protocols that anyone can implement. Projects which publish their protocols and state their intent to have interoperability across multiple independent projects, but in practice have few distinct implementations, include Tent [268] and RemoteStorage [75]. The latter bundles Webfinger, OAuth 2.0, JSON-LD and HTTP REST, together with a thin layer of additional requirements, and is published as an IETF draft [76]. Hubzilla [152] (a fork? of Friendica) publish the Zot! protocol [185], but it isn't clear there is uptake outside of Hubzilla implementations.

Implementations may be assembled from an existing set of more modular protocols. SMOB (Semantic Microblog) [223] and DSSN [109] are based on Semantic Web technologies, and use overlapping vocabularies for their data, but different mechanisms for content creation and federation. SMOB uses RDFa in published content, and propagates updates throughout the network with SPARQL/Update and HTTP POST requests. DSSN and OpenLink Open Data

Spaces (ODS) [154], both use WebID, FOAF, Semantic Pingback and PubSubHubbub. ODS is additionally LDP [200] (for content reading and writing) and LDN [59] (for notifications) compatible.

In some cases, existing sets of protocols have been identified and documented as complimentary. One example is OStatus [228], which combines Atom, ActivityStreams, PubSubHubbub, Salmon and Webfinger. OStatus has several independent implementations listed on its Wikipedia page at the time of writing, including GNU Social (previously StatusNet) and Mastodon, both of which there are hundreds to thousands of instances. Another is Solid [188], which uses LDP, WebID+TLS, FOAF and Web Access Control [258].

Conversely, some projects cultivate monocultures, either by not publishing their protocols at all or by gaining little adoption outside of a single core implementation. In computing, a monoculture is when a group of computers all run the same software. Several FOSS Social Web projects meet this description, with large and active developer communities who all contribute to the same codebase (Pump.io [229], NextCloud [211], Friendica [108], Hubzilla). One particular piece of software is expected to be downloaded and installed on multiple different servers, and different instances can interoperate with each other so it runs in a decentralised manner. This has the advantage that, assuming a straightforward setup process, new instances can be set up quickly and easily, aiding adoption. However, a security vulnerability in the core code base would be present across all instances; there is overhead for developers working on the project to agree its direction and realisation; and focusing on this type of architecture can mean that optimisations for same-software interoperability come at the expense of potentially more open protocols, as is the case with diaspora\* [132]. Monocultures can emerge when one implementation of an open protocol is particularly successful; an example is Known [168], which implements the open standards Micropub [221] and Webmention [248] but also has an easy setup process, good community support, and a for-profit company behind it which has driven wide adoption through managed and white-labelled instances for educational institutions, massively increasing the dependence upon a single codebase.

Friendica is based on DFRN [185], but has code which bridges to a wide variety of decentralised protocols, including OStatus and diaspora\*, as well as the APIs of centralised services. Systems like diaspora\* and GNU Social are moving in this direction too, desiring to federate with instances of software implementing totally different (or, sometimes overlapping) protocols.

## 4. Social (and) personal datastores

Research into Personal Data Stores (PDS) spans back a long way in the Personal Information Management Systems (PIMS) field. Once a PDS can communicate with another PDS, grant access to a third-party, or PDS owners are able to spread their data around across multiple separate systems or instances, we find many parallels with decentralised social networking [209]. MyDex [205] is a commercial offering which provides software to organisations who wish to securely hold data on behalf of their customers. With an emphasis on consent-based data sharing through their proprietary API, datastore owners can authorise trusted third-parties to access their data in exchange for a useful service. indx (previously known as WebBox) [166, 286, 288] is similar, but based on Web standards rather than proprietary technology; it is expected that individual users will host their own instances rather than relying on a provider (and an academic rather than commercial project). PrPl [247] on the other hand indexes personal data aggregated from other services, and provides a 'Personal Cloud Butler' to negotiate data sharing with other parties who may be able to provide something in return. Data is indexed with RDF, but they created a new query language (Socialite) for federating and searching across instances. The team behind PrPl outline a deployment plan which uses devices already existing in people's homes such as media centers to allow self-hosting with low setup cost.

NextCloud (the actively maintained fork of ownCloud) is a free and open source PDS which lets users manage files and media as well as data such as calendars and contacts. Social features such as file sharing, tagging, and commenting are possible across separate instances of NextCloud through the OCS (Open Collaboration Services <<http://open-collaboration-services.org/ocs/>>) federation API. NextCloud provides hosted services, as well support for self-hosted instances. OpenLink Open Data Spaces (ODS) is similarly hosted and for similar purposes; it is developed by a commercial organisation, and instances federate using existing open standards.

The RemoteStorage protocol provides users with the opportunity to trust their data to a third-party provider, or to set up an instance of the software on their own server. It is a grassroots rather than commercial effort, and intended as a generic personal data store. Solid is an open source project with its roots in academia; multiple server implementations exist, and are known as PODS

(Personal Online Data Stores). The data stores are generic, with little logic built into the servers, and use a combination of existing open standards with small specialisations to communicate with other instances by means of smart client applications and access control rules.

The data stored in a PDS may be distinctly personal, shared with third-parties only in specific contexts for a particular purpose (eg. sharing medical data with a doctor's office); or it may be inherently social, expected to be broadcast on creation to at least one other person and possibly subject to further reshaping, sharing, and propagation throughout a network of people. Many datastores are optimised for one or the other of these scenarios and may place constraints on their users accordingly. Others, in particular RemoteStorage and Solid, attempt to meet requirements for both types of use and encounter different implementation and usability challenges as a result.

## 5. From status updates to microblogging, and beyond

Many systems are designed with the emphasis on social networking over personal data storage. Even systems which emphasise privacy or access control are still inherently about sharing content through a network. Several of these attempt to replicate the functionality of centralised SNS, and a common focus was around the 'status update' feature of the likes of Facebook. Twitter, with its continuous feeds of short posts, changed the discourse somewhat to 'microblogging'. SMOB augments ordinary text publishing with machine-readable metadata, tags as links to other concepts on the Web, and data portability through RDF. [223] describes microblogging as a "hybrid of blogging, instant messaging, and status notifications."

The original 'Twitter clone' was StatusNet, an implementation of the OStatus protocol which is now known as GNU Social. Mastodon is another OStatus implementation, and pump.io has the same origins. Friendica, Hubzilla, diaspora\* [82] and Known provide similar user interfaces for posting to and reading from a timeline of notes, articles and messages from the profiles you friend or follow, but each use different (combinations of) protocols to pass content between instances. The architecture of Tent is also based on passing around 'posts' (pieces of text with various metadata).

Now, many centralised services offer different types of updates, like check-ins, or multimedia sharing experiences, and decentralised social networks are somewhat lagging behind on that front. The most advanced in terms of media is MediaGoblin <<http://mediagoblin.org/>>, an implementation of pump.io specifically oriented around sharing images and video. Known also provides different 'post types' such as audio and location. DSSN is similar to SMOB in that it is based on Semantic Web technologies, however the reference implementation is integrated into Semantic Media Wiki, so users see changes to wiki pages federated between instances, rather than status updates.

The types of content which social networks enable users to create is pertinent to understanding the self-presentation possibilities of profiles within certain systems. Decentralised systems which cannot accommodate a wide variety of types of content or interaction may in the end not meet users' needs for

expressiveness. Leaning on extensible technologies and protocols could be one way to ease customisation of basic systems by communities with particular needs.

## 6. If the app fits

A common focus of decentralised Social Web projects is decoupling application logic from data storage. Whilst most of the protocols have client-to-server APIs which would technically allow this architecture (like Pump.io), they don't necessarily focus on it as a strength; others however are designed specifically around this model.

RemoteStorage servers are developed by the same community behind Unhosted web apps [280], at the core of which is the idea that applications should run entirely in the browser so that the entire source code of the application is made available to the user; they can download it and run it locally if desired, and have no need to depend on a third party to run the service. This alone has limited use as data cannot be persisted, so coupling these kinds of Web apps with RemoteStorage servers allows users to authenticate with their own datastore and have the application operate on their data, without the application developer processing that data at a third-party, potentially untrusted, server. The Unhosted community provide several sample applications and a JavaScript library to help with app development.

Solid works in a similar manner, with the explicit aim to reduce the logic of a server, making them simple storage devices which can write data and serve it up again. Applications are expected to be smart, including managing access control for resources. Data is stored on the servers as RDF, so applications using shared vocabularies can reuse the same data, giving users even more freedom of choice. Again, the core Solid team are developing sample applications and helper libraries to encourage development of applications by others. OpenLink ODS has several protocols in common with Solid, and there is at least one clientside application ([dokieli <https://dokie.li>](https://dokie.li)) which works with both Solid PODS and ODS datastores [282, 188]. PrPl is similar, describing the potential of empowering users through choice of applications.

NextCloud encourages community development of applications which operate on user data, but these exist as plugins to the NextCloud architecture so they're a little more tightly coupled. SMOB, indx and Tent also purport to prefer this kind decoupled-application of architecture, however I was unable to find samples or galleries of applications which have been developed to work with the accompanying servers, and these projects are now largely retired.

There is some overlap between the mindset behind the PDS-type systems and the decoupled-application architectures. Where client-to-server APIs exist for systems like Pump.io, diaspora\* and GNU Social, these are typically employed to build applications on different platforms (ie. desktop and native mobile clients for different operating systems) rather than with the idea that multiple Web apps should exist to do different things with the same data stored on someone's instance. This is perhaps due to the broader notion of what kinds of data are stored on a server for PDSs compared to more focussed microblogging and social sharing sites.

## 7. The Social Web Working Group

[209] advocates for open standards, not as the solution to decentralisation problems, but as a prerequisite for interoperability success. Standards "needs serious developer commitment as well as the involvement of standards bodies with significant authority."

Ultimately, the Social Web Working Group has produced the following recommendations:

- ActivityStreams 2.0 Core and Vocabulary: A JSON data model and syntax for social content and interactions, with extensibility based on JSON-LD.
- ActivityPub: An API for creating content, delivering notifications, and managing common side-effects of interactions within social networks, based on ActivityStreams 2.0.
- Linked Data Notifications: An API for delivering and reusing generic notifications, based on JSON-LD.
- Micropub: An API for creating content, based on microformats2.
- Webmention: An API for delivering notifications when a resource on the Web refers to the URL of another.
- WebSub: An API for managing and fulfilling subscription requests to Web content (formally known as PubSubHubbub).

And the following Working Group notes:

- JF2: An alternative JSON syntax and vocabulary for representing social objects, based on microformats2.
- Post-type Discovery: An algorithm for converting between implicitly typed social objects (like JF2) to explicitly typed ones (like ActivityStreams 2.0).
- Social Web Protocols: A guide to the specifications produced by the group, including how they overlap or complement each other.

### 7.1. Implementations

Each of the recommendations has received a number of implementations during the Working Group's lifetime. Some were formally submitted to the Working Group as implementation reports whilst the specifications were undergoing review. This section briefly describes the implementations to date, and notes relationships to work previously described in this chapter where applicable.

ActivityStreams 2.0 has a number of publishing and consuming implementations by members of the Social Web Working Group and the Web Annotations Working Group (who rely on one portion of the AS2 model for their Annotations Vocabulary). There is ongoing work in upgrading ActivityStreams 1.0 implementations to AS2.

Webmention and Micropub have a number of implementations from the IndieWeb (<https://indieweb.org>) community, which is where these specifications were incubated prior to being standardised by the Working Group. The implementations span a variety of different programming languages; many are part of people's personal websites, but there are also implementations in the form of helper libraries and plugins for popular blogging or content management systems (like Wordpress).

ActivityPub has been implemented to some degree by several members of the Working Group, and integration into pump.io and various OStatus based systems are in progress.

Linked Data Notifications has also been integrated into personal software projects of Working Group members, as well as a number of academic projects or commercial to do with dataset and resource management. Applications and helper libraries have also been created as part of the Solid project. Existing Linked Data Platform servers can by default serve as one part of the LDN protocol.

WebSub has received implementations from the IndieWeb community, and has semi-compatible historical implementations from when it was PubSubHubbub.

## 8. Discussion

Here we uncover common modules of decentralised social systems from previous discussion, and look at where there are overlaps and things that stand out.

**Identifiers:** Most of the systems mentioned use URIs to identify and locate individual resources (like blog posts). Users of the systems, or their accounts, are identified by a combination of Webfinger URIs and HTTP URLs. A useful point of alignment between different systems might be a standard for mapping between Webfinger URIs and HTTP URLs for user profiles.

**Data storage:** Users of decentralised systems must be acutely more aware of the location(s) of their data. When faced with choices between multiple possible third-party providers (datastore providers or particular hosted software instances), or self-hosting (either on one's own local hardware, or a rented webhosting service), ones personal data gains a sense of concreteness, somewhat divorced from activities which might create the data, of interacting and socialising online.

**Trust:** In most cases, instances federate by passing data from one server or PDS to another. Unless the data is end-to-end encrypted and able to be interpreted by authorised *clients* only, users ultimately need to trust the servers of their entire social network in the cases where their posts are broadcast to all of their followers, and their followers responses are broadcast to all of *their* followers. This creeps into the territory of imagined vs. actual audience, which individuals might not be fully aware of.

**Content:** Content and activities being passed around between different instances and implementations of protocols means that it is likely to be seen in different user interfaces, and as part of different streams of other data, and even perhaps with additional affordances or missing features. Systems which mix and match functionality of different protocols, and are extended or otherwise customised, will have a significant impact on how people interact with and consume content from others, and in a way that will be fairly unpredictable to the content publisher.

In the next chapter we take a deeper look at the Social Web Protocols and the context in which they were developed.

# Chapter 5

## Standards for the Social Web

## 1. Introduction

Many of the systems discussed in the previous chapter have proven foundational for ongoing efforts to create standards for decentralised social interactions on the Web. Previously we gave an overview of the standards being produced by the W3C Social Web Working Group; now we discuss in more depth notable decisions and debates of the group, describe the resulting standards in more detail, and outline how to actually go about building decentralised social systems which empower self-presentation using these standards. The contributions of this chapter are as follows:

- A critical analysis of contemporary standards for decentralised social interaction on the Web, taking into account social dynamics of collaborative projects and the W3C consensus model, as well as the technical considerations.
- A characterisation of the problems being solved by the Social Web WG, and how these relate to the more specific problem of online self-presentation, by means of the conceptual framework from chapter 3.
- A technical primer for the work produced by the group (published by the W3C as a Working Group Note: *Social Web Protocols*).
- Prototype implementations of standards produced by the group, and a report on their interoperability with implementations produced by others.

This chapter brings together the qualitative research from earlier with concrete technical outcomes in the form of protocol designs. The work of the Social Web Working Group is in effect a case study for designing decentralised Social Web systems, but what is presented here is more than a survey or observational study. Since I was first a member of the Working Group, and then the W3C Team Contact, I was immersed in every part of the decision making and contributed in some form to all of the specifications produced. The *Social Web Protocols* document contributes a deeper understanding of the various protocols, and importantly how they can complement or contradict each other. This document is particularly useful *because* of the complicated social dynamics of the group, and produced as an introductory piece for developers entering this space anew. Technical decisions that have been made by the Working Group over the past (almost) three years were not made in a vacuum, nor dictated by cold logic, but rarely backed up with truly meaningful data. Most decisions were made on the - admittedly well-honed - gut instinct of experts, data from small samples, and

anecdotal evidence. The specifications that resulted were influenced by this, so it is important to examine the development processes. *Social Web Protocols* contains fine technical details of the Working Group's recommendations, which are important for a complete picture and analysis of the current cutting-edge of decentralised Social Web standards. An overview of Social Web Protocols is presented in this section, and the specifics can be found in Appendix SWP.

## 2. Standards and self-presentation

We have in previous chapters established that online self-presentation is greatly more complex than listing attributes and a photo on a Web page. For decentralised systems to emulate the multitude of self-presentation possibilities provided by mainstream centralised systems today they must coordinate across a range of potential user activities and interactions. Common protocols enable disparate systems to communicate without any prior arrangements. Specifications describing such protocols must be agreed upon and published in such a way that makes them easy to find, and appear authoritative enough that developers of decentralised systems are persuaded to implement them.

This section documents and furthers the standardisation process, as part of the work of the W3C Social Web Working Group. Chapter 4 includes a survey of the specifications and their implementations at the time of writing; here I provide a behind-the-scenes look at and contribution to their development.

### 2.1. Standardisation as Context

The benefits and costs of standardisation are a prominent socio-technical factor [209]

The standards developed by the Social Web Working Group will be used as the basis for systems which incorporate social features, and as such, create the opportunity for users of the system to present some aspects of their personhood in an online space. This chapter goes into detail about the non-technical parts of development of these standards.

The reason for this is as follows: the formation of the Working Group and its charter design; the individual members of the Group and their particular interests and experiences; and the processes of the W3C, all serve to make up part of the **context** (one of the 5 Cs from Chapter 3) of any systems built from these standards. This is an example of things to analyse, or at least take note of, with respect to the industrial or organisational context in which users of social systems are engaging in self-presentation.

### 2.2. The standardisation process

Once a group is formed and participants are in place, the W3C has many processes in place to facilitate standards development. These processes have ramifications on the end result of work produced by Working Groups, so I will outline key processes here.

Specifications advance usually over the course of one to two years, through a number of stages of increasing stability, to bring them to a final status of *REC* (recommendation). Each stage (see also [193]) is designed to elicit development, input, and review from different qualified parties to iron out bugs, correct omissions, and generally make sure the specification will do what it is intended to do. While direct input is limited to Working Group members, specification development is carried out in public. Each date-stamped draft is online for anyone to review, mailing list archives, meeting minutes, wiki pages, and (at least in the case of the Social Web Working Group) IRC logs are publicly visible. Working Groups tend to take public comments over a specific mailing list, or as GitHub issues, and are obliged to be responsive and reach a considered resolution on how to handle all feedback so that commenters feel heard.

Specifications are maintained as Editor's Drafts (ED) throughout their life cycle. An ED is the most up to date version of the specification, and updated at the editor's discretion. Working Groups do not have the authority to publish specifications unsupervised. Each Working Group is supported by one or two W3C employees (Team Contacts), and at each transition from one specification maturity level to the next, a request is sent to the W3C Director, who reviews the relevant information, checks that the Working Group have been handling feedback from commenters appropriately, clarifies any points of confusion, and grants or denies the request.

The first formal iteration (hosted at the W3C domain) is the First Public Working Draft (FPWD). An ED need not be perfect, or even complete, but when it is sufficiently outlined the Working Group participants vote to publish the FPWD. This is the first stage of the Working Group committing to progress the document towards recommendation. As the specification is discussed and implemented, and feedback comes in, features are added, removed and refined. After each batch of major changes to the ED, the Working Group may vote to publish updated Working Drafts (WD). WDs are essentially official snapshots at particular points in time. As a specification becomes stable (ie. it receives fewer and fewer major changes) the Working Group reaches out further to solicit 'wide review' from relevant communities. These may be outside the W3C as necessary,

but there are specific groups inside the W3C who are expected to review all specifications along particular dimensions; namely: security and privacy, internationalisation, and accessibility.

When the specification is sufficiently stable, and wide review has been achieved, the Working Group may vote to advance to Candidate Recommendation (CR). The CR phase lasts for a minimum of four weeks. This commences a broader call for implementations from outside of the W3C, begins the window in which W3C members must disclose patent conflicts, and prompts W3C Advisory Committee members to review the specification. If major (non-editorial) changes are made to the specification during this phase (which is not uncommon as a result of third-party implementation feedback), then a new CR must be published, which restarts the four week time period. During this time, the editors and their collaborators should be polishing up official test suites, and soliciting implementation reports. The specification can advance to Proposed Recommendation (PR) when it meets a CR exit criteria previously defined by Working Group consensus. In the case of the Social Web Working Group, specifications are expected to have tests and reports for at least two independent implementations of each feature of a specification (where 'feature' is defined per specification). During PR, which must also last a minimum of four weeks, Advisory Committee representatives are re-prompted to review the spec. This is the last time during which anyone can make a Formal Objection to the specification's progression, or raise patent conflicts. *Finally*, if enough positive Advisory Committee reviews are received, the W3C Director approves the specification to transition to REC. It is carved in stone.

Why am I telling you all this?

This is an example of organisational processes having impact on technology design long before the technology is in the hands of end users. The specifications of the Social Web Working Group were not only moulded by their editors and Working Group participants, but reshaped and influenced by W3C staff and by representatives of paying W3C Members who weren't participating in the Group directly. Specifications were poked and tweaked by other Working Groups who do not specialise in the Social Web (most contentious input came from the Internationalisation (i18n) Working Group), as well as critiqued by complete outsiders at every step of the way.

Most specification editors in the Social Web Working Group were invited experts, and thus not paid by an organisation for their contributions. They were working on these specifications, attending weekly telecons, and often quarterly face-to-face meetings, on their own time, and own dime. Editors are also burdened with test suite development; no small task. The W3C process imposes structure, deadlines and deliverables to the specification development process that may be missing (or certainly different) were the specs advanced elsewhere. These deadlines and review processes ultimately affect what is included in a specification, and what is removed. Smaller specs are easier to review, easier to test, and therefore faster to progress. This tended to mean that when in doubt, features were dropped or marked as 'at risk' rather than have them hold up progress. In particular, ActivityStreams 2.0 was brutally trimmed down over the years, and requests for additions to the vocabulary were automatically rejected after a certain point for fear of slowing things down. I wonder how these exclusions will impact future software designed around AS2.

Something else worth bearing in mind is that for all of this *process*, it is commonly held that most 'regular Web developers' don't know about or don't care about (or both) the difference between the different maturity levels of W3C specifications, or even the difference between a Recommendation and a Note. This makes it fairly easy for developers to be implementing software on the basis of an out of date document, or giving weight to a protocol design that was ultimately rejected or even unfinished. Not everything with the W3C stamp on has been fully thought through or passed quality tests, but not every developer realises this.

### 2.3. The Social Web Working Group charter

Technical specifications, at least those produced by the W3C, are intended for software developers. A mark of the success of a standard is when multiple developers, who have no knowledge of each others' activities, can independently implement the specification into code and have their systems interoperate successfully.

Interoperability: ability of a system ... to work with or use the parts or equipment of another system - *Merriam Webster Dictionary*

Contributions to W3C standards may be made by individuals representing themselves (if invited and approved by Working Group chairs), but more so by representatives of organisations which pay for membership to the W3C. As the

Web is an open platform on which anyone can build, there is a lot of space for many ways to solve the same problems. This is a virtue in that it promotes innovation and competition amongst Web services, but a problem if it results in technical fragmentation of solutions, whereby end users are forced to choose one and forgo (or uncomfortably juggle) interaction with others (remember the 'browser wars' of the 1990s and early 2000s? [308]). Organisations join the W3C so that their interests may be represented as they collaborate to produce standard ways of interacting with Web technologies in order to reduce the negative impact of technical fragmentation on end users.

As has been raised on multiple occasions by this thesis, the Social Web is presently in a state of technical fragmentation [132]. End users, also known as people or human beings, are living with the effects of this on a daily basis. Beyond being a mere inconvenience (not being able to port one's friends from Twitter to Facebook), the competition between social platforms has developed in such a way that people are *locked in* to services. Once one depends upon a particular social networking platform for communication and content creation it is almost impossible to change provider. Platforms like this have come to be known amongst decentralisation advocates as *silos*, in that they do not let information flow in or out [313].

Silo: a trench ... usually sealed to exclude air and used for making and storing silage - *Merriam Webster Dictionary*

These closed systems bring a plethora of social, cultural, political and economic issues, which have been touched upon at various stages elsewhere in this thesis and we will not detail further here. The *Social Activity*<sup>act</sup>, which includes the Social Web Working Group, was created in W3C with an eye to promoting interoperability between social systems, and breaking down some of the walls of silos.

The focus of the Social Activity is on making "social" a first-class citizen of the Open Web Platform by enabling standardized protocols, APIs, and an architecture for standardized communication among Social Web applications.  
- W3C Social Activity <<https://www.w3.org/Social/>>

Objectives of the Social Web Working Group were to produce standards for:

- a syntax and vocabulary for describing social data;
- an API for reading and writing social data (create, read, update, delete);

- a federation API for passing social data between disparate systems (subscriptions and notifications).

<sup>act</sup> a W3C 'Activity' is a framework for clustering related Working Groups and Interest Groups together.

## 2.4. Working Group participants and audience

Working Groups may be chartered with the agreement of 5% of the W3C membership [196] and the Social Working Group was convened in July 2014<sup>wg-me</sup>. One of the W3C Members which helped found the Working Group was the Open Social Foundation, which was a collaboration between several large companies and expected to use their influence to drive participation in the Working Group. The Open Social Foundation dissolved upon the beginning of the Social Web Working Group, handing its assets to W3C [296].

It is noteworthy that no major social networking companies are members of the W3C Social Web Working Group. Big companies who joined were those motivated primarily by producing social standards for use in business. Many organisations use proprietary, and often in-house, social networking platforms for their employees to communicate, organise, and share information. The benefit of standards in this scenario is to enable inter-organisational social interaction, to better smooth partnership and other business processes. This outlook set the tone for much of the early discussions in the group.

However, over the past two and a half years, active participation in the group has dwindled to such an extent that the group consists of mostly invited experts<sup>wg-ie</sup>. As time progressed, representatives of organisations interested in social business were reassigned and their participation in the group diminished. This dramatically (although it wasn't noticeable at the time) altered the tone of the group.

Several invited experts currently in the Group are representing their own interests, passionate about social standards they can implement for themselves. Others are from open source or free software communities, and want existing or emerging projects to interoperate with regard to social interactions, as a way to add value for users but also to uphold principles of their users' freedom to choose to take their data elsewhere. These two groups are by no means mutually exclusive.

How can Social Web standards possibly be adopted widely enough to have any impact without the support of major social networking platforms? An informal hypothesis by various members of the Working Group is that there are many more (e-)industries who can benefit from social networking than the ones who make advertising- and data-mining-supported social platforms. Such businesses either produce tailored in-house solutions to the very specific corner of social media they need (think Amazon reviews) or embed functionality provided by a major platform (think adding a Facebook Like button). Some have moved all of their publishing and customer interaction to one or more mainstream social networking platforms (some newspapers and magazines; restaurants and cafes). Yet other services have been designed from the ground up to depend on a major platform for the provision of their service at all (marketing and customer analytics software; many games).

Complete dependence is risky, as has been shown on countless occasions when, for example, Twitter changes its developer Terms of Service so that existing third-party applications are suddenly in violation [302, 162, 10, 11]; or Facebook changes its API, resulting in an endless cycle of unpredictable code maintenance [19, 225].

Depending on centralised platforms for a customer base results in either excluding non-users of the chosen platform, or having to manage a presence on multiple platforms. The circle continues with the availability of services designed to help manage broad social media presence over multiple platforms; these are in the category of social-platform-dependent business models.

Organisations which do not want to depend on existing services tend to have to build their own, creating a burden of storing data (securely and privately) on behalf of their customers when this may not even be central to their business process. Standards which allow their customers to point to a personal data store that they already have would be an advantage in this case.

It is thus organisations for which 'social' is an enhancement (albeit in many cases a critical one) rather than the core of their revenue stream that may be persuaded to invest time in implementing open Web standards. A result of this adoption can contribute towards normalising expectations of more decentralised social interactions from the perspective of end users, too. One could argue that situation with the Social Web is at a maturity level analogous to that of the software industry in the decade before the free software movement was re-launched in the early 1980s.

Unfortunately there has been low participation from this category of organisation as well, so the Working Group has not necessarily managed to appreciate their needs.

In summary, the Social Web WG specifications are targeted at:

- individual developers, hobbyists, hackers;
- open source projects with principles around data ownership;
- companies which enhance their core offering with social functionality;
- organisations which produce social systems to facilitate business operations amongst employees.

Further, with increasing public awareness of the privacy and freedom implications of handing all data to a select few organisations, these organisations may seek new business models (beyond selling consumer data) and innovate on other fronts in order to retain user trust.

wg-me One of these was the University of Edinburgh, ie. myself representing my own curiosity at the time, having no idea how deep I'd get. That's right, I joined W3C WGs as a hobby.

wg-ie Invited experts are individuals who cannot pay W3C membership fees but have valuable insight to contribute to a WG. They apply to join, and must be approved by WG chairs.

## 2.5. API Requirements

An early activity of the Working Group was to write 99 "user stories" describing actions that people should be able to carry out using systems based on standards produced by the group. The goal of this was to focus efforts on the most important standards to be worked on, to meet the needs that appeared most critical to members of the group. Most group members contributed one or more user stories, and they varied based on the perspective of the individual or organisation they represented. The group then voted (-1, -0, 0, +0 or +1) on every user story, and selected a top eight to prioritise.

My subsequent contribution was to derive API requirements from the shortlisted user stories.

### 2.5.1. Process

1. Read each user story and straightforwardly list required functionality.

2. Cluster related functionality, find overlap between stories.
3. Label the clusters.
4. Organise labels into shortlist of requirements.

### 2.5.2. Results

The simplified story requirements and their respective labels are listed in table 12. The labelled requirements, with descriptions derived from the requirements of the combined user stories, are as follows:

- **Read content** (read): social content should be consumable in a standard way, may be restricted according to the permissions of the viewer, and should be distinguishable by type, author, and associations with groups or other content.
- **Publish content** (pub): users should be able to create, update and delete social content, including metadata, and relationships with other users or content.
- **Notifications** (notif): users are alerted when content is created that somehow targets them (ie. as recipient or subject).
- **Subscribe to content** (sub): users can request notifications about updates certain streams content, eg. by a certain user, posted to a certain group; and users can undo a subscription.

Table 1. Table 12. The top eight Social Web Working Group user stories

| Story name                    | Details   | Labels  |
|-------------------------------|---|---|
| Social Web Acid Test (SWAT0)  | <ol style="list-style-type: none"> <li>1. user A takes a photo of B from their phone and posts it</li> <li>2. user A explicitly tags the photo with B</li> <li>3. B gets notified that they are in a photo</li> <li>4. C who follows A gets the photo</li> <li>5. C makes a comment on the photo</li> <li>6. A and B get notified of the comment</li> </ol>   | <ul style="list-style-type: none"> <li>• Publishing a photo. pub</li> <li>• Push notifications.</li> <li>notif</li> <li>• Subscription to a user. sub</li> <li>• Commenting on / replying to a post. pub</li> </ul>   |
| User posts a note             | <ol style="list-style-type: none"> <li>1. Eric writes a short note to be shared with his followers.</li> <li>2. After posting the note, he notices a spelling error. He edits the note and re-posts it.</li> <li>3. Later, Eric decides that the information in the note is incorrect. He deletes the note.</li> </ol>  | <ul style="list-style-type: none"> <li>• Publishing text content. pub</li> <li>• Updating published text.</li> <li>pub</li> <li>• Deleting published text.</li> <li>pub</li> </ul>  |
| Reading a user's recent posts | <ol style="list-style-type: none"> <li>1. Iris finds a comment by Sam on one of her photos funny. She'd like to read more posts by Sam.</li> <li>2. Iris reads the latest notes by Sam. She also reviews his latest photos.</li> </ol>  | <ul style="list-style-type: none"> <li>• Reading comments / replies. read</li> <li>• Reading posts from a particular user. read</li> <li>• Viewing photos from a particular user. read</li> </ul>   |
| Following a person            | <ol style="list-style-type: none"> <li>1. Delano meets Beth at a company meeting. They are both user interface designers. He finds her ideas interesting.</li> <li>2. Delano follows Beth on their company social network.</li> <li>3. Beth posts a photo from a whiteboarding session at a company retreat.</li> <li>4. Delano sees the photo in his inbox stream.</li> <li>5. Ted, Delano's coworker, wants to find new people to follow. He looks at the list of people that Delano follows. He finds Beth in the list, reads her stream, enjoys it, and decides to follow her, too.</li> <li>6. Beth posts frequently. Delano is having a hard time reading his inbox stream because Beth's activities drown out everyone else's. He stops following Beth.</li> </ol> | <ul style="list-style-type: none"> <li>• Subscription to a user. sub</li> <li>• Publishing a photo. pub</li> <li>• See a user's subscriptions.</li> <li>sub</li> <li>• Read posts from a particular user. read</li> <li>• Unsubscribe from a user. sub</li> </ul> |

---

The first two columns are the work of members of the Social Web Working Group (see wiki history for attribution [swwg-user-stories]) and the third column is my own work.

|                        |  |  |
|------------------------|--|--|
| Read social stream     | <ol style="list-style-type: none"> <li>1. Jake is bored at work. He checks his social inbox stream to see what his friends, family, and coworkers are up to.</li> <li>2. Jake sees in his social stream a note by Tammy about her new apartment. Tammy is his friend.</li> <li>3. Jake sees in his social stream a photo by Edith from her concert last night. Jake follows Edith but Edith doesn't know Jake. Edith has thousands of followers.</li> <li>4. Jake sees in his social stream a video from Damon. Damon and Jake are both in the "Boxing Fans" group. Damon posted the video to the group.</li> <li>5. Jake sees in his social stream a sound file from Carol. Carol is Jake's wife. The sound file is a reminder to stop for groceries after work. Carol posted the sound file only for Jake.</li> <li>6. Jake sees in his social stream that his friend Tammy has added a new friend, Denise. Jake remembers Denise from high school.</li> <li>7. Jake requests to add Denise as a friend, too.</li> </ol> | <ul style="list-style-type: none"> <li>• Subscribe to a user (one way).</li> <li>sub</li> <li>• Join a group</li> <li>~= subscribe to group content.</li> <li>sub</li> <li>• Publish a video. pub</li> <li>• Publish audio. pub</li> <li>• Private sharing. read</li> <li>• Publish subscriptions.</li> <li>sub</li> </ul> |
| Adding recommendations | <ol style="list-style-type: none"> <li>1. James maintains an application for managing architectural designs</li> <li>2. Maggie, a senior architect would like to recommend many of the better designs</li> <li>3. James uses an existing liking service which allows him to post any recommendations, to provide this</li> <li>4. This service also allows James to present existing likes for the design in question</li> <li>5. Maggie gets to like specific designs, and her followers see these as do viewers of these designs</li> <li>6. James achieves this with a simple inclusion on the associated web page, but could have chosen a more detailed integration if greater control was needed over the user interface</li> </ol>  | <ul style="list-style-type: none"> <li>• Liking / recommending a post. pub</li> <li>• Subscribe to a user. sub</li> <li>• View likes of a post. read</li> <li>• Posting from one interface to another system (scope?)</li> </ul>   |

---

The first two columns are the work of members of the Social Web Working Group (see wiki history for attribution [swwg-user-stories]) and the third column is my own work.

|                  |  |   |
|------------------|--|---|
| Adding comments  | 1. Maria, an IT Architect, has been tasked with encouraging better collaboration on the development of her company's Industrial Processes  | • Reply to posts. pub                                   |
|                  | 2. As these Processes are tightly controlled (though generally visible) an associated discussion and evangelisation capability is required   | • Notifications. notif                                  |
|                  | 3. Maria integrates with an existing comment capability to store and retrieve comments rather than redeveloping  | • Subscribe to a user. sub                              |
|                  | 4. May-Ling sees the comment area with the Processes and suggests changes, as she herself does not have rights to update   | • Posting from one interface to another system (scope?) |
|                  | 5. The Process owner gets a notification that someone has commented on this Process  |   |
|                  | 6. Followers of both the Process owner and May-Ling will see this comment event  |   |
|                  | 7. Maria achieves this with a simple inclusion on the associated web page, but could have chosen a more detailed integration if greater control was needed over the user interface |   |
| Direct messaging | 1. Kyle wants to tell Lisa something privately.  | • Publishing text content. pub                          |
|                  | 2. Kyle sends her a message that no one else can view.   | • Private sharing;                                      |
|                  | 3. Lisa is notified she has a message.   | specifying audience. read                               |
|                  | 4. Lisa reads the message and responds privately.  | • Notifications. notif                                  |

---

The first two columns are the work of members of the Social Web Working Group (see wiki history for attribution [swwg-user-stories]) and the third column is my own work.

## 2.6. Competing specifications

As this section is concerned with providing insight into the process that resulted in the outcomes of the Social Web Working Group, I will now provide background for a key technical direction that was taken. The work of the Group commenced with some guidelines about deliverables in the charter, but not a specific list; this was something the participants had to figure out in order to meet the previously described API Requirements.

The technologies promoted by active participants of the Working Group settled into roughly three categories: microformats-based <<http://microformats.org/wiki/microformats2>>, JSON-based <<http://json.org/>>, and RDF-based <<https://www.w3.org/2001/sw/wiki/RDF>>, with some small points of overlap. The proponents of JSON-based technologies tended to come

from the Open Social Foundation background, with experience in open source social systems designed to support multiple users per server/instance. The microformats supporters brought the perspective of individuals running their own personal implementations of social systems, federating with other individuals on a small scale. The RDF advocates brought experience with large-scale data modelling, open data publishing and data integration, often in an academic or proprietary business context. Producing JSON-based protocols was a requirement of the Working Group charter; the other technologies had the potential to still meet this requirement through the microformats2 parsing algorithm <<http://microformats.org/wiki/microformats2-parsing>> in the former case, and JSON-LD <<https://json-ld.org>> in the latter case. These three perspectives are valuable and in theory complementary, but in practice caused drawn out arguments, ideological disagreements, and frequent misunderstandings.

The participants set about bringing their preferred solutions up to standard, and submitted them to the Working Group as Editor's Drafts. I started work on documenting the commonalities between the specifications with the intent that we'd manage to converge them into a single optimal protocol (this is the origin of Social Web Protocols). After many months of work, many hours of telecons, and several face-to-face meetings, technical disagreements and general unwillingness to compromise (all captured for posterity in meeting minutes, mailing list discussions, and GitHub issues) resulted in the convergence effort stalling.

Even leaning on participants' past experiences (see Chapter 4) of prior systems, this was still relatively untrodden ground, so it was never clear (to me, and other more neutral parties) which technology or ideology was most likely to succeed. In most disagreements, it was never obvious that one party was right and the other was wrong. Eventually the Working Group as a whole acknowledged this, and resolved to move forward *all* of the prospective standards separately, and to stop trying to force convergence.

This decision was controversial in the eyes of other members of the wider W3C community who were not members of Social Web Working Group, and potentially confusing for developers looking for *the* solution to decentralised Social Web protocols. However the effect was that specification editors stopped arguing about why their way was better, and were free to move their work forward without needing to defend their decisions from people who fundamentally disagreed with their underlying assumptions. Specification

editors who had accepted their differences began to help each other, and to share findings and experiences (because they are all working towards the same end goal, after all).

Ultimately the Working Group has produced a suite of specifications that is not as coherent as it might have been had the participants been united around fundamental technical decisions. However, we also have a better understanding of how to bridge these different perspectives (in terms of writing code, as well as in terms of discussions) than we would have if one perspective had dominated the group and the others had continued their own work elsewhere. Whilst the "glue code" approach is advocated by [209], it's too early to tell if this means we increase the chance of these standards being adopted (because we have something to please a broader spectrum of developers out there) or decrease the chances (because we look indecisive and nobody will take the outputs seriously). Similarly, if we see wide uptake of these standards, will we get three (or more) fragmented decentralised Social Webs because developers are opinionated, and writing bridging code is too complicated; or will the efforts towards bridging the approaches be taken up so that completely different protocol stacks can interoperate on some level at least?

The trials and tribulations of the Social Web Working Group have the potential to have far-reaching consequences for the future of the decentralised Social Web, and as such, on how people are able to present and express themselves online. Even if the Working Group's final outputs are not an ultimate solution, the authority given by the W3C standardisation process means that they will at least be referred to, and probably built upon, as the decentralised Social Web grows.

## 2.7. Social Web Protocols

Having covered the context of the specifications' development, we now dive into their actual functionality. This section introduces the *Social Web Protocols*, a description of the specifications produced by the Social Web Working Group, and is arranged by the previously derived API requirements (read, publish, notify, subscribe). Some content in this section is published as a W3C Working Group Note [130]. Systems can be built with these protocols in great variety. Incorporating the standards produced by the Social Web Working Group into a system does not automatically mean the system is empowering to its users; the protocols provide only a skeleton, leaving much open for specialisation by developers. Conformance to these protocols does however imply that users are

able to move their data between systems; that clients and servers are somewhat decoupled so users have more flexibility with regards to tools; and that users are not compelled to follow their network or locked into the system where their friends are.

### 2.7.1. Overview

People and the content they create are the core components of the Social Web; they make up the social graph. The Social Web Protocols describe standard ways in which people can:

- connect with other people and subscribe to their content;
- create, update and delete social content;
- interact with other peoples' content;
- be notified when other people interact with their content;

regardless of *what that content is or where it is stored*.

These components are core building blocks for interoperable social systems.

Each of these components can be implemented independently as needed, or all together in one system, as well as extended to meet domain-specific requirements. Users can store their social data across any number of compliant servers, and use compliant clients hosted elsewhere to interact with their own content and the content of others. Put simply, Social Web Protocols tells you, according the recommendations of the Social Web Working Group:

- how to expose/consume social content (reading).
- what to post, and where to, in order to create, update or delete content (publishing).
- how to ask for notifications about content (subscribing).
- how to deliver notifications about content or users (delivery).

The following table shows the high level requirements according to the Social Web Working Group charter  
[<https://www.w3.org/2013/socialweb/social-wg-charter.html>](https://www.w3.org/2013/socialweb/social-wg-charter.html) and the Social API Requirements  
[<https://www.w3.org/wiki/Socialwg/Social\\_API/Requirements>](https://www.w3.org/wiki/Socialwg/Social_API/Requirements), and how the specifications of the Working Group overlap with respect to each.

Table 2. Table 13. The API requirements and which specs they are met by

|                     | Vocabulary | Syntax | Read | Create | Update | Delete | Subscription | Delivery |
|---------------------|------------|--------|------|--------|--------|--------|--------------|----------|
| ActivityPub         |            |        | X    | X      | X      | X      | X            | X        |
| ActivityStreams 2.0 | X          |        | X    |        |        |        |              |          |
| Linked Data         |            |        | X    |        |        |        |              | X        |
| Notifications       |            |        |      |        |        |        |              |          |
| Micropub            |            |        | X    | X      | X      |        |              |          |
| WebSub              |            |        |      |        |        | X      |              |          |
| Webmention          |            |        |      |        |        |        | X            |          |

The specifications may be implemented alongside each other in various configurations. Figure 17 shows a high level view of how different parties in a social system may be connected together. The arrows show data flowing through the system, and the labels of the arrows are the protocols by which data is enabled to flow.

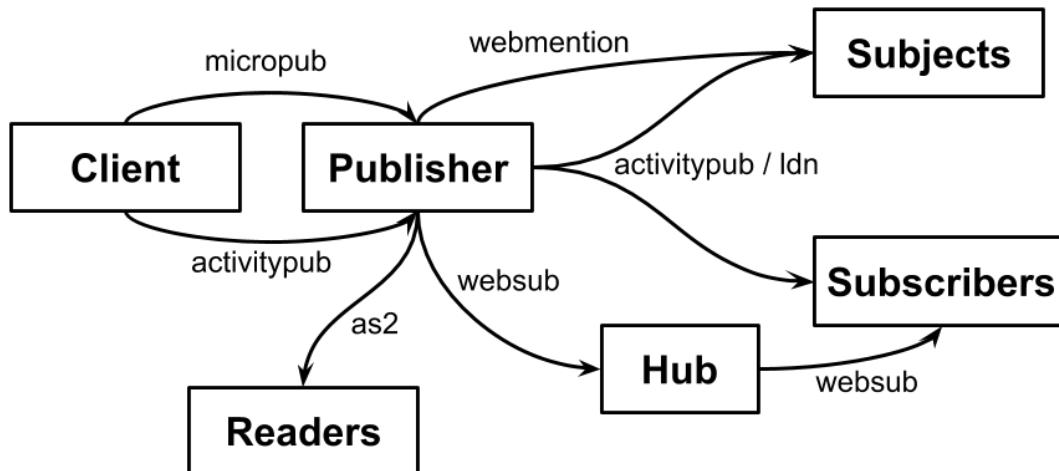


Figure 1. Figure 17. How the Social Web Protocols specifications connect different (high level) parties together.

Some of the specifications overlap in functionality, or complement each other explicitly. This list provides detail on some key relations between different specifications, and table 14 provides a summary.

- **ActivityPub and ActivityStreams 2.0:** ActivityPub uses the AS2 syntax and vocabulary for the payload of all requests.
- **ActivityPub and Linked Data Notifications:** ActivityPub specialises LDN as the mechanism for delivery of notifications by requiring that payloads are AS2. *Inbox* endpoint discovery is the same. LDN receivers can understand

requests from ActivityPub federated servers, but ActivityPub servers can't necessarily understand requests from generic LDN senders.

- **ActivityStreams 2.0 and Linked Data Notifications:** LDN MAY use the AS2 syntax and vocabulary for the payload of notification requests.
- **Webmention and Linked Data Notifications:** Overlapping functionality that needs to be bridged due to different content types of requests. An LDN request MAY contain the equivalent data as a Webmention request, but not necessarily vice versa.
- **ActivityPub and Micropub:** Overlapping functionality that needs to be bridged due to different vocabularies and possibly different content types of requests. Micropub specifies client-to-server interactions for content creation; ActivityPub specifies this, *plus* side-effects and server-to-server interactions.
- **Micropub and Webmention:** Are complementary but independent. Content could be created with Micropub, then Webmention discovery can be commenced on any URLs in the content.
- **Micropub and Linked Data Notifications:** Are complementary but independent. Content could be created with Micropub, then LDN discovery can be commenced on any relevant resources identified by the server.
- **Micropub and WebSub:** Are complementary but independent. Content could be created with Micropub, then passed to a WebSub *hub* for delivery to subscribers.

Table 3. Table 14. How each spec relates to the others

| AS2                                    | AP   | MP  | LDN                                 | WM     | WS                                |
|--|--|-----|-------------------------------------|--------|-----------------------------------|
| AS2                                    | represent content for CRUD, delivery and subscriptions | n/a | MAY represent notification contents | n/a    | MAY represent publishers' content |
| AP pass it around with                 |  |     | bridge use to trigger notifications | n/a    | n/a                               |
| MP n/a                                 | bridge   |     | n/a                                 | n/a    | n/a                               |
| LDN MAY pass notifications around with | use for delivery                                       | n/a |                                     | bridge | use for delivery                  |
| WM n/a                                 | n/a  | n/a | bridge                              |        | n/a                               |
| WS MAY pass feeds to subscribers       | n/a  | n/a | use to trigger notifications        | n/a    |                                   |

**To read (approx):** if I have x, it uses y to \_.

*n/a* means there is no explicit relation between the specs, but does not mean to suggest they can't be used together.

*bridge* means these specs have overlapping functionality and bridging code is needed for interoperability.

### 2.7.2. Reading

An individual's self presentation online can be partially composed of content they produce and interact with. The *read* label covers how these are exposed for consumption by others. This may include permissions or access control, which could require the reader to identify themselves before content is made available. Different types of content and interactions should be discoverable, perhaps according to criteria like the type of content, a group or individual with which it is associated, or through its association with other content (eg. through replies).

#### 2.7.2.1. Content representation

**ActivityStreams 2.0** (AS2) models content and interactions as objects and activities. AS2 includes a *vocabulary* for modelling different types of objects and activities as well as various relations they might have with other objects (including user profiles) and activities. The AS2 *syntax* describes a consistent structure for objects and activities including sets of objects and activities as collections. Collections can be explicitly created and updated by a user (like adding photos to an album) or generated automatically as a result of other user

actions or the properties of certain objects/activities (eg. a list of followers of a user). AS2 does not specify how objects, activities, or collections come into existence, only what they look like once they do.

AS2 content must be served with the Content-Type `application/activity+json` or, if necessary for JSON-LD extended implementations, `application/ld+json; profile="https://www.w3.org/ns/activitystreams"`. Content must be described using the AS2 vocabulary, and may use other vocabularies in addition or instead, per the extension mechanism.

To make content available as ActivityStreams 2.0 JSON, one could do so directly when requested with an appropriate Accept header (eg. `application/activity+json` or `application/ld+json`), or indirectly via a `rel="alternate" type="application/activity+json"` link. This link could be to a different domain, for third-party services which dynamically generate ActivityStreams 2.0 JSON on behalf of a publisher.

AS2 builds upon ActivityStreams 1.0 [12] and is not fully backwards compatible; the relationship between AS1 and AS2 is documented in the AS2 spec <<http://www.w3.org/TR/activitystreams-core#activitystreams-1.0>>.

Because AS2 is a data model, it does not recommend how data should be displayed. Its utility is in enabling a consistent representation of social objects and activities to be passed between potentially disconnected systems (eg. from Alice's social network to Bob's). The systems consuming the data are responsible for rendering it appropriately. This means that system designers can provide their users with options for customising the presentation data that may constitute their profiles.

### *Extending AS2*

AS2 specifies a finite set of object and activity types and properties. These are a baseline set of common social interactions which can be extended upon by systems which need additional terms or more specific variations. The extension mechanism is based on Linked Data, via JSON-LD. Developers are expected to publish documentation and an RDF representation of their terms at a domain under their control, and refer to terms by URI in the normal manner. Further, the Social Web Working Group describes a mechanism by which well-used extensions are included with the AS2 namespace document by means of a W3C Community Group vetting process. The advantage of this is that

implementations can adopt common extensions easily, without needing to include additional namespaces. It also makes extensions more discoverable for newcomers to AS2.

The first such extension is in fact **ActivityPub**. ActivityPub uses ActivityStreams 2.0 for all data in all requests, and also adds additional terms to the AS2 namespace. Thus, ActivityPub requires requests have the Content-Type `application/ld+json;`  
`profile="https://www.w3.org/ns/activitystreams".`

#### 2.7.2.3. Other ways of representing content

Despite AS2 being the recommended syntax and vocabulary of the Working Group, some specifications use different or broader mechanisms. This helps to let developers pick and choose different specifications for different tasks even if they prefer not to use AS2:

- **Linked Data Notifications** notification contents can use any vocabulary, so long as the data is available in JSON-LD. Thus notifications may use ActivityStreams 2.0, but don't have to. Clients and servers can negotiate between themselves (using Accept and Accept-Post HTTP headers) about using different RDF syntaxes, as well.
- **Micropub** clients which expect to read data (this would usually be clients for *updating* content) are expecting it as JSON in the parsed microformats2 syntax [<http://microformats.org/wiki/microformats2-parsing>](http://microformats.org/wiki/microformats2-parsing).
- **WebSub** is agnostic as to the Content-Type used by publishers; hubs are expected to deliver the new content to subscribers as-is produced by the publisher, at the publisher's topic URL.

#### 2.7.2.4. Objects and streams

Whichever syntax and vocabulary is used, there are some general recommendations for representing objects (individual entities) and streams (sets or collections of objects).

- All objects must have URLs in their `id` property. This URL should resolve to return the properties of an object; what is returned may depend on the requester's right to access the content, determined by authentication and/or authorisation.

- Each stream must have a URL, which must resolve to return the contents of the stream (according to the requester's right to access, and could be paged). The data returned may include additional metadata about the stream (such as title, description).
- Each object in a stream must contain at least its URL, which can be dereferenced to retrieve all properties of the object, and may contain other properties of the object.
- One user may publish one or more streams of content. Streams may be generated automatically or manually, and might be segregated by post type, topic, audience, or any arbitrary criteria decided by the curator of the stream. A user profile MAY include links to multiple streams, which a consumer could follow to read or subscribe to.

#### *2.7.2.5. Special streams*

Streams are represented in AS2 as a `Collection` or `OrderedCollection`. ActivityPub defines some special usages; two streams that must be accessible from a user's profile, and four which are optional, via the following properties:

- `inbox`: A reference to an AS2 collection comprising all the objects sent to the profile's owner.
- `outbox`: An AS2 collection comprising all the objects produced by the profile's owner.
- `following`: An optional AS2 collection of the users that this user is following.
- `followers`: An optional AS2 collection of the users that follow this user.
- `likes`: An optional AS2 collection of every object from all of the user's Like activities (generated automatically by the server).
- `streams`: An optional list of supplementary AS2 collections which may be of interest.

ActivityPub permits arbitrary streams to be updated through specifying special behavior for the server when it receives activities with types `Add` and `Remove`. When a server receives such an activity in the `outbox`, and the target is a stream, it must add the object to the target (for `Add`) or remove the object from the target (for `Remove`).

Two kinds of special streams are `inbox` and `outbox`. When read (ie. with an HTTP GET request) they return ordinary streams of objects, but they also double as endpoints which can be POSTed to directly to add objects, for delivery of notifications and publishing new content respectively.

The `inbox` is a notion shared by ActivityPub and Linked Data Notifications, however in order to be *read* by both ActivityPub *and* LDN clients, publishers must relate the stream to the objects it contains using both the `as:items` and `ldp:contains` predicates. This is an unfortunate discord, but since ActivityPub is immovably tied to AS2 and LDN is immovably tied to compatibility with the vocabulary of existing Linked Data Platform servers, there was really no compromise to be had. Fortunately this bridge is relatively minor in terms of coding, once a developer is aware of it.

### 2.7.3. Publishing

Publishing in this context incorporates creating new content, and updating or deleting existing content. The ability to publish content and generate (and update or remove) new data is critical to online self-presentation, in particular it is part of **customisability** in building an online profile.

Content generated through a client (such as a web form, mobile app, sensor, smart device) is created when it is sent to a server for processing, where it is typically stored and usually published (either publicly or to a restricted audience, in human- and/or machine-readable forms). Clients and servers may independently support creating, updating and deleting; there are no dependencies between them.

Authentication and authorization between clients and servers for creating content are not included in these specifications, as they are considered orthogonal problems which should be solved elsewhere.

The two specifications recommended by the Social Web Working Group for publishing are Activitypub and Micropub. They use similar high level mechanisms, but differ in requirements around both the vocabularies and content types of data. ActivityPub contains a client-to-server API for creating ActivityStreams 2.0 objects and activities, and specifies additional responsibilities for clients around addressing objects, and for servers around the side-effects of certain types of objects. Micropub provides a basic client-to-server API for creating blog-post type content which can be implemented alone and is intended as a quickstart for content creation, on top of which more complex (but optional) actions can be layered.

Both provide similar media endpoints for uploading files.

Neither ActivityPub nor Micropub define APIs for publishing based on HTTP verbs, and thus differ from the more RESTful Linked Data Platform (LDP) <<https://www.w3.org/TR/ldp>>.

### 2.7.3.1. Creating

The publishing endpoint of **ActivityPub** is the `outbox`. Clients are assumed to have the URL of a (ideally authenticated) user profile as a starting point, and discover the value of the `https://www.w3.org/ns/activitystreams#outbox` property found at the profile URL (which should be available as JSON[-LD]). The client then makes an HTTP POST request with an ActivityStreams 2.0 activity or object as a JSON[-LD] payload with a content type of `application/ld+json; profile="https://www.w3.org/ns/activitystreams"`. The URL of the created resource is generated at the discretion of the server, and returned in the `Location` HTTP header. This is an appropriate protocol to use when:

- You want to send/receive a JSON or JSON-LD payload.
- Your data is described with AS2 (optionally extensible via JSON-LD).
- You want serve to carry out a known set of actions upon content creation.

Side-effects of creating content with ActivityPub are for the most part adding things to various different collections collections (likes, follows, etc); but also include requirements about blocking users, and a hook to enable federated servers.

The publishing endpoint for **Micropub** is the `micropub` end point. Clients discover this from a user's URL via a `rel="micropub"` link (in an HTTP Link header, or an HTML element). Clients make a `x-www-form-urlencoded` POST request containing the key-value pairs for the attributes of the object being created. The URL of the created resource is generated at the discretion of the server, and returned in the `Location` HTTP header. Clients and servers must support attributes from the Microformats 2 h-entry vocabulary. Micropub also defines special reserved attributes (prefixed with `mp-`) which can be used as commands to the server. Any additional key names sent outside of these vocabularies may be ignored by the server.

Micropub requests may alternatively be sent as a JSON payload, the syntax of which is derived from the Microformats 2 parsing algorithm. This is an appropriate protocol to use when:

- You want to send/receive a form-encoded or JSON payload.

- Your data is described with the h-entry syntax and vocabulary.
- You can rely on out-of-band agreements between clients and servers for vocabulary extensibility.

### 2.7.3.2. Updating

Content is updated when a client sends changes to attributes (additions, removals, replacements) to an existing object. If a server has implemented a delivery or subscription mechanism, when an object is updated, the update MUST be propagated to the original recipients using the same mechanism.

**ActivityPub** clients send an HTTP POST request to the `outbox` containing an AS2 Update activity. The object of the activity is an existing object, and the fields to update should be nested. If a partial representation of an object is sent, omitted fields are *not* deleted by the server. In order to delete specific fields, the client can assign them a `null` value. However, when a federated server passes an `Update` activity to another server's `inbox`, the recipient must assume this is the *complete* object to be replaced; partial updates are not performed server-to-server.

**Micropub** clients perform updates, as either form-encoded or JSON POST requests, using the `mp-action=update` parameter, as well as a `replace`, `add` or `delete` property containing the updates to make, to the Micropub endpoint. `replace` replaces all values of the specified property; if the property does not exist already, it is created. `add` adds new values to the specified property without changing the existing ones; if the property does not exist already, it is created. `delete` removes the specified property; you can also remove properties by value by specifying the value.

### 2.7.3.3. Deleting

Content is deleted when a client sends a request to delete an existing object. If a server has implemented a delivery or subscription mechanism, when an object is deleted, the deletion MUST be propagated to the original recipients using the same mechanism.

**ActivityPub** clients delete an object by sending an HTTP POST request containing an AS2 Delete activity to the `outbox` of the authenticated user. Servers MUST either `replace` the object of this activity with a tombstone and return a `410 Gone` status code, or return a `404 Not Found`, from its URL.

**Micropub** delete requests are two key-value pairs, in form-encoded or JSON: `mp-action: delete` and `url: url-to-be-deleted`, sent to the Micropub endpoint

.

## 2.7.4. Subscribing

An agent (client or server) may *ask* to be notified of changes to a content object (eg. edits, new replies) or stream of content (eg. objects added or removed from a particular stream). This is *subscribing*. This is part of the process of creating links between individuals in a social network, and other individuals or resources; part of **connectivity**. Specifications which contain subscription mechanisms are ActivityPub and WebSub.

Nothing should rely on implementation of a subscription mechanism. That is, implementations may set themselves up to receive notifications without always being required to explicitly ask for them from a sender or publisher: see delivery.

### 2.7.4.1. Subscribing with `as:Follow`

**ActivityPub** servers maintain a *Followers* collection for all users. This collection may be directly addressed, or addressed automatically or by default, in the `to`, `cc` or `bcc` field of any Activity, and as a result, servers deliver the Activity to the `inbox` of each user in the collection.

Subscription requests are essentially requests to be added to this collection. They are made by the subscriber's server POSTing a `Follow` Activity to the target's `inbox`. This request should be authenticated, and therefore doesn't need additional verification. The target server then SHOULD add the subscriber to the target's `Followers` collection. Exceptions may be made if, for example, the target has blocked the subscriber.

This is a suitable subscription mechanism when:

- The subscriber wants to request updates from a specific actor (rather than objects, streams or threads).
- The subscriber and publisher both speak ActivityStreams 2.0.
- The publisher is aware of who has subscribed, and capable of delivering content to subscribers itself.

Since delivery is only a requirement for federated servers, prospective subscribers will not be able to POST their Follow activity to the inbox of a profile which is on a non-federated server (expect a 405 Method Not Allowed), and thus are not able to subscribe to these profiles. In this case, prospective subscribers may wish to periodically *pull* from the publisher's outbox instead.

#### 2.7.4.2. Delegating subscription handling

**WebSub** provides a mechanism to delegate subscription handling and delivery of content to subscribers to a third-party, called a *hub*. All publishers need to do is link to their chosen hub(s) using HTTP Link headers or HTML <link> elements with rel="hub", and then notify the hub when new content is available. The mechanism for notifying the hub is left deliberately unspecified, as publishers may have their own built in hub, and therefore use an internal mechanism.

Hubs and publishers which would like to agree on a standard mechanism to communicate might consider employing an existing delivery mechanism, namely Linked Data Notifications (fig. [subscription-notification]) or Webmention<sup>sub-wm</sup>.

```

POST /inbox HTTP/1.1
Host: hubbub.example
Content-type: application/ld+json

{
  "@context": "https://www.w3.org/ns/activitystreams",
  "id": "",
  "type": "Announce",
  "object": "https://rhiaro.co.uk/tags/socialwg",
  "target": "https://hubbub.example/"
}

```

Listing 1. Notifying a hub of new content with LDN, using an AS2 Announce in the notification body. The object is the topic URL and the target is the hub itself. The hub can use this information to fetch new content for subsequent delivery to subscribers.

The subscriber discovers the hub from the publisher, and sends a form-encoded POST request containing values for hub.mode ("subscribe"), hub.topic (the URL to subscribe to) and hub.callback (the URL where updates should be sent to, which should be 'unguessable' and unique per subscription). The hub and subscriber carry out a series of exchanges to verify this request.

When the hub is notified of new content by the publisher, the hub fetches the content of the topic URL, and delivers this to the subscriber's callback URL.

This is a suitable subscription mechanism when:

- The subscriber wants to request updates from any resource (not just user profiles), and of any content type.
- Subscription requests are not authenticated, so you need a way to verify them.
- The publisher wants to delegate distribution of updates to another service (the hub) instead of doing it itself.

LDN Receivers can receive deliveries from WebSub hubs by using the `inbox` URL as the `hub.callback` URL and *either* only subscribing to resources published as JSON-LD *or* accepting content-types other than JSON-LD.

swp-sub-wm though this seems to be me to be somewhat outside of the spirit of 'mentioning'.

### 2.7.5. Delivering

A user or application may wish to push a notification to another user that the receiver has *not explicitly asked* for. For example to send a message or some new information; because they have linked to (replied, liked, bookmarked, reposted, etc) their content; because they have linked to (tagged, addressed) the user directly; to make the recipient aware of a change in state of some document or resource on the Web. This is also part of **connectivity**, as well as a potential way to make individuals aware of how their content is **cascaded** throughout a network, and how they are connected to others of whom they were previously unaware. The Social Web Working Group specifications contain several mechanisms for carrying out delivery; they are listed here from general to specialised.

#### 2.7.5.1. Targeting and discovery

The target of a notification is usually the addressee or the subject, as referenced by a URL. The target may also be someone who has previously requested notifications through a subscription request. Once you have determined your target, you need to discover where to send the notification for that particular

target. Do this by fetching the target URL and looking for a link to an endpoint which will accept the type of notification you want to send (read on, for all of your exciting options).

Bear in mind that many potential targets will *not* be configured to receive notifications at all. To avoid overloading unsuspecting servers with discovery-related requests, your application should employ a "back-off" strategy when carrying out discovery multiple times to targets on the same domain. This could involve increasing the period of time between subsequent requests, or caching unsuccessful discovery attempts so those domains can be avoided in future. You may wish to send a User-Agent header with a reference to the notification mechanism you are using so that recipient servers can find out more about the purpose of your requests.

Your application should also respect relevant cache control and retry headers returned by the target server.

#### 2.7.5.2. *Generic notifications*

**LDN** provides a protocol for sending, receiving and consuming notifications which may contain any content, or be triggered by any person or process. Senders, receivers and consumers can all be on different domains, thus this meets the criteria for a federation protocol. This is a suitable notification mechanism when:

- Notifications need to be identifiable with their own URLs and exposed by the receiver for other applications to discover and re-use.
- Notifications are represented as a JSON-LD payload (ie. a 'fat ping').
- You need to advertise constraints on the type or contents of notifications accepted by a receiver.

LDN functionality is divided between *senders*, *receivers* and *consumers*. The endpoint to which notifications are sent is the `inbox`. Any resource (a user profile, blog post, document) can advertise its `inbox` so that it may be discovered through an HTTP Link header or the document body in any RDF syntax (including JSON-LD or HTML+RDFa). To this `Inbox`, senders make a POST request containing the JSON-LD (or other RDF syntax per Accept-Post negotiation with the receiver) payload of the notification. The receiver returns a URL from which the notification data can be retrieved, and also adds this URL to a list which is returned upon a GET request to the `Inbox`. Consumers can

retrieve this Inbox listing, and from there the individual notifications, as JSON-LD (optionally content negotiated to another RDF syntax). An obvious type of consumer is a script which displays notifications in a human-readable way.

An existing LDP implementation can serve as an LDN receiver; publishers simply advertise any `ldp:Container` as the `inbox` for a resource.

The payload of notifications is deliberately left open so that LDN may be used in a wide variety of use cases. However, receivers with particular purposes are likely to want to constrain the types of notifications they accept. They can do this transparently (such that senders are able to attempt to conform, rather than having their requests rejected opaquely) by advertising data shapes constraints such as SHACL <<https://www.w3.org/TR/shacl/>>. Advertisement of such constraints also allows consumers to understand the types of notifications in the Inbox before attempting to retrieve them. Receivers may reject notifications on the basis of internal, undisclosed constraints, and may also access control the Inbox for example by requiring an `Authorization` header from both senders and consumers.

**WebSub** publishers deliver content to their *hub*, and *hubs* to their *subscribers* using HTTP POST requests. The body of the request is left to the discretion of the sender in the first case, and in the latter case must match the Content-Type of and contain contents from the topic URL.

#### 2.7.5.3. Activity notifications

**ActivityPub** uses LDN to send notifications with some specific constraints. These are:

- The notification payload MUST be a single AS2 Activity.
- The notification payload MUST be compact JSON-LD.
- The receiver MUST verify the notification by fetching its source from the origin server.
- All notification POST requests are authenticated.

ActivityPub specifies how to define the *target(s)* to which a notification is to be sent (a pre-requisite to LDN sending), via the AS2 audience targeting and object linking properties.

ActivityPub also defines side-effects that must be carried out by the server as a result of notification receipt. These include:

- Creating, updating or deleting new objects upon receipt of Create, Update and Delete activities.
- Reversing the side-effects of prior activities upon receipt of the Undo activity.
- Updating specialised collections for Follow, Like and Block activities.
- Updating any other collections upon receipt of Add and Remove activities.
- Carrying out further delivery to propagate activities through the network in the case of federated servers.

ActivityPub actor profiles are linked to their inboxes via the <https://www.w3.org/ns/activitystreams#inbox> property. This is an *alias* (in the AS2 JSON-LD context) for LDN's <http://www.w3.org/ns/ldp#inbox>. Applications using a full JSON-LD processor to parse these documents will see these terms as one and the same. Applications doing naive string matching on terms may wish to note that if you find an ldp:inbox it will accept POST requests in the same way as an as:inbox.

#### 2.7.5.4. Mentioning

**Webmention** provides an API for sending and receiving notifications when a relationship is created between two documents by including the URL of one document in the content of another. It works when the two documents are on different domains, thus serving as a federation protocol. This is a suitable notification mechanism when:

- You have a document (source) which contains the URL of another document (target).
- The owner of the endpoint has access to view the source (so the request can be verified).
- The only data you need to send over the wire are the URLs of the source and target documents (ie. a 'thin ping').

There are no constraints on the syntax of the source and target documents. Discovery of the Webmention endpoint (a script which can process incoming webmentions) is through a link relation (`rel="webmention"`), either in the HTTP Link header or HTML body of the target. This endpoint does not need to be on the same domain as the target, so webmention receiving can be delegated to a third party.

Webmentions are verified by the server dereferencing the source and parsing it to check for the existence of the target URL. If the target URL isn't found, the webmention MUST be rejected.

Webmention uses `x-www-form-urlencoded` for the source and target as parameters in an HTTP POST request. Beyond verification, it is not specified what the receiver should do upon receipt of a Webmention. What the webmention endpoint should return on a GET request is also left unspecified.

#### 2.7.5.5. Delivery interop

This section describes how receiver implementations of either Webmention or LDN may create bridging code in order to accept notifications from senders of the other. This can also be read to understand how a sender of either Webmention or LDN should adapt their discovery and payload in order to send to a receiver of the other.

**Webmention receivers** wishing to also accept LDN POSTs at their Webmention endpoint MUST:

- Advertise the webmention endpoint via `rel="http://www.w3.org/ns/ldp#inbox"` in addition to `rel="webmention"` (in the Link header, HTML body or JSON body of a target).
- Accept POST requests with the Content-Type `application/ld+json`. Expect the body of the request to be:

```

1 {
2   "@context": "http://www.w3.org/ns/webmention#",
3   "@id": "",
4   "source": { "@id": "https://waterpigs.example/post-by-barnaby" },
5   "target": { "@id": "https://aaronpk.example/post-by-aaron" }
6 }
```

- Use the `source->@id` and `target->@id` values as the `source` and `target` of the Webmention, and proceed with verification.
- If returning a `201 Created`, it MUST return a `Location` header with a URL from which the contents of the request posted can be retrieved. `202 Accepted` is still fine.
- Note than when verifying the source, there's a good chance you can request/parse it as RDF.

**LDN receivers** wishing to also accept Webmentions to their Inbox MUST:

- Advertise the Inbox via `rel="webmention"` in addition to `rel="http://www.w3.org/ns/ldp#inbox"` (in the Link header, HTML body or JSON body of a target).
- Accept POST requests with a content type `application/x-www-form-urlencoded`. Convert these requests from:

```
1 source=https://waterpigs.example/post-by-barnaby&
target=https://aaronpk.example/post-by-aaron
```

to:

```
1 {
2   "@context": "http://www.w3.org/ns/webmention#",
3   "@id": "",
4   "source": { "@id": "https://waterpigs.example/post-by-barnaby" },
5   "target": { "@id": "https://aaronpk.example/post-by-aaron" }
6 }
```

and proceed per LDN; receivers MAY add other triples at their discretion.

- Receivers MUST return a 201 Created with a Location header or 202 Accepted.
- Receivers MUST verify the request by retrieving the source document and checking a link to the target document is present. If the Webmention is not verified, receivers MUST NOT keep it.

#### 2.7.5.6. Webmention as AS2

A webmention may be represented as a persistent resource with AS2. This could come in handy if a Webmention sender *mentions* a user known to be running an ActivityPub federated server. In this case, the sender can use an AS2 payload and carry out delivery of the notification per ActivityPub/LDN.

```
{
  "@context": "https://www.w3.org/ns/activitystreams#",
  "type": "Relationship",
  "subject": "https://waterpigs.example/post-by-barnaby",
  "object": "https://aaronpk.example/post-by-aaron"
}
```

A receiver or sender may want to augment this representation with the relationship between the two documents, and any other pertinent data. In the receiver's case, this could be gathered when they parse the source during the

verification process. For example:

```

1 {
2   "@context": "https://www.w3.org/ns/activitystreams",
3   "type": "Relationship",
4   "subject": {
5     "id": "https://waterpigs.example/post-by-barnaby",
6     "name": "Hi Aaron, great post."
7   },
8   "object": {
9     "id": "https://aaronpk.example/post-by-aaron",
10    "name": "Aaron's first post."
11  },
12  "relationship": "inReplyTo"
13}

```

## 2.8. Protocols for Customisability and Connectivity

The protocols produced by the Social Web Working group deal with creating content and social interactions, and propagating them around a network. They give varying degrees of freedom to implementors about when and how to pass data between servers, and say little to nothing about the presentation of the content or user interface associated with interactions. The core types of social objects and interactions indicate an initial constraint how users will be able to behave within a system, though implementations can extend from this baseline as they see fit. How and whether they do so remains to be seen<sup>ext</sup>. The protocols do *not* deal with changes to attributes of a user profile, however.

Engaging in particular types of public or partially public social interaction online is a way to shape one's self-presentation. Being able to choose which of these interactions are used and presented outwardly is part of **customisability**. Building social interactions from a common base of standard ones means that the semantics of these activities can be shared across platforms.

Interoperable implementations based on these protocols increase the potential **connectivity** of individuals online, as they are no longer constrained to interacting with others within the confines of a single technical system. People can potentially find and connect with, follow and subscribe to, other people and content no matter where it is published. This brings with it further complications around how peoples' activities are presented across disparate

systems. Users have even less surety with regards to what their content will look like when it is seen by others if they have no way of knowing what kinds of systems their content is being transmitted to or through.

<sup>ext</sup> Development of extensions for ActivityStreams 2.0 documented at the namespace <<https://www.w3.org/ns/activitystreams#extensions>>; for ActivityPub on the SWWG wiki <[https://www.w3.org/wiki/ActivityPub\\_extensions](https://www.w3.org/wiki/ActivityPub_extensions)> and for Webmention in the Indieweb wiki <<https://indieweb.org/webmention#Extensions>>.

## 2.9. Remaining problems

Some of the problems which the Social Web Working Group did not address (and nor were chartered to) are listed here. This list is substantial, and many of the items seem critical for the future of the Social Web, however I posit that these are mostly unsolvable by technical means. As the SWWG has produced technical specifications rather than a code of ethics or policy recommendations, its reach is somewhat limited.

- Identity (authentication, authorisation, presentation of personal data).
- Abuse, spam, data misuse.
- Preventing mass surveillance.
- Decentralisation outside of HTTP protocols.
- Security or privacy of personal data and online interactions (though each specification was reviewed by W3C Security and Privacy specialists, and contains a section which explicitly addresses these kinds of concerns with regard to that particular protocol).
- Federated search.
- Economic incentives for building decentralised social software.

## 3. Personal data and self-presentation

Standards aren't really all that interesting until they're implemented and used. This section describes a single-user system, a personal social datastore, built around some of the protocols designed by the Social Web Working Group. I continually used and improved this system for over two years, as the core of my online presence and a public log of various digital and physical activities. I reflect upon the impact that doing so has had on me at a personal level, because in designing personal datastores and decentralised social systems we are encouraging this type of behaviour, and I believe as developers and system designers, we should experience it first-hand.

Further, in a decentralised Social Web, we cannot assume any kind of consistency between nodes in a network. Everyone's experience of the technology through which interactions are mediated may be completely different. I use my own experience with a personal social datastore to discuss how and why we need to take this into consideration when observing and understanding peoples' behaviour in future online social networks. To facilitate this discussion, I use the concept of a *Web Observatory*.

A Web Observatory is a system which gathers and links to data on the Web in order to answer questions about the Web, the users of the Web and the way that each affects the other. - webscience.org  
<<http://www.webscience.org/web-observatory/>>

This section has been adapted from work published as *Observing the Decentralised Social Web (one telescope at a time)* (2016, Proceedings of WWW, Perth).

### 3.1. Introduction

Studying communities through passive and active digital traces, as Web Observatories are designed to do [131, 272], brings with it a host of privacy, ethical, and methodological concerns. Attempts to address online privacy issues in general are being made with a push towards *re-decentralising* the Web [132], in part through open Web standards and work on promoting personal data stores as alternatives to centralised or third-party services. Using this momentum can benefit the Web Science community as well as their observees

(though it brings with it its own set of challenges). Decentralisation is applied to Web Observatories in particular through the idea of a *Personal Web Observatory* [289] as a user-controlled (as opposed to third-party imposed) means of collecting and tracking data.

I believe that the perspectives of both subject and experimenter are valuable when it comes to studying people in new ways. Taking inspiration from the reflective practices of the Quantified Self community, I progressively built features into my own personal website which allow monitoring and visualisation of day-to-day aspects of my life, and used it continuously for the period of two years and counting. Immersion in the ongoing effects of self-tracking in a decentralised manner led to greater insight in working towards Personal Web Observatories than either developing a personal data store or engaging in self-tracking through third-party services could do alone.

This section begins by outlining related work on Web Observatories and Quantified Self. It includes a summary of the architecture of the personal data store being studied, and the types of data collected. I discuss the results in terms of psychological impact of the experiment, evolving motivations, and expected and unexpected consequences. In conclusion I relate these results to new and existing challenges for Web Observatories and Web Scientists who want to study data generated by Personal Web Observatories or similarly decentralised systems.

## 3.2. Background

### 3.2.1. Personal Web Observatories

Web Observatories concern the use of peoples' digital footprints as the subject of academic inquiry [131, 272]. Such data encompasses all manner of online and offline activities, and it may be collected passively by systems with which individuals interact, or actively logged, or some combination of the two. In order to address negative privacy implications of collecting and analysing this data, [289] introduces the idea of a *Personal Web Observatory*. Personal data stores are presented as an architecture for a decentralised Web Observatory, to allow individuals to maintain control over their data whilst still participating in scientific studies or otherwise releasing their data for use by third parties.

A Personal Web Observatory relies on individuals opting in to self-tracking activities; connecting their personal data store to sensors or user interfaces.

### 3.2.2. Self-tracking and Quantified Self

There have been a small number of high profile instances of individuals collecting a large volume of data about themselves, then offering it up for auction [111, 254]. Even more common are those who track data about themselves in order to: orchestrate behavioural changes; monitor bodily functions; learn new things about themselves; discern cause-and-effect relations; aid memory; manage information and life events; make better decisions; or just for the fun of it (for surveys of self-tracking motivations and tools, see [177, 283, 122, 262]). This personal self-tracking is disruptive to traditional notions of big data and data science [262].

The Quantified Self (QS) community is large, global, and growing. In-person conferences and meetups are held regularly [203, 207]. Whilst they use a wide range and combination of DIY and off-the-shelf devices for self-tracking [207, 264], what is particularly interesting are their reflective practices. QS practitioners engage in creative ways to collect, visualise, and understand data about themselves. At meetups, talks are focussed around deep personal insights, optimisation and improvement, and active self-awareness, rather than technology or tools [283, 264, 207].

However, the psychological impact of quantifying day-to-day activities is more complex than detecting trends and setting goals. Self-tracking may work against our best interests as interpretation of data is both subjective, and prone to *re-interpretation* at each viewing. There is also evidence that attempting to enforce a particular behaviour can have the opposite effect. Peoples' memories and impressions are easily influenced by external and internal factors [57]. A participant in [177] expressed concern about becoming "compulsive" about data collection, and [177] also describes how many off-the-shelf self-tracking services do not provide adequate means to aid user reflection.

Another negative effect of QS tracking is poor security of sensor software and third-party storage which can compromise individuals' data [15]. Plus, using third-party software typically comes with terms of service which are problematic from a privacy perspective.

Quantified self tracking is gamification of non-play activities, and [305] describes gamification as having surveillance at its root. [305] emphasises that quantification is a tool for governance and control and [207] concurs that statistics are historically used to manage populations, and this form of control is internalised by individuals for management of self. Prevalence of QS devices

and habits can serve to normalise surveillance. Further, QS tracking in the workplace is being introduced in ways that are becoming increasingly difficult to opt-out of, and raises unrealistic expectations of workers "fostered by a quantified, machine-like image of human productivity" [204].

Nonetheless, [207] describes QS participants who pushing back against the expectations and categories of the companies whose devices they use with their own interpretations of their data, "calling into question who gets to do the aggregation and how".

### *3.2.3. Social Publishing*

Though most bloggers and social media users do not explicitly set out to log small events in their lives, other kinds of social publishing can also be considered a form of self-tracking. Indeed, it is these types of data which Web Observatories today are most commonly consuming and analysing. Blogs have been around since the early days of the Web, initially for collecting links and then for documenting ideas and experiences. More recently, such expressions are commonly published to centralised social networking sites, and as we have seen in Chapter 3, many systems facilitate deeper data logging than just prose content. For example, the Facebook 'status update' input presents preset options for mood, location, films, books, friends, and other activities. These are easily attached to a text post through a seamless user interface, where the *intent* feels quite different to services and devices specifically marketed for self-tracking, but the end result in terms of data collected is largely the same.

In Chapter 4 I described various efforts to address decentralisation of the Social Web, and earlier in this chapter I detailed the formal standardisation efforts at the W3C. This work exemplifies technical foundations on which Personal Web Observatories can be built.

## **3.3. Building a Personal Web Observatory**

In this section I describe my Personal Web Observatory setup. One notable constraint from the outset was that in order to minimise maintenance requirements the system needed to be no more complicated to run than a personal website (setup and maintenance being a concern called out in [289]). As such it is implemented in PHP and runs on standard shared Web hosting, with a MySQL backend.

### 3.3.1. Architecture

The system (which is named *sloph* <<https://rhiaro.co.uk/sloph>>) constitutes a central database which is an RDF quadstore (layered on top of a MySQL database by the ARC2 PHP library). Using a graph data model facilitates the addition of new data without the overhead of updating schema or models in the core code.

For incoming data, it uses two endpoints: *publication* and *notification*, which implement the server portions of ActivityPub and LDN respectively. To compensate for overlapping standards, the publication endpoint additionally includes bridging code which converts Micropub requests in ActivityPub requests before proceeding, and the notification endpoint converts Webmentions into LDN. Data is processed to examine its validity, and stored as-is in the quad store. Publication data all uses the AS2 syntax and vocabulary (or extensions thereof), and notification data is stored using whatever vocabulary is sent by the notifying party. Data from each endpoint are stored in different graphs in the quad store to manage provenance.

The notifications endpoint can receive unsolicited messages from any LDN Sender, which may be somebody else's personal data store, or a clientside tool. In addition, some third-party services have been configured to send notifications to the notification endpoint. Webmention.io <<https://webmention.io>> and GitHub <<https://github.com>> have webhook settings, which are set to the notifications endpoint. They forward JSON data, which is easily convertible into JSON-LD, used by the LDN standard. Brid.gy <<https://brid.gy>> is a service which runs in the background and monitors my social media profiles for replies to my posts, then sends these as Webmention notifications. Upon certain new incoming notifications, the endpoint sends a request to the *PushOver* <<https://pushover.net/>> API, which sends a push notification to my Android phone.

Data may be retrieved as individual items (AS2 *Objects* or *Activities*), or in sets (AS2 *Collections*); all are identified by URLs. Content negotiation is employed so that requesting clients may access the data in any desired RDF syntax, or HTML. For HTML display, simple templates are created for each 'type' (or shared between a set of similar types) of data item. It is expected that most requests come from Web browsers, so the HTML content is delivered most often. However, other applications or services may consume the data, including readers (which may mix together multiple streams of data for the user), aggregators (which read the data and perform some manipulation or calculation

over it to display the results) or publishing clients (which offer editing or combining of existing data). All data is public; I did not implement access control.

The publication endpoint performs additional functions for data enhancement and distribution. It automatically adds missing metadata to posts if necessary (for example, published date and author), as well as storing new relations between posts and relevant collections such as tags, as well as the specific collections required by ActivityPub. The publication endpoint forwards text posts to Twitter, if necessary truncating them and adding a link back to the original, which helps with reach of content (since I don't yet have a subscription mechanism implemented). The publication endpoint also scans the content and certain attributes of incoming data for URLs, and behaves as an LDN and Webmention Sender to deliver notifications to others, if possible.

Figure 2. Figure 18. The high level architecture for my Personal Web Observatory. Parts in black are the core, and parts in grey are external or third-party services which interact or interoperate.

### 3.3.2. Data

[289] suggests that the first feature of a PWO should be to allow individuals to consolidate data collected by third parties into one repository under their control. Rather than attempting to aggregate - or even find - *all* data about myself spread across the Web, I chose a handful of services which have particular value to me, that I have been using to actively log particular things. I exported data from Lastfm <<https://last.fm>> (over ten years of music listening history), Twitter <<https://twitter.com>> (7 years of short notes), Runkeeper <<https://runkeeper.com>> (1 year of runs, walks and hikes with GPS traces), and Github <<https://github.com>> and Bitbucket <<https://bitbucket.com>> (5 years of code commit history), and Firefox bookmarks (2 years). I also exported data from 750words.com (almost 7 years of intermittent use) but did not import this into my store due to private content and no reliable access control built in.

On top of these data dumps I created the following templates:

Table 4. Table 15. Post types

| Type of data   | Attributes displayed *  |
|--|---|
| Articles (blog posts)  | name, content   |
| Short notes (like Tweets)  | content   |
| Meals logged   | description, restaurant if applicable                                 |
| Travel plans   | start and end location (map), date and time, means of transport, cost |
| Checkins to specific places  | location (map)  |
| Checkins to categories of place, aka 'vague' checkins (eg. 'home', 'office') | location, duration of time there, associated colour                   |
| Likes  | URL of thing liked  |
| Bookmarks  | URL of bookmark, name of bookmark, optional comment or quote          |
| Reposts (aka shares aka retweets aka reblogs)                                | URL of post, optional comment   |
| Acquisitions (purchases and free stuff)                                      | description, source, cost, photo                                      |
| Additions to photo albums  | photos and URL of album   |
| Events and RSVPs   | location, date and time, name, description, event website             |
| Subscriptions / follows  | URL of profile followed   |
| Sleep times  | start and end date and time, optional comment                         |

\* All posts contain tags and a published date, and may contain a last modified date.

Templates were created not all at once, but as I decided to start tracking something new and wanted to visualise it. Templates were continually modified and improved over the course of the year.

At the time of writing, I display posts in three different formats on my homepage (figure 19): a feed of the most recent eight article and note posts displayed in full; a list of the most recent of each type of post, displayed as a sentence (eg. "the last thing I ate was toast with peanut butter, 25 minutes ago"); and the last 1600 posts of all kinds, visualised as a string of small coloured boxes with icons. The colours represent where I was at the time of making the post (according to the most recent prior 'vague' checkin) and the icons indicate the type of post. Clicking on any of these boxes takes you to the post itself. In addition, the background colour of the homepage changes according to where I am at the present time. I also show my top 128 tags, and the number of posts for each.

Another type of output is a /summary <<https://rhiaro.co.uk/summary>> page, which aggregates data between any two dates, defaulting to the past seven days. This is useful for producing a year- and week-in-review, and includes total

amount of money spent, top foods eaten, number of words written, and various averages.

Figure 3. Figure 19. A screenshot of my homepage at the time of writing. The purple background shows I was at home. Note: I do not profess to be a designer.

### 3.4. Using a Personal Web Observatory

#### 3.4.1. Recording data

All data is actively recorded; that is, I enter all of the data and no posts are triggered by some other process or sensor. Unfortunately, for the sources of data exported from other services mentioned previously, I did not (yet) implement a connection to their various APIs to post subsequent data directly to my site.

As I decided to log a new kind of data, I either created or sought out a suitable client. As an intermediary measure (eg. while a client was in development), I could insert data into the quadstore directly using my SPARQL endpoint, which turns out to be a pretty useful bootstrapping measure. Clients I developed myself are simple web forms, which post AS2 data as JSON-LD to my publication endpoint.

I typically logged data at the time of its occurrence if possible. If I didn't have an internet connection (a frequent occurrence whilst traveling) I used a local 'timestamp' Android app to record the time and note crucial details, then back-dated posts at the earliest opportunity.

#### 3.4.2. Practical uses for small data

[289] proposes that "small data analytics, while sparse, could be made statistically viable when gathered longitudinally over time." Whilst I haven't done any complex statistical analysis on my data, certainly aggregated results, counts, and some calculations of means have been personally insightful. To take some trivial examples: that I drank on average 0.8 cups of coffee per day in 2016 was lower than I expected; that I spent an average of \$28.71 per day was higher than I expected; and that I spent 1 month, 15 days, 6 hours, 41 minutes, and 15 seconds travelling between places is just kind of interesting (and

something I particularly wanted to find out when I started recording 'vague' checkins). I don't expect these statistics to be remotely captivating for anyone else; such is the value of personal "small data."

Such aggregations were able to be put to more focused uses. Logging all of my purchases did not raise my internal awareness of how much I was spending on a weekly or monthly basis, but when I realised I wasn't saving money after about six months it was trivial to write a small web app which consumed my data stream, aggregated the total cost by certain categories, and displayed the amount I spent in any given month or week. Adding a setting for a monthly budget goal let the app send me notifications to stop spending when it noticed I was on track to exceed this in the current week. The app is not coupled to my personal site or data, so it can be used by anyone who publishes a stream of their purchases using AS2.

An initial motivation of recording 'vague' checkins was so people could check my site to see whether I was on my way if we were planning to meet, or if I shouldn't be disturbed (eg. if I was in a meeting). Industrious computer science friends created an IRC bot that consumed my /where <<https://rhiaro.co.uk/where>> endpoint and responded when asked \$whereis rhiaro. Other users of the same IRC channel added their own location reporting endpoints for the bot to read, some down to GPS accuracy with a map, shortly thereafter. Other unanticipated uses of data I was recording include launching a travel blog which filtered travel-related posts and photos from my main feed, and a food blog which includes my food logs, food-related photos, and any posts or bookmarks about recipes or restaurant reviews.

### 3.4.3. *If my website is down, I do not exist*

Having visible output for each log on my website motivated me to keep logging. I am missing lots of checkins to specific cities because I had trouble with a maps API and didn't complete the template to display them. Similarly I never got around to creating a compelling view of sleep records, so logs for those are few and far between.

I felt an internal pressure to ensure my data stream was complete. If I was offline or out of battery for a length of time, I would keep logs on paper to back-date later. This was due both to wanting to ensure aggregate data was accurate, and fear of 'losing' associated memories. Relatedly, I looked (and in fact still do) through historical logs with surprising frequency, and found memories of events, people and places were triggered by descriptions of meals or photos of purchases that I might otherwise have forgotten. I worry that this is unhealthy, bordering on obsessive behaviour though. At times, particularly before I had a streamlined offline-logging plan in place, or if it was socially inappropriate to be writing or typing logs, I felt *anxious* that I would neglect to log something.

When my webhost experienced outages, leaving me unable to record data, I both noticed far more quickly and felt a far stronger personal impact than I previously would have when my site was mostly static and for infrequent blogging. I complained to my hosting provider more frequently, and projected a sense of urgency that was probably disproportionate. I was left with a feeling of *if my website is down, I do not exist*, and I found myself wondering if sysadmins in a data center the other side of the world could possibly know or care about the anxiety they were causing me.

Other, predicted, psychological impact was that publicly logging photos of all of my purchases made me more conscious about what I bought. *Even though I didn't actually think anyone looked at my feed*, I found being aware that someone *could* see it helped me to, for example, resist buying junk food at the supermarket.

## 3.5. Discussion

In this section I expand on some of the topics raised by the results of my creation and use of a PWO, and in particular the implications for Web Observatories, and Web Science as a whole. I think these results highlight many

open questions and future research challenges.

PWOs are at the intersection of Web Observatories and Personal Data Stores. Research on the former is concerned with passively collecting and analysing how individuals and communities already use the Web, in order to learn about society. Research on the latter is concerned with improving and often actively influencing the lives of individuals. When we think about integrating individual data stores into an architecture for mass observation, we need to find a way to account for the conflicts that arise.

### *3.5.1. Personal and social impact*

Good user interfaces for visualising data from logs and perhaps making recommendations for improvements to a particular aspect of someone's life can be a strong way to encourage people to engage in self-tracking. However given the potential for compulsive behaviour around self-tracking, we also need to enable people to be 'off the grid' without affecting their overall aggregates or statistics. The motivational write-every-day site 750words.com has a 'vacation' setting, by which you can tell the system when you'll be away so that it doesn't penalise you for 0-word days during this time. If self-tracking is going to become the norm (as implicit self-tracking via social media already is in certain parts of the world, and as would be beneficial for Web Observatories) then taking breaks from tracking should be part of that norm. The benefits of continuous self-tracking can be acknowledged, but disconnecting can also be accepted as a healthy practice. Building this in from the outset can perhaps help to mitigate against internalised normalisation of surveillance, too.

Relatedly, the effects of missing data or inaccurate logs for any reason may have a damaging psychological effect on people. I posit that an important part (and in fact an ethical responsibility) of Web Science research would be to examine these effects with at least as much priority as working on the best gamification techniques for encouraging people to self-track.

It is critical that interfaces for visualising data logs are effective in helping the subject of the data to interpret its meaning. On the one hand, simply displaying correlations may incline people towards inferring spurious cause-effect relationships; on the other hand, as members of the QS community have shown, individuals may be far more effective at interpreting their own data than a third party service - which is perhaps missing other relevant information - could be. A

challenge for PWOs would be to offer data visualisation interfaces which not only take into account all available data, but account for what is (or might be) missing as well.

Furthermore, there is no one-size-fits-all for causing positive behavioural change based on self-tracking. How we can use personal data logs to improve the lives of individuals might vary by personality as well as the social and cultural context in which the tracking takes place. Ways of *tailoring* PWOs may be extremely important when it comes to actually benefitting their users.

From a traditional social media Web Observatories perspective, subjects are not aware necessarily that their data is being observed and analysed by academics. People use social media for different reasons, and I suspect that enabling social scientists to understand the world is pretty far down the list. Especially given the assumption that users of PWOs are actively opting in to being observed, and perhaps have fine grained control over what can and cannot be used for the purposes of research, is going to impact how they engage in self-tracking. As people shape and are shaped by social media use, people will shape and be shaped by their use of PWOs. What are the tradeoffs when it comes to benefits to individuals, and benefits to academic research? How can we take this into account when drawing conclusions from collected data?

### 3.5.2. Technical considerations

People are likely to be discouraged from using a PWO if it comes with a high burden of maintenance or cost of running [289]. I was able to minimise the impact of this on myself by integrating it into a system I was already maintaining (my personal website). Doing this had a significant impact on the technologies I was able to use, which was in some ways restricting, but beneficial in the long term. Shared Web hosting, PHP and MySQL are widespread and well-supported; this demonstrates that a PWO need not be architected around specialist or niche technologies, and need not be difficult or burdensome to set up and use. Certainly lowering the barrier to entry to people who *already* run their own websites could help to springboard adoption.

Centralised services are frequently bought out, shut down, or change focus or terms of service. Whilst the technical burden of maintaining ones own personal data store may be higher than delegating this to a third-party comes with different, but not insubstantial, risks.

Serious review of common practices and formal Web standards can smooth the path to interoperability between different instances of WOs and PWOs. Though the standards discussed in this article were designed from a decentralised Social Web perspective rather than a Web Observatories one, the overlap is clear. Implicit self-tracking that makes up a part of ordinary social media use can be supported in the move towards decentralisation, and privacy-preserving PWOs may start to exist as a side-effect. In my implementation, I did not address the use case from [289] of aggregation of data from a crowd to produce net benefit. This is also something that shared use of open standards, in particular for data representation, subscription and notification, could facilitate.

Anticipation of future needs is raised as a challenge by [289]. With regard to my changing motivations and day-to-day requirements, I did not set out with a specific list of things I wanted to record and design the system around that. Instead, I used a flexible schemaless architecture which meant that for any new type of data I decided to log I had a minimum of new engineering to do: build a client (or potentially hook up a sensor) to generate the data; and (optionally) create a template to view it.

Working with open standards for creating, updating and deleting data helped here too, as I had the option to look for pre-existing standards compliant clients to post to my data store. Conversely, others whose data stores use the server portion of the publishing protocols can use my clients with their own storage.

Discussion of Personal Data Store architectures tends to revolve around reducing fragmentation and prioritises gathering together all kinds of data in one place [265]. This results in *generalised* tools and interfaces, which try to make it possible to do a variety of things in one place. I argue that more effective and appealing applications are *specialised*: particularly good at doing one thing. Whilst the data store itself is generic, standards for decentralisation permit the decoupling of clients - both for creating and displaying data - which is perhaps the best of both worlds.

### 3.5.3. Data context and integrity

Web Observatories which collect data from centralised social networking sites may be vulnerable to ingesting misinformation, ranging from subconscious selective disclosure to deliberate acts of protest against privacy infringement (as seen in The Many Dimensions of Lying Online in Chapter 3). It is difficult to

say whether proliferation of *personal* WOs would mitigate this, but it becomes increasingly important to find ways to capture contextual information when data is recorded.

If Web Observatories begin to collect data from both decentralised and centralised services, it is natural to want to align the data so it can be combined into the same data set. However the source of the data cannot simply be discarded. The same type of data from different types of sources is not necessarily equivalent or directly comparable. Just as different centralised social media sites (and the communities and subcommunities within) have their own cultural norms and expectations, as well as technical constraints or affordances, individual personal data stores come with their own unique contextual information as well. In the decentralised case, the context for data logs may be more difficult to discern, as well as capture.

To take a concrete example: for researchers to find more meaning in the logs in my PWO they need to look at in the context of *all* of the systems I interact with. My system forwards text notes and longer articles to Twitter, which in turn drives traffic back to my site where all my other kinds of logs can be seen. Researchers studying my data would need to consider how my awareness of my Twitter audience (directly through known followers and extended through *their* audiences in the case of retweets, plus how it may be used by Twitter itself, ie. the **cascade**) impacts all of the content I post. This might be different for someone who shares their PWO content with a different social network, or not at all. We see evidence of the impact of the network on the posting decisions of an individual in Social Media Makers in Chapter 3.

The interface used to log the data can also have an impact. Whilst I use a Web form based user interface to check in to a place, someone else might generate exactly the same data automatically by the GPS on their phone, making less of a conscious effort to record their movements. Researchers studying how people use, for example, Swarm, have the advantage of knowing that the interface used by everyone is consistent as well as being able to directly explore it themselves. I anticipate that PWOS will be far more diverse, personalised to fit into their users' day-to-day lives, and possibly inaccessible to researchers.

### 3.5.4. Limitations

Using data from QS activities for scientific research comes under fire for a number of reasons: self-reported data may be unreliable or biased; context is often lost when quantifying qualitative data; and data collection is limited to individuals who are inclined to record details about their lives [283]. Though I am not using the *contents* of my data logs in this report, the same issues apply to my recall of the effects of logging.

Perhaps most importantly, though the length of the study is significant, my sample size is 1. As such, I make no claims about generalisability or conclusivity of the results. I am documenting these experiences as a starting point, to begin to explore Personal Web Observatories in depth, and to highlight areas for focus in the future; this is similar to [265], in which the authors conducted their research on *MyLifeBits* with large amount of a single person's data as well.

I invested a considerable amount of time in building new features, fixing bugs, and making small improvements and adjustments to my PWO. As the only user and only developer, I was first to notice if something wasn't working properly, and unhindered by conflicting interests or opinions when it came to making changes. This has advantages for rapid prototyping of a somewhat novel system, as well as commitment to the ideology of *dogfooding*: if I don't want to use a system I've built, why should anybody else? It similarly meant that features I lost interest in fell by the wayside, whilst features I used regularly were well-maintained, so I didn't invest time in developing unnecessary functionality as I might have done had I been designing this for others. Whilst use cases and UI improvements were constrained strictly to my personal needs, and discussion of features was limited to a handful of like-minded developers, this approach was nonetheless appropriate for the purposes of this analysis.

### 3.6. Personal datastores for Control, Customisability and Context

This section documents over a year of developing and *using* a PWO. By taking the role of subject, not only developer, I have come to a better understanding of the day-to-day impact of PWO technology on an individual. 'Small data' is of interest to the academic community in terms of understanding how individuals engage with their own self-tracking activities and outputs, as well as to highlight the diversity of contexts in which data is logged and how this can impact analysis of an aggregation.

I believe that Personal Web Observatories are more useful when intimately personalised for the individual user. Small scale, pluggable components may help to enable this, and using open standards to integrate components can help with designing specialised logging clients or data interpretation interfaces. This gives users more choice to pick and choose the tools they use with their personal data store, as well as making it easier to add and remove components as desired.

Less explored here, but still pertinent, is enabling individuals to manage the relationship between different slices of their personal data or logs, and having control over who sees what. Most people do not want their logs entirely public, and may want to present different combinations of information to different audiences at different times.

### 3.7. Next steps

This does not mark the end of my self-tracking experiment, but the beginning of the next phase. Over the next twelve months, I expect to add to my repository: data about mood and health (specifically headaches); the people I spend time with offline, and amount of time spent chatting online; and to add more detail about exercise, and to re-start tracking sleep. I will continue to self-track publicly, but experiment with different views over my data for others, so that visitors to my site from different contexts (eg. professional) are not overwhelmed with data, nor left with an impression of inappropriate oversharing. For more detail about one approach to achieve this, read on...

Finally, I want to reiterate that *providers* of PWOs have a great responsibility to the individuals whose data they host. For people who engage in frequent self-tracking, a disruption in service can provoke a minor existential crisis. This is pertinent to bear in mind for researchers who wish to provide systems in order to study their users, as well as product developers building new services for personal data storage.

## 4. Audience and self-presentation

### 4.1. Introduction

Conflict or power imbalance arises when one party imposes frontstage expectations on another party's backstage behaviour. For example:

My landlady recently complained to me about an experience she had with a short-term guest in my shared apartment. On the day of the guest's checkout, my landlady had arranged to come over and collect the keys and return the deposit at a certain time. She arrived one hour early, without letting the guest know. She was shocked to find dishes unwashed, jars left open, food on the counter, and the guest watching a movie on the sofa in her underwear. She told me this was "not a good guest", bemoaning the untidiness.

Little does she know that that is often what the apartment looks like when I am home alone and she is not due for a visit. Nonetheless, because I have learned her tendency to turn up unannounced, as well as her penchant for tidiness, I make more effort to keep on top of cleaning day-to-day.

The problem here is that the guest was engaging in backstage activities, because she did not anticipate being observed for another hour. The guest fully intended to have the apartment clean and ready at the agreed time, and was entitled to behave as she pleased (enjoying her vacation) until then. Similarly, my landlady has some right to enter the apartment she owns when she chooses, however she does *not* have a right to expect that everyone therein behave constantly as if she were already present. She fails to acknowledge that her guests are going to behave differently when she is not around, and is holding her guests at fault for her flawed expectations.

One part of engaging in appropriate self-presentation is understanding and accounting for the expectations of those present: our audience. Even when thrown in to unexpected scenarios, we usually have some chance to react and accommodate in face-to-face interactions. When we are unable to do so, unpleasant social situations arise.

Online, we present a face but cannot see our audience. We may have no reliable information about who are audience are or their expectations, or we may *imagine* a different audience to the one(s) we really have. Furthermore, multiple

audiences may access our single 'face', and we have no opportunity to adapt and change our presentation to suit their expectations.

In this section I describe a possible user interface which attempts to address the problem of accommodating audience expectations with our unified online face. It's called *face*: facilitate audience control of expectations. I explain it in terms of a layer that can be added on top of a basic personal data store, and for the purposes of the following mock-up, I assume that datastore to be *sloph* (the system from the previous section).

## 4.2. System design

The *face* system is essentially a series of stylesheets which can be applied to adapt the presentation of any data object or stream of content viewed using *sloph*, and a set of controls to adjust the types of content which show up at all. I came up with five dimensions which can be used to describe different aspects of myself, and along which every data object in the datastore can be rated. I can input these ratings using a custom post creation/editing application which implements the same protocol (ActivityPub) which I use for creating any kind of content in the system, so it can talk directly to my existing serverside publishing endpoint. The stylesheet(s) for a particular data object is determined by its combination of ratings. In addition, a set of controls are presented to a visitor to the site which allow them to rate how interested they are in each of the five dimensions. Adjusting this rating changes which data objects are visible, as well as the overall appearance of the homepage.

The default view is a fairly neutral representation of my online presence, which can be adjusted to give a more or less personal or professional view, and gear it towards particular topics of interest. This way, the visitor still may be confronted with content they feel is inappropriate or uninteresting, however they take responsibility for adjusting the controls to give these results. These controls introduce a collaborative approach to impression management; since I cannot react to an audience I don't know is there, maybe I can prompt my audience to give enough clues to the system that my online presence can react on my behalf.

### 4.2.1. Vocabulary and data integration

The terms I use for the ratings are published as a Linked Data ontology, available in RDF from <https://terms.rhiaro.co.uk/view#> (prefix: view). This is effectively an extension to the core ActivityStreams 2.0 vocabulary which I use for most of my data. These terms, somewhat flowery, are named to represent aspects of myself which I think are useful for people to distinguish my different 'faces' by. I expect these are fairly unique to me, and for this system to be applicable to other people, they would need to choose their own dimensions. Fortunately the decentralised publishing expectations of Linked Data make it possible for people to reuse existing terms and augment them with their own as they see fit.

- **banality**: describes the more mundane things I record, like what I had for breakfast or how long I was in a cafe for.
- **intimacy**: describes posts of a personal nature, feelings, problems, hopes and dreams.
- **tastiness**: for food-related data objects, like meal logs and recipes.
- **informative**: for more formal academic output, technical comments or tutorials.
- **wanderlust**: for travel-related data objects.

To give you an idea of how different combinations of ratings along each of these axes line up: A post containing the lyrics to a song which is in my head would rate highly for **banality** and **intimacy** and low along other axes, unless the song happens to be food, travel or work-related. A restaurant review in an unfamiliar city would rate highly for **tastiness** and **wanderlust**. A complaint about bedbugs in a hostel has **intimacy** and **wanderlust**. Expressing my frustration at progress with my thesis is **intimacy** and **informative**, and a summary of a talk I did in a new city might be **informative** and **wanderlust**. My food logs are **banality**, **tastiness** and somewhat **intimacy**.

Each of these terms is a predicate, the value of which is an integer between 0 and 5, where 0 means 'contains no content of this nature' and 5 means 'is strictly only interesting to people who want content of this nature'.

I also use the following terms from the W3C Annotations Vocabulary [240] (prefix: oa):

- **CssStyle**: a class for stylesheets.
- **styleClass**: indicates which CSS classes should be applied to this resource when it is rendered.

- **styledBy:** indicates a stylesheet that should be used when rendering this resource.

A series of stylesheets are specialised to different combinations of ratings. These stylesheets exist in my triplestore as `oa:CssStyle` typed objects, and each is also associated with a rating along each of the face dimensions, to indicate the types of content it is most appropriate for, ie.:

```
</css/style.css> view:banality "5" .
```

Some examples in the form `style-name [banality, intimacy, tastiness, informative, wanderlust]:`

```
food [5,3,5,0,0] // Meal logs
lyric [5,5,0,0,0] // Posts which are just song lyrics
wg [0,0,0,0,4] // W3C Working Group related
phd [0,1,0,0,5] // PhD related posts
trek [4,4,0,0,0] // Posts about Star Trek
checkin [5,3,0,4,0] // Posts which announce my location
feels [0,5,0,0,1] // Posts about emotions or feelings
banal [3,0,0,0,0] // Boring posts
intimate [0,3,0,0,0] // Intimate posts
3 tasty [0,0,5,0,0] // Posts about food
1 wander [0,0,0,0,3] // Posts about travel
2 travel [3,5,0,0,5] // Travel plans and specific schedules
```

When a resource is rendered, the values for the face dimensions of the resource are compared with the values of the face dimensions for each stylesheet available, and the stylesheet which most appropriately matches the set of ratings is attached.

This way, if someone comes across an individual post out of context (eg. through seeing the results from a search engine) it is displayed in a default way which suits it the most. For example, a recipe might look like a post from a normal food blog.

#### *4.2.2. User interface*

The minimum viable interface to present to visitors to my homepage is a series of sliders, one for each dimension, set at neutral defaults. Visitors can move the sliders to increase or reduce the appearance of different types of posts. Turning

a slider up to "5" means "show me all the posts which have a value of at most 5 for this dimension." For example:

- My Mum: as perhaps the only person in the world with a genuine interest in what I had for breakfast, will probably crank everything up to 5.
- My PhD supervisor, wanting to see my latest informal thoughts about my topic: might look only for **informative** posts, and increase **intimacy** to see more heartfelt posts such as rants or complaints about technology.
- My PhD supervisor, wanting to know why I'm not replying to emails or making code commits: can increase **banality**, **tastiness** and **wanderlust** to see if I have been spending my time eating, exploring and blogging about it instead of working.
- A potential employer: may be interested in **informative** posts, but also where I am in the world through **wanderlust**.
- Someone interested in vegan food: a combination of 5 for both **banality** and **tastiness** will reveal both what I eat every day, and recipes and restaurant reviews.
- New friends I make whilst traveling: can maximise **wanderlust** and increase **intimacy** for a personal take on my current adventure.

I can configure preset options for common types of views people might want. To take it one step further, particularly useful preset views can have their own domain names, which are simply a particular view on a particular feed of posts, and I can give the most suitable of these URLs out to people I meet offline who want to track my activities online. (Specific examples already on the cards are a food blog, [whatdoveganseveneat.com](http://whatdoveganseveneat.com), and a travel blog, [homeiswherethehammockis.com](http://homeiswherethehammockis.com)).

I can also tailor the defaults based on for example, the referrer (did they click the link from Twitter) or the physical location in the world from where the traffic is coming. Similarly, if certain people I know are able to authenticate with sloph, actively making themselves known, I can default to the settings I'd prefer them to use.

Changing the default baseline depending on my current mood can give visitor to my site an immediate impression of how I'm feeling - perhaps more focussed or emotional (or hungry) - just as my facial expression and body language might give off this impression were someone to encounter me in person.

#### 4.2.3. *Limitations*

Obviously this does not provide a solution for strict access control, privacy protection, or any kind of concealment or separation of online personas. To hide content altogether I would still need some kind of authorisation flow, or to refrain from posting it entirely. My various personas are fluid and flexible, however they are clearly linked together.

### 4.3. Discussion

I will briefly discuss this system in terms of the 5 Cs framework outlined in chapter 3.

#### 4.3.1. *Connectivity and cascade*

The audience of my online presence is determined by both my **connections** and the **cascade**. In the case of sloph, I haven't implemented a subscription protocol so I do not push content out to people, and nor do I have an notion of friends or followers built in. I do cross-post content to other networks however, such as Twitter. Once content leaves my system, the audience I imagine I have is even less convincing; the likelihood of someone I do not know stumbling across my content and making their way back to my site increases. This is why it is particularly important to attempt to negotiate my impression management with visitors, lest someone on the trail of a technical blog post come across my latest opinions about Star Trek Voyager or vegan cheesecake, and leave in confusion.

#### 4.3.2. *Control and customisability*

This system does not provide a way for me to limit the connections between the different online personas it allows me to present, which is an element of **control**. However, I am able to greatly alter the appearance of different kinds of information, once the designs are developed, simply by adding ratings to each resource I create. This is important for the **customisability** of my online presence overall. If someone is looking for a food blog, and they express that through the input options I provide, then what they'll get is a food blog that both contains the right kind of content and is visualised appropriately.

#### 4.3.3. *Context*

Meaning is given to my data according to the **context** in which it is both produced and observed. The closest the observer has to understanding any of the context in which the data was produced is to look at other data logged

immediately before or after the post of interest. Of course, these may be filtered out by their preferences, and I do not capture everything necessary to provide an accurate personal context (yet). At present most of my posts display the application with which they were created (though the observer is not necessarily familiar with it), but they are missing which other networks I may have cross-posted to.

Given that my data is available in a standard, machine-readable format, it may also be consumed by other applications and re-displayed, potentially removing it from the context provided by my own system. As the face dimensions are a vocabulary of my own design, I cannot assume that other systems will understand and make use of them in any way.

What my audience member provides me of their context is through their interaction with the slider controls. This is neither precise nor accurate, but perhaps nonetheless better than nothing.

## 5. Conclusions

In this chapter I have described in detail key elements of standards development within the W3C Social Web Working Group, and then discussed my personal experiences of developing systems based on these standards.

The standardisation process and its participants are part of the background **context** for any system which are based on these standards in the future, as the protocols many influence user interfaces and interaction models, which in turn impacts end users. The protocols themselves provide key building blocks for **customisability** and **connectivity**, as they describe ways for different types of content to be created, and for individuals to connect to one another and propagate content around a decentralised network. The implementation work highlights **context** which is closer to the point of data production, including the different kinds of applications which can generate the data, and the individual's mindset whilst doing so.

The experimental content display system, *face*, ties together **connectivity** and **customisability** by creating an environment in which the subject and their audience can collaborate in forming an appropriate self-expression, despite neither being co-present.

This chapter outlines a journey from theoretical ideas about how to build systems, to working software, by way of a consensus-based formal standardisation process with many players and stakeholders, as well as personal experimentation and reflection. Many Social Web systems have their own stories which can be traced along these lines, though rarely are they told in completeness. This chapter contains only my perspective, and only one software implementation, when there are many others which likely differ.

In the next and final chapter, I will bring together my contributions throughout this thesis, and reflect on the research questions I asked at the beginning.

# Chapter 6

## Conclusions

## 1. Summary

The goal of this thesis was to reach an improved understanding of how people present themselves online, and how this may evolve in the future. I carried out five empirical studies which bring to light diverse identity behaviours in different types of social system. Against my results and a backdrop of existing social science studies, I developed the 5 Cs conceptual framework which can be used to organise ideas when studying representations of individuals in networked publics. I also produced the Social Web Protocols, a primer for the cutting edge formal standards work for the Social Web which is taking place at the W3C, demonstrated prototype implementations of systems which use these standards, and discuss how systems like this affect self-presentation behaviours.

In this chapter I review the research questions which I set out in chapter 1, and summarise my contributions to the field. I wrap up with some still unanswered questions, and new questions which have arisen through my work, as well as some suggestions for directions to take with future research in this area.

## 2. Review of research questions

### **How can we access the bigger picture when it comes to understanding the impact of networked publics on presentation of self?**

I provide a conceptual framework, the 5 Cs, which includes high level interconnected concepts which are critical to any studies of online self-presentation: **control**, **customisability**, **connectivity**, **cascade**, and **context**. Whilst it is difficult for studies to take into account the huge number of factors which influence presentation of self, we have seen that most don't even acknowledge the bigger picture, tending to focus on technical affordances or particular actions or feelings of people. We must acknowledge that people and technology affect each other in a continuous cycle, and are both similarly affected by - and affect in turn - external factors, social, political, and economic. Table 16 shows the studies reviewed in chapter 2 about Social Network Sites, blogging and privacy grouped according to which of the 5Cs are acknowledged in their background, results or analysis. None of them examine *all* of the 5Cs.

Table 1. Table 16. A set of studies reviewed in chapter 2, grouped by which of the 5 Cs they discuss.

| Concept         | SNS studies  | Blogging studies                               | Privacy studies                               |
|-----------------|--|--|---|
| Context         | [17] [39] [113] [34] [41] [190] [40] [270] [220] [172] [85] [36] [315] [104] [293] [99] [180] [252] [71] | [79] [34]                                      | [5] [106] [198] [105] [132] [209] [212] [213] |
| Control         | [113] [41] [33] [98] [34] [257] [85] [104]   | [79]   | [78] [279] [212] [213]                        |
| Customisability | [113] [14] [40] [269] [70] [285] [173] [246] [110] [270] [220] [172] [257] [33]                          | [236] [189] [79] [274] [103] [291] [219] [244] | n/a   |
| Connectivity    | [33] [98] [39] [190] [85] [315] [104] [293] [79] [99] [41] [192] [180] [71]                              |  | [78] [279] [212] [213]                        |
| Cascade         | [98] [34] [41] [39] [250] [94] [220] [85] [315] [192] [180] [303] [74] [95]                              | n/a  | [5] [106] [198] [105] [132] [209] [42]        |

**Control** and **customisability** for the most part relate to what technical systems enable people to do or prevent them from doing. **Control** additionally calls out the multi-faceted nature of identity, so we are forced to consider the relationships between people's different and potentially disconnected versions of themselves. **Connectivity** and **cascade** emphasise that self-presentation is

not a solo activity. We are influenced by audiences, real, imaginary, seen and unseen. The **cascade** requires us to acknowledge some of the important differences between offline publics and networked publics, including how our personal data is collected, stored, and used by unknown third parties, and how it may persist for long periods of time, and possibly be accessed in vastly different ways than we imagined when we created it. Finally, **context** captures the overall backdrop against which our online identity performances take place. It is imperative to consider, or at least document, the cultural, political, and economic surroundings of individuals; a lack of geographical bounds for online activities does not make these disappear. By no means least, personal feelings, motivations, desires, aspirations and troubles, as well as access to particular technologies and attitudes toward them, make up the immediate **context** of every single person. It may be impossible to know these details, yet these details may be at the core of explaining particular online behaviours, and it is easy to lose sight of this when analysing data about collectives.

Deeper work on **control**, **customisability** and **connectivity** stems from What is a profile?, whence I conclude that what constitutes a profile is context-dependent and varies with individual and community needs. This study focuses on technical platforms, but accommodates the perspectives on the platforms of users, developers, and the broader social landscape. Social Network Sites provide varying facilities to build profiles, from asking for explicit input from the profile owner, to a feed of owner- and other-generated content, to automated output from the system itself. Systems may be judged along dimensions of: flexibility; access control; portability; representation; and prominence. The combination of features can determine how much authority, in terms of **control** and **customisability**, a profile owner has over their online self-presentation in that particular system. This study provides a framework for describing technical systems as a backdrop for studies which focus primarily on the people using the systems.

A detailed description of the process of standards-making for the Social Web in Standards and self-presentation also provides **contextual** backdrop for any future work which studies systems which are built on top of the protocols produced by the W3C Social Web Working Group.

**How does self-presentation change depending on the power dynamics of the Social Web services they use?**

When software systems do not do what people need, and there are few realistic alternatives, people innovate. They find new ways to work around the constraints they face. Much mainstream Social Web software today is designed around representing and connecting 'real people', but without taking into account the nuances of what a real person actually is. In Constructing Online Identity and The Many Dimensions of Lying Online, we can see a wide variety of motivations for bending the truth on the Social Web, most of which are not malicious or even truly dishonest. We see people adapting to their audiences, and selectively disclosing parts of their identity in order to protect their own wellbeing. This is all part of how people **control** their self-presentation(s) online. In Turkle's early (mid-1990s) research into identity in digital spaces, she optimistically proclaimed these new media to be fora for creative expression, where people can safely explore to find their true selves. Turkle's view on technology may have soured since then, but we are still seeing playful, artistic, and empathic behaviours on the Social Web, much of which tends to take place outside of the rules and regulations of the underlying systems (which themselves provide important **context** against which to understand uses of the systems). This is a key insight that informs our understanding of what it is to be a social entity within rigid technological limitations, and something to bear in mind before we make assumptions about what people use a system for purely on the basis of the system's affordances or what it claims to be for.

An important takeaway from Computationally Mediated Pro-Social Deception is that if Social Network Sites demanded *less* of their users, people would be inclined to entrust them with *more*. People's perceptions of the **cascade** vary; in many cases either they know enough about it to not want to share, or they know they don't understand it, and are too suspicious to share. A second is that a world which is at all times revealing and accurate is not necessarily a social or humane world. This seems obvious from a social perspective, but we see Social Network Sites are still baking in expectations of their users based on a flawed understanding of integrity or authenticity which does not leave people space to maintain their relationships (and sanity) through the different levels of mild deception which are second nature in offline interactions.

My study of Social Media Makers is the first of its kind; I engaged with individuals who are building and using decentralised Social Web systems as alternatives and augmentations to mainstream centralised systems. These individuals, mostly starting from scratch, have completely changed the power dynamic by taking ownership of their personal data, moulding it to their needs, and selectively sharing this across different networked publics. Their responses

demonstrate the importance of flexible functionality for self-expression - **customisability** - as well as a desire to break away from the cultural norms which have arisen around centralised Social Network Sites. Nonetheless, they are still affected by outside systems thanks to their desire for **connectivity** to existing networks, and the **cascade** which results from this. There is a sense that once diversity in online self-presentation is the norm, people will be more able to be themselves on the Social Web.

### **What can developers do to adapt to or accommodate self-presentation needs of individuals?**

By following Web standards when designing new social systems (specifically the W3C Social Web Working Group ones in this case), developers can accommodate people who want to move between or spread themselves across different systems, without creating the burden of lock-in that we see with proprietary systems today. This opens the door to extensibility and more specialised social systems which can rise to the challenge of accommodating increasingly niche communities and individual quirks. I provide some examples of where to start with this in Social Web Protocols.

In direct application of Goffman's dramaturgical theories, I argue that it is important to find new ways of helping people to engage with their audience, across space and time. *face* is one possible approach, which does not address privacy or information access concerns, but relies on the idea that an audience member is willing to play their part in the performance, is somewhat aware from their own experiences of the complexities of online self-presentation, and will participate in helping their expectations to be met. Just as we politely pretend not to see someone stumble as they enter the room, or accept without question the actions of characters in a play on stage.

## 2.1. Methodological Recommendations

Web Science must coordinate engineering with a social agenda, policy with technical constraints and possibilities, analysis with synthesis - it is inherently interdisciplinary. - [25]

A further outcome of my work has been to find ways to link research and methods from social science and computer science fields in a way which benefits both, and serves to further our understanding of socio-technical systems in general. I have leaned on Goffman's dramaturgy throughout, and

highlighted ways in which theories about face-to-face interactions can be applied to the Social Web, as well as places where these theories need to be extended or amended for digital spaces.

I also advocate exploring more auto-ethnographic style approaches to system design and development. Researchers can immerse themselves in the systems they are studying, and gain greater insight into the motivations and activities of other participants. At the very least, this should enable a more comprehensive description and review of technical features than an outside overview could provide, which in itself is valuable for situating any particular study.

In software engineering, using yourself the software you produce is often known as "dogfooding" or "eating one's own dogfood." This comes with the ethic that one shouldn't build systems for other people which aren't useful to you. More extreme advocates argue that if you *rely* on the software you are building you're also much more likely to resolve problems, and see gaps for missing features. Many large software development companies already employ this policy [309], and for one-person or small teams developing experimental platforms to explore new ideas I think academics would do well to adopt this approach as well. It has the potential to reduce speculative features and cement a greater level of commitment to a project. I demonstrated this with *sloph*, the personal social datastore I built and described in Personal data and self-presentation. In terms of the highly personal Social Web systems I expect us to be moving towards as this field progresses, I think it is particularly presumptuous for us to theorise around systems which we aren't willing to engage with directly ourselves.

In conclusion, I have advanced the field of Web Science through bridging interdisciplinary approaches, and propose a new mode of research approach when it comes to experimental software systems of a personal or social nature.

### 3. Things to come

This thesis covers a mere snapshot of a point in time, in the history of the Web and online social interactions. By the time you read this, anything labeled 'current' or 'modern' will probably be obsolete. Even in the last month of writing up, *Mastodon* gained sudden popularity, and continues to grow (in terms of new instances and total user accounts) at an exponential rate. Mastodon is an implementation of OStatus and alternative to GNU Social, decentralised microblogging platform designed to compete with the monolithic Twitter. If you were paying attention in Chapter 4, you know that GNU Social is the community takeover of StatusNet. The founder of StatusNet and the primary driver of the OStatus architecture, Evan Prodromou, is a chair of the Social Web Working Group. Following StatusNet, Evan worked to simplify OStatus with Pump.io [132], a popular instance of which includes *identica*. A modest community of developers maintain the primary Pump.io codebase and keep federated instances running. The Social Web Working Group's ActivityPub used to be called ActivityPump, and is an evolution of the Pump.io specification, with Evan's oversight. As such, the suddenly popular Mastodon codebase is not one but two generations of protocol behind, if we assume Evan (and Social Web Working Group co-conspirators) know what they're talking about.

"I'm happy for Mastodon's success but disappointed they didn't use the modern protocol ActivityPub we developed at W3C. All I have to say." - Evan Prodromou, on various social platforms [Twitter  
<<https://mobile.twitter.com/evanpro/status/85115551325229058>>].

What followed were murmured complaints about the OStatus dependence on the *ancient* XML (developers these days supposedly prefer JSON). Some of us in the Working Group wondered where this will go next. Does its sudden popularity (at a scale not previously enjoyed by decentralised social efforts) mean that after all, the Social Web Working Group's efforts were for naught; OStatus was sufficient all along, it just needed the timing of the current political environment perhaps combined with the (comparatively) beautiful user interface that Mastodon provides? Or does it mean that thanks to its popularity, a flurry of open source developers will update the codebase to be compatible with ActivityPub, and the Working Group's work will see widespread success off the back of Mastodon after all?

Subsequently core Mastodon developers joined the Social Web Community Group (a less formal follow-on from a Working Group) and began raising issues and engaging in discussions with the ActivityPub specification authors and implementors. ActivityPub made small changes to the specification in response, and in September 2017, Mastodon announced <<https://hackernoon.com/mastodon-and-the-w3c-f75f376f422>> a release which uses ActivityPub mechanisms for key features, as well as an intention to deprecate OStatus in future versions.

Does this mean that other OStatus-based systems (such as PostActiv <<https://www.postactiv.com/>> and GNU Social <<https://gnu.io/social/>>) will follow Mastodon's example? Early to say, but at a minimum I hope for some level of bridging between the different protocols. After all, even if one particular combination of protocols is widely used there will always be use cases it doesn't quite meet, or developers who just don't really like the look of it. I think that an important indicator for the long-term potential of OStatus based systems will be whether we soon see implementations which do things *other* than Twitter-style microblogging. OStatus was designed with microblogging as its core use case in mind, but (as we have seen) there is a whole host of social software out there, from health tracking to service exchange. The extensibility-by-design of ActivityStreams 2.0 may be what gives ActivityPub an edge in terms of diversifying the possibilities of decentralised social systems. This diversity, as I have hinted at previously, brings a host new challenges of course.

### 3.1. Decentralisation considered harmful

Throughout this thesis I have assumed that decentralisation is a positive route forward for empowering individuals through Web technologies. I have done little to reflect on the new problems that arise with this kind of architecture. Here I outline a few areas where decentralisation might cause *new* issues or make things worse. I am barely scratching the surface here, and finding and solving the social problems associated with decentralised Social Web technologies is an important direction for future research.

**Smaller attack surfaces:** Large centralised systems have robust network architectures, and plenty of resources to keep things running under duress, or to recover from attacks. Many decentralised architectures imagine smaller 'pods', independent servers which federate. It's possible many of these servers will be run by volunteers, hobbyists, or small organisations, and could be easily taken down and kept down by malicious actors.

**Quieter takedowns:** We want it to be easier for small communities, perhaps vulnerable minorities, to create safe spaces in their own corner of the Web, and to be able to keep out those who jeopardise that. If these communities are 'disappeared' (perhaps made easier by the previous point) the rest of the Web might not notice until it's too late.

**Illusion of control 1:** We promote decentralisation as a way to **customise** who has access to your personal or social data, and to be able to move it somewhere else if you want. But a key part of decentralisation is federation, or enabling access to your data by other systems, ie. so that you and your friends can use a different applications for the same thing, without that getting in the way of your interactions. This involves open data formats and standard APIs and likely complex access control setups. We already see that people have difficulty managing their Facebook privacy settings, and these are for a single unified system. On top of that, not only must you trust the server where you host your data to correctly enforce access controls, depending on the architecture, there can be serious **connectivity** implications; you may need to trust your friends' servers, and their friends' too, as blobs of your information are passed through the network. Just because you *could* move your data to a different service, doesn't mean it's safe where it is.

**Illusion of control 2:** When I log and publish data about myself with my homemade personal datastore (described in chapter 5), I feel like I have more control over my expression given off. I provide data on my own terms, and I know that *my* software is not drawing inferences or aggregating my data with others in order to learn more about me than I'm sharing explicitly. However, related to the previous point, my data is all public and machine readable, using open standards; there's nothing to stop Facebook from connecting the dots and consuming this data about me as well, so the **cascade** is still present. If social media has normalised dangerous oversharing, and the general populace is starting to realise how their data is being used and carrying out countermeasures, then decentralised social media runs the risk of convincing people their oversharing is 'safe' again, setting us back a decade.

**The filter bubble:** The easier we make it for people to avoid abuse online (not that decentralised systems have necessarily solved this yet), the easier we make it for people to filter out diverse points of view. Last year, Twitter introduced the ability to filter out certain phrases from one's timeline. An immediate reaction from privileged Twitter users (people who have never been flooded by abusive posts) was to decry the new filter bubble this could create. If filtering abuse is

directly at odds with exposing people to different worldviews we might be suppressing one of the core potentials of the Web. At the very least, people need to be able to choose how selective to be with what they consume, whilst being made aware of the potential consequences.

**Lost in translation:** We have to assume that protocols used in decentralised social systems will at some point not meet all of the needs of system designers and users, and will need to be extended. Indeed, ActivityPub and other Social Web Working Group specifications are designed with extensibility in mind. As different implementations extend the core protocols in different ways, whilst continuing to federate with each other, mismatches will start to occur. One example is Mastodon's *content warning* feature, which allows people to hide the majority of a post behind a small label (from NSWF to politics to TV spoilers), so that receivers can opt into reading the content, or skim past if they'd rather not. Friendica rushed to implement this <<https://github.com/friendica/friendica/issues/3285>> so that their users wouldn't be negatively affected by inappropriate posts from Mastodon instances which the creators thought would be tactfully hidden, and there was also concern about how this would play with ActivityPub integration <<https://mastodon.social/users/Gargron/updates/3244985>> (a future extension <<https://github.com/w3c/activitypub/issues/231>> to ActivityPub was eventually agreed). As implementations and extensions diversify, we run the risk of content being miscommunicated between systems as well as conflicts about how particular features are expected to behave.

### 3.2. Long live decentralisation

None of the problems in the previous section are things that will foil mass adoption of decentralised social Web technologies. They're just things that may make this type of technology more harmful by not being resolved if there *is* adoption. So what will happen next?

I cannot herald a golden age, a revolution in human expressiveness. I unexcitedly predict gradual changes, in fits and starts, with intermittent controversy, but little fanfare. Even as social media becomes central to Western everyday lives, we are simultaneously turning away from it. In truth, far more people do not use the social layer of the Web than those who do, globally.

Over the next few years, distaste for social media will peak. Those with the resources to do so will pay for less obnoxious experiences. Those fortunate enough to not live in isolation or want for community support, will disconnect. The Web will drift into the background of their lives; something to check in an emergency, or on a special occasion. Their identities will recentre in their physicality and moments will pass unrecorded. They will nonetheless be accommodated in society, because they are already a privileged group.

For all of the people for whom social media was a detriment, there are perhaps a great many more whose lives have been vastly (or even slightly) improved. People who can't or won't disconnect even if they want to. Service providers will eventually realise that Stockholm Syndrome is not the most effective way of retaining users, and will become gentler and more outwardly respectful than they are now. Today people switch contexts by switching applications; over the next few years, service providers will stop fighting this, and work to make it more seamless. In understanding that we live life in different modes, interface for particular tasks will become more specialised and more personalised, and for someone to put on the right face at the right time will become second nature digitally, as it is in-person.

To enable this, service providers will share data under the hood. No longer competing to be the sole proprietor of an individuals' network and personhood (they never were, after all) systems will instead trade between each other, provide service integrations, and use this more complete contextual information to compete on utility and usability.

With each enormous data leak or trust-breaching design decision, a flurry of smaller more respectful services will emerge and gain adoption. More specialised applications will be more pluggable; their users benefit when they cooperate. As services share data under the hood, context switching becomes easier and smoother.

Data sharing deals will be exclusive, between the most popular and most successful platforms to begin with, but as people get more discerning, and as they get used to a landscape of better performing specialised applications with variety to choose from, the market will broaden. APIs for data exchange will stabilise and generalise. Maybe some of the standards of this decade will be picked up and adapted. Personal data consolidators will step in to broker between applications. Not as end-user products, but as services for service providers themselves. Everyone will have a context-aware personal data store

without knowing it. Personal data legislation in some parts of the world will make sure people *do* know it though. A side effect of service interoperability will be data portability. A small portion of people will take ownership of their accumulated online presence, but most won't.

As human beings, we will remain diverse. Acknowledging this diversity spawns *more* options for applications and services. But just like hundreds of brands in a grocery store which are all owned by two or three conglomerates lends an illusion of choice, the number of *providers* behind the options may not actually increase, just their offerings. But we will perfect our digital daily routine, tweak it so it suits our tastes, and it needn't look the same as that of anyone else in the world.

As ever, some people will register and check up on every step taken, every character typed. Some will knowingly or unknowingly log this data and pay it no mind. Every microsecond will be captured and some will be replayed, some will be lost forever on a harddrive. Portraits of your personas will be scary, accurate, and very convenient. They'll still get things wrong, and there will be things you just don't do over the network. It'll be harder and harder to get a device or use an application that doesn't already know some part of who you are - whether your personal data is in your own hands or in someone else's black box.

As much as I'd love to believe the data ownership revolution will be driven by those steadfastly building decentralised tooling right now, I think the history of fragmentation and infighting will have repercussions for years to come. Systems which only interoperate with other versions of themselves will have no place in a future of diverse tailored digital experiences, and neither will stand-alone systems which individuals have to set up and maintain all by themselves. Change will be driven by the big players, who must eventually realise that putting all of their users into matching boxes isn't actually in anyone's interests. Personalised, more diverse social applications which support context-switching may create a backdoor of data access through which those who care enough can claw back some ownership, but ultimately our online presence will still be largely out of our hands. What matters though, for most people, is the interface for expressing ourselves through our data; ownership alone is not empowering.

*(Alternatively, corporations and government will merge, personal data will be gathered and unified, and opting out will be at the cost of healthcare, jobs, and being recognised as human by passing self-driving cars).*

### 3.3. Future work

As is declared at the end of all theses but the most confident or self-important ones, there remains much to be done. I have demonstrated in the preceding chapters that there is a complex interplay between the personal, the social and the technical in considering the dynamics of the Social Web, and in engineering its future. I hope that one overarching impact of my work will be to help set a research agenda for better understanding the place of social media within the context of Web Science. The following list, albeit incomplete, highlights four key areas where I both recommend and expect to see future research.

**Cogs in the machine:** Social machines are to date largely studied in collective terms. It is important to acknowledge that this is one zoom level, which is a valid perspective, but I'd like to see comparisons between results of studies of the impression of a social machine as whole, with studies of individual participants. There is scope for work in differentiating identity performance as part of a crowd or collective with shared purpose, with how individuals in that crowd perform their identity on their own terms. When are these at odds and when are they complementary? How do reputation systems play into this? What affects the ability for applications to simultaneously be tools for individuals and tools for social coordination?

**Decentralisation and communities:** This thesis examined possible effects of decentralisation of the Social Web on individuals, but I have not looked at how this type of architecture will affect communities. I anticipate that distinct sub-communities will become more obvious in an instance-based architecture (that is, servers will host a particular community, which can nonetheless interact through federation protocols with other communities on other instances). One question to ask is how individuals will manage their participation in multiple communities in this case? Linkability of different identities may need to be carefully managed by individuals, and it is worth investigating possible user interfaces to help with this, to maximise convenience for individuals whilst minimising the effects of information leakage. In addition, interoperable protocols mean that people using diverse software implementations can interact. Certain things may need to be translated between systems, eg.: terminology; user interface design; cultures, norms and quirks; and features above and beyond what is specified in the official protocols. Right now, most Facebook users who see someone post a 'retweet' as a status update understand that this has likely been cross-posted automatically from Twitter, even if they are not Twitter users themselves. Decentralisation will enable an explosion of

different ideas around how to describe social interactions, and it will be impossible for an individual in one system to be aware of all of the others. One term may even be used in different ways by different communities. Just last week I called out to the ether from Mastodon to ask what the difference between a "bap" and a "boost" is. It turns out that they're the same thing, but the administrators of my chosen Mastodon instance, which is loosely cat-themed, has tweaked the UI so that "replies", "favorites" and "boosts" (like Twitter's retweets) are labelled "meows", "boops" and "baps". Anyone who saw my post from a different instance, unaware of the existence of "baps" likely had no idea what I was talking about; somebody replied to point out that "bap" is what the Scottish call bread rolls.

**Semantics of identity:** I am hesitant to suggest that it's feasible to model identity behaviours in terms of the formal semantics desired by Semantic Web advocates. However, a longitudinal study into emerging community descriptions, or folksonomies, of how people present themselves or how their different identities interact or intersect would be a useful step in terms of both better understanding presentation of self online on a theoretical level, as well as how we can engineer interfaces for decentralised systems which help rather than hinder individuals on the Social Web.

**The Web for the vulnerable:** It is important that we better understand the impact of decentralised social systems on minority communities and vulnerable people. Many safeguards developed as add-ons to centralised systems (such as shared or automated block lists) are themselves developed by the very people who need them the most. It is imperative that we (as scholars and developers) seek out diverse voices, listen to their needs and support their efforts without questioning their experiences, *and* that we put both research and development (funding, opportunities) into the hands of the people who are affected the most.

As for my personal continuation of this work: now I am addicted to life-logging with my personal data store, I expect that I will develop it further, and in particular improve the more social features, as well as experimenting with *face* for communicating with my audiences.

Stay tuned: I am rhiaro.co.uk <<https://rhiaro.co.uk>>.